Principles of RNA-based Gene Expression Control in *Vibrio cholerae*

Dissertation

Zur Erlangung des Doktorgrades der Naturwissenschaften
(Dr. rer. nat.)
an der Fakultät für Biologie der
Ludwig-Maximilians-Universität München

vorgelegt von
Mona Gräfin Hoyos

München, August 2020

Diese Dissertation wurde angefertigt
im Zeitraum von Februar 2016 bis August 2020
unter der Leitung von Prof. Dr. Kai Papenfort
an der Fakultät für Biologie, Department I
an der Ludwig-Maximilians-Universität München.

Erstgutachter:	Prof. Dr. Kai Papenfort
Zweitgutachterin:	Prof. Dr. Kirsten Jung
Tag der Abgabe:	24. August 2020
Tag der mündlichen Prüfung:	21. Dezember 2020

Eidesstattliche Erklärung

Ich versichere hiermit an Eides statt, dass diese Dissertation von mir selbstständig und ohne unerlaubte Hilfsmittel angefertigt wurde. Die vorliegende Dissertation wurde weder ganz noch teilweise einer anderen Prüfungskommission vorgelegt. Ich habe noch zu keinem früheren Zeitpunkt versucht, eine Dissertation einzureichen oder an einer Doktorprüfung teilzunehmen.

I declare that I have authored this thesis independently and that I have not used other than the declared sources. The present dissertation has not been presented to another examination board, neither entirely nor in parts. As well I declare that I have not submitted or defended a dissertation previously without success.

München, 24.08.2020

Mona Gräfin Hoyos

Contents

Abbreviations

asRNA	antisense RNA
BCM	bicyclomycin
CDS	coding sequence
CLASH	cross-linking, ligation, and sequencing of hybrids
CRISPR	clustered regularly interspaced short palindromic repeats
CRP	cAMP receptor protein
CS	cleavage site
FFL	feed-forward loop
GFP	green fluorescent protein
IGR	intergenic region
IPTG	Isopropyl beta-D-1-thiogalactopyranoside
kb	kilobase
miRNA	eukaryotic microRNA
mRNA	messenger RNA
NAR	negative autoregulation
nt	nucleotides
OMP	outer membrane protein
ORF	open reading frame
PNPase	polynucleotide phosphorylase
RBS	ribosome binding site
RIL-seq	RNA interaction by ligation and sequencing
RNA-seq	RNA sequencing
RNAP	RNA polymerase

Abbreviations

rRNA	ribosomal RNA
rut site	Rho utilization site
SD sequence	Shine-Dalgarno sequence
sRNA	small regulatory RNA
Term-seq	3' end-specific RNA-seq protocol
TF	transcription factor
TIER-seq	transiently inactivating an endoribonuclease followed by RNA-seq
TIR	translation initiation region
tRNA	transfer RNA
TSS	transcriptional start site
UTR	untranslated region

Publications and Contributions

Publications and Manuscripts Originating from this Thesis

Chapter 2:

Mona Hoyos, Michaela Huber, Konrad U. Förstner, and Kai Papenfort (2020). "Gene autoregulation by 3' UTR-derived bacterial small RNAs." *eLife* 9, e58836

Chapter 3:

Nikolai Peschek, Mona Hoyos, Roman Herzog, Konrad U. Förstner, and Kai Papenfort (2019). "A conserved RNA seed-pairing domain directs small RNA-mediated stress resistance in enterobacteria." *The EMBO Journal* 38.16, e101650

Not presented in this Thesis

Benjamin R. Wucher, Thomas M. Bartlett, Mona Hoyos, Kai Papenfort, Alexandre Persat, and Carey D Nadell (2019). "*Vibrio cholerae* filamentation promotes chitin surface attachment at the expense of competition in biofilms." *PNAS* 116.28, pp. 14216–14221

Manuscripts in preparation

Michaela Huber, Mona Hoyos, Anne Lippegaus, and Kai Papenfort. "A Qrr sponge modulates quorum sensing dynamics in *Vibrio cholerae*."

Mona Hoyos*, Kavyaa Venkat*, Liam Cassidy, Andreas Tholey, David Grainger, and Kai Papenfort. "The dual function regulator VcdR controls carbohydrate uptake and TCA cycle activity in *Vibrio cholerae*."

*Authors contributed equally

Contributions to Publications Presented in this Thesis

Chapter 2:

MHo and KP initialized and conceptualized the study. MHo constructed the majority of plasmids and strains and performed the majority of experiments and data analyses (Fig. 1A, Fig. 1-S1, Fig. 1-S4, Fig. 1-S5, Fig. 2A-C, Fig. 2-S1, Fig. 3B; D, Fig. 3-S1C-D, Fig. 3-S2C, Fig. 3-S3, Fig. 4, Fig. 4-S1, Fig. 5, Fig. 6, Fig. 7, Fig. 7-S1A-B, Fig. 7-S2, Fig. 8, Fig. 8-S1, Fig. 9). MHu contributed by plasmid and strain construction, by analyzing Hfq dependence of OppZ and CarZ through Co-IP and sRNA stability experiments, by measuring OppZ repression strength, by probing Northern blots and by initial work on CarZ regulation (Fig. 1S6, Fig. 2D, Fig. 2-S2, Fig. 3C, Fig. 5D, Fig. 7-S1C). MHo and MHu performed OppZ pulse expression and analyzed sequencing data (Fig. 3A, Fig. 3-S1A-B). KUF contributed by analyzing the TIER-seq data (Fig. 1B-D, Fig. 1-S2, Fig. 1-S3). MHo had technical assistance from Andreas Starick (Fig. 3-S2A-B, Fig. 7-S1C-D). MHo constructed the figures, KP and MHo wrote the manuscript.

Mona Gräfin Hoyos

Prof. Dr. Kai Papenfort

Chapter 3:

NP and KP initiated and conceptualized the study of MicV and VrrA. NP, MH and KP conceptualized the experiments employing the synthetic sRNA library. NP constructed the majority of plasmids and strains and performed the majority of experiments and data analyses on MicV and VrrA (Fig. 1, Fig. 2, Fig. 3, Fig. 4A; B; D, Fig. 7, Fig. EV1A-E, Fig. EV2, Appendix Fig. S1B-C, Appendix Fig. S2, Appendix Fig. S3). MH constructed and validated the synthetic sRNA library and performed the selection experiments, MH and NP analyzed and validated the resulting sequencing data (MH: Fig. 5B-C, Fig. EV4A-C, MH and NP: Fig. EV3, Fig. EV5A-B, NP: Fig. 5A, Fig. EV4D). NP and MH identified and validated OmpA as the key target for ethanol resistance (NP: Fig. 6D, Fig. EV5C-E, MH: Fig. 6C, Fig. EV4A, NP and MH: Fig. 6A; B). RH contributed by measuring PmicV promoter activity in E. coli and performing analyses of OMP composition in V. cholerae and E. coli cells carrying sRNA over-expression plasmids (Fig. 4C, Fig. EV1F, Appendix Fig. S1A). KUF contributed by analyzing the transcriptional start sites identified in (46) for sigma factor binding motifs. NP constructed the figures, KP, NP, MH, and RH wrote the manuscript. NP was assisted by the research students: Roman Herzog and Raphaela Götz.

Mona Gräfin Hoyos

Prof. Dr. Kai Papenfort

Summary

Post-transcriptional control of gene expression by small regulatory RNAs (sRNAs) is a widespread regulatory principle among bacteria. The sRNAs typically act in concert with RNA binding proteins such as the RNA chaperone Hfq to bind mRNA targets via imperfect base pairing. They affect translation initiation and/or transcript stability. Additionally, sRNAs can influence transcription termination of their targets or function indirectly as so-called sponges for other sRNAs. Regulation often involves the major endoribonuclease RNase E, which contributes to both sRNA biosynthesis and function.

In the first part of this thesis, we globally identified RNase E cleavage sites in the major human pathogen *Vibrio cholerae* by employing TIER-seq (transiently inactivating an endoribonuclease followed by RNA-seq). We validated the involvement of RNase E in the synthesis and maturation of several previously uncharacterized sRNAs. Two examples, OppZ and CarZ, were chosen for further study due to their unique regulatory mechanism. They are processed from the 3' untranslated regions (3' UTR) of the *oppABCDF* and *carAB* operons, respectively, and subsequently target mRNAs transcribed from the very same operons by binding to base pairing sites upstream of the second (*oppB*) or first (*carA*) cistrons. This leads to translational inhibition and triggers premature transcription termination by the termination factor Rho, thereby establishing an autoregulatory feedback loop involving both the protein-coding genes and the processed sRNAs. In the case of OppZ, the regulation is limited to the *oppBCDF* part of the operon in a discoordinate fashion due to the position of the OppZ base pairing site. This mechanism of target regulation by Opp and CarZ represents the first report of an RNA-based feedback regulation that does not rely on additional transcription factors.

The second study included in the thesis characterizes two sRNAs involved in the envelope stress response (ESR) of *V. cholerae*. Misfolded outer membrane proteins (OMPs) induce the σ^E-dependent transcriptional activation of the sRNAs MicV and VrrA, which reduce membrane stress by repressing the mRNAs of several OMPs and other abundant membrane protein. MicV and VrrA share a conserved seed region with their functionally analogous counterpart from *Escherichia coli*, RybB, indicating that this seed sequence might represent a universally functional RNA domain. To study the involvement of this seed domain in the ESR in an unbiased fashion, we constructed a complex library of artificial sRNAs and performed laboratory selection experiments under membrane-damaging conditions. We isolated the most highly enriched sRNA variants and indeed discovered a strong enrichment of the conserved seed-pairing domain. We were able to pinpoint the repression of *ompA* as the key factor responsible for the sRNA-mediated resistance

to ethanol-induced membrane damage.

Taken together, this thesis expanded the knowledge on the mechanisms of sRNA-dependent gene regulation by reporting a novel autoregulatory feedback loop. Additionally, it introduced a synthetic sRNA library as a tool to study complex microbial phenotypes and their underlying sRNA-target interactions.

Zusammenfassung

Post-transkriptionelle Kontrolle der Genexpression durch regulatorische kleine RNAs (sRNAs) ist ein weitverbreitetes regulatorisches Prinzip in Bakterien. Gewöhnlich binden sRNAs an ihre Ziel-mRNAs durch unperfekte Basenpaarung, meist in Zusammenarbeit mit RNA bindenden Proteinen wie Hfq. Dadurch kontrollieren sie die Initiation der Translation und/oder die Stabilität der mRNAs. Desweiteren können sRNAs auch die Termination der Transkription beeinflussen oder ihre Zielgene indirekt regulieren, in dem sie als sogenannte „sponges“ andere regulatorische sRNAs abfangen. Häufig ist auch die zentrale Endoribonuklease RNase E in die Regulation involviert, indem diese sowohl zur Biosynthese der sRNAs, als auch zu deren Funktionalität beiträgt.

Im ersten Teil dieser Dissertation wurden zunächst Schnittstellen für RNase E im Genom des bedeutenden humanpathogenen Bakteriums *Vibrio cholerae* untersucht. Für deren genomweite Identifizierung wurde das TIER-seq-Protokoll angewendet (transiently inactivating an endoribonuclease followed by RNA-seq). Die Bedeutung von RNase E für die Synthese und Prozessierung von sRNAs wurde anhand von mehreren bisher uncharakterisierten sRNAs bestätigt. Zwei dieser sRNAs, OppZ und CarZ, wurden aufgrund ihres einzigartigen Regulationsmechanismus eingehender analysiert. Beide sRNAs werden aus der 3‘ untranslatierten Region (3‘ UTR) ihres jeweiligen Operons (*oppABCDF-oppZ* bzw. *carA-carZ*) herausprozessiert und regulieren anschließend jeweils weitere mRNAs von genau diesen Operons, indem sie Basenpaarungen mit dem mRNA-Bereich vor dem zweiten (*oppB*) bzw. ersten (*carA*) Gen bilden. Dadurch inhibieren sie die Translation der folgenden Gene und induzieren die vorzeitige Termination der Transkription durch den Terminationsfaktor Rho. Es entsteht ein autoregulatorischer Feedback-Mechanismus, der sowohl die die proteinkodierenden Gene als auch die sRNAs selbst umfasst. Im Falle von OppZ ist diese Regulation, bedingt durch die Position der Bindestelle der sRNA, begrenzt auf den *oppBCDF* umfassenden Teil des Operons. Der für OppZ und CarZ beschriebene Mechanismus der Genregulation ist der erste Bericht über eine neuartige, RNA-basierte Autoregulation, die unabhängig von zusätzlichen Transkriptionsfaktoren agiert.

Die zweite in dieser Dissertation enthaltene Publikation charakterisiert zwei sRNAs als Teil der Membranstressreaktion von *V. cholerae*. Die Akkumulation von fehlgefalteten Protein der äußeren Membran (OMP) führt zur σ^E-abhängigen Aktivierung der Transkription von MicV und VrrA. Diese sRNAs wirken dem Membranstress entgegen, indem sie die Translation von OMPs und anderen abundanten Membranprotein inhibieren. MicV und VrrA haben einen konservierten Bereich (die sogenannte Seed-Region) gemeinsam mit der sRNA RybB, ihrem funktional analogen Gegenstück aus *Escherichia coli*. Diese Gemeinsamkeit deutet darauf hin, dass es sich bei der Seed-

Region um eine universell funktional RNA-Domäne handeln könnte. Um die Bedeutung dieser konservierten RNA-Domäne für die Membranstressreaktion zu untersuchen, wurde eine Bibliothek von synthetischen sRNAs konstruiert. Damit wurden Selektions-Experimente unter Membranstress durchgeführt und die am stärksten angereicherten sRNA-Varianten wurden isoliert. Dabei wurde in der Tat eine starke Selektion der zuvor beschriebenen, konservierten RNA-Domäne festgestellt. Die Repression der Translation eines bestimmten Membranproteins, *ompA*, wurde als zentraler Faktor für die sRNA-abhängige Resistenz gegen Ethanol-bedingten Membranstress identifiziert.

Zusammenfassend leistet diese Dissertation einen wertvollen Beitrag zum Forschungsstand über sRNA-abhängige Genregulation durch die Beschreibung eines neuen autoregulatorischen Feedbackmechanismus. Darüber hinaus wurde mit der synthetischen sRNA-Bibliothek ein neues Werkzeug vorgestellt, mit dem komplexe mikrobielle Phänotypen und die zugrunde liegenden sRNA-mRNA-Interaktionen untersucht werden können.

1 Introduction

1.1 Regulatory RNAs in bacteria

Our understanding of RNA and its importance has changed tremendously over the past decades. While initially considered as merely transferring information from DNA to protein, RNA is now assigned increasingly diverse regulatory functions. Similarly, bacterial gene expression had been described as almost exclusively regulated at the transcriptional level. Extensive research on bacterial regulatory RNAs has challenged this assumption and revealed a complex network of post-transcriptional control affecting basically all cellular processes [59, 348].

This thesis focuses specifically on the class of *trans*-encoded, base pairing small RNAs (referred to as sRNAs). There is a plethora of regulatory RNAs in bacteria with distinct modes of action that will not be discussed in this work, but are covered by several comprehensive reviews. These include among others housekeeping RNAs like RNase P [96, 98] or tmRNA [223], small RNAs acting by protein sequestration (like 6S or the CsrB/RsmZ family [346, 286]), antisense RNAs that are transcribed from the complementary strand of their targeted gene [179], *cis*-acting elements within mRNAs like riboswitches or thermosensors [45, 188], or CRISPR RNAs for the defense against foreign genetic elements [219].

The first example of a *trans*-acting, RNA-based regulator, MicF, was serendipitously discovered upstream of the *ompC* promoter sequence in 1984 [217]. It set the stage for a whole new field of research investigating the so far overlooked regulatory potential of bacterial RNAs. The sRNAs are commonly described as non-coding transcripts that interact with their *trans*-encoded target mRNAs through short and often imperfect base pairing to regulate their translation and/or stability [343]. They can form extensive regulatory networks, as most sRNAs regulate multiple targets and for example the *csgD* mRNA has been shown to be regulated by no less than seven different sRNAs [11, 176]. Bacterial sRNAs are often expressed under stress conditions and typically rely on protein partners such as the RNA chaperone Hfq and the endoribonuclease RNase E for their functionality [334]. But as there is no rule without an exception, this classic definition of sRNAs does not cover all studied regulators. The non-coding nature of sRNAs is one example for initial assumptions that had to be revised, as several regulators were reported to encode short peptides that may act independently from or in concert with the RNA function [33, 342]. Bacterial sRNAs are sometimes considered functional analogs to eukaryotic microRNAs (miRNAs), with whom they share some fundamental characteristics like their regulation via seed region-mediated base pairing

[121]. But while miRNAs are very uniform in their size and structure and always act in concert with a complex of protein partners, bacterial sRNAs come in lengths of of 50 to 500 nt with diverse secondary structures and not all of them require partners like Hfq [25, 343].

The majority of sRNA regulators have been studied in the Gram-negative enterobacterial model organisms *Escherichia coli* and *Salmonella typhimurium*, whereas the work presented in this thesis has been conducted in the major human pathogen *Vibrio cholerae*. While many underlying principles of post-transcriptional regulation are conserved throughout the bacterial tree, not all results will be fully transferable. For example, only a subset of the well-characterized sRNAs from *E. coli* is also known outside of the *Enterobacteriales*, as most sRNAs evolved after the split from the other γ-proteobacteria [258]. However, as this work focuses more on the general molecular mechanism of sRNA-based regulation rather than on the phenotypical characterization of individual regulators, most studied principles will apply equally well to the studied organism *V. cholerae*.

1.2 Various types of sRNAs in bacteria

1.2.1 sRNA-encoding loci

The first systematic, genome-wide searches in the early 2000s revealed the presence of dozens of new sRNA-encoding genes in *E. coli* [17, 284, 347, 68]. These studies were limited to intergenic regions (IGRs), which were at that time considered to be the only source for sRNA biogenesis. Candidate sRNAs were determined based on predictions of intergenic promoters and terminators or on phylogenetic conservation of RNA sequence or structure. These primary (unprocessed) transcripts from free-standing genes typically have well-defined promoters and terminators and are transcribed independently from their adjacent genes. However, in some cases the genetic context still provided some hints on their transcriptional control, as e.g. the sRNA GcvB is encoded adjacent to its transcriptional regulator GcvA [335]. Moreover, some well-defined sRNA promoters could be screened for the consensus motifs of known transcriptional regulators, revealing e.g. RybB as part of the σ^E regulon and indicating a functional connection [252].

Additional studies following shortly after the initial sRNA searches already identified abundant sRNAs derived from untranslated regions (UTRs) of mRNAs [163], giving rise to the concept of "parallel transcriptional output": protein-coding genes would produce both mRNAs and regulatory RNAs by sharing promoter or terminator sequences [340]. However, only the emergence of high-throughput, RNA-seq-based methods allowed the discovery of sRNAs in a more unbiased fashion compared to the initial, IGR-restricted searches and uncovered the vast extent of sRNAs derived from a variety of genomic locations (Fig. 1.1) [317, 306, 38, 251]. Especially 3' UTRs turned out to be a rich source for sRNAs biogenesis [61]. These 3' UTR-derived regulators depend on one of two pathways for their synthesis with large consequences for sRNA expression and the chemical nature of its 5' end [61]: sRNAs can either be transcribed from independent, mRNA-internal promoters (type I sRNAs like MicL [126]) or they are released from the longer mRNA transcript by endonucleolytic cleavage (type II sRNAs such as SroC [214]). Use of an independent

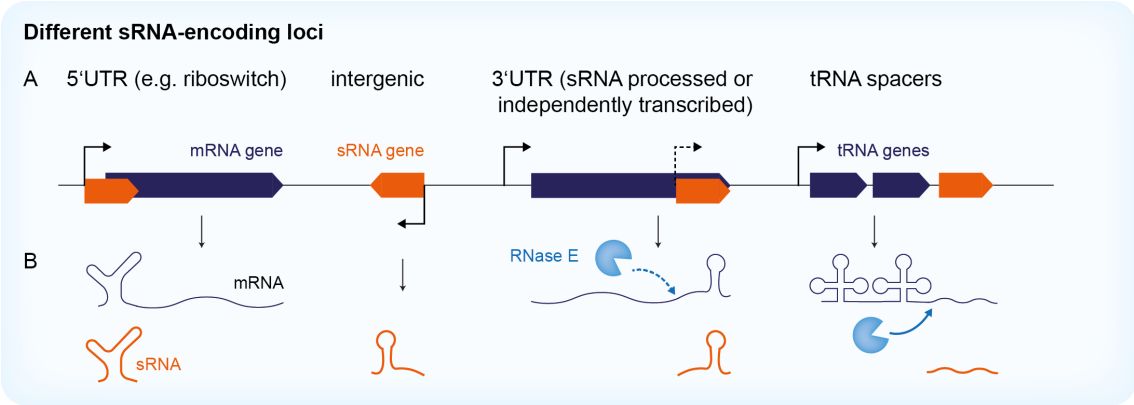


Figure 1.1: **Different sRNA-encoding loci.** (A) Schematic genomic organization of genes encoding mRNAs or tRNAs (dark blue) and sRNA genes (orange). Arrows denote transcriptional start sites. (B) RNA species transcribed from the genes in (A) with mRNAs and tRNAs in dark blue and regulatory RNAs in orange. RNase E or other ribonucleases (light blue) are important for the production of some sRNAs.

promoter uncouples sRNA and mRNA expression, thereby allowing independent regulation despite a potential overlap of the two genes, and equips the regulator with a triphosphate at its 5' end. In contrast, processed sRNAs carry a 5' monophosphate, depend on the upstream mRNA promoter for their expression and are often functionally connected to the corresponding regulatory networks [166, 62]. Either way, both mRNA and sRNA commonly share a Rho-independent terminator, which is often bound by Hfq to stabilize the released regulator [220, 242].

Fewer examples are known for 5' UTR-encoded sRNAs produced from the same promoter as their associated mRNAs [187, 350, 72, 209]. One example with an intriguing biosynthesis pathway is derived from a riboswitch in *Listeria monocytogenes* [187]: in its 'ON' state, the riboswitch allows transcription of the downstream operon, while the 'OFF' state causes premature transcription termination and accumulation of a short RNA species called SreA. This RNA, however, is not only a byproduct of riboswitch function but rather a *trans*-acting sRNA itself, which represses a virulence master regulator. Additional stable riboswitch-dependent transcripts have been detected in different species [163, 251], potentially encoding more functional regulators. While many bacterial sRNAs have been identified through Hfq cross-linking experiments, these methods might not be suitable for a comprehensive detection of sRNAs from 5' UTRs. Due to their overlap with the downstream-encoded parental mRNA transcripts, the known examples of 5' UTR-derived sRNAs lack the classic Rho-independent terminator that is involved in Hfq binding [242]. Consequently, most of them act independently of Hfq and cannot be captured by co-purification with Hfq. An exception is the recently reported *mgtC* leader that represses its target *in trans* in an Hfq-dependent fashion, similar to many classic sRNAs [72].

Finally, there are individual reports of sRNAs derived from other genomic locations. One of the most unusual sRNAs was reported to be excised from the external spacer of a polycistronic tRNA precursor. 3'ETS^{leuZ} is presumably constitutively expressed and binds to the two sRNAs RybB and RyhB. Most likely it inactivates the regulators generated by transcriptional noise and setting a threshold expression level for their stress response pathways [175]. Furthermore, a transposon-

derived sRNA regulates pathogenesis-related genes in *Salmonella* and may provide a selective advantage of transposon maintenance [97].

1.2.2 Control of sRNA expression

Apart from a few constitutively expressed exceptions like the housekeeping tmRNA, bacterial regulatory RNAs are made only under certain conditions or in response to distinct environmental signals [343]. They bear unique regulatory properties and are often important stress regulators (see section 1.4). Their expression can be regulated at the levels of transcription initiation and termination, by maturation from precursor transcripts, or modulation of sRNA stability.

Initiation of sRNA transcription

Many sRNAs encoded by free-standing genes within IGRs are transcriptionally controlled by alternative sigma factors [252, 107, 129], transcriptional activators and repressors [7, 198] or two-component systems [124, 222, 194]. Deciphering the inducing signals for the transcriptional regulators of an sRNA can give hints to its physiological function. Underlining their importance for the respective biological pathways, sRNAs can be found among the most strongly regulated genes within the corresponding regulons [230, 303].

This connection between expression conditions and functionality is more difficult to establish for sRNAs processed from untranslated regions, as they share the transcriptional control with their parental mRNA. Sometimes, mRNA and sRNA act as complementing arms of the same stress response, as it was described for the inner membrane stress-induced factors CpxP and CpxQ: the *cpxP* mRNA encodes a chaperone mediating the degradation of misfolded membrane proteins in the periplasm, while the CpxQ sRNA is cleaved from its 3' UTR and represses the *de novo* synthesis of the same proteins by binding to their mRNAs in the cytoplasm [62]. Thus, the simultaneous expression of both RNAs by the regulator CpxR results in a dual output within the Cpx stress pathway. In other cases, the connection is less clear, e.g. for the ProQ-dependent sRNA RaiZ processed from the *raiA* mRNA [321]. While *raiA* encodes a cold shock-inducible protein involved in ribosome inactivation, its 3' UTR-derived sRNA RaiZ acts to remodel the composition of the histone-like HU complexes. Is is not yet fully understood how *raiA-raiZ* expression is transcriptionally controlled or if processing of the sRNA always occurs with the same efficiency [321].

Although uncommon, some sRNAs actually do seem to be expressed constitutively, such as the tRNA-derived 3'ETS^{leuZ} described above [175]. Its continuous synthesis (at least under fast growth conditions) provides a constant pool of sponge RNAs that filter out transient bursts in RyhB or RybB expression. Thus, instead of being a stress-induced sRNA itself, it helps to modulate the response of other stress-related sRNA regulators.

Control of transcription termination

Transcription of sRNA genes typically stops at Rho-independent (intrinsic) terminators [186]. These are sequence-encoded elements consisting of a GC-rich palindrome followed by a T stretch (constituting a stem loop and a poly-U tail in the transcribed RNA), at which the RNA polymerase (RNAP) pauses and dissociates from the DNA [285]. The strength of the terminator hairpin and the length of the U stretch have been reported to determine the exact position of termination and thereby affect sRNA functionality [242, 227]. Intrinsic termination does not depend on additional proteins like the termination factor Rho (which dissociates RNAP from the DNA at so-called Rho-dependent terminators) and has long been considered a static process. However, recent results imply that sRNA termination is not always fully efficient, but can be remarkably increased under stress conditions, thereby enhancing sRNA production [228]. Intrinsic terminators not only act as transcription termination signals, but are also crucial for the ability of sRNAs to bind Hfq [152, 208]. Especially a stretch of at least six contiguous Us is involved in binding to the proximal face of Hfq [242, 297]. Nevertheless, some sRNAs that harbor a disrupted U stretch still regulate their targets in an Hfq-dependent way [276].

Maturation by ribonucleases

While the majority of IGR-derived sRNAs are produced in their functional form by transcription, some regulators undergo an additional maturation step [81, 254, 250]. The best-studied example is ArcZ from *Salmonella*, which is transcribed from its own promoter during aerobic growth, resulting in a full-length transcript of low abundance [194]. While it is unclear if this full-length ArcZ also fulfills a regulatory function in the cell, processing by RNase E generates a shorter form of the sRNA from the 3' end and places the conserved seed region at the 5' end of this shorter form. This maturation step is essential for the function of the short ArcZ species, most likely due to the increased accessibility of the now exposed seed region [60]. Similarly, MicX in *V. cholerae* is processed by RNase E from an sRNA precursor transcript into a shorter form that still includes the base pairing site. In this case, maturation is needed for sRNA stability, as the short MicX form exhibits a strongly increased half-life compared to the full-length precursor [81].

sRNA stability and turnover

Apart from its special role for MicX stability, cleavage by RNase E usually results in rapid turnover of the respective RNA [316]. Binding to RNA chaperones like Hfq or ProQ can protect sRNAs from both exo- and endoribonucleolytic decay [220, 10, 140]. Also features within the sRNA itself such as stem loops and other structured regions contribute to increased stability [9, 304, 262]. In contrast, mRNA binding often results in turnover of the sRNA together with its target [197, 266]. The GlmZ sRNA is even targeted for RNase E-dependent decay through interaction with its highly specific adapter protein RapZ [119]. More generally, sRNAs can be rapidly degraded in the absence of their targets, contributing to a 'proofreading' function that removes unpaired regulators [22]. Finally, accumulating studies report on sRNA sponges that bind to and destabilize

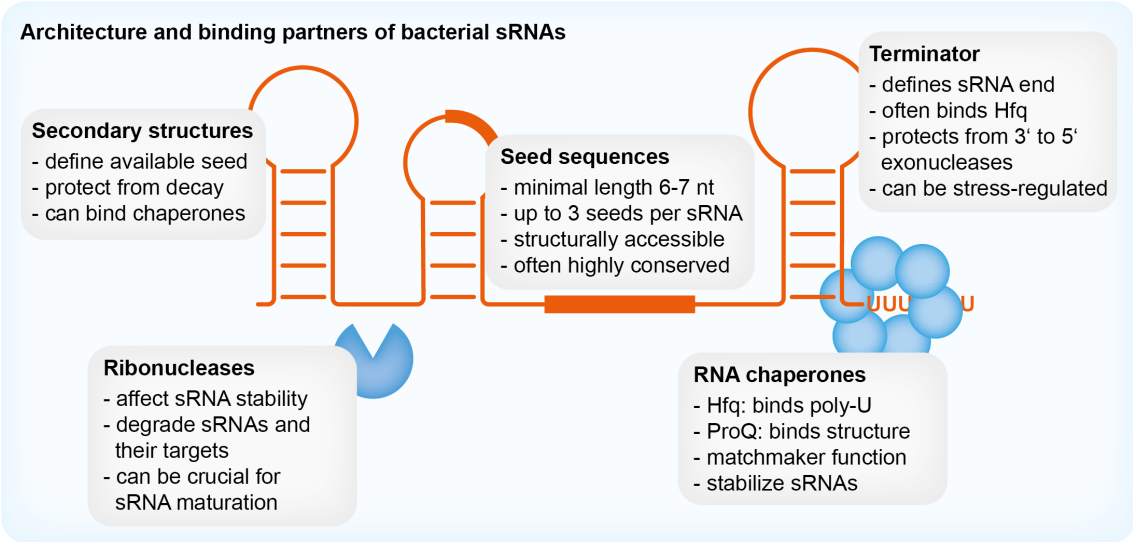


Figure 1.2: **Architecture and binding partners of bacterial sRNAs.** A prototypical sRNA (orange) consists of structured regions formed by stem loops and its terminator hairpin and of accessible seed sequences in single-stranded regions or loops. Interaction with accessory factors (blue) like ribonucleases and Hfq contribute to and modulate sRNA-mediated regulation. Additional key characteristics of the respective components are listed in the grey boxes.

their cognate sRNAs, often introducing cross-talk between different sRNA targets [214, 208]. In summary, sRNA stability is controlled by a plethora of different factors, which strongly influence the abundance of the active regulator.

1.3 Molecular Mechanisms of bacterial small RNAs

1.3.1 Components involved in sRNA-based regulation

Bacterial sRNAs need to meet several requirements to exert their designated functions within the cell. Generally speaking, they need to be expressed, stable and able to base pair and induce downstream effects on their targets. While the role of promoter and terminator sequences for sRNA expression and stability have been described above, the following chapter will introduce the properties of target-binding sequences within sRNAs, the importance of RNA chaperones such as Hfq, as well as the involvement of additional protein partners like the endoribonuclease RNase E (Fig 1.2).

Seed region

Base pairing RNAs need to identify their partners within a large pool of nucleic acids in a cell, reliably discriminating between true targets and non-specific interactions. To achieve this degree of specificity, regulatory RNAs have evolved to carry specialized subregions for target search and binding, referred to as “seed sequences” [333]. The general concept of seed pairing is wide-spread among regulatory RNAs, as bacterial sRNAs share this feature with both CRISPR RNAs and eukaryotic miRNAs and siRNAs [172, 121].

The minimal length of the essential seed region in bacterial sRNAs is commonly described as six to seven consecutive nucleotides, although the actual duplexes formed with targets *in vivo* might extend further including adjacent, probably non-essential base pairs [162, 21, 73]. Within this short stretch of complementarity, single mismatches can have drastic effects on target regulation and provide a basis to distinguish between targets [253, 289]. Increasing the seed length would likely reduce such discriminatory potential of individual positions, thereby increasing the risk of off-target effects [262, 121]. sRNAs do not necessarily regulate all their targets with the same seed region; up to three different target-binding sites have been described for a single sRNA [30, 135].

Identification of the seed region within an sRNA tremendously increases the performance of bioinformatic target predictions, since the search for potential RNA-RNA interactions can be restricted to the base pairing-competent regions within the regulator [165, 351]. Systematic analyses of up to 23 sRNAs in *E. coli* and *Salmonella* have revealed evolutionary conservation and structural accessibility as two key features of their seed regions [257, 283]. While conservation may be a consequence of various evolutionary constraints on both regulator and target side (reviewed by [333]), structural organization of many sRNAs indicates a division of labor between different parts of the regulator. Due to the limited chemical diversity of nucleic acids, RNA folding into secondary and tertiary structures can be important to define sequences involved in target regulation. Specificity is increased by sequestering sequences not relevant for base pairing in structurally inaccessible conformations, while exposing the seed sequences as unstructured regions or single-stranded loops [257, 121]. Additional studies show the importance of different sRNA parts for different functions such as target regulation, binding of RNA chaperones like Hfq or protection against endo- and exonucleolytic decay. This proposed modularity is also in agreement with the observation that seed regions transferred to unrelated sRNA scaffolds are sufficient to mediate target regulation by the chimeric RNAs [265, 249, 108].

RNA chaperones

The most common protein partner of bacterial sRNAs is the Sm/LSm-family protein Hfq, which is found in ~50% of all bacterial species and binds dozens of sRNAs [161, 144, 207]. It is often called a molecular matchmaker, as it provides a platform for the binding of sRNAs to their targets. Hfq fulfills many roles in sRNA-based regulation, including but not restricted to sRNA stabilization, melting of RNA structures to allow base pairing, increasing sRNA-mRNA duplex rate formation, or recruiting RNase E to induce target degradation [296, 293, 139, 334, 161]. Consistent with its central role in RNA metabolism, *hfq* deletions show pleiotropic phenotypes and Δ*hfq* cells are often more susceptible to environmental stresses [323]. Hfq has been shown to bind hundreds of mRNAs and sRNAs and co-immunoprecipitation experiments followed by RNA-seq have contributed strongly to the identification of new sRNA candidates in different bacterial species [61, 38, 133]. In accordance with this large target spectrum, Hfq is considered to be limited in the cell compared to the excess of binding partners and sRNAs compete for Hfq by active cycling [101, 147].

The homohexameric ring made out of six Hfq monomers can generally bind RNAs via four different sites [300]: its proximal site (binding the intrinsic terminators of most sRNAs [152]), its distal site (for binding A-rich sequences in mRNAs and some sRNAs [354]), the rim (providing a secondary binding site for UA-rich sequences [296]) and its unstructured C-terminal tail (important for some RNA contacts and hexamer stabilization [339]). While structural information about ternary Hfq-sRNA-mRNA complexes is limited, a recent study presented the first crystal structure of Hfq in conjunction with the sRNA RydC, proposing a model for the association of sRNA and mRNA to the chaperone [93]. Apparently, Hfq exposes the single-stranded seed region at the 5' end of RydC at its outer rim, near the conserved arginine residues that have been assigned an essential role in mediating base pairing [246]. It should be noted that in addition to its well-characterized matchmaker function, Hfq seems to play a variety of sRNA-independent roles in cellular processes, which are only beginning to be uncovered [294].

Hfq is the most common sRNA chaperone, but not all bacterial species encode an Hfq variant and not all sRNAs are Hfq-dependent. The recently discovered FinO-domain protein ProQ binds and stabilizes dozens of sRNA and has been shown to mediate target regulation in *trans* by the sRNA RaiZ [320, 321]. While Hfq preferably binds single-stranded, AU-rich sequences at the base of RNA hairpins [242, 152], ProQ seems to recognize RNAs by highly structured elements rather than by their primary sequence, but the detailed mechanism is still unknown [140]. Two other members of the FinO family, RocC and FinO itself, specifically associate with single sRNAs, while the vast majority of FinO-like proteins in diverse bacteria is still uncharacterized [19, 37]. Yet these results indicate that the FinO family might constitute a second major class of RNA chaperones [239].

RNase E and other ribonucleases

Alteration of mRNA stability is often part of sRNA-mediated gene regulation (see section 1.3.2). To this end, sRNAs in conjunction with Hfq can manipulate the access of RNases to their targets [177]. Typically, this blocks or induces cleavage by the major endoribonuclease RNase E [197, 58, 349], which will be briefly characterized below.

RNase E was initially discovered as the enzyme responsible for the processing of rRNA precursors [13]. It also mediates the crucial first step in tRNA maturation and is essential for cell viability under most growth conditions [244, 328]. Accordingly, the *rne* gene is present in ~80% of all sequenced bacterial genomes and organisms lacking RNase E like *Bacillus subtilis* typically encode functional homologs such as RNase J1/J2 [100, 5]. RNase E consists of a globular N-terminal domain harboring the catalytically active site and an unstructured C-terminal domain, which acts as a scaffold for a multi-protein complex called the degradosome [52, 53]. Preferred substrates for RNase E are single-stranded, AU-rich RNA regions with a degenerate consensus cleavage motif [60]. The central role of RNase E for sRNA-based regulation can be illustrated by some key findings from recent years: Hfq-bound sRNAs can form a ternary ribonucleoprotein complex with

RNase E, thereby destabilizing targeted mRNAs by locally increasing RNase E concentration or actively stimulating target cleavage through allosteric activation of RNase E [225, 22, 349]. The opposite effect, target stabilization, can be achieved by masking sensitive cleavage sites through sRNA-mRNA duplex formation [108, 255]. Moreover, accumulating reports of 3' UTR-derived sRNAs underline the importance of RNase E for sRNA biogenesis, while the majority of sRNAs is also turned over by the ribonuclease [61, 60, 86].

In addition to RNase E, other ribonucleases are involved in sRNA-based regulation as well. A central player in target inhibition by many sRNAs is RNase III, which specifically recognizes double-stranded RNA. Cleavage within the sRNA-mRNA duplexes degrades the target while simultaneously inactivating the sRNA, contributing to unique regulatory dynamics [197, 3, 146]. Additionally, the 3'-to-5' exonuclease polynucleotide phosphorylase (PNPase) is involved in sRNA base pairing, turnover and stabilization, most likely by interaction with Hfq and RNase E [84, 10, 23, 56]. Recently, PNPase was also shown to degrade short mRNA fragments that would otherwise sponge sRNAs by titrating them away from their true targets [57].

1.3.2 Mechanisms of target regulation

The standard mode of action for sRNAs has long been considered to be translational inhibition by direct competition for the ribosome binding site. However, sRNA studies from recent years have shown that there is a plethora of novel mechanisms to be characterized (and even more probably yet to be identified). Individual sRNAs can employ different mechanisms to regulate different targets (sometimes using the very same sRNA region for activation and repression, as described for SgrS; [39]) or can combine multiple modes for the control of a single target [268]. The following chapter provides an overview on the most common regulatory pathways and some atypical mechanisms, described along selected sRNA/*mRNA* examples (Fig. 1.3).

Control of translation initiation

Many well-characterized sRNAs regulate their mRNA targets by pairing close to the translation initiation region (TIR), thus inhibiting ribosome assembly at the Shine-Dalgarno (SD) sequence (Fig. 1.3A) [343]. Systematic analysis of the inhibition-competent region revealed that pairing from ~35 nt upstream to ~15 nt downstream of the AUG start codon (also called the "five codon window") allows sRNAs to block ribosome binding [43, 149]. However, inhibition of translation was also described for sRNAs pairing outside of this region. These can bind to other translational elements than the SD and start codon such as upstream enhancer elements (GcvB/*gltI* [305]) or ribosome standby sites (IstR1/*tisB* [79]). Repression can also occur indirectly, e.g. by the control of a translationally-coupled leader ORF (RyhB/*fur* [338]), by competition with the ribosomal S1 protein for binding of a translational enhancer element (SgrS/*manY* [20]), or by recruitment of Hfq to the TIR of the mRNA where the protein, not the sRNA, competes with initiating ribosomes (Spot 42/*sdhC* [89]).

In contrast to the above described examples, sRNAs may also positively influence translation

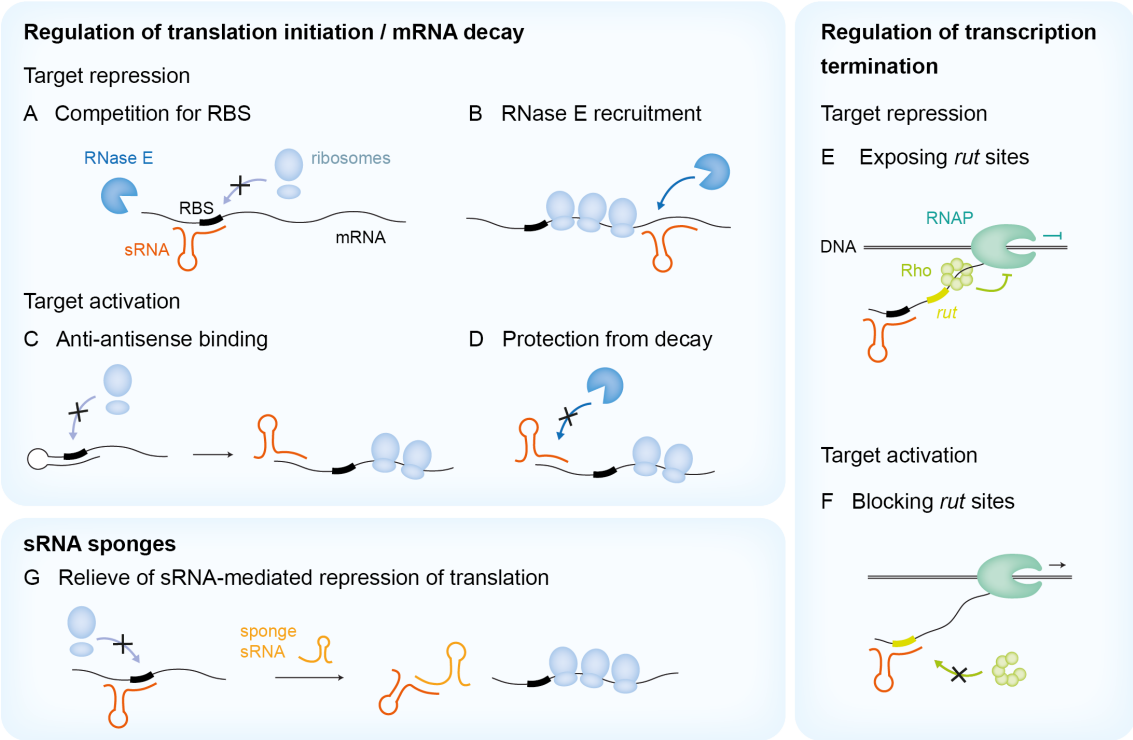


Figure 1.3: **Mechanisms of post-transcriptional gene regulation by sRNAs.** Most commonly, sRNAs repress their targets by blocking the ribosome binding site (A) and/or inducing mRNA decay (B). Increased mRNA translation and/or stability can be achieved by opening self-inhibitory structures in the mRNA (C) or protecting the target from ribonucleases (D). At the level of transcription termination, sRNAs can repress targets by allowing access of termination factor Rho to *rut* sites (E) or activate gene expression by blocking *rut* sites (F). sRNA-sRNA cross-talk through sponges can reduce the strength of mRNA regulation (G).

initiation. This typically involves an anti-antisense mechanism, where a long 5' UTR folds into a self-inhibitory structure sequestering the ribosome binding site (RBS) in a stable stem-loop structure (Fig. 1.3C). sRNA pairing to an upstream target site leads to remodeling of the mRNA structure and allows ribosome access to the TIR, as it was first reported for RNAIII/*hla* in *Staphylococcus aureus* [224] and DsrA/*rpoS* in *E. coli* [191].

Target degradation or stabilization as secondary effects

Translation and mRNA stability are often closely connected in bacteria due to the coupling of transcription and translation [18, 44, 87]. Accordingly, sRNA-mediated repression of translation is typically followed by rapid mRNA decay, as the "naked" transcript is no longer protected from endonucleolytic attacks by translating ribosomes [197, 225, 252]. Conversely, increased translation of sRNA targets is associated with mRNA stabilization [269, 205]. This secondary effect of altered target RNA stability can be important for sRNA function if target regulation is dependent on the combined effect of translational inhibition and RNA degradation [70]. In other cases, mRNA decay is dispensable for target repression, as studied in detail for SgrS/*ptsG* and RyhB/*sodB* [226].

Translation-independent regulation of mRNA stability

Changes in target stability are not necessarily a secondary consequence of translational deregulation, as multiple sRNAs have been shown to directly affect mRNA stability without altering translation initiation (Fig. 1.3B). One prominent example is MicC, the first sRNA discovered to repress its target by binding deep in the coding sequence(CDS) and thus far downstream of the TIR [265]. Mechanistically, MicC pairing to *ompD* promotes mRNA decay by recruiting the degradosome and stimulating RNase E activity through the sRNA's monophosphorylated 5' end [22]. Four different sRNAs are currently known to repress *ompD* under various stress conditions, with MicC and SdsR sharing the same mechanism of pairing within the CDS and recruitment of RNase E [107].

Once again, sRNA-based regulation can also have positive effects on target expression. RydC and SgrS are two examples for translation-independent target stabilization by interference with RNase E-mediated decay (Fig. 1.3D). RydC binds a *cfa* isoform with a long 5' UTR and base pairing at a region ~100 nt upstream of the start codon protects the transcript from degradation [108]. SgrS targets a decay intermediate of the dicistronic *pldB-yigL* operon by pairing within the CDS of the upstream *pldB* gene. Binding of the sRNA blocks the process of RNase E scanning along the mRNA from the monophosphorylated 5' end for downstream cleavage sites. Thereby it limits further processing and increases the abundance of the *yigL* mRNA and YigL protein without directly affecting its translation [255, 282].

sRNA-controlled transcription termination

While translation initiation and mRNA stability are arguably the hotspots of sRNA-based control, less common mechanisms have also been reported for several regulators. Of particular relevance for the present work is the interference with Rho-dependent transcription termination. This process is the second pathway for transcription termination in bacteria, in addition to the intrinsic termination described in section 1.2.2. It depends on so-called Rho utilization (*rut*) sites on the nascent transcript. Termination factor Rho recognizes these sequence motifs in newly synthesized RNA and subsequently terminates transcription [285]. Accessibility of *rut* sites is key for Rho-dependent termination and sRNAs can regulate this process (Fig. 1.3E).

This type of regulation was first reported for ChiX-mediated control of the *chiPQ* operon. At the first glance, ChiX affects *chiP* expression through the canonical mode of target repression by binding to its RBS, inhibiting translation and promoting mRNA decay [104]. However, also the distal *chiQ* gene in the operon was co-regulated by an unknown mechanism. In-depth analysis of this polarity effect revealed that inhibition of *chiP* translation exposes a normally hidden *rut* site in the ribosome-free mRNA, thereby promoting premature transcription termination within the *chiP* coding sequence [42]. A similar mechanism is described for the *galETKM* operon, where base pairing of Spot 42 to the *galK* leader inhibits expression of *galKM* by translational repression and Rho-dependent transcription termination [345]. Promoting intra-operonic transcription termination could be a general pathway for sRNAs to uncouple the expression of co-transcribed

genes.

While stripping an mRNA from translating ribosomes to allow Rho access to *rut* sites is a rather indirect process, sRNAs can also directly antagonize Rho function (Fig. 1.3F). The aforementioned positive regulation of *rpoS* translation by anti-antisense sRNA pairing is complemented by directly blocking Rho-dependent termination within the *rpoS* leader sequence through the action of RprA, DsrA and ArcZ [301]. Another example is protection of the *rho* leader itself by the sRNA SraL, implying that cells might modulate Rho protein levels through a second mechanism independent of the protein-based autoregulation [314].

RNA decoys

The complexity of RNA-based control is further increased by the emergence of RNA decoys (also called RNA predators, traps or sponges) as a new paradigm in post-transcriptional regulation (Fig. 1.3G) [103]. Detailed study of the above described ChiX sRNA revealed that the roles of regulator and target can get swapped depending on the sRNA/mRNA stoichiometry or the involvement of additional players: ChiX normally represses the chitoporin *chiP* and a chitobiose transporter encoded by parts of *chbBCARFG* in the absence of their substrate. Upon chitosugar-induced *chb* transcription, the mRNA is produced in excess and turns from a ChiX target into a decoy RNA as it binds and destabilizes ChiX, thereby indirectly de-repressing *chiPQ* [266].

Such sponges have been described to originate from various sources and exert diverse functions. For example, they can set a substrate threshold for transporter biosynthesis (ChiX/*chb*), contribute to feedback control of metabolic enzymes (RNA mimicry by GlmY/GlmZ to regulate *glmS* [119]), sequester sRNAs produced by transcriptional noise (tRNA spacer-derived 3'ETS^{leuZ}/RybB, RyhB [175]) and many more. If decoy RNAs act on multiple regulators (like 3'ETS^{leuZ}), they can also interconnect and balance the corresponding regulatory networks. New methodologies like RIL-seq (RNA interaction by ligation and sequencing [208]) or CLASH (cross-linking, ligation, and sequencing of hybrids [171]) allow the identification of RNA-RNA interactions at the transcriptome-wide level. The vast amount of potential interactions recently revealed by these techniques in *E. coli* implies that cross-talk and sponging between RNAs could be much more common than previously appreciated. However, most detected interactions and their regulatory relevance still need to be experimentally confirmed [208, 349].

1.3.3 Requirements for productive base pairing

The development of the RIL-seq technology has led to the discovery of many RNA-RNA interactions that have no effect on the expression of either of the two binding partners, thereby constituting so called non-productive base pairings. This stresses the importance of understanding which molecular features of an RNA-RNA interaction are required to mediate target regulation, a question which various studies have tried to answer for more than a decade. Despite great improvements in the sensitivity and accuracy of bioinformatic algorithms for target prediction, the amount of false positive predictions is still high [245]. Constraints on the sRNA side, which have

been discussed above, include the seed region, structural elements like hairpins, a tightly regulated sRNA expression pattern and interactions with protein partners such as Hfq or RNase E. In addition, also mRNAs require specific features to be an sRNA target, such as a sequence that is able to pair with the respective sRNA.
As described above, systematic mutational analyses have narrowed down the essential base pairing regions to as few as six to seven nucleotides [162, 21]. But adjacent bases often have the potential to pair as well and the formation of longer duplexes has been validated by structural probing of multiple sRNA-mRNA pairs [305, 141, 291]. As the length of a paired RNA duplex determines the thermodynamic stability of the interaction, different studies have analyzed the influence of binding strength and extended complementarity on target regulation. They report a correlation between calculated hybridization energy and strength of the regulation, supporting the use of free hybridization energy as a widely used criterion in computational target predictions tools [131, 31, 262, 39]. However, not all observed variability in regulation can be explained by thermodynamic differences. Especially, some nucleotide positions within a stretch of complementarity seem to be more crucial for target regulation than others [253]. These often represent the actual seed, i.e. those critical nucleotides that establish the initial contacts to the mRNA before pairing extends further to the adjacent nucleotides [121].

Some sRNAs can base pair with more than one region within their mRNA targets. These interaction sites can be in close proximity or even overlapping [16, 307] or at different genes within a polycistronic operon [31, 281]. Such multiplicity is assumed to increase the efficiency of target regulation, although the stoichiometry of sRNA binding is not always clear. One sRNA molecule can establish multiple connection to one mRNA transcript at the same time [16] or multiple regulators can bind to different parts of the mRNA independently [281]. It is also hypothesized that binding to either of two target sites in close proximity could be mutually exclusive, while together they still increase the affinity of the sRNA for its target [31].

Motivated by the obvious discrepancy between predicted and truly regulated sRNA targets, Beisel and colleagues performed a systematical search for requirements on the mRNA site that allowed regulation by the sRNA Spot 42 [31]. Increasing the strength of RNA-RNA interactions by extending the complementary duplex resulted in stronger target repression, as long as the interacting sequences were located in unstructured regions of the regulator. Similarly, secondary structures within the mRNA that enclose the predicted pairing site to Spot 42 prohibited regulation and mutations opening these inhibitory structures established Spot 42-dependent control [31]. These results are agreement with the observation that interaction regions are typically structurally accessible in both sRNAs and targets [257, 283] and with a recent study reporting a role of Hfq in melting an mRNA secondary structure to free the sRNA-binding site [139]. Accordingly, most bioinformatic tools for target prediction include calculations of structural accessibility in addition to hybridization energy and other features like sequence conservation [229, 49, 352, 165]. The majority of known sRNAs to date depend on Hfq for their functionality [341]. While this can

sometimes be due to the need of Hfq for sRNA stability rather than for actually mediating target regulation [220, 131], many mRNAs are also bound directly by Hfq [317, 38, 144]. In the above mentioned study on Spot 42 targets, non-regulated mRNAs with the potential to base pair to Spot 42 but lacking an Hfq site could be converted into true targets by transferring such an Hfq site from another target [31]. Additionally, mRNA targets have to bind Hfq in the right distance and in the correct orientation relative to the base pairing site [247, 259, 300].

These demands that mRNAs need to fulfill to be regulated by an sRNAs indicate that probably not all mRNAs can be sRNA targets. In this regard, it is noteworthy that many mRNAs are regulated by multiple sRNAs to different extent and with distinct outcomes. For example, *rpoS* is activated by direct base pairing of three sRNAs (DsrA, RprA and ArcZ) and repressed by pairing of another sRNA (CyaR) [168], while no less than seven sRNAs (OmrA/B, McaS, RprA, RybB, GcvB and RydC) directly repress the biofilm master regulator *csgD* [11]. Different sRNAs often use at least slightly variable binding sites on their target, but even the use of the very same binding site by an activating and a repressing sRNAs has been reported [85]. In general, such mRNA hubs of post-transcriptional regulation might reflect a potential evolutionary pathway of sRNA targets: establishing the necessary features for the first sRNA-mRNA interaction could increase the likelihood of an mRNA to acquire additional regulatory connections [333].

1.4 Physiological roles of sRNAs

The large diversity of sRNAs in many bacterial species is reflected by the variety of cellular processes that are subject to their post-transcriptional control (Fig. 1.4). Bacterial sRNAs provide fast and efficient means to adapt gene expression over a large range in response to sudden stimuli [312, 206]. Thus they are very suitable for fast adaptation of the bacterial transcriptome in fluctuating environmental conditions. Indeed, the majority of known sRNAs are expressed under very specific conditions and many are found in regulatory networks connected to cellular stress response pathways [138, 143]. For example, the Fur-dependent sRNA RyhB is upregulated under iron limitation [198] and replenishes the cellular iron pool, mainly by repressing non-essential iron-utilizing proteins [199, 269, 90]. Regulation of iron homeostasis is critical for cell integrity and many bacterial species encode either RyhB homologs or other sRNAs acting as functional analogs [12, 238]. Moreover, the general stress regulator RpoS is both controlling the expression of at least four sRNAs [240, 315, 106, 129] and is itself controlled by three different sRNAs [117, 194, 192], which help to fine-tune the RpoS response [26, 122].

In addition to stress responses, core cellular processes are also subject to sRNA-dependent control. Spot 42 and SgrS remodel the carbohydrate metabolism by contributing to CRP-dependent carbon catabolite repression [120, 30] or by counter-acting phosphosugar stress [336, 342, 255], respectively. Regulation of quorum sensing and virulence gene expression by sRNAs are best studied in *Vibrionaceae* and Gram-positive *Staphylococci*, where the Qrr sRNAs [332, 304, 248] or the dual function regulator RNAIII [237, 235, 46] constitute the heart of the respective quorum sensing

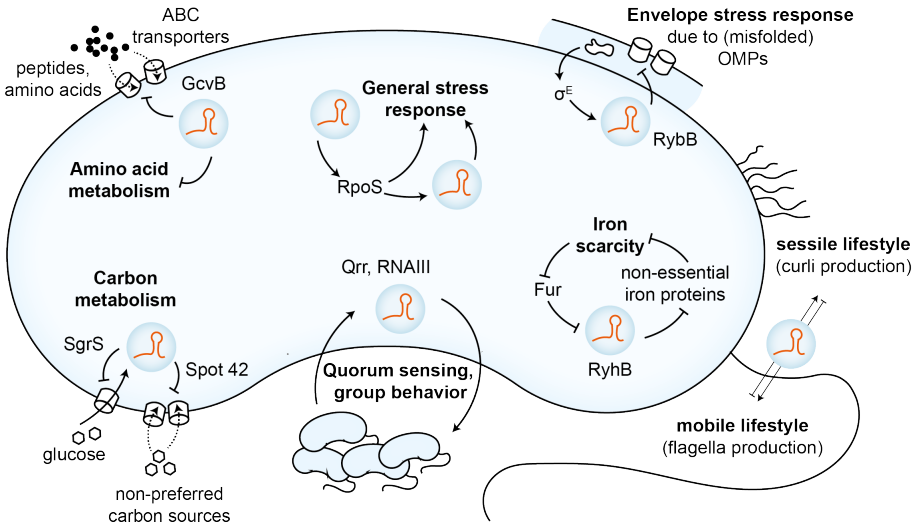


Figure 1.4: **Physiological processes involving sRNA-based control.** Regulation by sRNAs influences many cellular pathways ranging from stress response systems like the ones based on σ^E and σ^S to amino acid, carbon and iron metabolism. Also quorum sensing, group behavior and the decision between sessile and mobile lifestyles are mediated by and modulated through sRNAs.

systems. But also other group behaviors like motility or biofilm formation are tightly controlled by numerous sRNAs [85, 158]. This incomplete set of cellular processes involving sRNAs emphasizes their global importance for bacterial physiology [343]. The two sRNA-controlled processes that are most relevant for the present work are the regulation of amino acid metabolism by GcvB and the envelope stress response mediated by σ^E-dependent sRNAs, which will be described in the following section in more detail.

1.4.1 Regulation of amino acid metabolism

GcvB is one of the few sRNAs that are conserved beyond the *Enterobacteriales*, indicating an early evolutionary origin [258]. It is also among those sRNAs with the largest validated targetome, as it directly affects the expression of up to 1-2% of the *Salmonella* transcriptome [307, 208]. This unusually large set of direct targets is further extended through GcvB-dependent regulation of the key transcriptional regulator Lrp, which in turn regulates ~10% of all genes in *E. coli* [218, 329]. So far, 31 direct mRNA targets of GcvB have been validated *in vivo*, all of them repressed by the sRNA [176]. Through one of them, *cycA* encoding a glycine transporter, GcvB establishes a negative-feedback loop and limits its own transcriptional activation by the glycine-responsive transcription factor GcvA [335, 271]. Mechanistically, GcvB represses most of its targets through the same conserved seed region, which is very G/U-rich and binds to C/A-rich translational enhancer elements on the mRNA targets [305]. Deletion of the *gcvB* gene leads to strongly increased expression of its targets, *e.g.* of the periplasmic substrate-binding proteins of the major peptide transporters, DppA and OppA. GcvB is mainly expressed when cells are growing fast in rich media and it targets many additional ABC transporters for amino acids and peptides as well as genes

involved in amino acid biosynthesis [305, 307]. Thus, its major function has been described as the limitation of energy-intensive amino acid uptake and biosynthesis under nutrient-rich conditions. However, the physiological rationale behind this GcvB-mediated shutdown of amino acid metabolism is still not fully understood [176]. An additional layer of complexity is added by the discovery of a target-derived sponge of GcvB called SroC. It is processed from the *gltIJKL* mRNA (which is repressed by GcvB) and strongly destabilizes GcvB [214]. This establishes a feed-forward loop, in which transcriptional activation of *gltIJKL* also produces the sponge RNA SroC that blocks the inhibitor GcvB. At the same time, SroC indirectly de-represses the other GcvB targets, leading to a coordinate response of the GcvB regulon.

Another sRNA involved in the regulation of amino acid metabolism has been discovered in the 3' UTR of the *dapB* gene (encoding an essential protein for lysine biosynthesis) [61]. The sRNA DapZ is independently transcribed from a *dapB*-internal promoter and regulates the major ABC transporters encoded by the *dpp* and *opp* operons, which are responsible for the uptake of dipeptides and oligopeptides, respectively [83]. Remarkably, DapZ not only shares these targets with GcvB, but it also uses a very similar G/U-rich seed region to target the same mRNA regions in the first genes of the operons, *dppA* and *oppA* [305, 61].

1.4.2 The σ^E-dependent envelope stress response

The cell envelope of Gram-negative bacteria is a multilayered structure composed of the inner membrane, the aqueous periplasm, a thin peptidoglycan layer and the outer membrane (reviewed by [313]). Many (glyco)proteins are inserted into the outer membrane governing processes like nutrient uptake or surface attachment. One class of transmembrane proteins called OMPs (for Outer Membrane Proteins) are often highly abundant and mostly function as pores and channels [234]. But misfolding of OMPs can occur under unfavorable conditions (such as pH or redox stress) or stochastically in unstressed cells. As misfolded OMPs are highly toxic to the cell, their folding status is constantly monitored by the σ^E signaling system: misfolded OMPs release σ^E from the repression of its anti-sigma factor, allowing it to transcriptionally activate a large regulon including chaperones and proteases to counteract the experienced stress [279, 212].

However, as a transcriptional activator, σ^E itself is unable to repress gene expression. This function is exerted by the post-transcriptional arm of the response in the shape of multiple sRNAs such as MicA, MicL and RybB in *E. coli* and *Salmonella*. These are strongly upregulated by σ^E and repress many mRNAs encoding major OMPs by inhibiting translation and initiating decay of the unusually stable *omp* transcripts [252, 157, 249, 126]. Additional sRNA targets include the highly abundant lipoprotein Lpp and some other lipoproteins and transporters [118, 126]. By reducing the overall OMP synthesis, the sRNAs contribute to a fast relieve of the σ^E-inducing envelope stress, thereby providing a feedback loop. This regulation is not only relevant under σ^E-inducing, unfavorable conditions, but rather constitutes a permanent surveillance system. Thus, deletion of the sRNA genes leads to chronic envelope stress also in the absence of any external envelope damage [252].

While the general principle of the σ^E-dependent response is widely conserved in Gram-negative bacteria and essential for cell viability also in *V. cholerae*, its associated sRNAs are limited to *Enterobacteriales* [82, 212, 258]. However, *V. cholerae* encodes an unrelated but functionally analogous sRNA called VrrA, which is also controlled by σ^E and represses major OMPs as well as biofilm components [324, 325]. Through its repression of *ompA*, VrrA also increases the formation of outer membrane vesicles, which has been proposed as an additional envelope stress response system [204].

The σ^E-dependent envelope stress response has become a paradigm for mixed regulatory networks composed of transcription factors and regulatory RNAs. It exhibits many characteristic features of such mixed circuits like feedback regulation or inversion of the sign of transcription factor (TF) control by the sRNA, as described in the next section.

1.5 Concepts of bacterial gene regulation

Post-transcriptional control of gene expression by bacterial sRNAs is of course not isolated from other regulatory networks in the cell, but rather closely intertwined with *e.g.* transcriptional regulation. Research on bacterial gene expression in the past decades has uncovered some fundamental concepts that are widely conserved among and even beyond bacterial species, for instance certain regulatory motifs that are recurrently found within transcriptional networks [6]. They form autoregulatory loops, feedback circuits, feed-forward loops and many more. The discovery of pervasive sRNA-based control has revealed that these networks can also include sRNAs in addition to the well-studied transcriptional regulators and multiple regulons can be connected via sRNA hubs. The following chapter will highlight two concepts which are particularly relevant for this work: transcription factor-based autoregulation and mixed regulatory networks that combine regulation by both TFs and sRNAs.

1.5.1 Autoregulation

When studying bacterial transcription networks, one of the most abundant motifs is (direct) autoregulation: a transcription factor binds to its own promoter and thereby influences its own expression, typically in a repressive way leading to negative autoregulation (NAR) [330]. This regulatory architecture is employed by approximately 50% of all bacterial regulators [6]. One consequence of NAR (and probably one of the reasons for its widespread occurrence) is the accelerated response time of the system. NAR allows the use of a strong promoter for a fast initial rise of protein levels, as the synthesis of new proteins is efficiently slowed down when the protein concentration gets closer to its repression threshold [288, 55]. Additionally, NAR reduces the stochastic noise in gene expression, as high TF concentrations reduce the *de novo* synthesis, while low concentrations increase it. This dampens the amplitude of protein level fluctuations between different cells [28]. Negative autoregulation is also described for the termination factor Rho, which regulates the levels

of its own mRNA by inducing termination within its leader sequence [202]. For sRNAs, such direct autoregulation is naturally impossible, as they cannot bind DNA to regulate their transcription. However, indirect autoregulation by influencing their cognate transcription factors within feedback loops has been observed for some sRNAs (see below and [327, 142]).

1.5.2 Mixed regulatory networks

By their post-transcriptional mode of action, sRNAs add a second layer to the regulation of gene expression, acting in addition to the well-characterized transcriptional control (Fig. 1.5). This seemingly trivial observation opens up many possibilities for regulation that would not be possible with only transcriptional control. Moreover, many transcription factors are themselves regulated by sRNAs, giving rise to multi-layered and interconnected regulatory cascades.

Switching the sign of regulation

Most fundamentally, sRNAs can switch the sign of regulation of a transcriptional regulator (Fig. 1.5A): some regulators are intrinsically defined to act as either activators (such as sigma factors guiding RNAP to their designated promoters [134]) or as repressors (like the transcriptional repressor Fur that binds to *fur* boxes at promoters and blocks RNAP access [99]). The existence of σ^E-repressed genes or Fur-activated genes has long been puzzling, until the discovery of sRNAs within the respective regulons solved the riddle [95, 279]. While σ^E directly activates the transcription of ~100 genes in *E. coli*, its partner sRNAs MicA, MicF and RybB together repress ~30 genes encoding abundant OMPs, which need to be silenced under envelope stress conditions (see section 1.4.2 and [118, 126]). Conversely, the RyhB sRNA negatively regulates non-essential iron-utilizing genes during iron scarcity. By transcriptional repression of RyhB under iron-replete conditions, Fur acts as an activator for these genes [198].

Tight control of genes in an 'OFF' state

Transcription factors and sRNAs can also act simultaneously on the same target to establish an even tighter control (Fig. 1.5B). Transcriptional repression of an mRNA target keeps synthesis of new transcripts at a low level. However, short bursts of transcription can occur stochastically and are amplified during translation, giving rise to relatively large fluctuations in protein production despite the transcriptional 'OFF' state of the gene [51]. The additional repression by an sRNA helps to avoid these escapes from target repression by blocking translation of the mRNA transcripts produced during a transcriptional burst [15]. As the rate of mRNA synthesis is kept low by the transcriptional repressor, continuous moderate expression of the sRNA is sufficient to silence protein production [182]. Thus, the combined action of both repressors strongly enhances target regulation above the regulatory capacity of the individual players.

Such dual repression by TF and sRNA is for example predicted to control genes mediating bistable behavior such as the *csgD* master regulator of biofilm formation [123]. CsgD indirectly

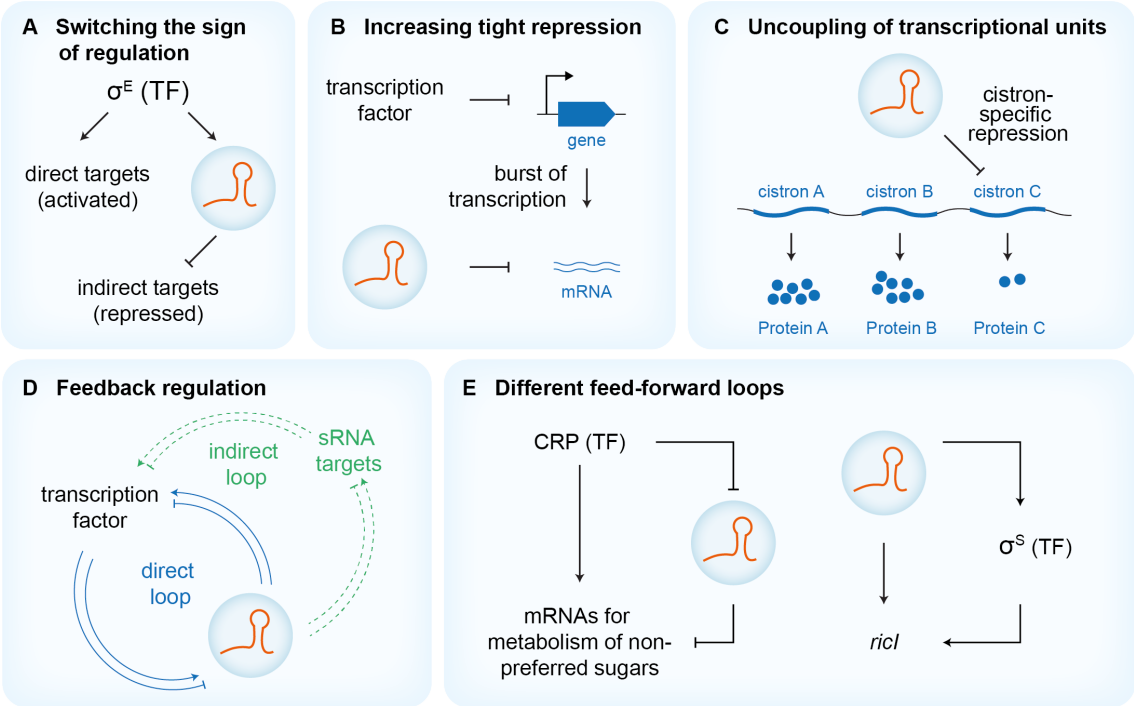


Figure 1.5: **Mixed regulatory networks of transcription factors and sRNAs.** (A) sRNAs can inverse the sign of regulation of a transcriptional regulator by e.g. repressing mRNA targets of the transcriptional activator σ^E. (B) Additional repression by an sRNA can tighten the control of a transcriptionally inactive gene. (C) Transcriptionally coupled genes can be differentially regulated by sRNAs at the post-transcriptional level. (D) Some sRNAs regulate their own transcription factors, either through direct binding of its mRNA or indirectly by modulating target gene expression that feeds back on the TF. (E) Feed-forward loops are constituted of two regulatory arms acting on the same target. sRNAs can be placed in the middle of the loop (left) or at its top (right).

enhances its own expression, thereby establishing a positive feedback loop and promoting the formation of multi-cellular aggregates important for environmental persistence [190]. The *csgD* gene is part of a complicated transcriptional network and additionally regulated by seven repressive sRNAs, although experimental evidence on how these sRNAs actually influence the CsgD bistability is currently lacking [190, 11].

Dynamics of target gene repression

In addition to increasing the strength of target repression, sRNAs can also contribute to a faster regulation. When synthesis of a certain protein is supposed to be turned off, the respective transcriptional regulator acts to terminate synthesis of new mRNAs. However, the existing transcripts can still be translated until they are degraded, which can substantially delay the response for mRNAs with long half-lives. In this case, sRNAs directly acting on the mRNAs allow an immediate block of protein synthesis, thereby speeding up the regulation. An example is the accelerated reduction of CRP targets by the sRNA Spot 42 upon CRP inactivation (see below and [30]).

Uncoupling of transcriptional units

Bacterial genes are often organized into operons to coordinate the expression of functionally linked genes. Binding of sRNAs to such polycistronic mRNAs can either lead to the regulation of all encoded genes simultaneously or to uncoupling of the operon (Fig. 1.5C). Regulation of the full operon is often achieved when the individual cistrons are translationally coupled and binding of the sRNA to the first RBS abolishes translation of all downstream genes (PhrS/*pqsR* [326]). In a different pathway, the above described ChiX sRNA indirectly represses the second gene in the *chiPQ* operon by blocking *chiP* translation, which exposes a normally hidden rut site in *chiP* and induces Rho-dependent transcription termination before *chiQ* [42]. Alternatively, sRNAs may pair at multiple binding sites along the mRNA, thereby blocking translation of multiple genes independently (SgrS/*manXYZ* [281]).

In contrast, differential expression of transcriptionally coupled genes can be advantageous when only a subset of the gene products is needed under a specific condition. Base pairing sRNAs can mediate this uncoupling by selectively stabilizing parts of the operon (SgrS/*yigL* [255]). Conversely, some cistrons can be protected from sRNA-induced decay of the remaining transcript by stable secondary structures (RyhB/*iscRSUA* [90]). Moreover, sRNA-based translational inhibition of a subset of genes within an operon does not necessarily reduce the stability of the full mRNA, whose other cistrons are then unaffected by sRNA binding (Spot 42/*galETKM* [221]). Apparently, sRNA binding at various positions along a polycistronic mRNA can lead to very different outcomes and regulation of the individual cistrons has to be studied on a case-to-case basis to uncover the respective regulation.

Feedback circuits

Under stress conditions, cells commonly induce the expression of gene products that counteract the stress and thereby reduce the signals that initially induced their own expression. One example for such an indirect example is the σ^E-dependent envelope stress response (see section 1.4.2 and Fig. 1.5D). But sRNAs also contribute to much simpler feedback loops consisting of only two components: a transcriptional regulator that controls synthesis of an sRNA, which in turn represses or activates the TF. Such loops can in principle lead to three outcomes [6]: (i) signal amplification if both regulations are positive (not yet observed for sRNAs), (ii) balanced expression of both partners if positive and negative regulation is combined (*e.g.* Qrr1-4/*luxO* [327]) or (iii) bistable behavior with inversely correlated expression levels if both regulations are negative (MicF/*lrp* [142]). However, these feedback loops never work in isolation, but are rather embedded into other regulatory connections that additionally modulate the output of the system. This becomes especially clear when looking at the complex quorum sensing response in *Vibrio*, where multiple direct and indirect feedback loops tightly control collective behavior [332, 290].

Various feed-forward loops

Feed-forward loops (FFLs) are network motifs consisting of three components and have been studied in detail for transcriptional regulators [6]. Briefly, component A regulates target C directly, but also indirectly through control of factor B (Fig. 1.5E). Depending on the signs of all their regulatory connections, FFLs are categorized into eight types and called either coherent (if both arms result in the same target regulation) or incoherent (if one arm is repressing and the other one is activating the target) [196]. A variety of such FFLs have been reported to include sRNAs as the factor B: they are controlled by a TF and regulate another direct target within the TF regulon [236]. One of the best-studied coherent FFLs is composed of the transcription factor CRP, the sRNA Spot 42 and various genes of the carbon catabolite repression pathway [30]. When cells grow on glucose, Spot 42 represses genes for the uptake and utilization of non-preferred carbon sources. During glucose starvation, CRP transcriptionally induces these genes and at the same time inhibits expression of the repressor Spot 42, thus contributing to the metabolism of alternative carbon sources via two ways [120, 30]

A more unusual FFL is the RprA-RpoS-*ricI* circuit, as it is one of the few described examples where an sRNA is placed at the top level of the circuit [250, 236]. RprA post-transcriptionally activates the general stress response regulator *rpoS* as well as *ricI*, a direct target of σ^S activation. Thus, it constitutes another coherent FFL, where both arms result in target activation. However, activation of *rpoS* alone, e.g. by other sRNAs, is not sufficient for full *ricI* induction, as RprA is needed to activate *ricI* translation by an anti-antisense mechanism. This AND-gate logic ensures that σ^S induction only leads to RicI synthesis if RprA is simultaneously present in the cell. Thus, RprA induces a specialized form of the RpoS response that cannot be mediated by other σ^S-inducing conditions. Physiologically, this circuit limits transfer of a virulence plasmid in *Salmonella* under unfavorable conditions [250].

1.6 Aim of this work

In recent years, bacterial sRNAs have emerged as powerful regulators that can rival transcription factors with regard to their regulatory scope and function. They establish extensive post-transcriptional networks by a variety of molecular mechanisms, with novel modes of action still being uncovered. While a number of sRNAs has been studied in great detail in the model organisms *E. coli* and *Salmonella*, less is known about the sRNA repertoire in the major human pathogen *V. cholerae*. In this organism, sRNAs constitute central components of multiple physiologically important processes such as quorum sensing and virulence [181, 260], natural competence [353], or maintenance of envelope integrity [324, 325]. An initial study from our group had identified a large set of previously unknown candidate sRNAs in *V. cholerae*, most of which still await functional characterization [251]. One goal of this study is to extend the knowledge on *V. cholerae* sRNAs encoded in varying genomic loci and produced via different biogenesis pathways. A special focus is set on their molecular characteristics and regulatory mechanisms.

Negative autoregulation is a central motif in regulatory networks, where a transcription factor controls its own expression by binding to its own promoter [6]. NAR has also been reported for sRNAs in a more indirect way through sRNA-dependent regulation of their cognate TFs which in turn control sRNA transcription [327, 142, 48]. However, such feedback loops are always dependent on the action of a transcriptional regulator. Given the increasing complexity of sRNA-based regulatory mechanisms that are explored, we speculated that sRNAs could also be able to establish protein-independent autoregulation. Indeed, the emergence of processed, 3' UTR-derived sRNAs opened up a novel possibility for feedback control: they depend on the expression of their parental mRNA, which could also be regulated at the post-transcriptional level, *i.e.* by sRNAs. Motivated by this idea, we started the first project of this thesis by studying the central ribonuclease RNase E in *V. cholerae* to find sRNAs that are processed from mRNA 3' UTRs. Subsequently, we screened these sRNAs for potential feedback regulation of their own operons by bioinformatic target predictions. We identified two potentially autoregulatory sRNAs and validated our hypothesis by characterizing the underlying molecular mechanism.

Our second study focuses on the mechanistic details of evolved sRNA-mRNA target pairs. As described in section 1.4.2, the σ^E-dependent envelope stress response is generally conserved between *E. coli* and *V. cholerae*, but the sRNAs constituting its repressive arm are not. Still they carry out the same function by regulating almost identical targets [118, 212]. It is yet unclear if this functional analogy is the consequence of a selective pressure for effective removal of stable *omp* mRNAs. It is also an open question if these sRNAs represent the only (or most efficient) way for the cells to cope with envelope stress. Moreover, the discovery of three σ^E-dependent sRNAs in *E. coli* suggests that VrrA might not be the only one in *V. cholerae*. We addressed these questions in the second study presented in this thesis by initially searching for other σ^E-dependent sRNAs in *V. cholerae*. Subsequently, we constructed a complex library of synthetic sRNAs based on a natural scaffold equipped with a randomized seed region. This library was transferred to *V. cholerae* to provide the cells with a large pool of potential regulators, from which the most suitable variants could be selected in the context of the σ^E response. We used laboratory selection experiments under membrane-damaging conditions to enrich beneficial regulators in an unbiased fashion and analyzed their regulatory properties and function. Finally, we compared the synthetic regulators to the native σ^E-dependent sRNAs from *V. cholerae* and *E. coli* to identify shared characteristics.

2 Gene autoregulation by 3' UTR-derived bacterial small RNAs

Mona Hoyos, Michaela Huber, Konrad U. Förstner, and Kai Papenfort (2020). "Gene autoregulation by 3' UTR-derived bacterial small RNAs." *eLife* 9, e58836

Source online: https://doi.org/10.7554/eLife.58836

This publication describes two examples of an RNA-based, autoregulatory loop by a 3' UTR-derived sRNA. OppZ and CarZ are processed from their parental mRNAs by RNase E and bind to upstream cistrons in their respective operons. Thereby, they inhibit mRNA translation and induce premature, Rho-dependent transcription termination.

RESEARCH ARTICLE

Gene autoregulation by 3′ UTR-derived bacterial small RNAs

Mona Hoyos[1,2], Michaela Huber[1,2], Konrad U Förstner[3,4], Kai Papenfort[1,2,5]*

[1]Friedrich Schiller University Jena, Institute of Microbiology, Jena, Germany; [2]Faculty of Biology I, Ludwig-Maximilians-University of Munich, Martinsried, Germany; [3]TH Köln - University of Applied Sciences, Institute of Information Science, Cologne, Germany; [4]ZB MED - Information Centre for Life Sciences, Cologne, Germany; [5]Microverse Cluster, Friedrich Schiller University Jena, Jena, Germany

Abstract Negative feedback regulation, that is the ability of a gene to repress its own synthesis, is the most abundant regulatory motif known to biology. Frequently reported for transcriptional regulators, negative feedback control relies on binding of a transcription factor to its own promoter. Here, we report a novel mechanism for gene autoregulation in bacteria relying on small regulatory RNA (sRNA) and the major endoribonuclease, RNase E. TIER-seq analysis (transiently-inactivating-an-endoribonuclease-followed-by-RNA-seq) revealed ~25,000 RNase E-dependent cleavage sites in *Vibrio cholerae*, several of which resulted in the accumulation of stable sRNAs. Focusing on two examples, OppZ and CarZ, we discovered that these sRNAs are processed from the 3′ untranslated region (3′ UTR) of the *oppABCDF* and *carAB* operons, respectively, and base-pair with their own transcripts to inhibit translation. For OppZ, this process also triggers Rho-dependent transcription termination. Our data show that sRNAs from 3′ UTRs serve as autoregulatory elements allowing negative feedback control at the post-transcriptional level.

***For correspondence:** kai.papenfort@uni-jena.de

Competing interests: The authors declare that no competing interests exist.

Funding: See page 23

Received: 12 May 2020
Accepted: 23 July 2020
Published: 03 August 2020

Reviewing editor: Joseph T Wade, Wadsworth Center, New York State Department of Health, United States

Introduction

Biological systems function on a mechanism of inputs and outputs, each triggered by and triggering a specific response. Feedback control (a.k.a. autoregulation) is a regulatory principle wherein the output of a system amplifies (positive feedback) or reduces (negative feedback) its own production. Negative feedback regulation is ubiquitous among biological systems and belongs to the most thoroughly characterized network motifs (**Nitzan et al., 2017**; **Shen-Orr et al., 2002**). At the gene regulatory level, negative feedback control has been qualitatively and quantitatively studied. Most commonly, a transcription factor acts to repress its own transcription by blocking access of RNA polymerase to the promoter region. This canonical mode of negative autoregulation is universally present in living systems and in *Escherichia coli* more than 40% of the known transcription factors are controlled by this type of regulation (**Rosenfeld et al., 2002**). Several characteristics have been attributed to negative autoregulatory circuits including an altered response time and improved robustness towards fluctuations in transcript production rates (**Alon, 2007**).

More recently, the mechanisms underlying RNA-based gene regulation have also been investigated for their regulatory principles and network functions (**Nitzan et al., 2017**; **Pu et al., 2019**). In bacteria, small regulatory RNAs (sRNAs) constitute the largest class of RNA regulators and frequently bind to one of the major RNA-binding proteins, Hfq or ProQ. Hfq- and ProQ-associated sRNAs usually act by base-pairing with *trans*-encoded target mRNAs affecting translation initiation and transcript stability (**Holmqvist and Vogel, 2018**; **Kavita et al., 2018**). The sRNAs frequently target multiple transcripts and given that regulation can involve target repression or activation, it has

become ever more clear that sRNAs can rival transcription factors with respect to their regulatory scope and function (**Hör et al., 2018**).

Another key factor involved in post-transcriptional gene regulation is ribonuclease E (RNase E), an essential enzyme in *E. coli* and related bacteria required for ribosome biogenesis and tRNA maturation (**Mackie, 2013**). RNase E's role in sRNA-mediated expression control is manifold and includes the processing of sRNAs into functional regulators (**Chao et al., 2017**; **Dar and Sorek, 2018a**; **Papenfort et al., 2015a**; **Updegrove et al., 2019**; **Chao et al., 2012**) as well as the degradation of target transcripts (**Massé et al., 2003**; **Morita et al., 2005**). Inhibition of RNase E-mediated cleavage through sRNAs can stabilize the target transcript and activate gene expression (**Fröhlich et al., 2013**; **Papenfort et al., 2013**; **Richards and Belasco, 2019**).

Global transcriptome analyses have revealed the presence of numerous sRNAs produced from 3′ UTRs (untranslated regions) of mRNAs, a significant subset of which requires RNase E for their maturation (**Adams and Storz, 2020**). These 3′ UTR-derived sRNAs can be produced from monocistronic (**Chao and Vogel, 2016**; **Grabowicz et al., 2016**; **Huber et al., 2020**; **Wang et al., 2020**) as well as long, operonic mRNAs (**Davis and Waldor, 2007**; **De Mets et al., 2019**; **Miyakoshi et al., 2019**) and typically act to regulate multiple target mRNAs in trans. The RNase E C-terminus also provides the scaffold for a large protein complex, called the degradosome, which in the major human pathogen, *Vibrio cholerae*, has recently been implicated in the turn-over of hypomodified tRNA species (**Kimura and Waldor, 2019**).

The present work addresses the regulatory role of RNase E in *V. cholerae* at a genome-wide level. To this end, we generated a temperature-sensitive variant of RNase E in *V. cholerae* and employed TIER-seq (transiently-inactivating-an-endoribonuclease-followed-by-RNA-seq) to globally map RNase E cleavage sites (**Chao et al., 2017**). Our analyses identified ~25,000 RNase E-sensitive sites and revealed the presence of numerous stable sRNAs originating from the 3′ UTR of coding sequences. Detailed analyses of two of these sRNAs, OppZ and CarZ, showed that 3′ UTR-derived sRNAs can act in an autoregulatory manner to reduce the expression of mRNAs produced from the same genetic locus. The molecular mechanism of sRNA-mediated gene autoregulation likely involves inhibition of translation initiation by the sRNA followed by Rho-dependent transcription termination. This setup directly links the regulatory activity of the sRNAs to their de novo synthesis, analogous to their transcription factor counterparts. However, we show that, in contrast to transcriptional regulators, autoregulatory RNAs can act at a subcistronic level to allow discoordinate operon expression.

Results

TIER-seq analysis of *V. cholerae*

The catalytic activity of RNase E (encoded by the *rne* gene) is critical for many bacteria, including *V. cholerae* (**Cameron et al., 2008**). To study the role of RNase E in this pathogen, we mutated the DNA sequence of the *V. cholerae* chromosome encoding leucine 68 of RNase E to phenylalanine (**Figure 1—figure supplement 1**). This mutation is analogous to the originally described N3071 rne^{TS} isolate of *E. coli* (**Apirion and Lassar, 1978**) and exhibits full RNase E activity at permissive temperatures (30˚C), but is rendered inactive under non-permissive temperatures (44˚C). We validated our approach by monitoring the expression of two known substrates of RNase E in *V. cholerae*: A) 5S rRNA, which is processed by RNase E from the 9S precursor rRNA (**Papenfort et al., 2015b**), and B) the MicX sRNA, which contains two RNase E cleavage sites (**Davis and Waldor, 2007**). For both RNAs, transfer of the wild-type strain to 44˚C only mildly effected their expression, whereas the equivalent procedure performed with the rne^{TS} strain led to the accumulation of the 9S precursor and the full-length MicX transcript (**Figure 1A**, lanes 1–2 vs. 3–4). Additionally, accumulation of the two RNase E-dependent processing intermediates of MicX was reduced in the rne^{TS} strain at the non-permissive temperature.

These results showed that we successfully generated a temperature-sensitive RNase E variant in *V. cholerae* and enabled us to employ TIER-seq to determine RNase E-dependent cleavage sites at a global scale. To this end, we cultivated *V. cholerae* wild-type and rne^{TS} strains at 30˚C to late exponential phase (OD_{600} of 1.0), divided the cultures in half and continued incubation for 60 min at either 30˚C or 44˚C. Total RNA was isolated and subjected to deep sequencing. We obtained ~187 million reads from the twelve samples (corresponding to three biological replicates of each strain

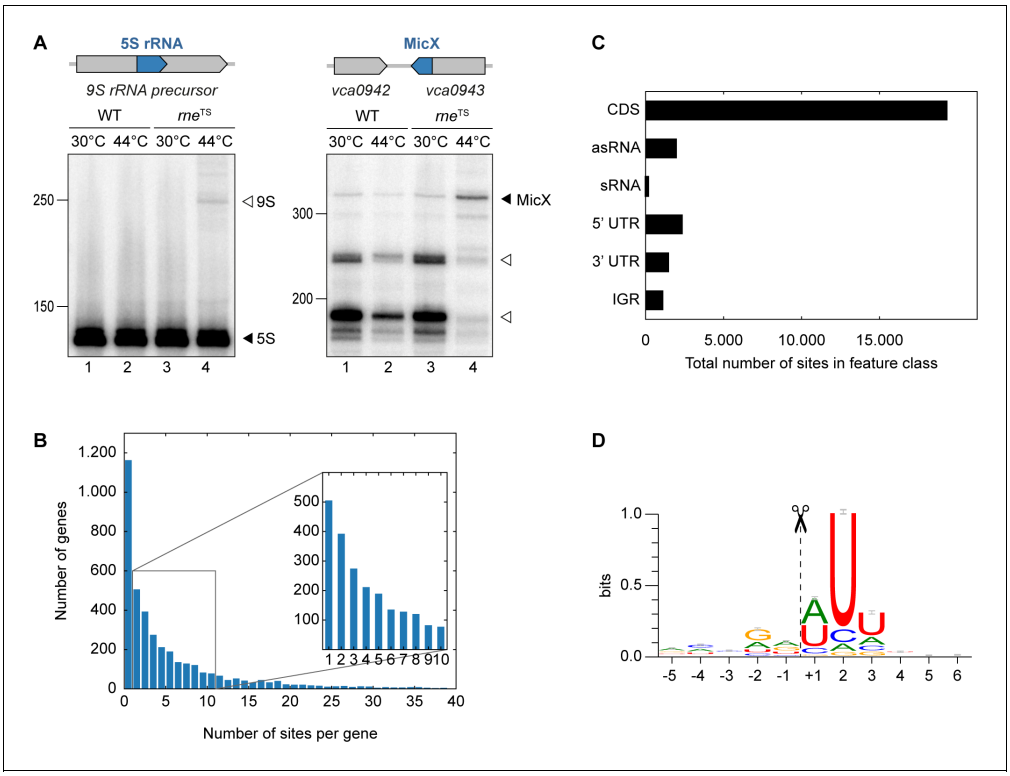


Figure 1. TIER-seq analysis of *V. cholerae*. (**A**) *V. cholerae* wild-type and *rne*^{TS} strains were grown at 30˚C to stationary phase (OD_{600} of 2.0). Cultures were divided in half and continuously grown at either 30˚C or 44˚C for 60 min. Cleavage patterns of 5S rRNA and 3' UTR-derived MicX were analyzed on Northern blots. Closed triangles indicate mature 5S or full-length MicX, open triangles indicate the 9S precursor or MicX processing products. (**B, C, D**) Biological triplicates of *V. cholerae* wild-type and *rne*^{TS} strains were grown at 30˚C to late exponential phase (OD_{600} of 1.0). Cultures were divided in half and continuously grown at either 30˚C or 44˚C for 60 min. Isolated RNA was subjected to RNA-seq and RNase E cleavage sites were determined as described in the materials and methods section. (**B**) Number of cleavage sites detected per gene. (**C**) Classification of RNase E sites by their genomic location. (**D**) The RNase E consensus motif based on all detected cleavage sites. The total height of the error bar is twice the small sample correction.
The online version of this article includes the following source data and figure supplement(s) for figure 1:

Source data 1. Full Northern blot images for the corresponding detail sections shown in ***Figure 1*** and RNase E cleavage site counts within genes or transcript categories.
Figure supplement 1. Conservation of RNase E between *E. coli* and *V. cholerae*.
Figure supplement 2. TIER-Seq read mapping statistics.
Figure supplement 2—source data 1. Number of obtained sequencing reads and Pearson correlation coefficients for library comparisons.
Figure supplement 3. Position and characteristics of RNase E cleavage sites.
Figure supplement 4. RNase E-mediated maturation of sRNAs from 3' UTRs.
Figure supplement 4—source data 1. Full Northern blot images for the corresponding detail sections shown in ***Figure 1—figure supplement 4***.
Figure supplement 5. RNase E-mediated maturation of sRNAs from IGRs.
Figure supplement 5—source data 1. Full Northern blot images for the corresponding detail sections shown in ***Figure 1—figure supplement 5***.
Figure supplement 6. Expression of RNase E-independent sRNAs.
Figure supplement 6—source data 1. Full Northern blot images for the corresponding detail sections shown in ***Figure 1—figure supplement 6***.

and condition; ***Figure 1—figure supplement 2A***), resulting in ~98 million unique 5' ends mapping to the *V. cholerae* genome. Comparison of the 5' ends detected in wild-type and *rne*^{TS} at 30˚C showed almost no difference between the two strains (Pearson correlation coefficients R^2 ranging from 0.82 to 0.99 depending on the compared replicates), whereas the same analysis at 44˚C revealed 24,962 depleted sites in the *rne*^{TS} strain (***Figure 1—figure supplement 2B–C***). Given that γ-proteobacteria such as *V. cholerae* do not encode 5' to 3' exoribonucleases (***Mohanty and Kushner, 2018***), we designated these positions as RNase E-specific cleavage sites (***Supplementary file 1***).

Next, we analysed the ~25,000 RNase E sites with respect to frequency per gene and their distribution among different classes of transcript. We discovered that RNase E cleavage sites occur with a frequency of 2.8 (median)/6.3 (mean) sites per kb (**Figure 1B**). The majority of cleavage events occurs in coding sequences (~69.1%), followed by 5' UTRs (~8.4%), antisense RNAs (~7.1%), 3' UTRs (~5.3%), intergenic regions (~4.0%), and sRNAs (~0.6%) (**Figure 1C**). RNase E sites were slightly enriched around start and stop codons of mRNAs (**Figure 1—figure supplement 3A**). Furthermore, cleavage coincided with an increase in AU-content (**Figure 1—figure supplement 3B**) and a rise in minimal folding energies (**Figure 1—figure supplement 3C**), suggesting reduced secondary structure around RNase E sites. Together, these data allowed us to determine a consensus motif for RNase E in *V. cholerae* (**Figure 1D**). This 5-nt sequence, *i.e.* 'RN↓WUU', is highly similar to previously determined RNase E motifs of *Salmonella enterica* (**Chao et al., 2017**) and *Rhodobacter sphaeroides* (**Förstner et al., 2018**), indicating that RNase E operates by a conserved mechanism of recognition and cleavage.

RNase E-mediated maturation of sRNAs

Earlier work on sRNA biogenesis in bacteria revealed that the 3' UTR of coding transcripts can serve as source for non-coding regulators and that RNase E is frequently required to cleave the sRNA from the mRNA (**Miyakoshi et al., 2015**). In *V. cholerae*, we previously annotated 44 candidate sRNAs located in the 3' UTR of mRNAs (**Papenfort et al., 2015b**). To analyse which of these sRNAs depend on RNase E for maturation, we searched for RNase E-cleavage sites matching with the first three bases of the annotated sRNAs. 17 sRNAs revealed potential RNase E-dependent maturation (**Supplementary file 2A**) and using Northern blot analyses of wild-type and *rne*^{TS} samples, we were able to confirm these results for 9 sRNAs (Vcr016, Vcr041, Vcr044, Vcr045, Vcr053, Vcr064, FarS, Vcr079, and Vcr084; **Figure 1—figure supplement 4**). In all cases, transfer of the *rne*^{TS} strain to non-permissive temperatures led to a change in mature sRNA levels and/or their upstream processing intermediates. We also discovered several sRNAs undergoing maturation by RNase E (**Supplementary file 2B**). Specifically, Northern blot analysis of Vcr043, Vcr065, and Vcr082 revealed that these sRNAs accumulate as multiple stable intermediates (**Figure 1—figure supplement 5**) that may contain different regulatory capacities as previously described for ArcZ and RprA of *S. enterica* (**Chao et al., 2017**; **Papenfort et al., 2015a**; **Soper et al., 2010**). In addition, we also analysed the expression of several RNase E-independent sRNAs (RyhB, Spot 42 and VqmR; **Figure 1—figure supplement 6**) on Northern blots. Inactivation of RNase E did not affect the levels of the mature sRNAs or any processed intermediates.

OppZ is produced from the *oppABCDF* 3' end

To understand the regulatory functions of 3' UTR-derived sRNAs in *V. cholerae*, we focussed on Vcr045, which is processed from the 3' end of the *oppABCDF* mRNA (encoding an oligopeptide transporter) and which we hence named OppZ. The *oppZ* gene is 52 bps long and conserved among the *Vibrios* (**Figure 2A**). RNase E-mediated cleavage of *oppABCDF* occurs immediately downstream of the *oppF* stop codon and using the *rne*^{TS} strain, we were able to validate RNase E-dependent processing of OppZ (**Figure 2B**). Northern and Western blot analysis of a *V. cholerae* strain carrying a 3XFLAG epitope at the C-terminus of the chromosomal *oppA* and *oppB* genes revealed that OppZ expression coincided with the expression of both proteins (**Figure 2C**, lanes 1–4). Previous transcriptome data showed that expression of *oppABCDF* is controlled by a single promotor located ~120 bps upstream of *oppA* (**Papenfort et al., 2015b**), indicating that the sRNA is co-expressed with all five *opp* genes. To test this prediction, we replaced the native promoter upstream of the chromosomal *oppA* gene with the L-arabinose-inducible pBAD promoter and monitored OppA, OppB, and OppZ expression under inducing and non-inducing conditions. In the absence of the inducer, expression of OppA/B and OppZ was strongly reduced (**Figure 2C**, lanes 5–8) and L-arabinose had no effect on the activity of the native *oppA* promoter (**Figure 2C**, lanes 9–10). In contrast, activation of the pBAD promoter led to a significant increase in OppA/B and OppZ (**Figure 2C**, lanes 11–12), indicating that expression of the *oppABCDF-oppZ* operon is indeed controlled by a single promoter.

To support these results and confirm production of OppZ from the longer precursor transcript, we generated two plasmids carrying either only *oppZ* or *oppF-oppZ* under the control of the

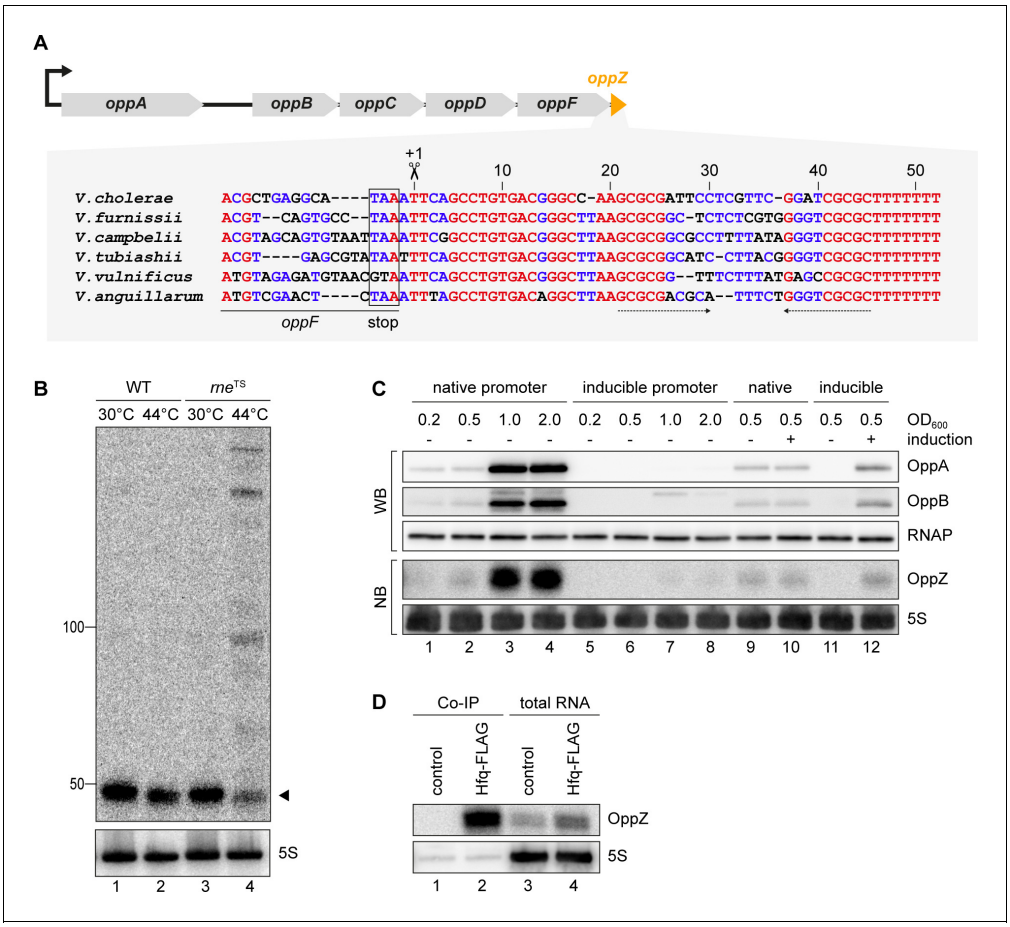


Figure 2. OppZ is produced from the *oppABCDF* 3′ end. (**A**) Top: Genomic organization of *oppABCDF* and *oppZ*. Bottom: Alignment of *oppZ* sequences, including the last codons of *oppF*, from various *Vibrio* species. The *oppF* stop codon, the RNase E cleavage site and the Rho-independent terminator are indicated. (**B**) *V. cholerae* wild-type and *rne*^{TS} strains were grown at 30˚C to stationary phase (OD_{600} of 2.0). Cultures were divided in half and continuously grown at either 30˚C or 44˚C for 30 min. OppZ synthesis was analyzed by Northern blot with 5S rRNA as loading control. The triangle indicates the size of mature OppZ. (**C**) Protein and RNA samples were obtained from *V. cholerae oppA*::3XFLAG *oppB*::3XFLAG strains carrying either the native *oppA* promoter or the inducible pBAD promoter upstream of *oppA*. Samples were collected at the indicated OD_{600} and tested for OppA and OppB production by Western blot and for OppZ expression by Northern blot. RNAP and 5S rRNA served as loading controls for Western and Northern blots, respectively. Lanes 1–8: Growth without L-arabinose. Lanes 9–12: Growth with either H_2O (-) or L-arabinose (+) (0.2% final conc.). (**D**) *V. cholerae* wild-type (control) and *hfq*::3XFLAG (Hfq-FLAG) strains were grown to stationary phase (OD_{600} of 2.0), lysed, and subjected to immunoprecipitation using the anti-FLAG antibody. RNA samples of lysate (total RNA) and co-immunoprecipitated fractions were analyzed on Northern blots. 5S rRNA served as loading control.

The online version of this article includes the following source data and figure supplement(s) for figure 2:

Source data 1. Full Northern and Western blot images for the corresponding detail sections shown in ***Figure 2***.
Figure supplement 1. Hfq dependence of OppZ processing.
Figure supplement 1—source data 1. Full Northern blot images for the corresponding detail sections shown in ***Figure 2—figure supplement 1***.
Figure supplement 2. Hfq dependence of OppZ stability.
Figure supplement 2—source data 1. Quantification of OppZ levels in wild-type and Δhfq cells from Northern blots.

constitutive P_{Tac} promoter (***Figure 2—figure supplement 1A***) and compared OppZ expression in wild-type and Δ*oppZ* cells. Expression of mature OppZ was readily detected from the precursor (***Figure 2—figure supplement 1B***, lane 1 vs. 4) and the size of the processed OppZ transcript was comparable to endogenously expressed OppZ (lane 1) and OppZ transcribed directly by the P_{Tac} promoter (lane 3). We also repeated these experiments in a *V. cholerae hfq* mutant (***Svenningsen et al., 2009***). Here, processing of the precursor into OppZ was still detected

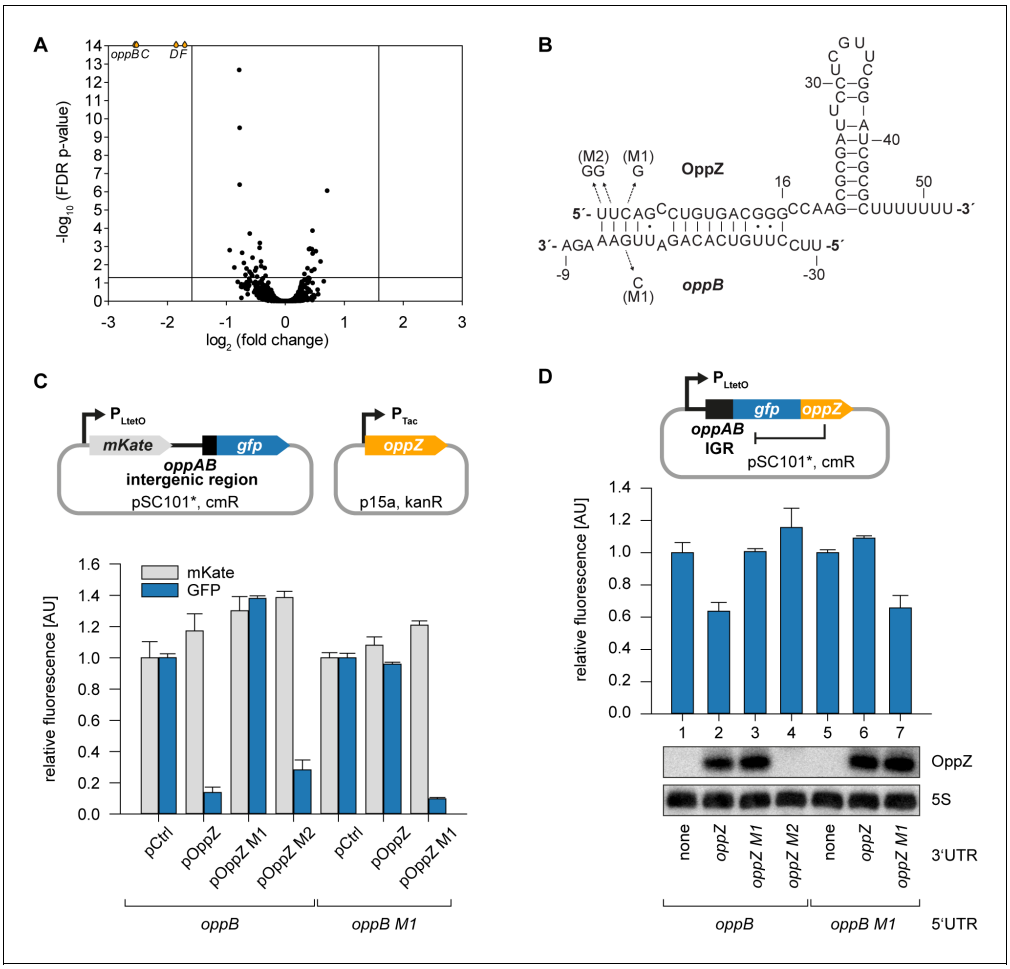


Figure 3. Feedback autoregulation at the suboperonic level. (**A**) Volcano plot of genome-wide transcript changes in response to inducible OppZ over-expression. Lines indicate cut-offs for differentially regulated genes at 3-fold regulation and FDR-adjusted p-value≤0.05. Genes with an FDR-adjusted p-value<10^{-14} are indicated as droplets at the top border of the graph. (**B**) Predicted OppZ secondary structure and base-pairing to *oppB*. Arrows indicate the mutations tested in (**C**) and (**D**). (**C**) *E. coli* strains carrying a translational reporter plasmid with the *oppAB* intergenic region placed between *mKate2* and *gfp* were co-transformed with a control plasmid or the indicated OppZ expression plasmids. Transcription of the reporter and *oppZ* were driven by constitutive promoters. Cells were grown to OD_{600} = 1.0 and fluorophore production was measured. mKate and GFP levels of strains carrying the control plasmid were set to 1. Error bars represent the SD of three biological replicates. (**D**) Single-plasmid regulation was measured by inserting the indicated *oppZ* variant into the 3' UTR of a translational *oppB::gfp* fusion. Expression was driven from a constitutive promoter. *E. coli* strains carrying the respective plasmids were grown to OD_{600} = 1.0 and GFP production was measured. Fluorophore levels from control fusions without an sRNA gene were set to one and error bars represent the SD of three biological replicates. OppZ expression was tested by Northern blot; 5S rRNA served as loading control.

The online version of this article includes the following source data and figure supplement(s) for figure 3:

Source data 1. Full Northern blot images for the corresponding detail sections shown in ***Figure 3*** and raw data for fluorescence measurements.

Figure supplement 1. Pulse expression of OppZ reduces *oppBCDF* transcript levels.

Figure supplement 1—source data 1. Full Northern blot images for the corresponding detail sections shown in ***Figure 3—figure supplement 1*** and raw data for transcript changes as determined by qRT-PCR.

Figure supplement 2. Hfq-dependent, post-transcriptional repression of OppBCDF by OppZ.

Figure supplement 2—source data 1. Full Northern blot images for the corresponding detail sections shown in ***Figure 3—figure supplement 2*** and raw data for fluorescence measurements.

Figure supplement 3. Mutational analysis of the RNase E site in *oppZ*.

Figure supplement 3—source data 1. Full Northern blot images for the corresponding detail sections shown in ***Figure 3—figure supplement 3***.

(lane 8), however, the steady-state levels of OppZ were lower, suggesting that OppZ binds Hfq. Indeed, stability experiments using rifampicin-treated *V. cholerae* showed that OppZ half-life is reduced in Δ*hfq* cells (***Figure 2—figure supplement 2***), and RNA co-immunoprecipitation experiments of chromosomal Hfq::3XFLAG revealed that OppZ interacts with Hfq in vivo (***Figure 2D***). Together, these data show that OppZ is an Hfq-dependent sRNA that is processed from the 3' UTR of the polycistronic *oppABCDF* mRNA by RNase E.

Feedback Autoregulation at the suboperonic level

Hfq-binding sRNAs control gene expression by base-pairing with *trans*-encoded target transcripts (***Kavita et al., 2018***). To determine the targets of OppZ in *V. cholerae*, we cloned the sRNA (starting from the RNase E cleavage site) on a plasmid under the control of the pBAD promoter. Induction of the pBAD promoter for 15 min resulted in a strong increase in OppZ levels (~30 fold, ***Figure 3—figure supplement 1A***) and RNA-seq experiments of the corresponding samples revealed four repressed genes (***Figure 3A*** and ***Figure 3—figure supplement 1B***). Interestingly, these genes were *oppBCDF*, *i.e.* the same transcript that OppZ is processed from. We validated OppZ-mediated repression of all four genes using qRT-PCR (***Figure 3—figure supplement 1C***), which also confirmed that the first gene of the operon, *oppA*, is not affected by OppZ. Despite the reduced transcript levels of *oppBCDF*, OppZ over-expression did not reduce the stability of the *oppB* messenger (***Figure 3—figure supplement 1D***). Using the *RNA-hybrid* algorithm (***Rehmsmeier et al., 2004***), we were able to predict RNA duplex formation of the *oppB* translation initiation site with the 5' end of the OppZ sRNA (***Figure 3B***). We confirmed this interaction using a variant of a previously reported post-transcriptional reporter system (***Corcoran et al., 2012***). Here, the first gene of the operon is replaced by the red-fluorescent mKate2 protein, followed by the *oppAB* intergenic sequence and the first five codons of *oppB,* which were fused to *gfp* (***Figure 3C***, top). Transfer of this plasmid into *E. coli* and co-transformation of the OppZ over-expression plasmid resulted in strong repression of GFP (~7 fold), while mKate2 levels remained constant. Mutation of either OppZ or *oppB* (mutations M1, see ***Figure 3B***) abrogated regulation of GFP and combination of both mutants restored control (***Figure 3C***, bottom). In contrast, OppZ-mediated repression of OppB::GFP was strongly reduced in *E. coli* lacking *hfq* (***Figure 3—figure supplement 2A–B***). We also generated three additional variants of the reporter plasmids in which we included the *oppBC*, *oppBCD*, and *oppBCDF* sequences fused to GFP (***Figure 3—figure supplement 2C***). In all cases, OppZ readily inhibited GFP but did not affect mKate2. These results confirm that OppZ promotes discoordinate expression of the *oppABCDF* operon.

Next, we aimed to reproduce OppZ-mediated repression from a single transcript. To this end, we compared GFP production of a translational *oppB::gfp* reporter with the same construct carrying the *oppZ* sequence downstream of *gfp* (***Figure 3D***, top). Northern blot analysis revealed that OppZ was efficiently clipped off from the *gfp* transcript in this construct and fluorescence measurements showed that OppZ also inhibited GFP expression (***Figure 3D***, bottom, lane 1 vs. 2). We confirmed that this effect is specific to base-pairing of OppZ with the *oppAB* intergenic sequence as we were able to recapitulate our previous compensatory base-pair exchange experiments using the single plasmid system (***Figure 3D***). In addition, mutation of the RNase E recognition site in *oppZ* (UU→GG, mutation M2; ***Figure 3—figure supplement 3A***) blocked OppZ maturation and abolished OppB::GFP repression (***Figure 3D***, lane 4; ***Figure 3—figure supplement 3B***), whereas expression of OppZ M2 from a separate plasmid efficiently reduced OppB:GFP levels (***Figure 3C***). Together, our data demonstrate that OppZ down-regulates protein synthesis from its own cistron. Furthermore, mutation M2 shows that this autoregulation is not mediated by long-distance intramolecular base-pairing of OppZ with the *oppB* 5' UTR, but rather requires RNase E-dependent maturation of the transcript followed by Hfq-dependent base-pairing.

Translational control of OppZ synthesis

The above experiments revealed that OppZ inhibits protein production through feedback control, however, it was not clear if OppZ would also inhibit its own synthesis. To address this question, we generated an OppZ over-expression plasmid in which we mutated the sequence of the terminal stem-loop at eight positions. We call this construct '*regulator OppZ*' (***Figure 4A***). These mutations are not expected to inactivate the base-pairing function of OppZ, but will allow us to differentiate

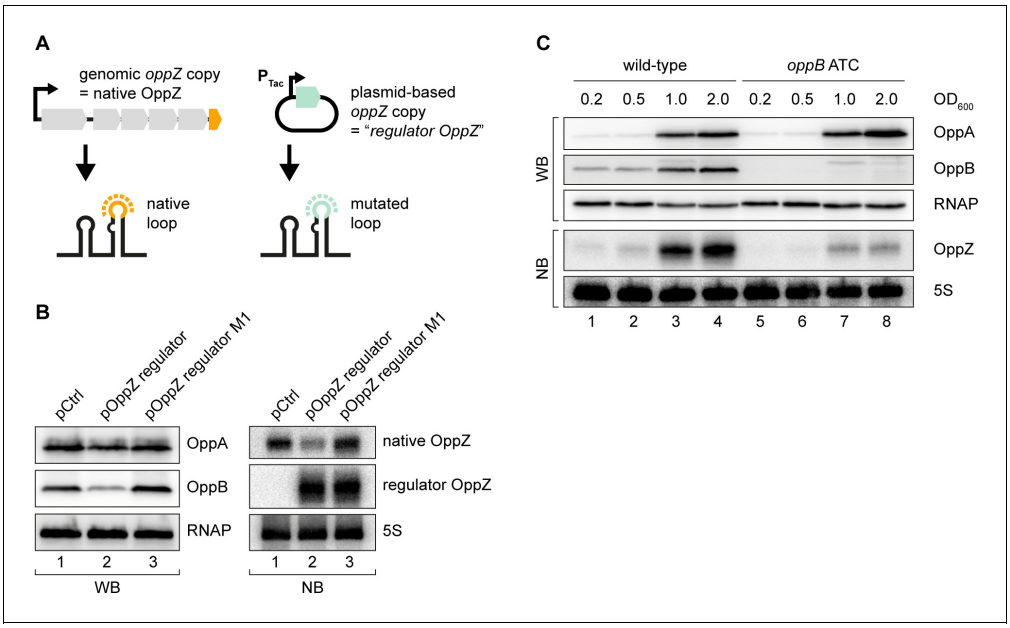


Figure 4. Translational control of OppZ synthesis. (**A**) Schematic of the analyzed OppZ variants containing the native stem loop sequence (produced from the genomic *oppZ* locus) or a mutated stem loop sequence ('*regulator OppZ*' produced from a plasmid-based constitutive promoter). (**B**) *V. cholerae oppA*::3XFLAG *oppB*::3XFLAG carrying a control plasmid (pCMW-1) or a plasmid expressing *regulator OppZ* (pMD194, pMD195) were grown to stationary phase (OD_{600} of 2.0). OppA and OppB production were tested by Western blot and expression of native OppZ and regulator OppZ was monitored on Northern blot using oligonucleotides binding to the respective loop sequence variants. RNAP and 5S rRNA served as loading controls for Western blot and Northern blot, respectively. (**C**) The *oppB* start codon was mutated to ATC in an *oppA*::3XFLAG *oppB*::3XFLAG background. *V. cholerae* strains with wild-type or mutated *oppB* start codon were grown in LB medium. Protein and RNA samples were collected at the indicated OD_{600} and tested for OppA and OppB production by Western blot and for OppZ expression by Northern blot. RNAP and 5S rRNA served as loading controls for Western and Northern blots, respectively.

The online version of this article includes the following source data and figure supplement(s) for figure 4:

Source data 1. Full Northern and Western blot images for the corresponding detail sections shown in ***Figure 4***.

Figure supplement 1. Translational control of OppZ synthesis.

Figure supplement 1—source data 1. Quantification of OppZ levels in wild-type and oppB ATC cells from Northern blots and full blot images for the corresponding detail sections shown in ***Figure 4—figure supplement 1***.

the levels of native OppZ and *regulator OppZ* on Northern blots. Indeed, when tested in *V. cholerae*, over-expression of *regulator OppZ* inhibited OppB::3XFLAG production, but did not affect OppA::3XFLAG levels (***Figure 4B***, left). Importantly, *regulator OppZ* also reduced the expression of native OppZ (***Figure 4B***, right) and introduction of the M1 mutation (see ***Figure 3B***) in *regulator OppZ* abrogated this effect. These results revealed that OppZ also exerts autoregulation of its own transcript.

Gene expression control by sRNAs typically occurs post-transcriptionally (***Gorski et al., 2017***) raising the question of how OppZ achieves autoregulation at the molecular level. Given that OppZ inhibits OppB production (***Figure 4B***), we hypothesized that OppZ synthesis might be linked to *oppB* translation. To test this prediction, we inactivated the chromosomal start codon of *oppB* (ATG→ATC) and monitored OppA/B and OppZ expression by Western and Northern blot, respectively. As expected, mutation of the *oppB* start codon had no effect on OppA::3XFLAG levels, but nullified OppB::3XFLAG production (***Figure 4C***, top). Lack of *oppB* translation also resulted in a strong decrease in OppZ levels (***Figure 4C***, bottom), however, did not change OppZ stability (***Figure 4—figure supplement 1A***). In addition, plasmid-based complementation of OppB::3XFLAG in the *oppB* start codon mutant failed to restore OppZ expression (***Figure 4—figure supplement 1B***), showing that OppZ production is independent of the cellular OppB levels. Based on these and the results above, we propose that autorepression of *oppBCDF-oppZ* must occur by a mechanism involving both translation inhibition, as well as transcription termination.

OppZ promotes transcription termination through Rho

To explain the reduction of OppZ expression in the absence of *oppB* translation, we considered premature transcription termination as a possible factor. This hypothesis was supported by our finding that OppZ over-expression efficiently reduced *oppB* mRNA levels without significantly affecting transcript stability (**Figure 3—figure supplement 1C–D**). In *E. coli*, Rho protein accounts for a major fraction of all transcription termination events (**Ciampi, 2006**) and has previously been associated with the regulatory activity of Hfq-dependent sRNAs (**Bossi et al., 2012**; **Sedlyarova et al., 2016**; **Wang et al., 2015**). Rho is specifically inhibited by bicyclomycin (BCM; **Zwiefka et al., 1993**) and consequently we tested the effect of the antibiotic on OppZ expression in *V. cholerae* wild-type and the *oppB* start codon mutant. Whereas BCM had no effect on OppZ synthesis in wild-type cells (**Figure 5A**, lane 1 vs. 2), it strongly increased OppZ and *oppBCDF* expression in the absence of *oppB* translation (**Figure 5A**, lane 3 vs. 4, and **Figure 5B**). We confirmed these results by employing Term-Seq analysis (**Dar et al., 2016**) to wild-type and *oppB* start codon mutants cultivated with or without BCM. Detailed inspection of transcript coverage at the *oppABCDF-oppZ* genomic locus showed that lack of *oppB* translation down-regulated the expression of *oppBCDF-oppZ*, while presence of BCM suppressed this effect (**Figure 5C** and **Supplementary file 3B**). Similarly, inhibition of the *oppBCDF* mRNA and OppZ by over-expression of *regulator OppZ* (see **Figure 4A**) was suppressed in the presence of BCM, whereas OppB protein levels remained low presumably due to continued repression of *oppB* translation initiation by OppZ (**Figure 5D–E**).

To map the position of Rho-dependent transcription termination in *oppB*, we generated five additional strains carrying a STOP mutation at the 2nd, 15th, 65th, 115th, or 215th codon of the chromosomal *oppB* gene (**Figure 6A**). In addition, we mutated the start codons of *oppC*, *oppD*, and *oppF* and probed OppZ levels on Northern blot (**Figure 6B**). In accordance with the data presented in **Figure 4C**, mutation of the *oppB* start codon resulted in strongly decreased OppZ levels (**Figure 6B**, lane 1 vs. 2) and we observed similar results when the STOP mutation was introduced at the 2nd, 15th, and 65th codon of *oppB* (**Figure 6B**, lanes 3–5). In contrast, a STOP mutation at codon 115 led to increased OppZ expression (lane 6) and OppZ levels were fully restored when the STOP was placed at codon 215 of *oppB* (lane 7). Likewise, mutation of the *oppC*, *oppD*, and *oppF* start codons had no effect on OppZ production (**Figure 6B**, lanes 8–10). To summarize, our data indicate that autorepression of the *oppBCDF-oppZ* genes relies on inhibition of *oppB* translation initiation by OppZ, which triggers Rho-dependent transcription termination in the distal part of the *oppB* sequence.

CarZ is another autoregulatory sRNA from *V. cholerae*

Our TIER-seq analysis revealed 17 3′ UTR-derived sRNAs produced by RNase E-mediated cleavage in *V. cholerae* (**Supplementary file 2A**). Detailed analysis of OppZ showed that this sRNA serves as an autoregulatory element inhibiting the *oppBCDF* genes as well as its own synthesis (**Figures 4–6**). We therefore asked how wide-spread RNA-mediated autoregulation is and if the other 16 3′ UTR-derived sRNAs might serve a similar function in *V. cholerae*. To this end, we searched for potential base-pairing sequences between the sRNAs and the translation initiation regions of their associated genes using the *RNA-hybrid* algorithm (**Rehmsmeier et al., 2004**). Indeed, we were able to predict stable RNA duplex formation between the Vcr084 sRNA (located in the 3′ UTR of the *carAB* operon; encoding carbamoyl phosphate synthetase) and the 5′ UTR of *carA*, which is the first gene of the operon (**Figure 7A–B**). In analogy to OppZ, we named this sRNA CarZ. Plasmid-borne expression of CarZ strongly inhibited GFP production from *carA::gfp* and *carAB::gfp* reporters in *E. coli* (**Figure 7—figure supplement 1A–B**) and we obtained similar results using a single transcript *carA::gfp::carZ* construct (**Figure 7C**). CarZ binds Hfq in vivo (**Figure 7—figure supplement 1C**) and repression of *carA::gfp* by CarZ requires Hfq, possibly due to reduced CarZ levels in the *hfq* mutant (**Figure 7—figure supplement 1D–E**). We validated the predicted interaction using compensatory base-pair exchange experiments (**Figure 7B–C**, **Figure 7—figure supplement 1A–B**). Transcription of *carAB-carZ* is controlled by a single promoter located upstream of *carA* and the three genes are co-expressed in vivo (**Figure 7D** and **Papenfort et al., 2015b**). These results suggested that CarZ provides feedback regulation and using an experimental strategy analogous to **Figure 4A**, we were able to show that CarZ inhibits CarA and CarB protein expression as well as its own synthesis (**Figure 7B,E**). Furthermore, introduction of a STOP codon at the 2nd codon of the chromosomal

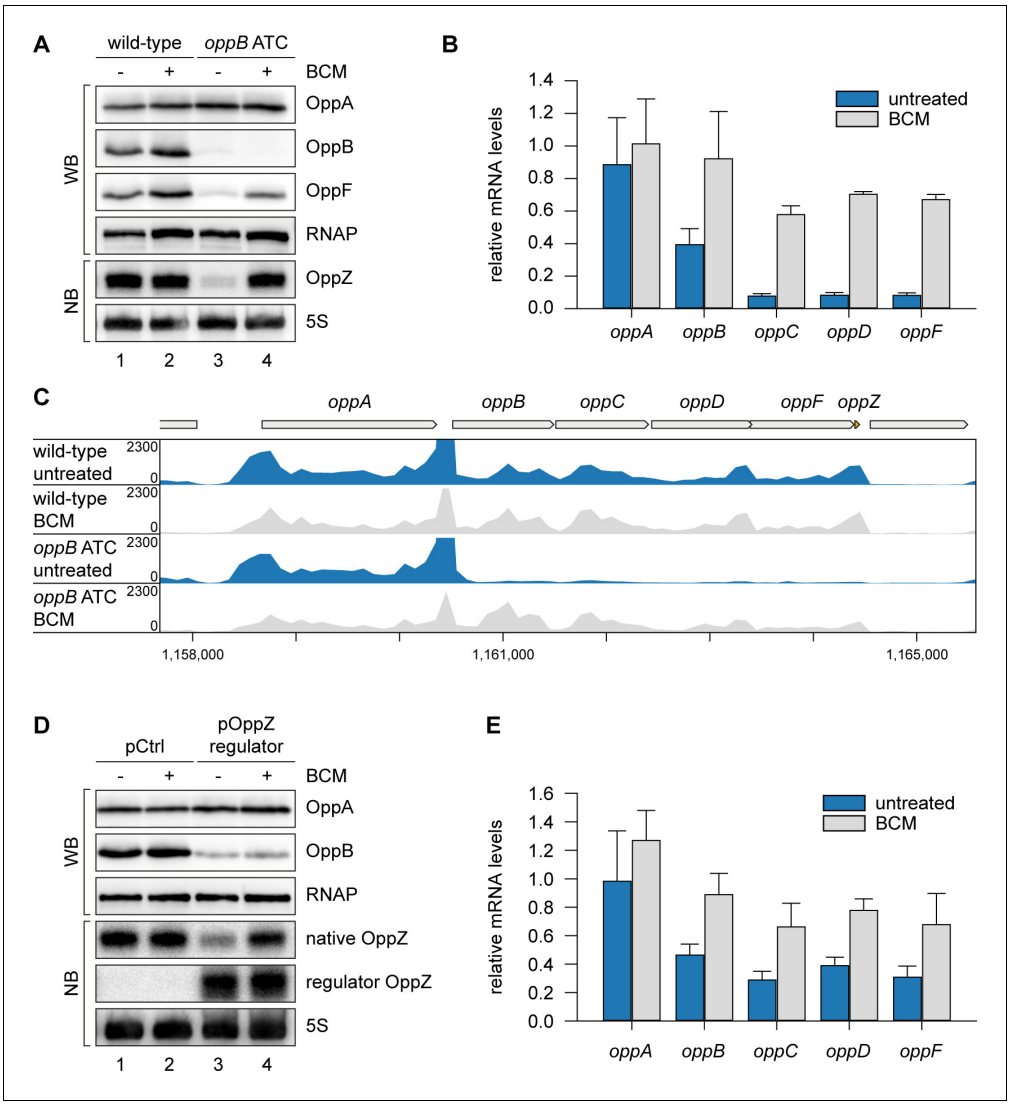


Figure 5. OppZ promotes transcription termination through Rho. (**A**) *V. cholerae oppA*::3XFLAG *oppB*::3XFLAG *oppF*::3XFLAG strains with wild-type or mutated *oppB* start codon were grown to early stationary phase (OD_{600} of 1.5). Cultures were divided in half and treated with either H_2O or BCM (25 µg/ml final conc.) for 2 hr before protein and RNA samples were collected. OppA, OppB and OppF production were tested by Western blot and OppZ expression was monitored by Northern blot. RNAP and 5S rRNA served as loading controls for Western and Northern blots, respectively. (**B**) Biological triplicates of *V. cholerae oppA*::3XFLAG *oppB*::3XFLAG strains with wild-type or mutated *oppB* start codon were treated with BCM as described in (**A**). *oppABCDF* expression in the *oppB* start codon mutant compared to the wild-type control was analyzed by qRT-PCR. Error bars represent the SD of three biological replicates. (**C**) Triplicate samples from (**B**) were subjected to Term-seq and average coverage of the *opp* operon is shown for one representative replicate. The coverage cut-off was set at the maximum coverage of annotated genes. (**D**) *V. cholerae oppA*::3XFLAG *oppB*::3XFLAG strains carrying a control plasmid (pMD397) or a plasmid expressing *regulator OppZ* (pMD398) were treated with BCM as described in (**A**). OppA and OppB production were tested by Western blot and expression of native OppZ and regulator OppZ was monitored on Northern blot using oligonucleotides binding to the respective loop sequence variants. RNAP and 5S rRNA served as loading controls for Western and Northern blots, respectively. (**E**) Levels of *oppABCDF* in the experiment described in (**D**) were analyzed by qRT-PCR. Error bars represent the SD of three biological replicates. The online version of this article includes the following source data for figure 5:

Source data 1. Full blot images for the corresponding detail sections shown in ***Figure 5*** and raw data for transcript changes as determined by qRT-PCR.

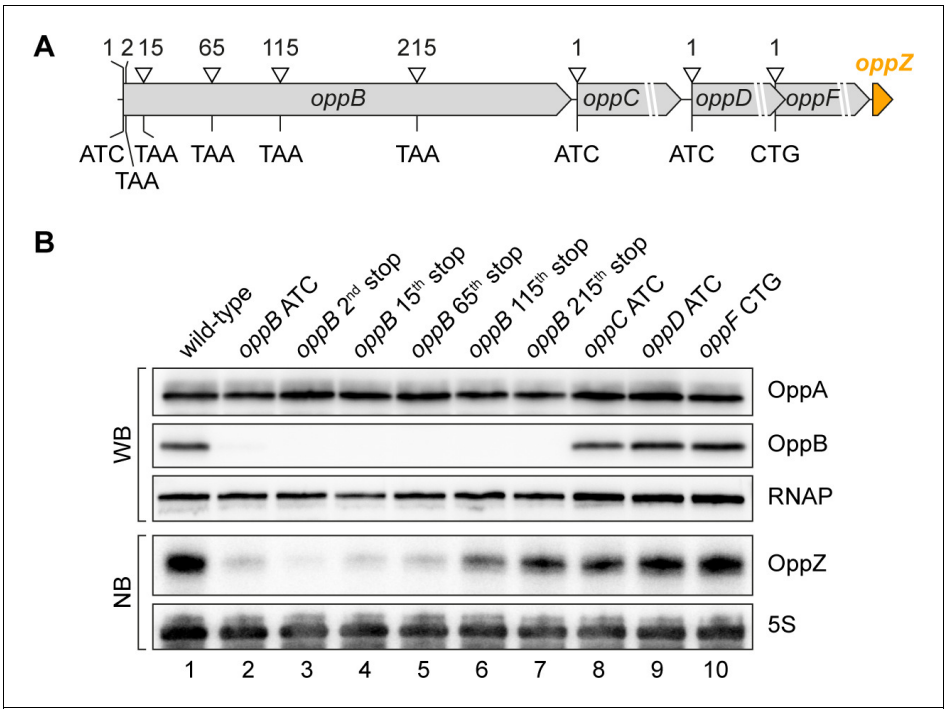


Figure 6. Influence of OppBCDF translation on OppZ expression. (**A**) The depicted mutations were individually inserted into the *opp* locus to inactivate the start codons of *oppB*, *oppC*, *oppD* or *oppF* or to insert STOP codons at the positions 2, 15, 65, 115 or 215 of *oppB*. (**B**) *V. cholerae oppA*::3XFLAG *oppB*::3XFLAG strains with the described *opp* mutations were grown: wild-type (lane 1), the *oppB* start codon mutated (lane 2), a STOP codon inserted at the 2nd, 15th, 65th, 115th or 215th codon of *oppB* (lanes 3–7) or mutated start codons of *oppC*, *oppD* or *oppF* (lanes 8–10). At stationary phase (OD_{600} of 2.0), protein and RNA samples were collected and tested for OppA and OppB production by Western blot and for OppZ expression by Northern blot. RNAP and 5S rRNA served as loading controls for Western and Northern blots, respectively.
The online version of this article includes the following source data for figure 6:

Source data 1. Full Northern and Western blot images for the corresponding detail sections shown in ***Figure 6***.

carA gene abrogated CarZ expression and similar results were obtained when the STOP codon was placed at the 2nd codon of *carB* (***Figure 7F***). Of note, inactivation of *carA* translation also blocked CarB production indicating, among other possibilities, that translation of the two ORFs might be coupled and that expression of CarZ relies on active translation of both ORFs. Together, these results provide evidence that CarZ is an autoregulatory sRNA and suggest that this function might be more wide-spread among the growing class of 3' UTR-derived sRNAs.

Autoregulatory sRNAs modify the kinetics of gene induction

Bacterial sRNAs acting at the post-transcriptional level have recently been reported to add unique features to gene regulatory circuits, including the ability to promote discoordinate operon expression (***Nitzan et al., 2017***). Plasmid-borne over-expression of OppZ resulted in decreased expression of the *oppBCDF* cistrons, while leaving *oppA* levels unaffected (***Figure 3—figure supplement 1B–C***). We therefore asked if OppZ expression had a similar effect on the production of their corresponding proteins. To this end, we cultivated wild-type and *oppZ*-deficient *V. cholerae* (both carrying a control plasmid), as well as Δ*oppZ* cells carrying an OppZ over-expression plasmid, to various stages of growth and monitored OppA and OppB levels on Western Blot (***Figure 8—figure supplement 1A***). Quantification of the results revealed a moderate increase in OppB expression (~1.8 fold) in cells lacking *oppZ* and ~5 fold decreased OppB levels when OppZ was over-expressed. Neither lack of *oppZ*, nor OppZ over-expression significantly affected OppA production (***Figure 8—figure supplement 1B–C***).

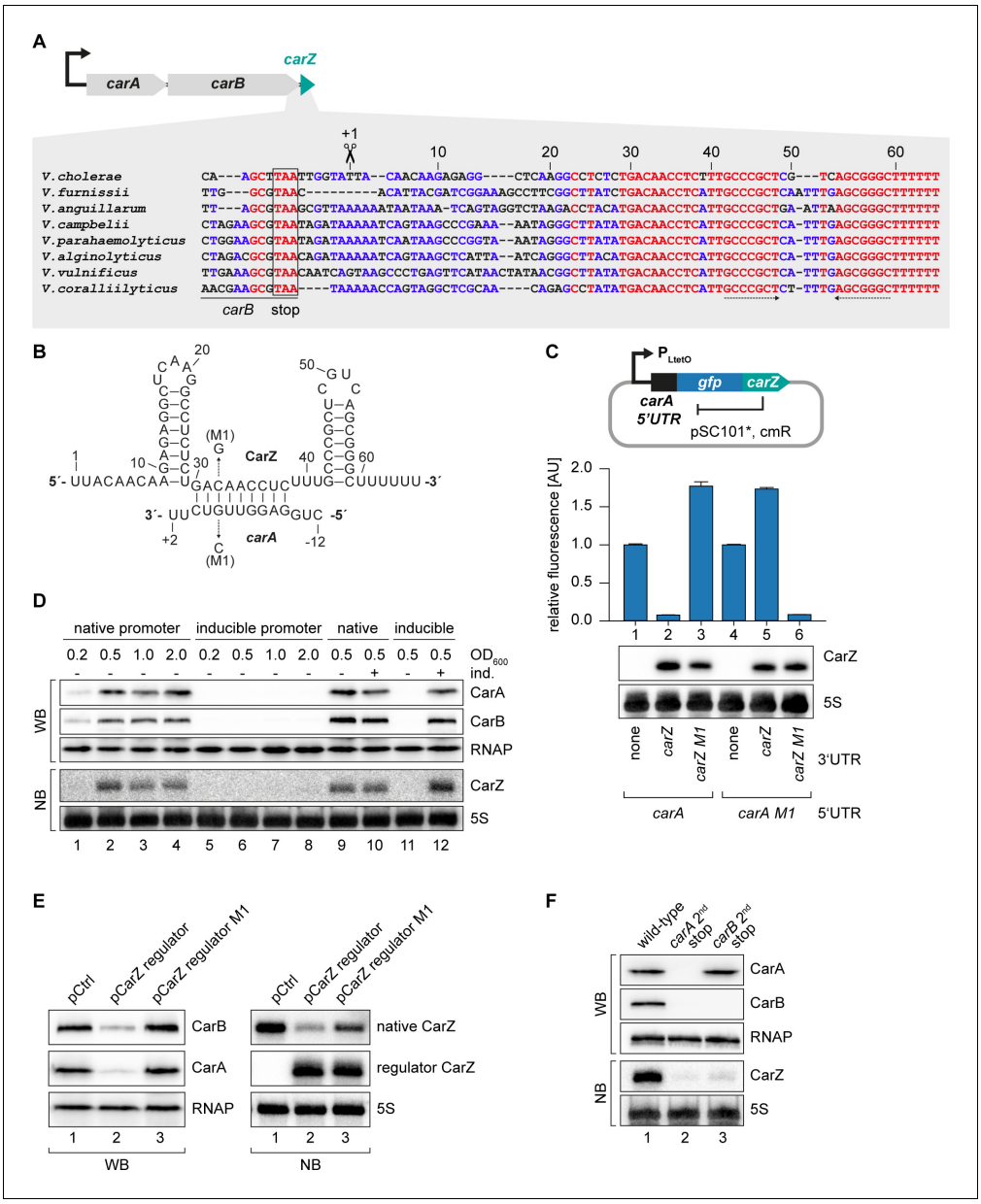


Figure 7. CarZ is another autoregulatory sRNA from *V. cholerae*. (**A**) Top: Genomic context of *carAB* and *carZ*. Bottom: Alignment of *carZ* sequences, including the last codons of *carB*, from various *Vibrio* species. The *carB* stop codon, the RNase E cleavage site and the Rho-independent terminator are indicated. (**B**) Predicted CarZ secondary structure and base-pairing to *carA*. Arrows indicate the single nucleotide mutations tested in (C). (**C**) Single-plasmid feedback regulation of *carA* by CarZ was measured by inserting the indicated *carZ* variant into the 3′ UTR of a translational *carA::gfp* fusion. Expression was driven from a constitutive promoter. *E. coli* strains carrying the respective plasmids were grown to OD_{600} = 1.0 and GFP production was measured. Fluorophore levels from control fusions without an sRNA gene were set to one and error bars represent the SD of three biological replicates. CarZ expression was tested by Northern blot; 5S rRNA served as loading control. (**D**) Protein and RNA samples were obtained from *V. cholerae carA*::3XFLAG *carB*::3XFLAG carrying either the native *carA* promoter or the inducible pBAD promoter upstream of *carA*. Samples were collected at the indicated OD_{600} and tested for CarA and CarB production by Western blot and for CarZ expression by Northern blot. RNAP and 5S rRNA served as loading controls for Western and Northern blots, respectively. Lanes 1–8: Growth without L-arabinose. Lanes 9–12: Growth with either H_2O (-) or L-arabinose (+) (0.2% final conc.). (**E**) *V. cholerae carA*::3XFLAG *carB*::3XFLAG strains carrying a control plasmid or a plasmid expressing a CarZ variant with a mutated stem loop (*regulator CarZ*) were grown to late exponential phase (OD_{600} of 1.0). CarA and CarB production were tested by Western blot and expression of native CarZ or regulator CarZ was monitored on Northern blot using oligonucleotides binding to the respective loop sequence

Figure 7 continued on next page

Figure 7 continued

variants. RNAP and 5S rRNA served as loading controls for Western blot and Northern blot, respectively. (**F**) *V. cholerae carA*::3XFLAG *carB*::3XFLAG strains with the following *carA* or *carB* mutations were grown: wild-type (lane 1) or a STOP codon inserted at the 2nd codon of *carA* (lane 2) or *carB* (lane 3), respectively. At late exponential phase (OD_{600} of 1.0), protein and RNA samples were collected and tested for CarA and CarB production by Western blot and for CarZ expression by Northern blot. RNAP and 5S rRNA served as loading controls for Western and Northern blots, respectively.
The online version of this article includes the following source data and figure supplement(s) for figure 7:

Source data 1. Full blot images for the corresponding detail sections shown in ***Figure 7*** and raw data for fluorescence measurements.
Figure supplement 1. Hfq-dependent, post-transcriptional repression of CarA and CarB by CarZ.
Figure supplement 1—source data 1. Full Northern blot images for the corresponding detail sections shown in ***Figure 4—figure supplement 1*** and raw data for fluorescence measurements.
Figure supplement 2. CarZ induces *carAB* degradation.
Figure supplement 2—source data 1. Raw data for transcript changes as determined by qRT-PCR.

Given the relatively mild effect of *oppZ* deficiency on steady-state OppB protein levels (***Figure 8—figure supplement 1A***), we next investigated the role of OppZ on the dynamics of OppABCDF expression. Specifically, transcription factor-controlled negative autoregulation has been reported to affect the response time of regulatory networks (***Rosenfeld et al., 2002***) and we speculated that sRNA-mediated feedback control could have a similar effect. To test this hypothesis, we employed a *V. cholerae* strain in which we replaced the native promoter upstream of the chromosomal *oppA* gene with the L-arabinose-inducible pBAD promoter (see ***Figure 2C***) and monitored the kinetics of OppA and OppB production in wild-type and Δ*oppZ* cells before and at several time-points post induction (***Figure 8A***). Whereas OppA protein accumulated equally in wild-type and *oppZ* mutants (***Figure 8B***), expression of OppB was significantly increased in Δ*oppZ* cells (***Figure 8C***). This effect was most prominent at later stages after induction (>30 min) and coincided with accumulation of OppZ (***Figure 8A***). Calculation of the OppB response time (50% of the maximal expression value) showed a significant delay in Δ*oppZ* cells (~78 min), when compared to the wild-type control (~52 min). We therefore conclude that alike transcription factors, autoregulatory sRNAs change the dynamics of their associated genes, however, in contrast to transcription factors, sRNAs act at the post-transcriptional level and can direct this effect towards a specific subgroup of genes within an operon.

Discussion

Base-pairing sRNAs regulating the expression of *trans*-encoded mRNAs are a major pillar of gene expression control in bacteria (***Gorski et al., 2017***). Transcriptomic data obtained from various microorganisms have shown that sRNAs are produced from almost all genomic loci and that the 3′ UTRs of coding genes are a hotspot for sRNAs acting through Hfq (***Adams and Storz, 2020***). Expression of 3′ UTR-derived sRNAs can either occur by independent promoters, or by ribonucleolytic cleavage typically involving RNase E (***Miyakoshi et al., 2015***). In the latter case, production of the sRNA is intimately connected to the activity of the promoter driving the expression of the upstream mRNA, suggesting that the regulatory function of the sRNA is linked to the biological role of the associated genes. Indeed, such functional interdependence has now been demonstrated in several cases (***Chao and Vogel, 2016***; ***De Mets et al., 2019***; ***Huber et al., 2020***; ***Miyakoshi et al., 2019***; ***Wang et al., 2020***), however, it remained unclear if and how these sRNAs also affected their own transcripts. In this regard, OppZ and CarZ provide a paradigm for 3′ UTR-derived sRNAs allowing autoregulation at the post-transcriptional level. This new type of feedback inhibition is independent of auxiliary transcription factors and we could show that autoregulation by sRNAs can either involve the full transcript (CarZ), or act at the suboperonic level (OppZ).

Features of RNase E-mediated gene control

RNase E is a principal factor for RNA turnover in almost all Gram-negative bacteria (***Bandyra and Luisi, 2018***). The protein forms a tetramer in vivo and serves as the scaffold for the degradosome, a large, multi-enzyme complex typically containing the phosphorolytic exoribonuclease PNPase, the RNA-helicase RhlB, and the glycolytic enzyme enolase (***Ait-Bara and Carpousis, 2015***). Substrates of RNase E are preferentially AU-rich and harbor a 5′ mono-phosphate. Thus, the enzyme relies on

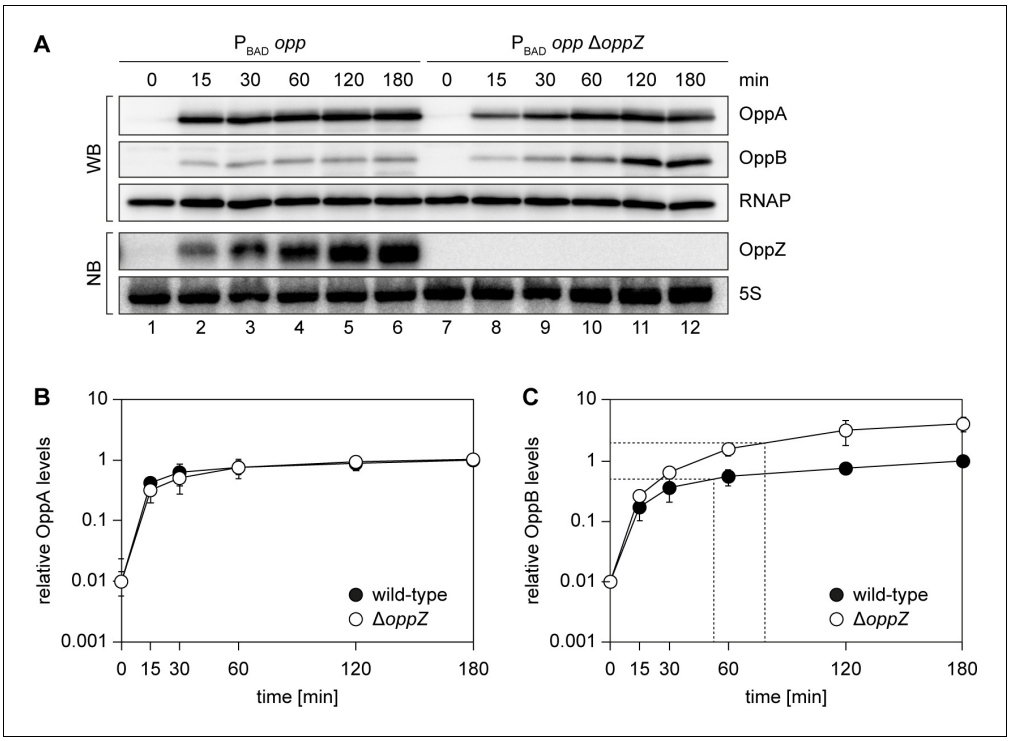


Figure 8. Modified kinetics of gene induction by autoregulatory OppZ. (**A**) Expression of the *opp* operon including the *oppA*::3XFLAG and *oppB*::3XFLAG genes and the native *oppZ* gene (lanes 1–6) or an *oppZ* deletion (lanes 7–12) was induced from the pBAD promoter at late exponential phase (OD_{600} of 1.0) by the addition of L-arabinose (0.2% final conc.). Protein and RNA samples were obtained at the indicated time points and tested for OppA and OppB production by Western blot and for OppZ expression by Northern blot. RNAP and 5S rRNA served as loading controls for Western and Northern blots, respectively. (**B, C**) Quantification of OppA (**B**) or OppB (**C**) levels from the experiment in (**A**); error bars represent the SD of three biological replicates. Data are presented as fold regulation of OppA or OppB in Δ*oppZ* compared to the wild-type. Dashed lines in (**C**) indicate the time points of half-maximum OppB expression.

The online version of this article includes the following source data and figure supplement(s) for figure 8:

Source data 1. Quantification of OppAB protein levels from Western blots and full blot images for the corresponding detail sections shown in ***Figure 8***.

Figure supplement 1. OppZ-dependent repression of OppA and OppB protein levels.

Figure supplement 1—source data 1. Quantification of OppAB protein levels from Western blots and full blot images for the corresponding detail sections shown in ***Figure 8—figure supplement 1***.

RNA pyrophosphohydrolases such as RppH, which convert the 5' terminus from a triphosphate to a monophosphate, before transcript degradation can be initiated (***Deana et al., 2008***). Recognition of a substrate is followed by scanning of RNase E for suitable cleavage sites along the transcript (***Richards and Belasco, 2019***). TIER-seq-based identification of a consensus sequence for RNase E target recognition revealed highly similar motifs for *V. cholerae* (***Figure 1D***) and *S. enterica* (***Chao et al., 2017***). These results further support the previously proposed 'U_{+2} Ruler-and-Cut' mechanism, in which a conserved uridine located two nts down-stream of the cleavage site is key for RNase E activity. However, in contrast to the data obtained from *S. enterica*, we discovered only a mild enrichment of RNase E cleavage sites occurring at translational stop codons (***Figure 1—figure supplement 3A***). This observation might be explained by differences in stop codon usage between *V. cholerae* and *S. enterica* (***Korkmaz et al., 2014***) and could point to species-specific features of RNase E activity.

The role of termination factor Rho in sRNA-mediated gene expression control

Approximately 25–30% of all genes in *E. coli* depend on Rho for transcription termination (**Cardinale et al., 2008**; **Dar and Sorek, 2018b**; **Peters et al., 2012**). BCM treatment of *V. cholerae* wild-type cells revealed 699 differentially regulated genes (549 upregulated and 150 repressed genes; **Supplementary file 3A**), suggesting an equally global role for Rho in this organism. Rho-dependent transcription termination is modulated by various additional factors (**Mitra et al., 2017**). This includes anti-termination factors such as NusG, as well as Hfq and its associated sRNAs (**Bossi et al., 2020**). For sRNAs, the effect on Rho activity can be either activating or repressing. Previous work has shown that sRNAs can mask Rho-dependent termination sites and thereby promote transcriptional read-through (**Lin et al., 2019**; **Sedlyarova et al., 2016**). Negative gene regulation involving sRNAs and Rho typically includes translation inhibition by the sRNA resulting in separation of transcription and translation complexes (**Figure 9**). Coupling of transcription and translation normally protects the nascent mRNA from Rho action and loss of ribosome binding supports

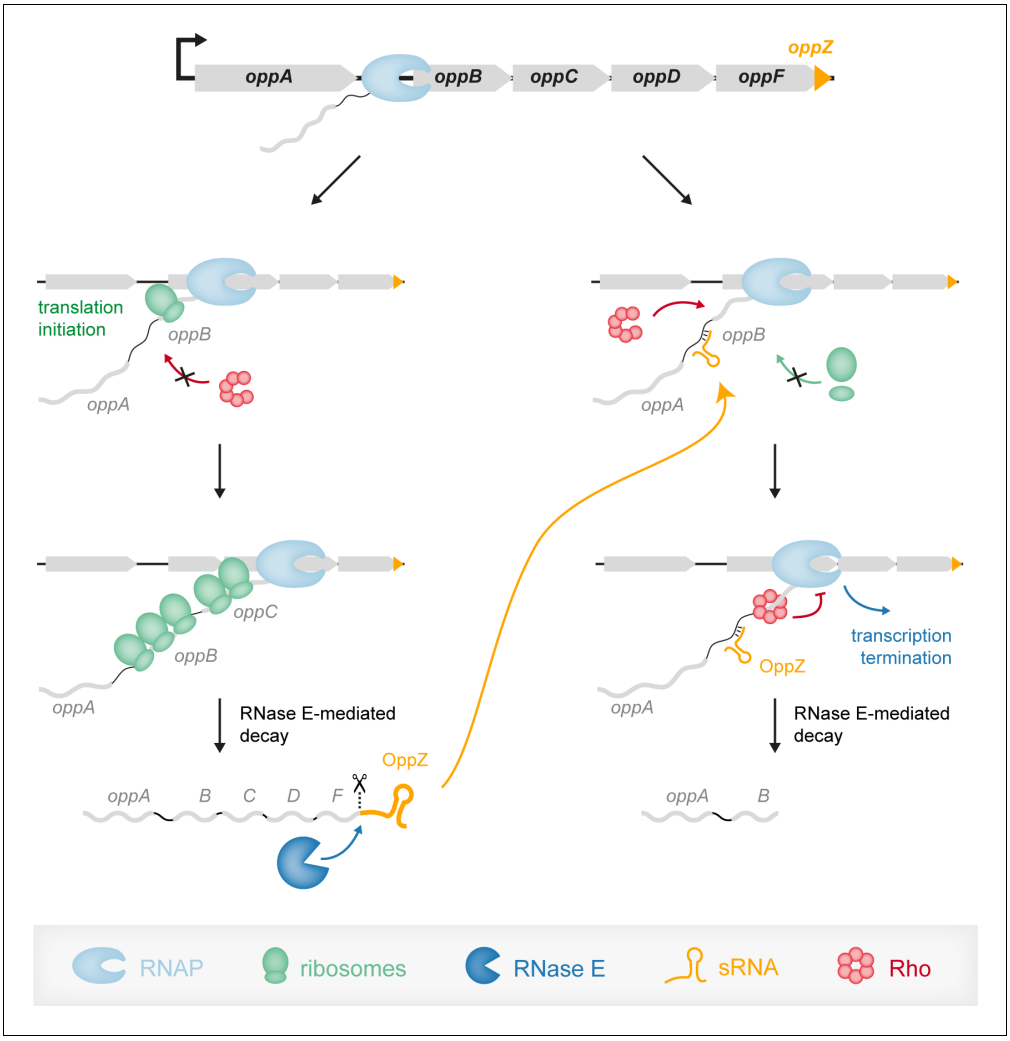


Figure 9. Model of the OppZ-dependent mechanism of *opp* regulation. Transcription of the *oppABCDF* operon initiates upstream of *oppA* and in the absence of OppZ (left) involves all genes of the operon as well as OppZ. In this scenario, all cistrons of the operon are translated. In the presence of OppZ (right), the sRNA blocks translation of *oppB* and the ribosome-free mRNA is recognized by termination factor Rho. Rho catches up with the transcribing RNAP and terminates transcription pre-maturely within *oppB*. Consequently, *oppBCDF* are not translated and OppZ is not produced.

transcription termination (**Bossi et al., 2012**). In addition, lack of ribosome-mediated protection can render the mRNA target vulnerable to ribonucleases, *e.g.* RNase E, which can also lead to the degradation of the sRNA (**Feng et al., 2015**; **Massé et al., 2003**). Which of these mechanisms are at play for a given sRNA-target mRNA pair is most often unknown and it is likely that both types of regulation can occur either independently or in concert. For example, over-expression of OppZ did not affect *oppB* transcript stability (**Figure 3—figure supplement 1D**), suggesting that induction of Rho-mediated transcription termination is the main mechanism for gene repression in this sRNA-target mRNA pair. In contrast, analogous experiments testing the stability of the *carA* and *carB* transcripts upon CarZ over-expression revealed a significant drop in transcript stability for both mRNAs (**Figure 7—figure supplement 2A–B**). These results suggest that translation inhibition of *carA* by CarZ has two outcomes: 1st) accelerated ribonucleolytic decay of the *carAB* transcript and 2nd) Rho-mediated transcription termination. Using two regulatory mechanisms (CarZ-*carA*) instead of one (OppZ-*oppB*) might explain the strong inhibition of *carA::gfp* by CarZ (~10 fold, **Figure 7C**), when compared to the relatively weak repression (1.8-fold) of *oppB::gfp* by OppZ (**Figure 3D**).

Employing multiple regulatory mechanisms on one target mRNA might have led to an underestimation of the prevalence of Rho-mediated transcription termination in sRNA-mediated gene control. In fact, sRNAs frequently repress genes that are downstream in an operon with their base-pairing target, which could point to a possible involvement of Rho (**Bossi et al., 2020**). Rho is known to bind cytosine-rich RNA elements (**Allfano et al., 1991**), however, due to the strong variability in size and composition of these sequences, predicting Rho binding sites (a.k.a. *rut* sites) from genomic or transcriptomic data has been a difficult task (**Nadiras et al., 2018**). Indeed, while our transcriptomic data of the *oppB* start codon mutant did not allow us to pinpoint the position of the *rut* site in *oppB* (**Figure 5C**), evidence obtained from genetic analyses using various *oppB* STOP codon mutants revealed that Rho-dependent termination likely occurs at or close to codon 115 in *oppB* (**Figure 6B**). We attribute the lack of this termination event in the transcriptomic data to the activity of 3′−5′ acting exoribonucleases (*e.g.* RNase II or PNPase **Bechhofer and Deutscher, 2019**; **Mohanty and Kushner, 2018**), which degrade the untranslated *oppB* sequence. Identifying the relevant exonucleases might well allow for an advanced annotation of global Rho-dependent termination sites and cross-comparison with documented sRNA-target interaction could help to clarify the relevance of Rho-mediated termination in sRNA-based gene control.

Dynamics of RNA-based feedback regulation

Transcription factors and sRNAs are the principal components of gene networks. While the regulatory outcome of sRNA and transcription factor activity is often very similar, the underlying regulatory dynamics are not (**Hussein and Lim, 2012**). Regulatory networks involving sRNAs and transcription factors are called mixed circuits and have now been studied in greater detail. Similar to systems relying on transcription factors, feedback regulation is common among sRNAs (**Nitzan et al., 2017**). However, unlike the examples presented in this study, these circuits always involve the action of a transcription factor, which has implications for their regulatory dynamics. For example, the OmpR transcription factor activates the expression of the OmrA/B sRNAs, which repress their own synthesis by inhibiting the *ompR-envZ* mRNA (**Guillier and Gottesman, 2008**). This constitutes an autoregulatory loop, however, given that transcription of OmrA/B ultimately relies on OmpR protein levels, this regulation will only become effective when sufficient OmpR turn-over has been achieved (**Brosse et al., 2016**). In contrast, autoregulatory circuits involving 3′ UTR-derived sRNAs are independent of such auxiliary factors and therefore provide a more rapid response. In case of OppZ-*oppB*, we showed that the sRNA has a rapid effect on OppB expression levels (**Figure 8C**) and given the involvement of Rho-mediated transcription termination in this process, we expect similar dynamics for OppZ autoregulation (**Figure 9**).

Another key difference between feedback regulation by transcription factors and 3′ UTR-derived sRNAs is the stoichiometry of the players involved. In transcription factor-based feedback loops, the mRNA coding for the autoregulatory transcription factor can go through multiple rounds of translation, which will lead to an excess of the regulator over the target promoter. The degree of autoregulation is then determined by the cellular concentration of the transcription factor and the affinity towards its own promoter (**Rosenfeld et al., 2002**). In contrast, autoregulatory sRNAs which are generated by ribonucleolytic cleavage come at a 1:1 stoichiometry with their targets. However, this situation changes when the sRNA controls multiple targets. For OppZ, we have shown that

oppBCDF is the only transcript regulated by the sRNA (**Figure 3A**) and we currently do not know if CarZ has additional targets besides *carAB*. In addition, not all sRNA-target interactions result in changes in transcript levels as previously reported for the interaction of the Qrr sRNAs with the *luxO* transcript (**Feng et al., 2015**). New technologies, for example RIL-Seq (**Melamed et al., 2020**; **Melamed et al., 2016**), capturing the global interactome of base-pairing sRNAs independent of their regulatory state could help to address this question and clarify the stoichiometric requirements for sRNA-mediated autoregulation.

Possible biological relevance of autoregulatory sRNAs

Autoregulation by 3′ UTR-derived sRNAs allows for discoordinate operon expression, which is in contrast to their transcription factor counterparts. This feature might be particularly relevant for long mRNAs containing multiple cistrons, such as *oppABCDF*. The *oppABCDF* genes encode an ABC transporter allowing high affinity oligopeptide uptake (**Hiles et al., 1987**). OppBCDF constitute the membrane-bound, structural components of the transport system, whereas OppA functions as a periplasmic binding protein. The overall structure of the transporter requires each one unit of OppB, OppC, OppD, and OppF, while OppA does constitutively interact with the complex and typically accumulates to higher concentrations in the periplasm (**Doeven et al., 2004**). Given that transcription of *oppABCDF* is controlled exclusively upstream of *oppA* (**Figure 2C** and **Papenfort et al., 2015b**), OppZ-mediated autoregulation of *oppBCDF* (rather than the full operon) might help to achieve equimolar concentrations of OppB, OppC, OppD, and OppF in the cell without affecting OppA production.

The *carAB* genes, which are repressed by CarZ, encode carbamoyl phosphate synthetase; an enzyme complex catalyzing the first step in the separate biosynthetic pathways for the production of arginine, and pyrimidine nucleotides (**Castellana et al., 2014**). Similar to OppBCDF, the CarAB complex contains one subunit of CarA and one subunit of CarB. Transcriptional control of *carAB* is complex and controlled by several transcription factors integrating information from purine, pyrimidine, and arginine pathways (**Charlier et al., 2018**). While the exact biological role of CarZ-mediated feedback regulation of *carAB* requires further investigation, transcription factor-based feedback regulation has been reported to reduce transcriptional noise (**Alon, 2007**), which could also be an important feature of sRNA-mediated autoregulation. The OppZ and CarZ sRNAs identified in this study now provide the framework to test this prediction.

Orthogonal use of gene autoregulation by 3′ UTR-derived sRNAs

Regulatory RNAs have now been established as powerful components of the synthetic biology toolbox (**Qi and Arkin, 2014**). RNA regulators are modular, versatile, highly programmable, and therefore ideal candidates for synthetic biology approaches. Similarly, autoregulatory loops using transcriptional repressors find ample use in synthetic regulatory circuits (**Afroz and Beisel, 2013**). While it might be counterintuitive for a transcript to also produce its own repressor, negative feedback regulation has been reported to endow regulatory networks with improved robustness when disturbances to the system are imposed. Hfq-binding sRNAs providing feedback control have recently also been demonstrated to efficiently replace transcriptional regulation in artificial genetic circuits (**Kelly et al., 2018**). However, these sRNAs were produced from separate genes and therefore required additional transcriptional input, which increases noise. In contrast, the autoregulatory sRNAs presented here are produced by ribonucleolytic cleavage and we have shown that both OppZ and CarZ are efficiently clipped off from foreign genes, such as *gfp* (**Figure 3—figure supplement 3**, **Figure 7C**). We therefore propose that autoregulatory sRNAs can be attached to the 3′ UTR of other genes as well, offering a simple and highly modular concept to introduce autoregulation into a biological system. These circuits can be further tuned by modifying the base-pairing strength of the RNA duplex formed between the sRNA and the target, as well as the introduction of Rho-dependent termination events. The latter could be used to avoid over-production of the sRNA, which will further shape the regulatory dynamics of the system. Given that transcriptomic analyses have revealed thousands of stable 3′ UTR RNA tails derived from human transcripts (**Gruber and Zavolan, 2019**; **Malka et al., 2017**), we believe that RNA-based gene autoregulation also could be present and find applications in higher organisms.

Materials and methods

Key resources table

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Strain, strain background (*Escherichia coli*)	See ***Supplementary file 4***	This study		See ***Supplementary file 4***
Strain, strain background (*Vibrio cholerae*)	See ***Supplementary file 4***	This study		See ***Supplementary file 4***
Recombinant DNA reagent (plasmids)	See ***Supplementary file 4***	This study		See ***Supplementary file 4***
Sequence-based reagent (oligonucleotides)	See ***Supplementary file 4***	This study		See ***Supplementary file 4***
Antibody	ANTI-FLAG M2 antibody (mouse monoclonal)	Sigma-Aldrich	Cat#F1804; RRID:AB_262044	(Western blot 1:1.000)
Antibody	RNA Polymerase alpha antibody 4RA2 (rabbit monoclonal)	BioLegend	Cat#WP003; RRID:AB_2687386	(1:10.000)
Antibody	anti-mouse IgG HRP (goat polyclonal)	ThermoFischer	Cat#31430; RRID:AB_228307	(1:10.000)
Antibody	anti-rabbit IgG HRP (goat polyclonal)	ThermoFischer	Cat#A16104; RRID:AB_2534776	(1:10.000)
Commercial assay or kit	TURBO DNA-free Kit	Invitrogen	Cat#AM1907	
Commercial assay or kit	NEBNext Ultra II Directional RNA Library Prep Kit for Illumina	NEB	Cat#E7760	
Commercial assay or kit	Ribo-Zero rRNA Removal Kit (Gram-Negative Bacteria)	Illumina	Cat#MRZGN126	
Chemical compound, drug	Protein G Sepharose	Sigma-Aldrich	Cat##P3296	
Chemical compound, drug	Bicyclomycin (BCM)	SantaCruz Biotech.	Cat#sc-391755; CAS ID: 38129-37-2	
Software, algorithm	MultAlin	***Corpet, 1988*** (PMID:2849754)		http://multalin.toulouse.inra.fr/multalin
Software, algorithm	RNAhybrid	***Rehmsmeier et al., 2004*** (PMID:15383676)		http://bibiserv2.cebitec.uni-bielefeld.de RRID:SCR_003252
Software, algorithm	CLC Genomics Workbench	Qiagen		https://qiagenbioinformatics.com RRID:SCR_011853
Software, algorithm	SigmaPlot	SYSTAT		https://systatsoftware.com RRID:SCR_003210

Continued on next page

Continued

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Software, algorithm	GelQuantNET	biochemlabsolutions		http://biochemlabsolutions.com/GelQuantNET.html RRID:SCR_015703
Software, algorithm	BIO-1D	VILBER		http://vilber.de/en/products/analysis-software
Software, algorithm	ImageJ	***Schneider et al., 2012*** (PMID:22930834)		https://imagej.nih.gov/ij/ RRID:SCR_003070
Software, algorithm	cutadapt	***Martin, 2011***		https://doi.org/10.14806/ej.17.1.200
Software, algorithm	READemption	***Förstner et al., 2014*** (PMID:25123900)		https://doi.org/10.5281/zenodo.591469
Software, algorithm	DESeq2	***Love et al., 2014*** (PMID:25516281)		http://www.bioconductor.org/packages/release/bioc/html/DESeq2.html
Software, algorithm	RNAfold	***Lorenz et al., 2011*** (PMID:22115189)		http://www.tbi.univie.ac.at/RNA
Software, algorithm	WebLogo	***Crooks et al., 2004*** (PMID:15173120)		http://weblogo.threeplusone.com/
Software, algorithm	BEDTools	***Quinlan and Hall, 2010*** (PMID:20110278)		http://code.google.com/p/bedtools

Strains, plasmids, and growth conditions

Bacterial strains, plasmids and DNA oligonucleotides used in this study are listed in ***Supplementary file 4***. Throughout the study, *V. cholerae* C6706 (***Thelin and Taylor, 1996***) was used as the wild-type strain. *V. cholerae* and *E. coli* strains were grown aerobically in LB medium at 37˚C except for temperature-sensitive strains. For stationary phase cultures of *V. cholerae*, samples were collected with respect to the time point when the cells reached an OD_{600} >2.0, i.e., 3 hr after cells reached an OD_{600} reading of 2.0. For transcript stability experiments, rifampicin was used at 250 µg/ml. To inhibit Rho-dependent transcription termination, bicyclomycin (BCM; sc-391755; Santa Cruz Biotechnology, Dallas, Texas) was used at 25 µg/ml. Other antibiotics were used at the following concentrations: 100 µg/ml ampicillin; 20 µg/ml chloramphenicol; 50 µg/ml kanamycin; 50 U/ml polymyxin B; and 5,000 µg/ml streptomycin.

For transient inactivation of RNase E, *V. cholerae* wild-type and a temperature-sensitive strain harboring the *rne-3071* mutation were grown at 30˚C to the indicated cell density. Cultures were divided in half and either continuously grown at 30˚C or shifted to 44˚C. RNA samples were collected from both strains and temperatures at the indicated time points after the temperature shift.

RK2/RP4-based conjugal transfer was used to introduce plasmids into *V. cholerae* from *E. coli* S17λpir plasmid donor strains (***Simon et al., 1983***). Subsequently, transconjugants were selected using appropriate antibiotics and polymyxin B to specifically inhibit *E. coli* growth. *V. cholerae* mutant strains were generated as described previously (***Papenfort et al., 2015b***). Briefly, pKAS32 plasmids were transferred into *V. cholerae* strains by conjugation and cells were screened for ampicillin resistance. Single colonies were streaked on streptomycin plates for counter-selection and colonies were tested for desired mutations by PCR or sequencing. Strain KPEC53467 was generated by phage P1 transduction to transfer the Δ*hfq*::KanR allele (***Baba et al., 2006***) into *E. coli* Top 10 and subsequent removal of the KanR cassette using plasmid pCP20 ***Datsenko and Wanner, 2000*** following standard protocols.

Plasmid construction

The plasmids used in this study are listed in **Supplementary file 4B**, used DNA oligonucleotides are listed in **Supplementary file 4C**. For pMD004, the *rrnB* terminator from pKP8-35 (**Papenfort et al., 2015b**) was amplified with KPO-1484/1485 and cloned by Gibson assembly into pKP-331 (**Papenfort et al., 2015b**) linearized with KPO-0196/1397. pMD089 was generated by amplification of *oppZ* from KPS-0014 chromosomal DNA using KPO-2552/2553 and Gibson assembly with pMD004 linearized with KPO-0196/1397. pMD373 was constructed by amplification of *oppB::3XFlag* from KPVC11709 chromosomal DNA using KPO-5878/5879 and Gibson assembly with pMD004 linearized with KPO-2789/pBAD-ATGrev. pCMW-2 was obtained by removing the promoterless *gfp* from pCMW-1 (**Waters and Bassler, 2006**) by amplification with KPO-2757/5421. pMD090 was generated by amplification of *oppZ* from KPS-0014 chromosomal DNA using KPO-2568/2553 and Gibson assembly with pEVS143 (**Dunn et al., 2006**) linearized with KPO-0092/1397. The M1 point mutation was introduced into pMD090 by site-directed mutagenesis with KPO-2619/2620, yielding pMD118. pMD194 and pMD195 were obtained by site-directed mutagenesis of pMD090 and pMD118, respectively, with KPO-3190/3191. pMD397 and pMD398 were obtained by replacing the p15a origin of replication in pCMW-1 and pMD194, respectively, by the pSC101 origin including an E93K mutation in the *repA* sequence. To this end, pCMW-1 and pMD194 were linearized with KPO-2041/2049, the pSC101 origin was amplified from pXG10-SF (**Corcoran et al., 2012**) in three parts (with KPO-6490/6493, KPO-6492/6495 and KPO-6494/6491) and fragments were joined with Gibson assembly. pMD173 and pMD174 were generated by amplification of the pBR322 origin from pBAD-Myc-His (Invitrogen) with KPO-2042/2043 and Gibson assembly with pCMW-1 or pMD090, respectively (both linearized with KPO-2041/2049). pMD197 was obtained by replacing the *oppZ* gene in pMD174 with a longer *oppF-oppZ* fragment (amplified from KPS-0014 chromosomal DNA using KPO-3197/2553) by Gibson assembly. pNP015 was constructed by amplification of *carZ* from KPS-0014 chromosomal DNA using KPO-1013/1014 and subcloning into linearized pEVS143 (KPO-0092/1023) with XbaI. Again, the M1 point mutation was introduced into pNP015 by site-directed mutagenesis with KPO-1782/1783, yielding pMH013. pMD361 and pMD362 were obtained by site-directed mutagenesis of pNP015 and pMH013, respectively, with KPO-5686/5687.

For translational GFP reporters, pMD093 was generated by amplification of the *oppAB* intergenic region and the first 5 codons of *oppB* from KPS-0014 chromosomal DNA using KPO-2580/2583 and Gibson assembly with pXG10-SF linearized with KPO-1702/1703. Site-directed mutagenesis of pMD093 with KPO-2615/2616 yielded pMD125. Accordingly, pMH010 and pMD374 were generated by amplification of the *carA* 5'UTR and the first 20 codons of *carA* with KPO-1674/1675 (for pMH010) or a fragment including the *carA* 5' UTR, the complete *carA* gene and the first 20 codons of *carB* with KPO-1674/5874 (for pMD374) from KPS-0014 chromosomal DNA, followed by Gibson assembly with pXG10-SF linearized with KPO-1702/1703. Site-directed mutagenesis of pMH010 and pMD374 with KPO-1778/1779 yielded pMH012 and pMD375, respectively. For discoordinate translational reporters for *oppB* to *oppF*, fragments from the *oppAB* intergenic region to the first 5 codons of *oppB* or the first 20 codons of *oppC*, *oppD* or *oppF* were amplified from KPS-0014 chromosomal DNA using KPO-2622 and KPO-2583 (*oppB*), KPO-2577 (*oppC*), KPO-2578 (*oppD*) or KPO-2579 (*oppF*). *mKate2* was amplified from pMD079 (**Herzog et al., 2019**) with KPO-2511/2625 and the pXG10-SF backbone was linearized with KPO-2621/1703. Gibson assembly was used to join the pXG10-SF backbone, *mKate2* and the respective *opp* fragment to generate pMD120, pMD352, pMD353 and pMD354. Site-directed mutagenesis of pMD120 and pMD354 with KPO-2615/2616 yielded pMD129 and pMD355, respectively.

pMD091 and pMD112 were constructed by amplification of *oppZ* from KPS-0014 chromosomal DNA using KPO-2585/2586 and Gibson assembly with pXG10-SF (for pMD091) or pMD093 (for pMD112), both linearized with KPO-2584/2508. The M1 mutations in the *oppAB* IGR or *oppZ* were obtained by site-directed mutagenesis of pMD112 with KPO-2615/2616 or KPO-2617/2618, respectively, to construct pMD117, pMD127 and pMD128. Site-directed mutagenesis of pMD91 and pMD93 with KPO-2665/2666 to introduce the M2 mutation into *oppZ* yielded pMD124 and pMD126, respectively. Accordingly, pMD294 and pMD297 were constructed by amplification of *carZ* from KPS-0014 chromosomal DNA using KPO-4815/4817 and Gibson assembly with pMH010 (for pMD294) or pMH012 (for pMD297), both linearized with KPO-2584/2508. Site-directed mutagenesis of pMD294 and pMD297 with KPO-1782/1783 yielded pMD296 and pMD298, respectively.

All pKAS32-derived plasmids (**Skorupski and Taylor, 1996**) were constructed by Gibson assembly of the respective up and down flanks with the pKAS32 backbone (linearized with KPO-0267/0268) and an additional fragment containing the 3XFLAG sequence or an *araC*-pBAD fragment where appropriate. Flanks were amplified from KPS-0014 chromosomal DNA unless otherwise stated. Plasmids for gene deletions or chromosomal point mutations are listed in the following with the respective primer pairs for up and down flanks indicated: pMD003 (KPO-1440/1443 and KPO-1441/1442), pMD160 (KPO-2753/1199 and KPO-1200/2754), pMD350 (KPO-1429/1289 and KPO-1290/1430), pMD349 (KPO-5243/5244 from KPVC11709 chromosomal DNA and KPO-5245/5246), pMD357 (KPO-5243/5672 and KPO-5673/5246, both from KPVC11709 chromosomal DNA), pMD358 (KPO-5243/5674 and KPO-5675/5246, both from KPVC11709 chromosomal DNA), pMD370 (KPO-5880/5884 and KPO-5885/5881, both from KPVC11709 chromosomal DNA), pMD371 (KPO-5880/5886 and KPO-5887/5881, both from KPVC11709 chromosomal DNA), pMD372 (KPO-5882/5890 and KPO-5891/5883, both from KPVC11709 chromosomal DNA), pMD356 (KPO-3183/5670 and KPO-5671/3186, both from KPVC11709 chromosomal DNA), pMD367 (KPO-4395/5824 from KPVC11709 chromosomal DNA and KPO-5823/4400), pMD369 (KPO-4379/5828 and KPO-5827/4384), pMD385 (KPO-5235/6029 and KPO-6030/5238, both from KPVC12872 chromosomal DNA) and pMD386 (KPO-5223/6031 and KPO-6032/5226, both from KPVC12872 chromosomal DNA). For pMD199 and pMD200, flanks were amplified with KPO-3179/3180 and KPO-3181/3182 (for pMD199) or with KPO-3183/3184 and KPO-3185/3186 (for pMD200). The 3XFLAG fragment was obtained by annealing of the oligonucleotides KPO-3157/3158. Flanks and 3XFLAG tag for pMD269, pMD346 and pMD347 were amplified with the following oligonucleotides: KPO-4385/4386, KPO-4387/4388 and KPO-4389/4390 (for pMD269); KPO-5223/5224, KPO-5225/5226 and KPO-5231/5232 (for pMD346); KPO-5227/5228, KPO-5229/5230 and KPO-5233/5234 (for pMD347). pMD199 was used as template for the 3XFLAG fragments. For pMD280 and pMD351, a fragment containing the *araC* gene and the pBAD promoter was amplified from pMD004 using 4529/0196. Flanks were amplified with KPO-4527/4528 and KPO-4530/4531 (for pMD280) or with KPO-5235/5236 and KPO-5237/5238 (for pMD351).

RNA isolation, Northern blot analysis and quantitative real-time PCR

For Northern blot analyses, total RNA was prepared and blotted as described previously (**Papenfort et al., 2017**). Membranes were hybridized in Roti-Hybri-Quick buffer (Carl Roth, Karlsruhe, Germany) with [^{32}P]-labeled DNA oligonucleotides at 42˚C or with riboprobes at 63˚C. Riboprobes were generated using the MAXIscript T7 Transcription Kit (Thermo Fisher Scientific, Waltham, Massachusetts). Signals were visualized using a Typhoon Phosphorimager (GE Healthcare, Chicago, Illinois) and quantified using GelQuant (RRID:SCR_015703; BioChemLabSolutions, San Francisco, California). Oligonucleotides for Northern blot analyses are provided in **Supplementary file 4C**. For qRT-PCR, total RNA was isolated with the SV Total RNA Isolation System (Promega, Fitchburg, Wisconsin). qRT–PCR was performed in three biological and two technical replicates using the Luna Universal One-Step RT-qPCR Kit (New England BioLabs, Ipswich, Massachusetts) and the MyiQ Single-Color Real-Time PCR Detection System (Bio-Rad, Hercules, California). 5S rRNA and *recA* were used as reference genes; oligonucleotides used for all qRT-PCR analyses are provided in **Supplementary file 4C**.

Hfq co-immunoprecipitation

Hfq co-immunoprecipitations were performed as previously described (**Huber et al., 2020**). Briefly, *V. cholerae* wild-type (KPS-0014) and *hfq::3XFLAG* (KPS-0995) (**Peschek et al., 2019**) strains were grown in LB medium to OD_{600} of 2.0. Lysates corresponding to 50 OD_{600} units were subjected to immunoprecipitation using monoclonal anti-FLAG antibody (#F1804; Sigma-Aldrich, St. Louis, Missouri) and Protein G Sepharose (#P3296; Sigma-Aldrich).

Western blot analysis and fluorescence assays

Total protein sample preparation and Western blot analyses were performed as described previously (**Papenfort et al., 2017**). Signals were visualized using a Fusion FX EDGE imager (Vilber Lourmat, Marne-la-Vallée, France) and band intensities were quantified using the BIO-1D software (Vilber Lourmat). 3XFLAG-tagged fusions were detected using mouse anti-FLAG antibody (#F1804; RRID:

AB_262044; Sigma-Aldrich) and goat anti-mouse HRP-conjugated IgG antibody, (#31430; RRID:AB_228307; Thermo Fisher Scientific). RNAPα served as a loading control and was detected using rabbit anti-RNAPα antibody (#WP003; RRID:AB_2687386; BioLegend, San Diego, California) and goat anti-rabbit HRP-conjugated IgG antibody, (#16104; AB_2534776; Thermo Fisher Scientific). Fluorescence assays of *E. coli* strains to measure mKate and GFP expression were performed as previously described (**Urban and Vogel, 2007**). Cells were washed in PBS and fluorescence intensity was quantified using a Spark 10 M plate reader (Tecan, Männedorf, Switzerland). Control strains not expressing fluorescent proteins were used to subtract background fluorescence.

RNA-seq analysis: TIER-seq

V. cholerae wild-type and *rne*^{TS} strains were grown in biological triplicates at 30˚C to OD_{600} of 1.0. Cultures were divided in half and either continuously grown at 30˚C or shifted to 44˚C. Cells were harvested from both strains and temperatures at 60 min after the temperature shift by addition of 0.2 volumes of stop mix (95% ethanol, 5% (v/v) phenol) and snap-frozen in liquid nitrogen. Total RNA was isolated and digested with TURBO DNase (Thermo Fisher Scientific). cDNA libraries were prepared by vertis Biotechnology AG (Freising, Germany): total RNA samples were poly(A)-tailed and 5'PPP structures were removed using RNA 5'Polyphosphatase (Epicentre, Madison, Wisconsin). An RNA adapter was ligated to the 5' monophosphate and first-strand cDNA synthesis was performed using an oligo(dT)-adapter and M-MLV reverse transcriptase. The resulting cDNAs were PCR-amplified, purified using the Agencourt AMPure XP kit (Beckman Coulter Genomics, Chaska, Minnesota) and sequenced using a NextSeq 500 system in single-read mode for 75 cycles.

After quality trimming and adapter clipping with cutadapt (version 2.5, DOI: https://doi.org/10.14806/ej.17.1.200) the sequencing reads were mapped to the *V. cholerae* reference genome (NCBI accession numbers: NC_002505.1 and NC_002506.1) including annotations for Vcr001-Vcr107 (**Papenfort et al., 2015b**) using READemption's (**Förstner et al., 2014**, v0.5.0, https://doi.org/10.5281/zenodo.591469) sub-command 'align' (building on segemehl version 0.3.4, **Hoffmann et al., 2009**) and nucleotide-specific coverage values were calculated with the sub-command 'coverage' based on the first base of the reads. Positions with a coverage of 20 reads or more were used to perform an enrichment analysis using DESeq2 (v.1.20.0, **Love et al., 2014**) comparing the WT to the mutant libraries. Nucleotides for which DESeq2 calculated an absolute fold-change of 3.0 or more and an adjusted (Benjamini-Hochberg corrected) p-value of 0.05 were treated in following analysis steps as bona fide cleavage sites.

The Minimum free energy (MFE) of sequence windows was computed with RNAfold (version 2.4.14) of the Vienna package (**Lorenz et al., 2011**). Sequence logos were created with WebLogo (version 3.7.4; **Crooks et al., 2004**). Overlaps of cleavage sites with other features were found by BEDTools' (version 2.26.0, **Quinlan and Hall, 2010**) sub-command 'intersect'. Pair-wise Pearson correlation coefficients between all samples were calculated based on the above mentioned first-base-in read coverages taking positions with a total sum of at least 10 reads in all samples combined into account. Positions that represent outliers with coverage values above the 99.99 percentile in one or more read libraries were not considered. The values were computed using the function 'corr' of the pandas Dataframe class (https://doi.org/10.5281/zenodo.3509134). For further details, please see the analysis scripts linked in the data and code availability section.

RNA-seq analysis: Identification of OppZ targets

V. cholerae strains carrying either pBAD1K-ctrl or pBAD1K-*oppZ* were grown in biological triplicates to OD_{600} of 0.5 and treated with 0.2% L-arabinose (final conc.). Cells were harvested after 15 min by addition of 0.2 volumes of stop mix (95% ethanol, 5% (v/v) phenol) and snap-frozen in liquid nitrogen. Total RNA was isolated and digested with TURBO DNase (Thermo Fisher Scientific). Ribosomal RNA was depleted using the Ribo-Zero kit for Gram-negative bacteria (#MRZGN126; Illumina, San Diego, California) and RNA integrity was confirmed with an Agilent 2100 Bioanalyzer. Directional cDNA libraries were prepared using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (#E7760; NEB). The libraries were sequenced using a HiSeq 1500 System in single-read mode for 100 cycles. The read files in FASTQ format were imported into CLC Genomics Workbench v11 (RRID:SCR_011853; Qiagen, Hilden, Germany) and trimmed for quality and 3' adaptors. Reads were mapped to the *V. cholerae* reference genome (NCBI accession numbers: NC_002505.1 and

NC_002506.1) including annotations for Vcr001-Vcr107 (**Papenfort et al., 2015b**) using the 'RNA-Seq Analysis' tool with standard parameters. Reads mapping in CDS were counted, and genes with a total count cut-off >15 in all samples were considered for analysis. Read counts were normalized (CPM), and transformed (log2). Differential expression was tested using the built-in tool corresponding to edgeR in exact mode with tagwise dispersions ('Empirical Analysis of DGE'). Genes with a fold change ≥3.0 and an FDR-adjusted p-value≤0.05 were considered as differentially expressed.

RNA-seq analysis: Bicyclomycin-dependent transcriptomes

V. cholerae oppA::3XFLAG *oppB*::3XFLAG *oppF*::3XFLAG strains with wild-type or mutated *oppB* start codon were grown in biological triplicates to OD_{600} of 1.5, divided in half and treated with either bicyclomycin (25 µg/ml final conc.) or water. Cells were harvested after 120 min by addition of 0.2 volumes of stop mix (95% ethanol, 5% (v/v) phenol) and snap-frozen in liquid nitrogen. Total RNA was isolated and digested with TURBO DNase (Thermo Fisher Scientific). cDNA libraries were prepared by vertis Biotechnology AG in a 3' end-specific protocol: ribosomal RNA was depleted and the Illumina 5' sequencing adaptor was ligated to the 3' OH end of RNA molecules. First strand synthesis using M-MLV reverse transcriptase was followed by fragmentation and strand-specific ligation of the Illumina 3' sequencing adaptor to the 3' end of first-strand cDNA. Finally, 3' cDNA fragments were amplified, purified using the Agencourt AMPure XP kit (Beckman Coulter Genomics) and sequenced using a NextSeq 500 system in single-read mode for 75 cycles. The read files in FASTQ format were imported into CLC Genomics Workbench v11 (Qiagen) and trimmed for quality and 3' adaptors. Reads were mapped to the *V. cholerae* reference genome (NCBI accession numbers: NC_002505.1 and NC_002506.1) including annotations for Vcr001-Vcr107 (**Papenfort et al., 2015b**) using the 'RNA-Seq Analysis' tool with standard parameters. Reads mapping in CDS were counted, and genes with a total count cut-off >8 in all samples were considered for analysis. Read counts were normalized (CPM), and transformed (log2). Differential expression was tested using the built in tool corresponding to edgeR in exact mode with tagwise dispersions ('Empirical Analysis of DGE'). Genes with a fold change ≥3.0 and an FDR-adjusted p-value≤0.05 were considered as differentially expressed.

TIER-seq input data, analysis scripts and results are deposited at Zenodo (https://doi.org/10.5281/zenodo.3750832). Further information and requests for resources and reagents should be directed to and will be fulfilled by the corresponding author, Kai Papenfort (kai.papenfort@uni-jena.de).

Acknowledgements

We thank Helmut Blum for help with the RNA sequencing experiments and Andreas Starick for excellent technical support. We thank Jörg Vogel, Gisela Storz, and Kathrin Fröhlich for comments on the manuscript and all members of the Papenfort lab for insightful discussions and suggestions.

Additional information

Funding

Funder	Grant reference number	Author
Deutsche Forschungsgemeinschaft	GRK2062/1	Mona Hoyos Kai Papenfort
Vallee Foundation	Vallee Scholars program	Kai Papenfort
H2020 European Research Council	StG-758212	Michaela Huber Kai Papenfort
Human Frontier Science Program	CDA00024/2016-C	Mona Hoyos Michaela Huber Kai Papenfort
Deutsche Forschungsgemeinschaft	EXC 2051	Mona Hoyos Kai Papenfort
Deutsche Forschungsgemeinschaft	390713860	Mona Hoyos Kai Papenfort

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

Mona Hoyos, Conceptualization, Data curation, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing - review and editing; Michaela Huber, Data curation, Validation, Investigation, Methodology, Writing - review and editing; Konrad U Förstner, Resources, Data curation, Software, Formal analysis, Writing - review and editing; Kai Papenfort, Conceptualization, Data curation, Supervision, Funding acquisition, Investigation, Methodology, Writing - original draft, Project administration, Writing - review and editing

Author ORCIDs

Mona Hoyos https://orcid.org/0000-0003-1085-4723
Kai Papenfort https://orcid.org/0000-0002-5560-9804

Decision letter and Author response

Decision letter https://doi.org/10.7554/eLife.58836.sa1
Author response https://doi.org/10.7554/eLife.58836.sa2

Additional files

Supplementary files

- Supplementary file 1. TIER-seq sites in *Vibrio cholerae.*
- Supplementary file 2. RNase E-mediated maturation of sRNAs.
- Supplementary file 3. BCM-sensitive transcripts in *Vibrio cholerae.*
- Supplementary file 4. Bacterial strains, plasmids and DNA oligonucleotides.
- Transparent reporting form

Data availability

All high-throughput sequencing data was deposited at GEO: GSE148675 (TIER-seq), GSE144479 (OppZ target identification) and GSE144478 (Term-Seq analysis). We have uploaded source data for all figures and figure supplements showing the numerical data from our TIER-seq analysis, the raw data from GFP and mKate fluorescence measurements, fold changes obtained from qRT-PCR experiments and fold changes obtained by the quantification of Western and Northern blots. Additionally, for all figures showing cropped images of Western or Northern blots, we show the full image and indicate the cropped area and the antibody or labelled oligonucleotide used to detect the signal.

The following datasets were generated:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Hoyos M, Huber M, Förstner K, Papenfort K	2020	Identification of bicyclomycin-sensitive transcripts in Vibrio cholerae	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE144478	NCBI Gene Expression Omnibus, GSE144478
Hoyos M, Huber M, Förstner KU, Papenfort K	2020	Global identification of RNase E sites in Vibrio cholerae	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE148675	NCBI Gene Expression Omnibus, GSE148675
Hoyos M, Huber M, Förstner K, Papenfort K	2020	OppZ target identification	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE144479	NCBI Gene Expression Omnibus, GSE144479

References

Adams PP, Storz G. 2020. Prevalence of small base-pairing RNAs derived from diverse genomic loci. *Biochimica Et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **1863**:194524. DOI: https://doi.org/10.1016/j.bbagrm.2020.194524

Afroz T, Beisel CL. 2013. Understanding and exploiting feedback in synthetic biology. *Chemical Engineering Science* **103**:79–90. DOI: https://doi.org/10.1016/j.ces.2013.02.017

Aït-Bara S, Carpousis AJ. 2015. RNA degradosomes in Bacteria and chloroplasts: classification, distribution and evolution of RNase E homologs. *Molecular Microbiology* **97**:1021–1135. DOI: https://doi.org/10.1111/mmi.13095, PMID: 26096689

Allfano P, Rivellini F, Limauro D, Bruni CB, Carlomagno MS. 1991. A consensus motif common to all rho-dependent prokaryotic transcription terminators. *Cell* **64**:553–563. DOI: https://doi.org/10.1016/0092-8674(91)90239-U

Alon U. 2007. Network motifs: theory and experimental approaches. *Nature Reviews Genetics* **8**:450–461. DOI: https://doi.org/10.1038/nrg2102, PMID: 17510665

Apirion D, Lassar AB. 1978. A conditional lethal mutant of *Escherichia coli* which affects the processing of ribosomal RNA. *The Journal of Biological Chemistry* **253**:1738–1742. PMID: 342528

Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H. 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the keio collection. *Molecular Systems Biology* **2**:0008. DOI: https://doi.org/10.1038/msb4100050, PMID: 16738554

Bandyra KJ, Luisi BF. 2018. RNase E and the High-Fidelity orchestration of RNA metabolism. *Microbiology Spectrum* **6**:ch1. DOI: https://doi.org/10.1128/9781683670247.ch1

Bechhofer DH, Deutscher MP. 2019. Bacterial ribonucleases and their roles in RNA metabolism. *Critical Reviews in Biochemistry and Molecular Biology* **54**:242–300. DOI: https://doi.org/10.1080/10409238.2019.1651816, PMID: 31464530

Bossi L, Schwartz A, Guillemardet B, Boudvillain M, Figueroa-Bossi N. 2012. A role for Rho-dependent polarity in gene regulation by a noncoding small RNA. *Genes & Development* **26**:1864–1873. DOI: https://doi.org/10.1101/gad.195412.112, PMID: 22895254

Bossi L, Figueroa-Bossi N, Bouloc P, Boudvillain M. 2020. Regulatory interplay between small RNAs and transcription termination factor rho. *Biochimica Et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **1863**:194546. DOI: https://doi.org/10.1016/j.bbagrm.2020.194546

Brosse A, Korobeinikova A, Gottesman S, Guillier M. 2016. Unexpected properties of sRNA promoters allow feedback control via regulation of a two-component system. *Nucleic Acids Research* **177**:642–9666. DOI: https://doi.org/10.1093/nar/gkw642

Cameron DE, Urbach JM, Mekalanos JJ. 2008. A defined transposon mutant library and its use in identifying motility genes in *Vibrio cholerae*. *PNAS* **105**:8736–8741. DOI: https://doi.org/10.1073/pnas.0803281105, PMID: 18574146

Cardinale CJ, Washburn RS, Tadigotla VR, Brown LM, Gottesman ME, Nudler E. 2008. Termination factor rho and its cofactors NusA and NusG silence foreign DNA in *E. coli*. *Science* **320**:935–938. DOI: https://doi.org/10.1126/science.1152763, PMID: 18487194

Castellana M, Wilson MZ, Xu Y, Joshi P, Cristea IM, Rabinowitz JD, Gitai Z, Wingreen NS. 2014. Enzyme clustering accelerates processing of intermediates through metabolic channeling. *Nature Biotechnology* **32**:1011–1018. DOI: https://doi.org/10.1038/nbt.3018, PMID: 25262299

Chao Y, Papenfort K, Reinhardt R, Sharma CM, Vogel J. 2012. An atlas of Hfq-bound transcripts reveals 3′ UTRs as a genomic reservoir of regulatory small RNAs. *The EMBO Journal* **31**:4005–4019. DOI: https://doi.org/10.1038/emboj.2012.229, PMID: 22922465

Chao Y, Li L, Girodat D, Förstner KU, Said N, Corcoran C, Śmiga M, Papenfort K, Reinhardt R, Wieden HJ, Luisi BF, Vogel J. 2017. In Vivo Cleavage Map Illuminates the Central Role of RNase E in Coding and Non-coding RNA Pathways. *Molecular Cell* **65**:39–51. DOI: https://doi.org/10.1016/j.molcel.2016.11.002, PMID: 28061332

Chao Y, Vogel J. 2016. A 3′ UTR-Derived small RNA provides the regulatory noncoding arm of the inner membrane stress response. *Molecular Cell* **61**:352–363. DOI: https://doi.org/10.1016/j.molcel.2015.12.023, PMID: 26805574

Charlier D, Nguyen Le Minh P, Roovers M. 2018. Regulation of carbamoylphosphate synthesis in *Escherichia coli*: an amazing metabolite at the crossroad of arginine and pyrimidine biosynthesis. *Amino Acids* **50**:1647–1661. DOI: https://doi.org/10.1007/s00726-018-2654-z, PMID: 30238253

Ciampi MS. 2006. Rho-dependent terminators and transcription termination. *Microbiology* **152**:2515–2528. DOI: https://doi.org/10.1099/mic.0.28982-0, PMID: 16946247

Corcoran CP, Podkaminski D, Papenfort K, Urban JH, Hinton JC, Vogel J. 2012. Superfolder GFP reporters validate diverse new mRNA targets of the classic porin regulator, MicF RNA. *Molecular Microbiology* **84**:428–445. DOI: https://doi.org/10.1111/j.1365-2958.2012.08031.x, PMID: 22458297

Corpet F. 1988. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Research* **16**:10881–10890. DOI: https://doi.org/10.1093/nar/16.22.10881, PMID: 2849754

Crooks GE, Hon G, Chandonia JM, Brenner SE. 2004. WebLogo: a sequence logo generator. *Genome Research* **14**:1188–1190. DOI: https://doi.org/10.1101/gr.849004, PMID: 15173120

Dar D, Shamir M, Mellin JR, Koutero M, Stern-Ginossar N, Cossart P, Sorek R. 2016. Term-seq reveals abundant ribo-regulation of antibiotics resistance in Bacteria. *Science* **352**:aad9822. DOI: https://doi.org/10.1126/science.aad9822, PMID: 27120414

Dar D, Sorek R. 2018a. Bacterial noncoding RNAs excised from within Protein-Coding transcripts. *mBio* **9**:01730-18. DOI: https://doi.org/10.1128/mBio.01730-18
Dar D, Sorek R. 2018b. High-resolution RNA 3′-ends mapping of bacterial Rho-dependent transcripts. *Nucleic Acids Research* **46**:6797–6805. DOI: https://doi.org/10.1093/nar/gky274, PMID: 29669055
Datsenko KA, Wanner BL. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *PNAS* **97**:6640–6645. DOI: https://doi.org/10.1073/pnas.120163297, PMID: 10829079
Davis BM, Waldor MK. 2007. RNase E-dependent processing stabilizes MicX, a *Vibrio cholerae* sRNA. *Molecular Microbiology* **65**:373–385. DOI: https://doi.org/10.1111/j.1365-2958.2007.05796.x, PMID: 17590231
De Mets F, Van Melderen L, Gottesman S. 2019. Regulation of acetate metabolism and coordination with the TCA cycle via a processed small RNA. *PNAS* **116**:1043–1052. DOI: https://doi.org/10.1073/pnas.1815288116, PMID: 30591570
Deana A, Celesnik H, Belasco JG. 2008. The bacterial enzyme RppH triggers messenger RNA degradation by 5′ pyrophosphate removal. *Nature* **451**:355–358. DOI: https://doi.org/10.1038/nature06475, PMID: 18202662
Doeven MK, Abele R, Tampé R, Poolman B. 2004. The binding specificity of OppA determines the selectivity of the oligopeptide ATP-binding cassette transporter. *Journal of Biological Chemistry* **279**:32301–32307. DOI: https://doi.org/10.1074/jbc.M404343200, PMID: 15169767
Dunn AK, Millikan DS, Adin DM, Bose JL, Stabb EV. 2006. New *rfp*- and pES213-derived tools for analyzing symbiotic *Vibrio fischeri* reveal patterns of infection and *lux* expression in situ. *Applied and Environmental Microbiology* **72**:802–810. DOI: https://doi.org/10.1128/AEM.72.1.802-810.2006, PMID: 16391121
Feng L, Rutherford ST, Papenfort K, Bagert JD, van Kessel JC, Tirrell DA, Wingreen NS, Bassler BL. 2015. A qrr noncoding RNA deploys four different regulatory mechanisms to optimize quorum-sensing dynamics. *Cell* **160**:228–240. DOI: https://doi.org/10.1016/j.cell.2014.11.051, PMID: 25579683
Förstner KU, Vogel J, Sharma CM. 2014. READemption-a tool for the computational analysis of deep-sequencing-based transcriptome data. *Bioinformatics* **30**:3421–3423. DOI: https://doi.org/10.1093/bioinformatics/btu533, PMID: 25123900
Förstner KU, Reuscher CM, Haberzettl K, Weber L, Klug G. 2018. RNase E cleavage shapes the transcriptome of *Rhodobacter sphaeroides* and strongly impacts phototrophic growth. *Life Science Alliance* **1**:e201800080. DOI: https://doi.org/10.26508/lsa.201800080, PMID: 30456366
Fröhlich KS, Papenfort K, Fekete A, Vogel J. 2013. A small RNA activates CFA synthase by isoform-specific mRNA stabilization. *The EMBO Journal* **32**:2963–2979. DOI: https://doi.org/10.1038/emboj.2013.222, PMID: 24141880
Gorski SA, Vogel J, Doudna JA. 2017. RNA-based recognition and targeting: sowing the seeds of specificity. *Nature Reviews Molecular Cell Biology* **18**:215–228. DOI: https://doi.org/10.1038/nrm.2016.174, PMID: 28196981
Grabowicz M, Koren D, Silhavy TJ. 2016. The CpxQ sRNA negatively regulates skp to prevent mistargeting of β-Barrel outer membrane proteins into the cytoplasmic membrane. *mBio* **7**:00312-16. DOI: https://doi.org/10.1128/mBio.00312-16
Gruber AJ, Zavolan M. 2019. Alternative cleavage and polyadenylation in health and disease. *Nature Reviews Genetics* **20**:599–614. DOI: https://doi.org/10.1038/s41576-019-0145-z, PMID: 31267064
Guillier M, Gottesman S. 2008. The 5′ end of two redundant sRNAs is involved in the regulation of multiple targets, including their own regulator. *Nucleic Acids Research* **36**:6781–6794. DOI: https://doi.org/10.1093/nar/gkn742, PMID: 18953042
Herzog R, Peschek N, Fröhlich KS, Schumacher K, Papenfort K. 2019. Three autoinducer molecules act in concert to control virulence gene expression in *Vibrio cholerae*. *Nucleic Acids Research* **47**:3171–3183. DOI: https://doi.org/10.1093/nar/gky1320, PMID: 30649554
Hiles ID, Gallagher MP, Jamieson DJ, Higgins CF. 1987. Molecular characterization of the oligopeptide permease of *Salmonella typhimurium*. *Journal of Molecular Biology* **195**:125–142. DOI: https://doi.org/10.1016/0022-2836(87)90332-9, PMID: 2821267
Hoffmann S, Otto C, Kurtz S, Sharma CM, Khaitovich P, Vogel J, Stadler PF, Hackermüller J. 2009. Fast mapping of short sequences with mismatches, insertions and deletions using index structures. *PLOS Computational Biology* **5**:e1000502. DOI: https://doi.org/10.1371/journal.pcbi.1000502, PMID: 19750212
Holmqvist E, Vogel J. 2018. RNA-binding proteins in bacteria. *Nature Reviews Microbiology* **16**:601–615. DOI: https://doi.org/10.1038/s41579-018-0049-5, PMID: 29995832
Hör J, Gorski SA, Vogel J. 2018. Bacterial RNA Biology on a Genome Scale. *Molecular Cell* **70**:785–799. DOI: https://doi.org/10.1016/j.molcel.2017.12.023
Huber M, Fröhlich KS, Radmer J, Papenfort K. 2020. Switching fatty acid metabolism by an RNA-controlled feed forward loop. *PNAS* **117**:8044–8054. DOI: https://doi.org/10.1073/pnas.1920753117, PMID: 32193348
Hussein R, Lim HN. 2012. Direct comparison of small RNA and transcription factor signaling. *Nucleic Acids Research* **40**:7269–7279. DOI: https://doi.org/10.1093/nar/gks439, PMID: 22618873
Kavita K, de Mets F, Gottesman S. 2018. New aspects of RNA-based regulation by hfq and its partner sRNAs. *Current Opinion in Microbiology* **42**:53–61. DOI: https://doi.org/10.1016/j.mib.2017.10.014, PMID: 29125938
Kelly CL, Harris AWK, Steel H, Hancock EJ, Heap JT, Papachristodoulou A. 2018. Synthetic negative feedback circuits using engineered small RNAs. *Nucleic Acids Research* **46**:9875–9889. DOI: https://doi.org/10.1093/nar/gky828, PMID: 30212900
Kimura S, Waldor MK. 2019. The RNA degradosome promotes tRNA quality control through clearance of hypomodified tRNA. *PNAS* **116**:1394–1403. DOI: https://doi.org/10.1073/pnas.1814130116, PMID: 30622183

Korkmaz G, Holm M, Wiens T, Sanyal S. 2014. Comprehensive analysis of stop Codon usage in Bacteria and its correlation with release factor abundance. *Journal of Biological Chemistry* **289**:30334–30342. DOI: https://doi.org/10.1074/jbc.M114.606632, PMID: 25217634

Lin P, Pu Q, Wu Q, Zhou C, Wang B, Schettler J, Wang Z, Qin S, Gao P, Li R, Li G, Cheng Z, Lan L, Jiang J, Wu M. 2019. High-throughput screen reveals sRNAs regulating crRNA biogenesis by targeting CRISPR leader to repress rho termination. *Nature Communications* **10**:3728. DOI: https://doi.org/10.1038/s41467-019-11695-8, PMID: 31427601

Lorenz R, Bernhart SH, Höner Zu Siederdissen C, Tafer H, Flamm C, Stadler PF, Hofacker IL. 2011. ViennaRNA package 2.0. *Algorithms for Molecular Biology* **6**:26. DOI: https://doi.org/10.1186/1748-7188-6-26, PMID: 22115189

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**:550. DOI: https://doi.org/10.1186/s13059-014-0550-8, PMID: 25516281

Mackie GA. 2013. RNase E: at the interface of bacterial RNA processing and decay. *Nature Reviews Microbiology* **11**:45–57. DOI: https://doi.org/10.1038/nrmicro2930, PMID: 23241849

Malka Y, Steiman-Shimony A, Rosenthal E, Argaman L, Cohen-Daniel L, Arbib E, Margalit H, Kaplan T, Berger M. 2017. Post-transcriptional 3´-UTR cleavage of mRNA transcripts generates thousands of stable uncapped autonomous RNA fragments. *Nature Communications* **8**:2029. DOI: https://doi.org/10.1038/s41467-017-02099-7, PMID: 29229900

Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**:10–12. DOI: https://doi.org/10.14806/ej.17.1.200

Massé E, Escorcia FE, Gottesman S. 2003. Coupled degradation of a small regulatory RNA and its mRNA targets in *Escherichia coli*. *Genes & Development* **17**:2374–2383. DOI: https://doi.org/10.1101/gad.1127103, PMID: 12975324

Melamed S, Peer A, Faigenbaum-Romm R, Gatt YE, Reiss N, Bar A, Altuvia Y, Argaman L, Margalit H. 2016. Global mapping of small RNA-Target interactions in Bacteria. *Molecular Cell* **63**:884–897. DOI: https://doi.org/10.1016/j.molcel.2016.07.026, PMID: 27588604

Melamed S, Adams PP, Zhang A, Zhang H, Storz G. 2020. RNA-RNA interactomes of ProQ and hfq reveal overlapping and competing roles. *Molecular Cell* **77**:411–425. DOI: https://doi.org/10.1016/j.molcel.2019.10.022

Mitra P, Ghosh G, Hafeezunnisa M, Sen R. 2017. Rho protein: roles and mechanisms. *Annual Review of Microbiology* **71**:687–709. DOI: https://doi.org/10.1146/annurev-micro-030117-020432, PMID: 28731845

Miyakoshi M, Chao Y, Vogel J. 2015. Regulatory small RNAs from the 3′ regions of bacterial mRNAs. *Current Opinion in Microbiology* **24**:132–139. DOI: https://doi.org/10.1016/j.mib.2015.01.013, PMID: 25677420

Miyakoshi M, Matera G, Maki K, Sone Y, Vogel J. 2019. Functional expansion of a TCA cycle operon mRNA by a 3′ end-derived small RNA. *Nucleic Acids Research* **47**:2075–2088. DOI: https://doi.org/10.1093/nar/gky1243, PMID: 30541135

Mohanty BK, Kushner SR. 2018. Enzymes involved in posttranscriptional RNA metabolism in Gram-Negative Bacteria. *Microbiology Spectrum* **6**:2017. DOI: https://doi.org/10.1128/microbiolspec.RWR-0011-2017

Morita T, Maki K, Aiba H. 2005. RNase E-based ribonucleoprotein complexes: mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs. *Genes & Development* **19**:2176–2186. DOI: https://doi.org/10.1101/gad.1330405, PMID: 16166379

Nadiras C, Eveno E, Schwartz A, Figueroa-Bossi N, Boudvillain M. 2018. A multivariate prediction model for Rho-dependent termination of transcription. *Nucleic Acids Research* **46**:8245–8260. DOI: https://doi.org/10.1093/nar/gky563, PMID: 29931073

Nitzan M, Rehani R, Margalit H. 2017. Integration of bacterial small RNAs in regulatory networks. *Annual Review of Biophysics* **46**:131–148. DOI: https://doi.org/10.1146/annurev-biophys-070816-034058, PMID: 28532217

Papenfort K, Sun Y, Miyakoshi M, Vanderpool CK, Vogel J. 2013. Small RNA-mediated activation of sugar phosphatase mRNA regulates glucose homeostasis. *Cell* **153**:426–437. DOI: https://doi.org/10.1016/j.cell.2013.03.003, PMID: 23582330

Papenfort K, Espinosa E, Casadesús J, Vogel J. 2015a. Small RNA-based feedforward loop with AND-gate logic regulates extrachromosomal DNA transfer in *Salmonella*. *PNAS* **112**:E4772–E4781. DOI: https://doi.org/10.1073/pnas.1507825112, PMID: 26307765

Papenfort K, Förstner KU, Cong JP, Sharma CM, Bassler BL. 2015b. Differential RNA-seq of *Vibrio cholerae* identifies the VqmR small RNA as a regulator of biofilm formation. *PNAS* **112**:E766–E775. DOI: https://doi.org/10.1073/pnas.1500203112, PMID: 25646441

Papenfort K, Silpe JE, Schramma KR, Cong JP, Seyedsayamdost MR, Bassler BL. 2017. A *Vibrio cholerae* autoinducer-receptor pair that controls biofilm formation. *Nature Chemical Biology* **13**:551–557. DOI: https://doi.org/10.1038/nchembio.2336, PMID: 28319101

Peschek N, Hoyos M, Herzog R, Förstner KU, Papenfort K. 2019. A conserved RNA seed-pairing domain directs small RNA-mediated stress resistance in enterobacteria. *The EMBO Journal* **38**:e101650. DOI: https://doi.org/10.15252/embj.2019101650, PMID: 31313835

Peters JM, Mooney RA, Grass JA, Jessen ED, Tran F, Landick R. 2012. Rho and NusG suppress pervasive antisense transcription in *Escherichia coli*. *Genes & Development* **26**:2621–2633. DOI: https://doi.org/10.1101/gad.196741.112, PMID: 23207917

Pu M, Chen J, Tao Z, Miao L, Qi X, Wang Y, Ren J. 2019. Regulatory network of miRNA on its target: coordination between transcriptional and post-transcriptional regulation of gene expression. *Cellular and Molecular Life Sciences* **76**:441–451. DOI: https://doi.org/10.1007/s00018-018-2940-7, PMID: 30374521

Qi LS, Arkin AP. 2014. A versatile framework for microbial engineering using synthetic non-coding RNAs. *Nature Reviews. Microbiology* **12**:341–354. DOI: https://doi.org/10.1038/nrmicro3244, PMID: 24736794

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**:841–842. DOI: https://doi.org/10.1093/bioinformatics/btq033, PMID: 20110278

Rehmsmeier M, Steffen P, Hochsmann M, Giegerich R. 2004. Fast and effective prediction of microRNA/target duplexes. *RNA* **10**:1507–1517. DOI: https://doi.org/10.1261/rna.5248604, PMID: 15383676

Richards J, Belasco JG. 2019. Obstacles to scanning by RNase E govern bacterial mRNA lifetimes by hindering access to distal cleavage sites. *Molecular Cell* **74**:284–295. DOI: https://doi.org/10.1016/j.molcel.2019.01.044

Rosenfeld N, Elowitz MB, Alon U. 2002. Negative autoregulation speeds the response times of transcription networks. *Journal of Molecular Biology* **323**:785–793. DOI: https://doi.org/10.1016/S0022-2836(02)00994-4, PMID: 12417193

Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH image to ImageJ: 25 years of image analysis. *Nature Methods* **9**:671–675. DOI: https://doi.org/10.1038/nmeth.2089, PMID: 22930834

Sedlyarova N, Shamovsky I, Bharati BK, Epshtein V, Chen J, Gottesman S, Schroeder R, Nudler E. 2016. sRNA-Mediated control of transcription termination in *E. coli*. *Cell* **167**:111–121. DOI: https://doi.org/10.1016/j.cell.2016.09.004

Shen-Orr SS, Milo R, Mangan S, Alon U. 2002. Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nature Genetics* **31**:64–68. DOI: https://doi.org/10.1038/ng881, PMID: 11967538

Simon R, Priefer U, Pühler A. 1983. A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in gram negative Bacteria. *Bio/Technology* **1**:784–791. DOI: https://doi.org/10.1038/nbt1183-784

Skorupski K, Taylor RK. 1996. Positive selection vectors for allelic exchange. *Gene* **169**:47–52. DOI: https://doi.org/10.1016/0378-1119(95)00793-8, PMID: 8635748

Soper T, Mandin P, Majdalani N, Gottesman S, Woodson SA. 2010. Positive regulation by small RNAs and the role of hfq. *PNAS* **107**:9602–9607. DOI: https://doi.org/10.1073/pnas.1004435107, PMID: 20457943

Svenningsen SL, Tu KC, Bassler BL. 2009. Gene dosage compensation calibrates four regulatory RNAs to control *Vibrio cholerae* quorum sensing. *The EMBO Journal* **28**:429–439. DOI: https://doi.org/10.1038/emboj.2008.300, PMID: 19165149

Thelin KH, Taylor RK. 1996. Toxin-coregulated pilus, but not mannose-sensitive hemagglutinin, is required for colonization by *Vibrio cholerae* O1 el Tor biotype and O139 strains. *Infection and Immunity* **64**:2853–2856. DOI: https://doi.org/10.1128/IAI.64.7.2853-2856.1996, PMID: 8698524

Updegrove TB, Kouse AB, Bandyra KJ, Storz G. 2019. Stem-loops direct precise processing of 3′ UTR-derived small RNA MicL. *Nucleic Acids Research* **47**:1482–1492. DOI: https://doi.org/10.1093/nar/gky1175

Urban JH, Vogel J. 2007. Translational control and target recognition by *Escherichia coli* small RNAs in vivo. *Nucleic Acids Research* **35**:1018–1037. DOI: https://doi.org/10.1093/nar/gkl1040, PMID: 17264113

Wang X, Ji SC, Jeon HJ, Lee Y, Lim HM. 2015. Two-level inhibition of *galK* expression by spot 42: degradation of mRNA mK2 and enhanced transcription termination before the *galK* gene. *PNAS* **112**:7581–7586. DOI: https://doi.org/10.1073/pnas.1424683112, PMID: 26045496

Wang C, Chao Y, Matera G, Gao Q, Vogel J. 2020. The conserved 3′ UTR-derived small RNA NarS mediates mRNA crossregulation during nitrate respiration. *Nucleic Acids Research* **48**:2126–2143. DOI: https://doi.org/10.1093/nar/gkz1168, PMID: 31863581

Waters CM, Bassler BL. 2006. The *Vibrio harveyi* quorum-sensing system uses shared regulatory components to discriminate between multiple autoinducers. *Genes & Development* **20**:2754–2767. DOI: https://doi.org/10.1101/gad.1466506, PMID: 17015436

Zwiefka A, Kohn H, Widger WR. 1993. Transcription termination factor rho: the site of bicyclomycin inhibition in *Escherichia coli*. *Biochemistry* **32**:3564–3570. DOI: https://doi.org/10.1021/bi00065a007, PMID: 8466900

3 A conserved RNA seed-pairing domain directs small RNA-mediated stress resistance in enterobacteria

Nikolai Peschek, Mona Hoyos, Roman Herzog, Konrad U. Förstner, and Kai Papenfort (2019). "A conserved RNA seed-pairing domain directs small RNA-mediated stress resistance in enterobacteria." *The EMBO Journal* 38.16, e101650

Source online: https://doi.org/10.15252/embj.2019101650

This publication characterizes the σ^E-dependent sRNAs MicV and VrrA as non-coding arm of the envelope stress response. A complex library of synthetic sRNAs is employed in laboratory selection experiments to identify post-transcriptional repression of the porin *ompA* as key factor for cell survival under membrane-damaging conditions.

Article

SOURCE DATA TRANSPARENT PROCESS OPEN ACCESS

THE EMBO JOURNAL

A conserved RNA seed-pairing domain directs small RNA-mediated stress resistance in enterobacteria

Nikolai Peschek[1,2], Mona Hoyos[1], Roman Herzog[1], Konrad U Förstner[3,4] & Kai Papenfort[1,2,*]

Abstract

Small regulatory RNAs (sRNAs) are crucial components of many stress response systems. The envelope stress response (ESR) of Gram-negative bacteria is a paradigm for sRNA-mediated stress management and involves, among other factors, the alternative sigma factor E (σ^E) and one or more sRNAs. In this study, we identified the MicV sRNA as a new member of the σ^E regulon in *Vibrio cholerae*. We show that MicV acts redundantly with another sRNA, VrrA, and that both sRNAs share a conserved seed-pairing domain allowing them to regulate multiple target mRNAs. *V. cholerae* lacking σ^E displayed increased sensitivity toward antimicrobials, and over-expression of either of the sRNAs suppressed this phenotype. Laboratory selection experiments using a library of synthetic sRNA regulators revealed that the seed-pairing domain of σ^E-dependent sRNAs is strongly enriched among sRNAs identified under membrane-damaging conditions and that repression of OmpA is crucial for sRNA-mediated stress relief. Together, our work shows that MicV and VrrA act as global regulators in the ESR of *V. cholerae* and provides evidence that bacterial sRNAs can be functionally annotated by their seed-pairing sequences.

Keywords Hfq; MicV; seed pairing; sigma E; sRNA
Subject Categories Microbiology, Virology & Host Pathogen Interaction; RNA Biology
DOI 10.15252/embj.2019101650 | Received 28 January 2019 | Revised 31 May 2019 | Accepted 9 June 2019 | Published online 17 July 2019
The EMBO Journal (2019) 38: e101650

Introduction

Regulatory RNAs are key factors for efficient gene expression control in all domains of life. It is now clear that RNA regulators can rival transcription factors with respect to their regulatory scope, as many regulatory RNAs control multiple and sometimes dozens of transcripts (Hor *et al*, 2018). Various RNA-sequencing-based technologies have led to the discovery of RNA regulators from almost all regions of the genome. However, while these approaches provided a great deal of information about the expression, conservation, and overall distribution of regulatory RNAs, they allowed only limited conclusions toward their physiological roles (Cruz & Westhof, 2009; Storz *et al*, 2011).

Regulatory RNAs have now been found by the hundreds in bacterial genomes (Sorek & Cossart, 2010). The largest and most thoroughly studied group of bacterial RNAs are called small regulatory RNAs (sRNAs) and frequently associate with the RNA chaperone Hfq (Kavita *et al*, 2018). Hfq belongs to the large family of RNA-binding Lsm/Sm-like proteins and is required for efficient stabilization and annealing of sRNAs to their transcript targets. In analogy to their miRNA (microRNA) and crRNA (CRISPR RNA) counterparts, the sRNAs recognize cognate targets by a short stretch of base-pairing nucleotides, called the "seed" sequence (Gorski *et al*, 2017). Seed sequences are ~6–12 nucleotides long and structurally accessible. Hfq-dependent sRNAs of γ-proteobacteria have been reported to carry up to three seed-pairing domains, and mutation of either of these domains results in loss of regulation for a subset of target mRNAs (Herzog *et al*, 2019). Initial base-pairing by the seed typically relies on Watson–Crick base-pairing; however, it is not fully understood how these interactions discriminate against off-target interactions involving non-canonical G-U base-pairs (Papenfort *et al*, 2012).

Several sRNAs have been investigated for their seed-pairing capacities (Gorski *et al*, 2017). Here, the RybB sRNA, controlling envelope homeostasis of Gram-negative bacteria, has emerged as a model to study the mechanisms underlying seed-pairing regulatory RNAs (Bouvier *et al*, 2008; Balbontin *et al*, 2010; Papenfort *et al*, 2010). Transcription of RybB is controlled by the alternative sigma factor σ^E (encoded by the *rpoE* gene; Johansen *et al*, 2006; Papenfort *et al*, 2006; Thompson *et al*, 2007). σ^E belongs to the large class of extracytoplasmic function σ factors (ECFs), which are negatively controlled by a corresponding anti-sigma factor (Sineva *et al*, 2017). Under regular growth conditions, σ^E activity is weak as the protein is tethered to the inner membrane-bound anti-sigma factor, RseA. Misfolded outer membrane proteins (OMPs) trigger a cascade of regulated proteolysis events degrading RseA and releasing σ^E into the cytoplasm. σ^E associates with the core RNA polymerase and directs transcription toward specific promoters. Besides RybB, σ^E activates ~100 genes in *Escherichia coli* and related bacteria (Rhodius *et al*, 2006), including two additional Hfq-dependent

1 Faculty of Biology I, Department of Microbiology, Ludwig-Maximilians-University of Munich, Martinsried, Germany
2 Munich Center for Integrated Protein Science (CIPSM), Munich, Germany
3 Institute of Information Science, TH Köln – University of Applied Sciences, Cologne, Germany
4 ZB MED – Information Centre for Life Sciences, Cologne, Germany
*Corresponding author. Tel: +49 8921 8074502; E-mail: kai.papenfort@lmu.de

sRNAs, MicA and MicL. RybB and MicA regulate multiple target mRNAs in response to activation of σ^E. Targets of MicA and RybB are enriched for mRNAs encoding OMPs, suggesting that the sRNAs function to reduce the production of newly synthesized OMPs when the outer membrane is damaged (Brosse & Guillier, 2018). In contrast, MicL acts to inhibit translation of the highly abundant Lpp protein, which tethers the outer membrane to the peptidoglycan layer (Guo *et al*, 2014). Given that σ^E-bound RNA polymerase is restricted to act as a transcriptional activator, the σ^E-activated sRNAs have been suggested to provide an important inhibitory function to the system (Gogol *et al*, 2011).

σ^E-dependent sRNAs have also been described in other organisms. For example, in the major human pathogen *Vibrio cholerae*, the VrrA sRNA is activated by σ^E and inhibits the expression of four mRNA targets: the transcripts of the major OMPs, OmpA and OmpT; the biofilm matrix protein, RbmC; and the ribosome hibernation protein, Vrp (Song *et al*, 2008, 2010, 2014; Sabharwal *et al*, 2015). VrrA has also been shown to promote the production of outer membrane vesicles and to modulate virulence (Song *et al*, 2008).

In this study, we harnessed transcriptomic data to search for σ^E-dependent genes in *V. cholerae*. Our analysis identified the MicV sRNA as a new member of the σ^E regulon and we show that MicV associates with Hfq to regulate multiple target transcripts, including several mRNAs encoding OMPs. Global identification of target mRNAs revealed that MicV and VrrA control at least 32 mRNAs. While each sRNA controls a set of specific transcripts, the majority of targets are shared by the two sRNAs. We discovered that a conserved seed-pairing sequence present in MicV and VrrA accounts for the overlapping target regulation and that combined mutation of *micV* and *vrrA* impairs survival of *V. cholerae* under membrane-damaging conditions. This phenotype can be overcome by overexpression of MicV, VrrA, or RybB from *E. coli*, which also carries the conserved seed-pairing sequence of σ^E-dependent sRNAs. By employing an sRNA library carrying randomized base-pairing sequences, we show that the seed-pairing domain of σ^E-dependent sRNAs is strongly enriched during laboratory selection experiments and high-throughput sequencing of the selected seed sequences revealed a strong prevalence for sRNAs capable of repressing the OmpA protein. Indeed, deletion of *ompA* efficiently alleviated stress sensitivity of *rpoE*-deficient *V. cholerae*. Our data highlight the crucial role of seed-pairing domains in regulatory RNAs and describe a novel sRNA-based approach to study complex bacterial phenotypes in an unbiased fashion.

Results

MicV is a σ^E-dependent sRNA

We have recently determined the transcriptomes of *V. cholerae* under conditions of low and high cell densities and identified a total of 7,240 transcriptional start sites (TSS; Papenfort *et al*, 2015). However, these analyses did not provide information on how the activities of these TSS are controlled and which sigma factors could be involved. To address this question, we used a bioinformatics approach and searched for the σ^E consensus motif upstream of the 7,240 TSS in *V. cholerae*. We discovered 73 TSS associated with the σ^E motif (Appendix Table S1), including several TSS of genes previously linked with σ^E. For example, the 73 TSS included the promoters for *lptD*, *rpoH*, and *rpoE* itself, which have been documented to be activated by σ^E in *E. coli* (Rhodius *et al*, 2006), as well as the promoter for the VrrA sRNA. Thus, our approach allows the identification of σ^E-controlled genes in *V. cholerae*.

In addition to VrrA, we discovered that the promoter of another 68-nucleotide sRNA, Vcr089 (Papenfort *et al*, 2015), carried a sequence that aligned with the σ^E consensus (Fig 1A). The *vcr089* sRNA is conserved among *Vibrios* (Fig 1A). In analogy to the σ^E-dependent MicA sRNA (Udekwu *et al*, 2005), we renamed this sRNA MicV. In *V. cholerae*, the *micV* gene is located in the intergenic region of the *vc2640* (encoding a hypothetical protein) and *vc2641* (encoding argininosuccinate lyase) genes (Fig EV1A). Northern blot analysis showed that MicV expression peaks in stationary phase and that two MicV isoforms can be detected: the full-length transcript and a processed shorter variant (Fig EV1B). A similar expression pattern is observed for the VrrA sRNA (Fig EV1B). We also recovered both MicV isoforms in Hfq co-immunoprecipitation experiments (Fig EV1C), and MicV stability was strongly reduced in *hfq*-deficient *V. cholerae* (Fig EV1D) showing that MicV is a Hfq-dependent sRNA. For comparison, VrrA stability was only mildly affected in cells lacking *hfq* (Fig EV1D).

To test whether *micV* is controlled by σ^E, we generated an *rpoE* deletion mutant in *V. cholerae*. The *rpoE* gene is considered essential in *V. cholerae* and to avoid unpredictable suppressor mutations, we first deleted the *vchM* gene, encoding a known suppressor of σ^E (Chao *et al*, 2015), followed by deletion of *rpoE*. The *vchM* deletion did not affect the expression of MicV or VrrA (Fig EV1B). Northern blot analysis of MicV and VrrA showed that both sRNAs are undetectable in cells lacking *vchM* and *rpoE* (from here on referred to as Δ*rpoE*), whereas plasmid-borne production of σ^E from the inducible P_{BAD} promoter strongly activated the expression of both sRNAs (Fig 1B). To compare MicV and VrrA expression detected by Northern blot analysis with the activity of their associated promoters, we generated mKate2-based transcriptional reporters for both sRNAs and monitored production of the fluorescent protein at various points in growth. In wild-type *V. cholerae*, activity of the *micV* promoter was weak in exponentially growing cells and strongly increased when cells entered stationary phase growth (Fig 1C). Comparable levels were found for the P*vrrA*::mKate2 reporter (Fig EV1E). Mutation of *vchM* did not have a significant effect on the performance of both reporters; however, mKate2 production was hardly detectable in Δ*rpoE* cells (Figs 1C and EV1E). Given the conserved role of σ^E in enterobacteria, we also monitored the activity of P*micV*::GFP in wild-type and Δ*rpoE E. coli*. Again, GFP production in wild-type cells reached a maximum in stationary phase growth, whereas promoter activity was strongly reduced in *rpoE*-deficient cells (Fig EV1F). P_{BAD}-driven production of *V. cholerae* and *E. coli* σ^E rescued and further elevated GFP production in *E. coli*, indicating that the *micV* promoter is recognized and activated by σ^E.

VrrA is required for ethanol resistance in *Vibrio cholerae*

Exposure to ethanol has been reported to induce σ^E-mediated gene expression in *V. cholerae* (Chatterjee & Chowdhury, 2013). To test the effect of ethanol on *vrrA* and *micV* expression, we cultivated wild-type and Δ*rpoE V. cholerae* carrying the P*micV*::mKate2 or

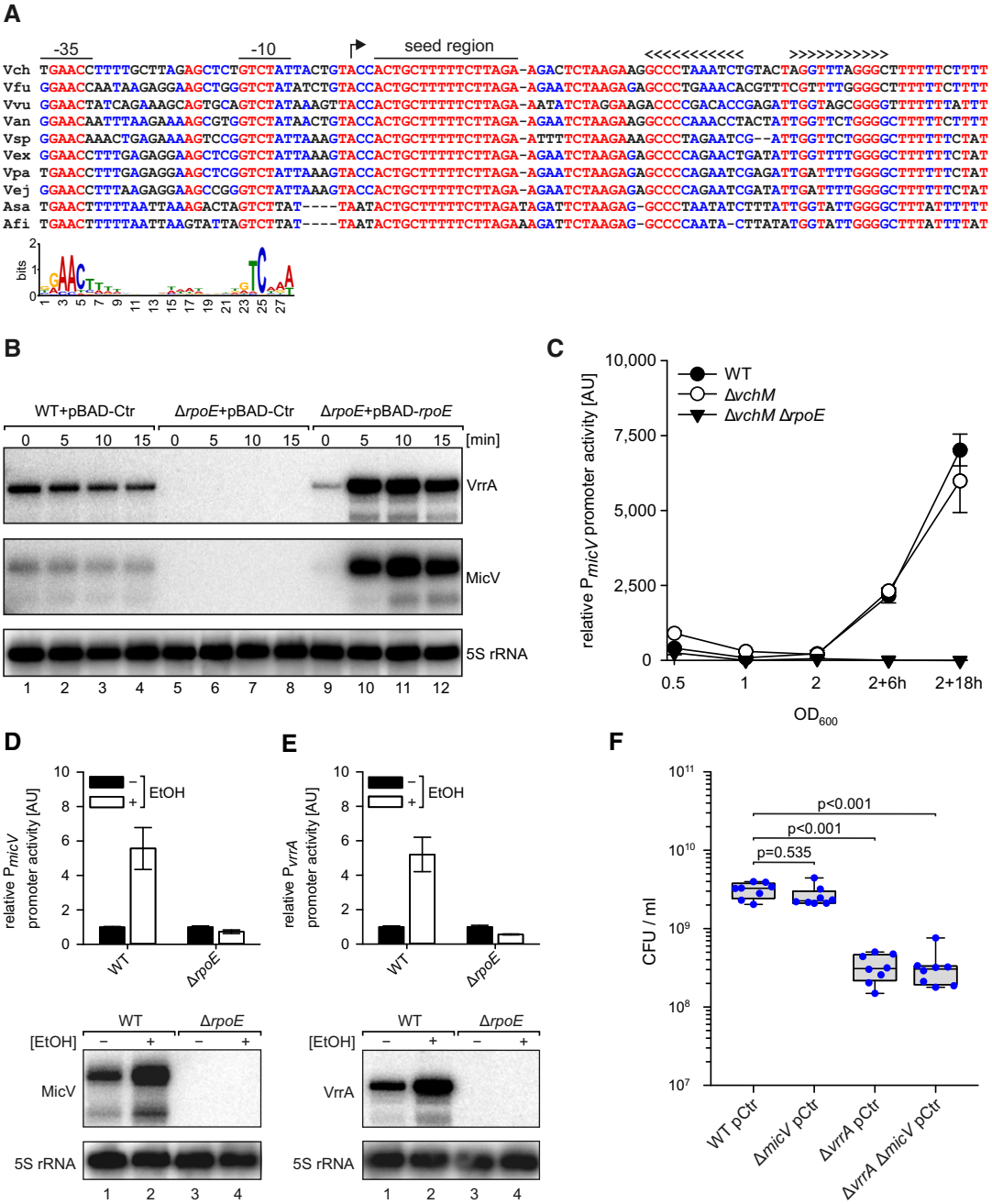


Figure 1. Transcriptional regulation of *micV*.

A Alignment of *micV* sequences, including the promoter regions, from various *Vibrio* species. The -35 box, -10 box, the TSS, the highly conserved seed region, and the rho-independent terminator are indicated. Lower part: consensus motif of *E. coli* σ^E-dependent promoters.

B *Vibrio cholerae* wild-type and Δ*rpoE* strains carrying the indicated plasmids were grown in LB medium to early stationary phase (OD_{600} of 1.5) and induced with L-arabinose (0.2% final conc.). Expression of MicV and VrrA was monitored on Northern blots. 5S rRNA served as loading control.

C *Vibrio cholerae* wild-type, Δ*vchM*, and Δ*vchM* Δ*rpoE* strains harboring the P*micV*::*mKate2* plasmid were grown in M9 minimal medium. Samples were collected at various stages of growth and analyzed for fluorescence.

D, E *Vibrio cholerae* wild-type and Δ*rpoE* strains carrying P*micV*::*mKate2* (D) or P*vrrA*::*mKate2* (E) plasmids were cultivated in LB medium to exponential phase (OD_{600} of 0.4) and treated with ethanol (3.5% final conc.) or water. Fluorescence was determined 180 min after ethanol treatment, and mKate2 levels of the mock-treated samples were set to 1. Corresponding Northern blot analyses of MicV and VrrA expression are shown at the bottom. 5S rRNA served as loading control.

F *Vibrio cholerae* wild-type, Δ*micV*, Δ*vrrA*, or Δ*vrrA* Δ*micV* strains were grown in LB medium to OD_{600} of 0.2 and treated with ethanol (3.5% final conc.). After 5 h of treatment, serial dilutions were prepared, recovered on agar plates, and CFU/ml were determined.

Data information: In (C–E), data are presented as mean ± SD, $n = 3$. In (F), the box plots indicate the median, 75th and 25th percentiles (boxes), and 90th and 10th percentiles (whiskers), $n = 8$. Statistical significance was determined using one-way ANOVA and post hoc Holm–Sidak test.
Source data are available online for this figure.

P*vrrA*::mKate2 reporters to exponential phase (OD_{600} of 0.4) and treated cells with ethanol (3.5% final conc.). After 180 min of exposure, we detected ~5-fold elevated mKate2 levels from the *micV* and *vrrA* promoters in wild-type cells, which we also confirmed at the transcript levels using Northern blot analysis (Fig 1D and E). In contrast, *rpoE*-deficient *V. cholerae* failed to significantly activate VrrA and MicV expression.

These results motivated us to investigate the role of *micV* and *vrrA* in *V. cholerae* challenged with ethanol. To this end, we treated exponential cultures (OD_{600} of 0.2) of wild-type, Δ*micV*, Δ*vrrA*, and Δ*vrrA* Δ*micV V. cholerae* with ethanol (3.5% final conc.) and determined CFU (colony forming units) after 5 h of incubation. While we discovered no significant difference in CFU for wild-type and Δ*micV V. cholerae*, Δ*vrrA* and Δ*vrrA* Δ*micV* cells displayed ~10-fold reduced CFU, when compared to the other two strains (Fig 1F). These data show that *micV* and *vrrA* are activated in response to membrane perturbations in *V. cholerae*, but only VrrA confers ethanol resistance.

MicV inhibits OmpT protein production

To study the role of MicV in gene regulation in *V. cholerae*, we constructed a MicV over-expression plasmid (pMicV) from which MicV production is driven by the constitutive P_{Tac} promoter. We transformed this plasmid into *V. cholerae* lacking *micV* and compared global changes in protein expression with wild-type and Δ*micV V. cholerae* carrying a control plasmid (pCtr) using SDS–PAGE (Appendix Fig S1A). Our data showed increased levels of a ~40 kDa protein in Δ*micV* cells and repression of the same protein when MicV was over-expressed (Appendix Fig S1A, compare lanes 3, 4 vs. 7, 8 vs. 11, 12). We excised the band from the gel and identified the protein as OmpT (VC1854) by mass spectrometry.

These data indicated that MicV inhibits OmpT expression in *V. cholerae*, as was also previously reported for the VrrA sRNA (Song *et al*, 2010). Given that both sRNAs are controlled by σ^E (Fig 1 and Song *et al*, 2008), we aimed to determine the contribution of each of the sRNAs to OmpT repression. To this end, we added a 3XFLAG epitope to the chromosomal *ompT* locus and monitored OmpT protein expression in wild-type, Δ*vrrA*, Δ*micV*, and Δ*vrrA* Δ*micV V. cholerae* at several stages of growth (Appendix Fig S1B, top). In accordance with the data presented in Appendix Fig S1A, OmpT production increased under stationary phase growth conditions and was elevated up to ~6-fold in cells lacking *micV*, when compared to wild-type *V. cholerae* (Appendix Fig S1B, lanes 1 vs. 3, 5 vs. 7, 9 vs. 11, and 13 vs. 15). OmpT was over-produced in Δ*micV* cells under all tested growth conditions (~6-fold at OD_{600} of 0.5 and 1.0, ~2.5-fold under stationary phase conditions), whereas lack of *vrrA* did not increase OmpT levels (lanes 2, 6, 10, and 14). However, mutation of both sRNAs (lanes 4, 8, 12, and 16) revealed an additive effect of the two sRNAs resulting in more than 12-fold higher OmpT levels when cells were cultivated to late exponential phase (OD_{600} of 1.0). Together, these results show that both MicV and VrrA repress OmpT production in *V. cholerae* with MicV being the dominant regulator under the tested conditions.

Over-production of OMPs has previously been reported to increase σ^E activity (Mecsas *et al*, 1993). Consequently, we predicted that elevated OmpT levels produced in Δ*micV* and Δ*vrrA* Δ*micV* cells would also increase σ^E activity in *V. cholerae*. To test this hypothesis, we collected total RNA samples using our previous experimental setup (Appendix Fig S1B) and probed VrrA and MicV levels on Northern blots. Indeed, we discovered increased VrrA accumulation in cells lacking *micV*, while mutation of *vrrA* had only a minor effect on *micV* expression (Appendix Fig S1B, bottom). In accordance with the hypothesis that increased OmpT production activates the ESR in *V. cholerae*, VrrA production was highest when OmpT levels were most strongly induced (~6-fold) in the *micV* mutant (Appendix Fig S1B, bottom, lanes 5 vs. 7). To corroborate these data with the status of the σ^E response, we used the P*micV*::mKate2 reporter as a proxy for σ^E activation in wild-type, Δ*vrrA*, Δ*micV*, and Δ*vrrA* Δ*micV V. cholerae* at several stages of growth (Appendix Fig S1C). When compared to wild-type *V. cholerae*, mKate2 levels did not change significantly in the *vrrA* mutant and were moderately induced in cells lacking *micV* (~1.5-fold). In contrast, mutation of both sRNAs had an additive effect on the activity of the *micV* promoter (~ 2.5-fold), suggesting that MicV and VrrA act redundantly to control σ^E activation in *V. cholerae*.

Global target profiles of MicV and VrrA in *Vibrio cholerae*

The VrrA sRNA has previously been shown to regulate multiple mRNAs through direct base-pairing, including *ompT* (Song *et al*, 2010). Likewise, we suspected that MicV inhibits *ompT* at the post-transcriptional level through translation repression and transcript degradation. To test whether MicV-mediated repression of *ompT* involves transcript degradation, we constructed an L-arabinose-inducible pBAD-*micV* plasmid. To avoid cross-regulation from chromosomal MicV and VrrA production, we transferred this plasmid into a *V. cholerae* mutant lacking *vrrA* and *micV* and investigated *ompT* levels by Northern blot analysis. Indeed, induction of MicV from this plasmid resulted in a rapid reduction of *ompT* mRNA (Fig 2A, lanes 3–7), whereas L-arabinose did not affect *ompT* levels in the same strain carrying a control vector (lanes 1–2). Moreover, the dynamics of MicV-mediated *ompT* repression were comparable to an equivalent experiment using a pBAD-*vrrA* plasmid (lanes 8–12), indicating that both sRNAs act post-transcriptionally on *ompT*.

These observations prompted us to design an experimental setup for the identification of MicV and VrrA target mRNA candidates at a genome-wide level. To this end, we cultivated Δ*vrrA* Δ*micV V. cholerae* carrying either pBAD-*micV*, pBAD-*vrrA*, or a control plasmid to early stationary phase (OD_{600} of 1.5) and induced sRNA expression from the P_{BAD} promoter for 10 min. Differentially expressed genes were determined by RNA-sequencing comparing cells induced for MicV or VrrA to the empty vector control. Transcripts displaying ≥ 3-fold change in abundance by either of the two sRNAs were considered potential target mRNAs. In total, we discovered 28 and 27 differently regulated genes for MicV and VrrA, respectively (Appendix Table S2 and Appendix Fig S2). Importantly, 23 of these targets, including *ompT*, were regulated by both sRNAs (Fig 2B).

Next, we tested whether the newly identified targets are regulated at the post-transcriptional level by MicV and/or VrrA. Specifically, we employed a well-established GFP-based reporter system tailored to determine post-transcriptional gene control in bacteria (Corcoran *et al*, 2012). In this system, the 5′ UTR (untranslated region) and the sequence corresponding to the first 20 amino acids of the target genes were fused to *gfp* under the control of the P_{TetO} promoter. These plasmids were transferred into *V. cholerae* along

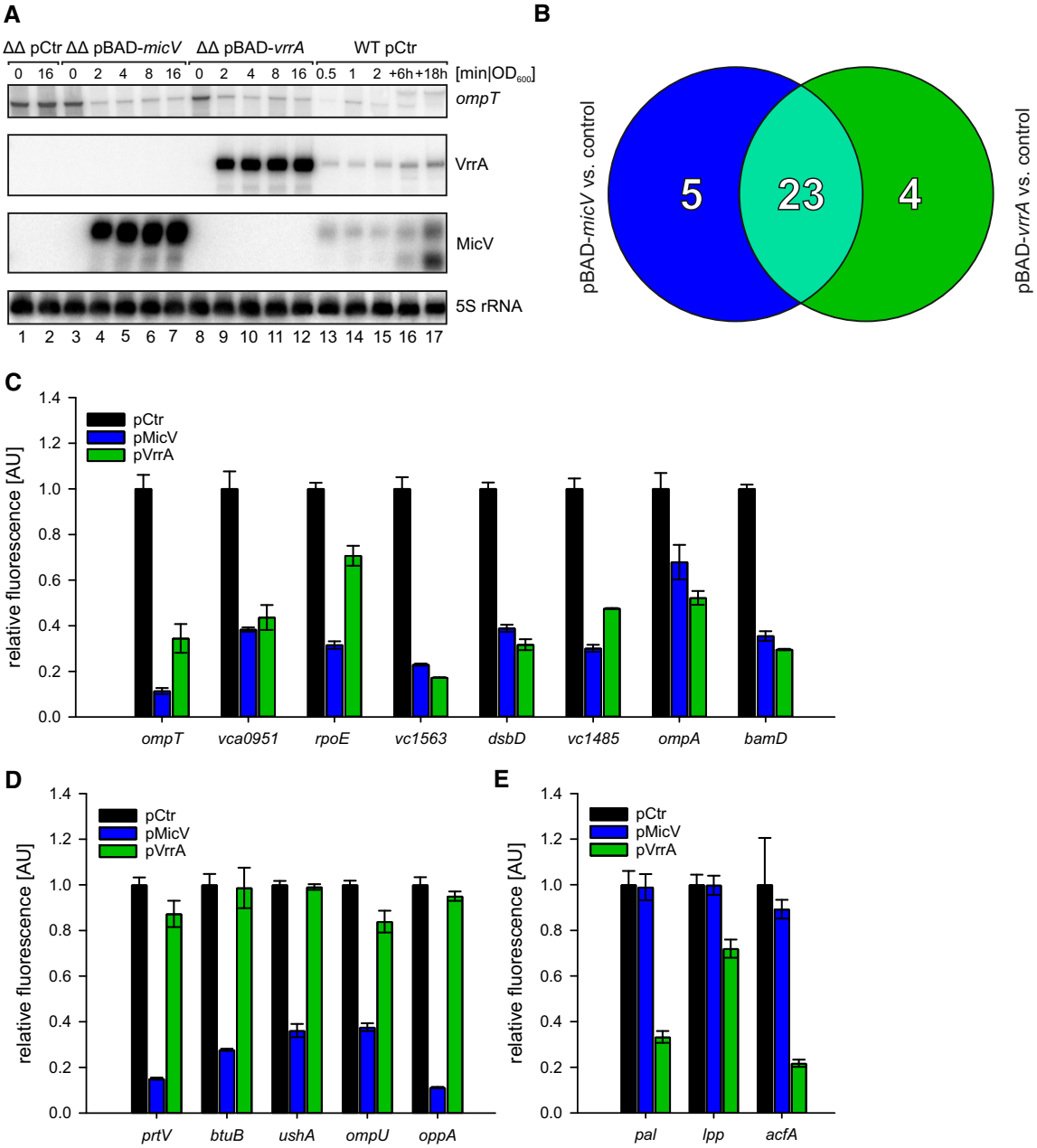


Figure 2. Target profiles of MicV and VrrA.

A *Vibrio cholerae* Δ*vrrA* Δ*micV* strains carrying pBAD-*micV*, pBAD-*vrrA*, or an empty vector control (pCtr) were cultivated to early stationary phase (OD_{600} of 1.5) in LB medium. Cells were treated with L-arabinose (0.2% final conc.), and RNA samples were collected at the indicated time points after induction. Northern blot analysis was performed to determine VrrA, MicV, and *ompT* levels. 5S rRNA served as loading control. For comparison, RNA samples of a wild-type strain carrying pCtr were collected during various growth phases, which indicated ~18-fold and ~7-fold higher levels of VrrA and MicV expressed from the pBAD plasmids, respectively (see Source data for quantifications).

B Venn diagram summarizing the RNA-Seq results: RNA samples were collected from *V. cholerae* Δ*vrrA* Δ*micV* strains carrying pBAD-*micV*, pBAD-*vrrA*, or an empty vector control. Depicted are genes displaying a fold change of ≥ 3 and FDR-adjusted p-value ≤ 1E-8 obtained from MicV-expressing conditions (blue) or *vrrA*-expressing conditions (green). Genes regulated by both sRNAs (fold change ≥ 3 in one condition, fold change ≥ 2.0 in the other) are depicted in light green.

C–E *Vibrio cholerae* Δ*vrrA* Δ*micV* strains carrying the indicated reporter plasmids (x-axis) and either an empty vector control (pCtr), the pMicV, or the pVrrA plasmid were cultivated in M9 minimal medium, and GFP fluorescence was measured. Fluorescence of the control strains was set to 1. The target genes were classified according to (B): regulated by both sRNAs (C), regulated only by MicV (D), or regulated only by VrrA (E).

Data information: In (C–E), data are presented as mean ± SD, $n = 3$.
Source data are available online for this figure.

with a second plasmid transcribing either *micV* or *vrrA* from a P_{Tac} promoter. We confirmed post-transcriptional control of 16 target mRNAs: Eight of these targets were regulated by both sRNAs, five targets were specific to MicV, and three targets were only regulated by VrrA (Fig 2C–E). Although gene expression changes obtained by RNA-sequencing were confirmed by qRT–PCR (Appendix Fig S2), we have not been able to further validate post-transcriptional control of *dsbA*, *vc1743-vc1744*, *vca0996*, *vc2240*, *vc1485*, *vc0429*, *vca0447*, *vca0845*, and *vca0789* by MicV and/or VrrA with the reporter assays (Appendix Table S2), suggesting that these genes might not be directly controlled by the sRNAs or that the relevant base-pairing sequences are lacking in the GFP reporter constructs. In agreement with our initial hypothesis, we discovered that VrrA and MicV both post-transcriptionally regulate *ompT* and that the protein products of several of the newly identified target genes are predicted to localize to the outer membrane or periplasmic space of *V. cholerae* (*e.g.*, OmpA, OmpU, Pal, Lpp, BamD, DsbD, BtuB, TolC, AcfA, and UshA). In addition, the operon encoding σ^E, the anti-sigma factor RseA, and the auxiliary regulators RseB and RseC is repressed at the post-transcriptional level by MicV and VrrA (Fig 2C, Appendix Fig S2A and Appendix Table S2). These data indicate that the σ^E response of *V. cholerae* comprises an auxiliary autoinhibitory loop that involves the base-pairing capacity of the two sRNAs.

Molecular basis for target mRNA recognition by MicV and VrrA

To understand how MicV and VrrA distinguish between shared and unique target genes, we selected three representative examples showing stable RNA duplexes (Appendix Fig S3) for further analysis: MicV-*ompT* (shared target), MicV-*ushA* (MicV-specific), and VrrA-*lpp* (VrrA-specific). We used the RNA hybrid algorithm (Rehmsmeier *et al*, 2004) to search for potential RNA duplexes formed between the sRNAs and their targets (Fig 3A–C and Appendix Fig S3A–C). For the MicV-*ompT* interaction, we predicted a 16-bp-long consecutive interaction involving the sequence of the MicV 5′ end and the ribosome binding site (RBS) of *ompT* (Fig 3A). The MicV and *ushA* RNA duplex was also predicted to involve the 5′ end of MicV; however, the interaction was shorter (10 bp) and required a sequence located upstream of the RBS in the 5′ UTR of *ushA* (Fig 3B). Interaction of VrrA and *lpp* was predicted to involve the RBS of *lpp* and a conserved sequence element located in the distal part of VrrA (nucleotides 90–107; Figs 3C and EV2A), which is separated from the sequence required to form the VrrA-*ompT* interaction (Song *et al*, 2010).

Next, we tested these predictions by mutational analysis (Figs 3D–F and EV2B). Point mutations in MicV (M1) abrogated repression of *ompT::gfp*, and conversely, mutation of *ompT* blocked target regulation by native MicV. Combination of the two dinucleotide mutations restored regulation and confirmed the predicted interaction (Fig 3A and D). The MicV M1 mutation also abrogated repression of *ushA::gfp*, which was restored by the compensatory change in the *ushA* mRNA (Fig 3E), showing that MicV uses a seed sequence located at the 5′ end of the sRNA to interact with *ompT* and *ushA*. To test the interaction between VrrA and *lpp*, we mutated three consecutive nucleotides in *vrrA* (M2, Fig 3C). Indeed, this mutation blocked repression of *lpp::gfp* and conversely mutation of the *lpp* interaction site prevented repression by native VrrA (Fig 3F). Regulation was restored and further increased when the two mutated variants were co-transformed, validating the predicted RNA duplex formation.

To study the relevance of shared versus specific base-pairing by MicV and VrrA in the context of the ESR, we introduced an inducible pBAD-*rpoE* plasmid into Δ*rpoE*, Δ*rpoE* Δ*vrrA*, Δ*rpoE* Δ*micV*, and Δ*rpoE* Δ*vrrA* Δ*micV V. cholerae*, cultivated these strains to early stationary phase (OD_{600} of 1.5), and induced the σ^E response by adding L-arabinose. We collected total RNA samples before and at several time points after σ^E induction and followed *ompT*, *ushA*, and *lpp* expression by qRT–PCR (Fig 3G–I). We also probed VrrA and MicV expression by Northern blot analysis, which validated the expected induction of these sRNAs (Fig EV2C). In all three cases, production of σ^E resulted in target mRNA repression. However, while *ompT* was inhibited by MicV or VrrA (Fig 3G), downregulation of *ushA* was significantly delayed in the absence of MicV (Fig 3H). Conversely, *lpp* repression relied on the presence of VrrA (Fig 3I). Together, our data show that MicV and VrrA both control a set of shared and specific targets, which are repressed upon activation of σ^E.

A conserved seed-pairing sequence in σ^E-dependent sRNAs

A comparison of the VrrA-*ompT* (Fig EV2D) and MicV-*ompT* (Fig 3A) RNA duplexes showed that both sRNAs sequester the RBS of *ompT*. Indeed, compensatory bp exchange experiments showed that region R1 of VrrA is required for *ompT* repression, while region R2 is dispensable (Fig EV2E). These data suggested that VrrA and MicV use similar seed-pairing domains to interact with *ompT*. An alignment of the VrrA R1 sequence with the 5′ end of MicV revealed a conserved sequence element of ten consecutive base-pairs, CRCUGCUUUU (R = purine), all of which engage in base-pairing with *ompT* (Fig 4A). In addition, the identical sequence was also found in the seed sequence of *rybB*, a σ^E-dependent sRNA conserved among enterobacteria but lacking in *V. cholerae* (Fig 4A and Papenfort *et al*, 2010). RybB acts analogous to MicV and VrrA by reducing the levels of OMP mRNAs, when the ESR is activated (Brosse & Guillier, 2018). Consequently, we hypothesized that all three sRNAs employ one conserved domain to mediate OMP repression. To test this idea, we performed three complementary experiments: First, we introduced a constitutive RybB plasmid (pRybB) into *V. cholerae* and compared OmpT production with strains carrying the VrrA plasmid, the MicV plasmid, or a vector control. In all three cases, sRNA over-expression resulted in strong OmpT repression (Fig 4B). Second, in the reciprocal experiment, we transferred the pMicV, pVrrA, or pRybB plasmids and a relevant control vector into a heterologous host, *i.e.*, *E. coli*. We cultivated these cells to stationary phase (OD_{600} of 2.0) and investigated total protein samples using SDS–PAGE and Coomassie blue staining. For all three sRNAs, we discovered repression of OmpA and OmpC (Fig 4C), which are previously reported targets of RybB (Papenfort *et al*, 2010). Third, we tested the effect of MicV, VrrA, and RybB over-expression on the survival of *rpoE*-deficient *V. cholerae* when challenged with ethanol. In agreement with previous observations (Kovacikova & Skorupski, 2002), treatment with ethanol (3.5% final conc.) drastically reduced the CFU of Δ*rpoE V. cholerae* when compared to wild-type cells (Fig 4D). In contrast, over-expression of either of the three sRNAs strongly suppressed this phenotype, with

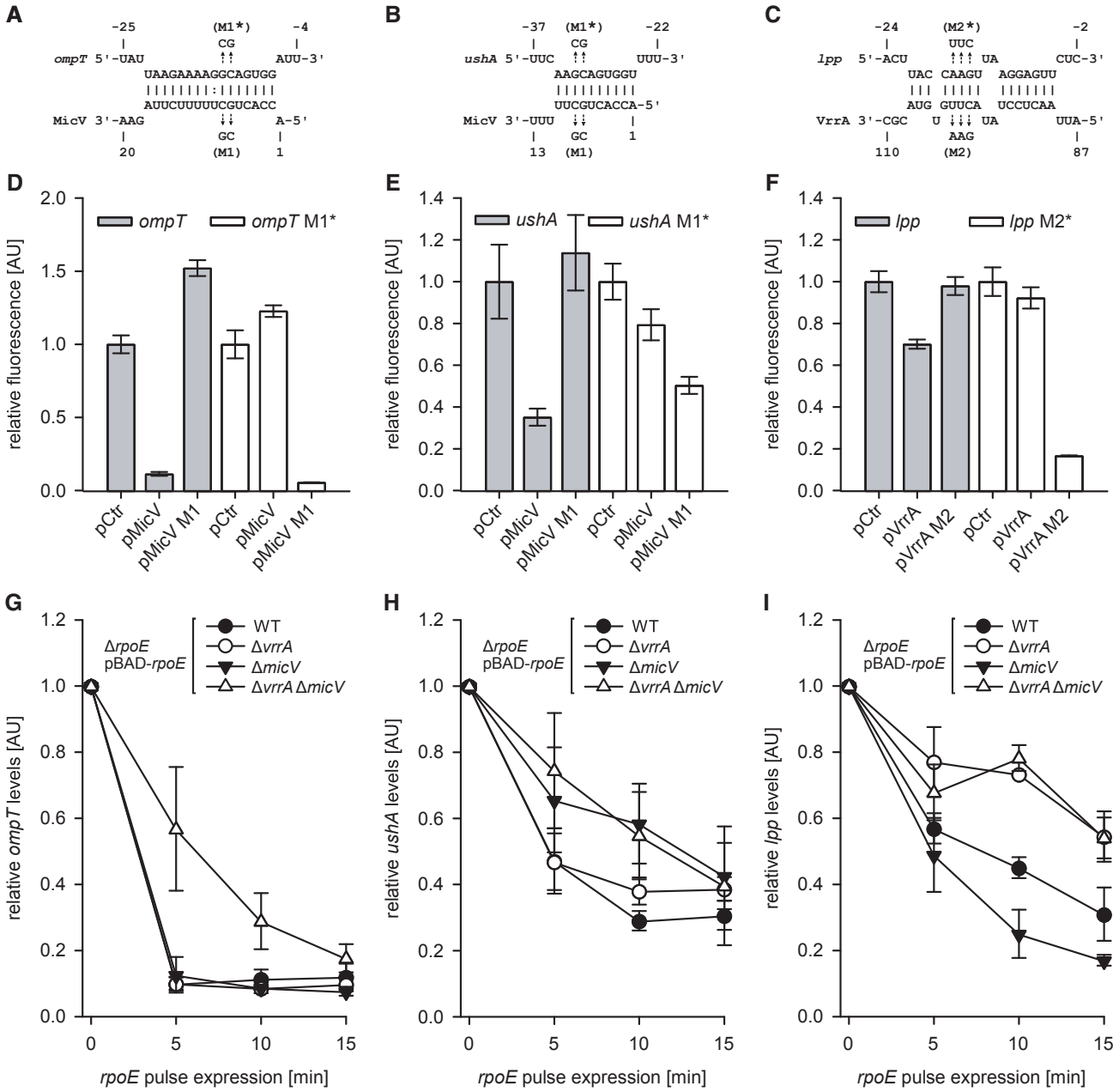


Figure 3. Patterns of target regulation by VrrA and MicV.

A–C Predicted base-pairings of MicV with the 5′UTR of *ompT* (A) and with the 5′UTR of *ushA* (B) or VrrA with the 5′UTR of *lpp* (C). Mutations tested in (D, E, F) are indicated.

D–F *Vibrio cholerae* Δ*vrrA* Δ*micV* strains carrying the *ompT::gfp* or *ompT* M1*::*gfp* fusions (D), *ushA::gfp* or *ushA* M1*::*gfp* fusions (E), or *lpp::gfp* or *lpp M2*::gfp* fusions (F) and an empty vector control (pCtr), the *micV* expression plasmids (pMicV, pMicV M1), or the *vrrA* expression plasmids (pVrrA, pVrrA M2) were grown in M9 minimal medium, and GFP fluorescence was measured. M1 and M2 denote the mutations indicated in (A, B, C). Fluorescence of the control strains was set to 1.

G–I *Vibrio cholerae* Δ*rpoE*, Δ*rpoE* Δ*vrrA*, Δ*rpoE* Δ*micV*, or Δ*rpoE* Δ*vrrA* Δ*micV* strains carrying pBAD-*rpoE* or an empty vector control (pCtr) were grown in LB medium to OD_{600} of 1.5, and L-arabinose (0.2% final conc.) was added. RNA samples were collected at the indicated time points and monitored for *ompT* (G), *ushA* (H), or *lpp* (I) levels using qRT–PCR.

Data information: In (D–I), data are presented as mean ± SD, $n = 3$.
Source data are available online for this figure.

VrrA and RybB supporting cell survival ~10-fold more efficiently than MicV. These data show that, when over-expressed, σ^E-dependent sRNAs can bypass the requirement of a conditionally essential transcriptional regulator, *i.e.*, σ^E, and suggested that a conserved seed sequence present in MicV, VrrA, and RybB is responsible for this phenotype.

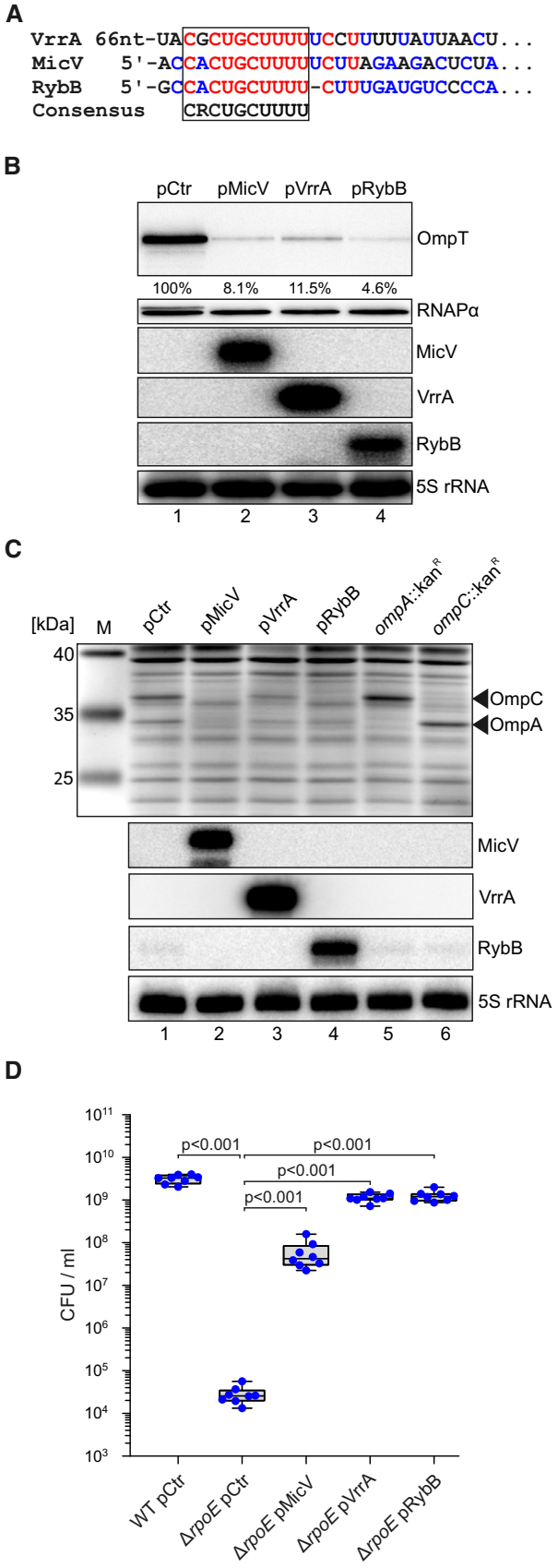


Figure 4. A conserved sRNA seed sequence inhibits OMP production.

A Alignment of the seed-pairing sequences of VrrA, MicV, and RybB.

B *Vibrio cholerae* Δ*vrrA* Δ*micV* strains carrying the *ompT*::3XFLAG gene and pMicV, pVrrA, pRybB, or an empty vector control (pCtr) were cultivated in LB medium to an OD_{600} of 2.0. RNA and protein samples were collected and analyzed for MicV, VrrA, and RybB expression on Northern blots. OmpT::3xFLAG production was tested on Western blots. RNAPα and 5S rRNA served as loading controls for Western and Northern blots, respectively.

C *Escherichia coli* wild-type strains carrying pMicV, pVrrA, pRybB, or an empty vector control (pCtr) were grown in LB medium to an OD_{600} of 2.0. RNA and protein samples were collected and investigated on Northern blots and SDS–PAGE, respectively. For comparison, we included the *E. coli* insertional mutant strains *ompA*::*kan*^{R} and *ompC*::*kan*^{R} for specific assignment of OmpA and OmpC bands.

D *Vibrio cholerae* wild-type and Δ*rpoE* strains carrying pMicV, pVrrA, pRybB, or an empty vector control (pCtr) were cultivated in LB medium to OD_{600} of 0.2 and treated with ethanol (3.5% final conc.). After 5 h of treatment, serial dilutions were prepared, recovered on agar plates, and CFU/ml were determined.

Data information: In (D), the box plots indicate the median, 75th and 25th percentiles (boxes), and 90th and 10th percentiles (whiskers), $n = 8$. Statistical significance was determined using one-way ANOVA and post hoc Holm–Sidak test. Source data are available online for this figure.

Strong enrichment for the seed-pairing sequence of σ^E-dependent sRNAs by selection experiments with randomized libraries

These results prompted us to develop an unbiased method to test the *in vivo* relevance of the conserved seed sequence present in MicV, VrrA, and RybB under stress conditions (Fig 5A). Specifically, we used the 3′ end of *E. coli* RybB, including the Hfq binding domain and the rho-independent terminator (Sauer *et al*, 2012), as a sRNA scaffold and randomized the first nine nucleotides of the seed-pairing sequence using a gene synthesis approach (see Materials and Methods for details). These constructs were cloned into a multi-copy plasmid and transferred into *V. cholerae* Δ*rpoE* cells. High-throughput sequencing of these plasmids revealed the presence of 253,570 sequence variants representing ~97% of all possible (262,144) permutations. Importantly, no single sequence variant constituted more than 0.0029% of the complete sRNA library (Fig EV3A), suggesting no major biases occurred during the construction process. Moreover, the nucleotide distribution was similar at the nine randomized positions (Fig EV3B). To select for sRNA variants providing improved stress resistance, we cultivated Δ*rpoE* cells containing the sRNA library to low cell density (OD_{600} of 0.2) and added ethanol (3.5% final conc.) to induce cell envelope stress. Following 6 h of incubation, cell dilutions were spotted on agar plates and screened for survival (Sel1; Fig 5B). Indeed, we observed a ~10-fold increase in survival of Δ*rpoE* cells carrying the sRNA library, when compared to *V. cholerae* Δ*rpoE* transformed with a control plasmid. Next, we collected the surviving cells and performed two additional rounds of selection. When compared to the Δ*rpoE* control carrying a control plasmid, survival was improved by ~1,000-fold in the second selection (Sel2) and by ~10,000-fold in the final selection (Sel3). Together, these data suggest that our approach allowed for the selection of sRNA variants providing ethanol resistance in *V. cholerae*.

To further investigate this possibility, we isolated the sRNA-containing plasmids from all three rounds of selection and

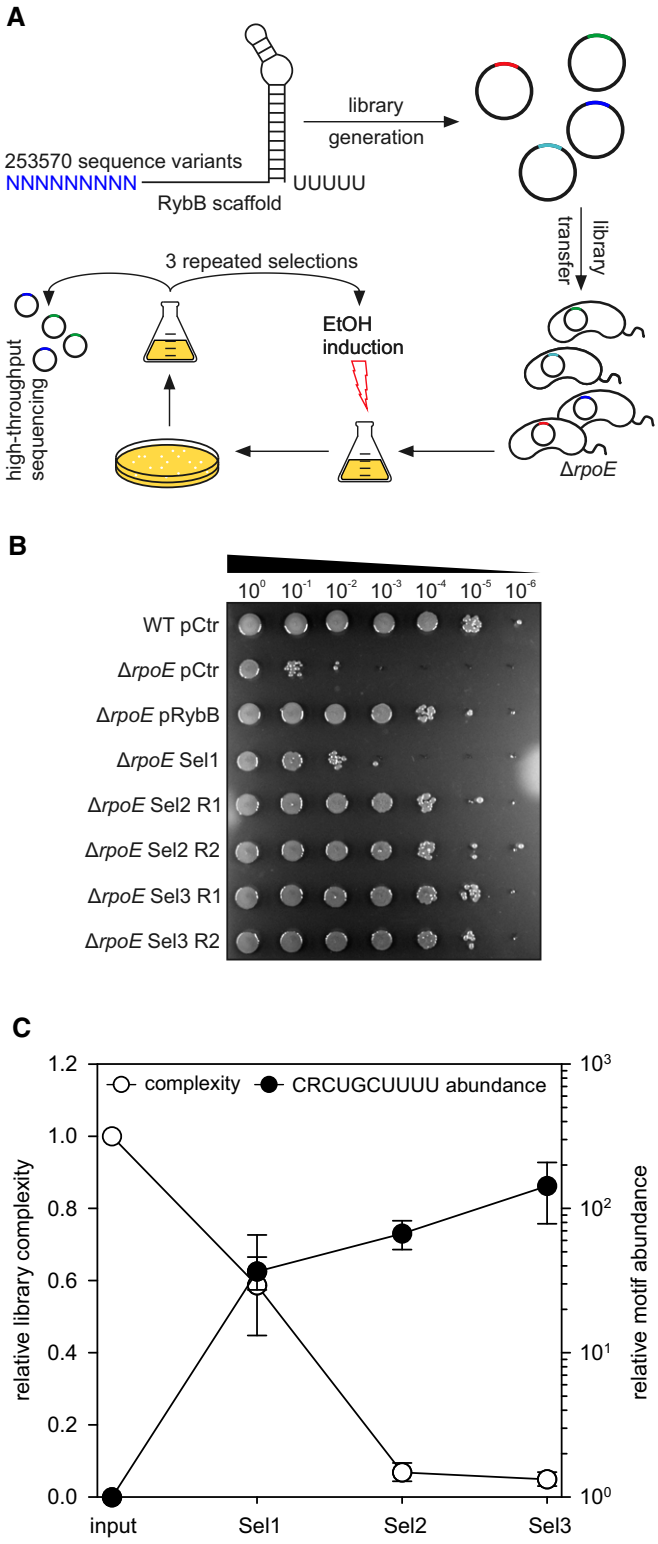
Figure 5. A conserved sRNA motif is enriched in laboratory selection experiments.

A Experimental strategy of the laboratory selection experiments: An sRNA library was generated using the *rybB* scaffold with nine randomized nucleotides at the 5′ end, cloned into a broad-range plasmid backbone, and transferred into *V. cholerae* Δ*rpoE* cells. These colonies were pooled, grown to OD_{600} of 0.2, and treated with ethanol (3.5% final conc.) for 6 h. Surviving cells were recovered on agar plates, pooled, and subjected to another round of selection (3 selections total). After each selection, the plasmids of surviving cells were analyzed using high-throughput sequencing.

B *Vibrio cholerae* wild-type and Δ*rpoE* strains carrying an empty vector control (pCtr), pRybB, or the sRNA library after consecutive selection experiments (Sel1, Sel2, and Sel3) were grown in LB medium to OD_{600} of 0.2. Cells were treated with ethanol (3.5% final conc.) for 6 h. Serial dilutions were prepared and spotted onto agar plates. R1 and R2 indicate two independent biological replicates.

C Plasmid contents of the strains carrying the sRNA libraries before selection (input) and after consecutive ethanol treatments (Sel1, Sel2, and Sel3) were analyzed using high-throughput sequencing. Relative library complexity (left *y*-axis) was determined by counting sequence variants present in the normalized samples. To test for the enrichment of possible sequence motifs, the sequence variants present in each sample were counted and normalized for sequencing depth. The resulting data were analyzed for the enrichment of the conserved CRCUGCUUUU motif (right *y*-axis).

Data information: In (C), data are presented as mean ± SD, $n = 2$.
Source data are available online for this figure.



determined the number of detectable sRNA sequence variants in two biological replicates using high-throughput sequencing. After the first round of selection, the number of detected sequence variants dropped by ~40% relative to the initial sRNA library and was further reduced to ~7% and ~5% in the following two selection steps, respectively (Fig 5C). We note that the steep drop in library complexity from the first to the second selection step coincided with a substantial increase in cell survival (compare Fig 5B and C). At the same time, we also discovered a very strong enrichment (~140-fold) of the conserved seed-pairing domain present in the MicV, VrrA, and RybB sRNAs (Figs 4A and 5C; for more details on enriched variants, see below and Figs EV4 and EV5), further documenting that this motif provides protection from ethanol-induced membrane damage in *V. cholerae*.

OmpA repression mediates ethanol resistance in *Vibrio cholerae*

To investigate the molecular basis of sRNA-mediated ethanol resistance in *V. cholerae*, we hypothesized that, analogous to the native RybB, VrrA, and MicV sRNAs, the selected sRNA variants could act by modulating the accumulation of OMPs in *V. cholerae*. To test this idea, we cultivated the initial and selected sRNA libraries in LB medium to stationary phase (OD_{600} of 2.0) and isolated membrane fractions to monitor OMP production (Fig 6A). We discovered a significant decrease in the abundance of two bands in the selected sRNA libraries (lanes 5–7), when compared to the initial library (lane 4). Similarly, over-expression of the RybB sRNA, which we have shown to mediate ethanol resistance (Fig 4D), also reduced these two bands (lane 3), and mutation of Δ*rpoE* resulted in increased protein levels when compared to a wild-type control (compare lanes 1 and 2). Using mass spectrometry, we determined that both bands corresponded to OmpA, which is detectable as a premature and mature variant (Freudl *et al*, 1986). We also discovered that the abundances of OmpT and OmpU, which are targets of MicV and VrrA (Fig 2C and D), did not change during these experiments. Similarly, qRT–PCR analysis of total RNA isolated during the selection process revealed that the selected sRNAs specifically repressed *ompA*, while the mRNA levels of additional MicV/VrrA targets encoding major OMPs (*ompT*, *lpp*, *pal*, and

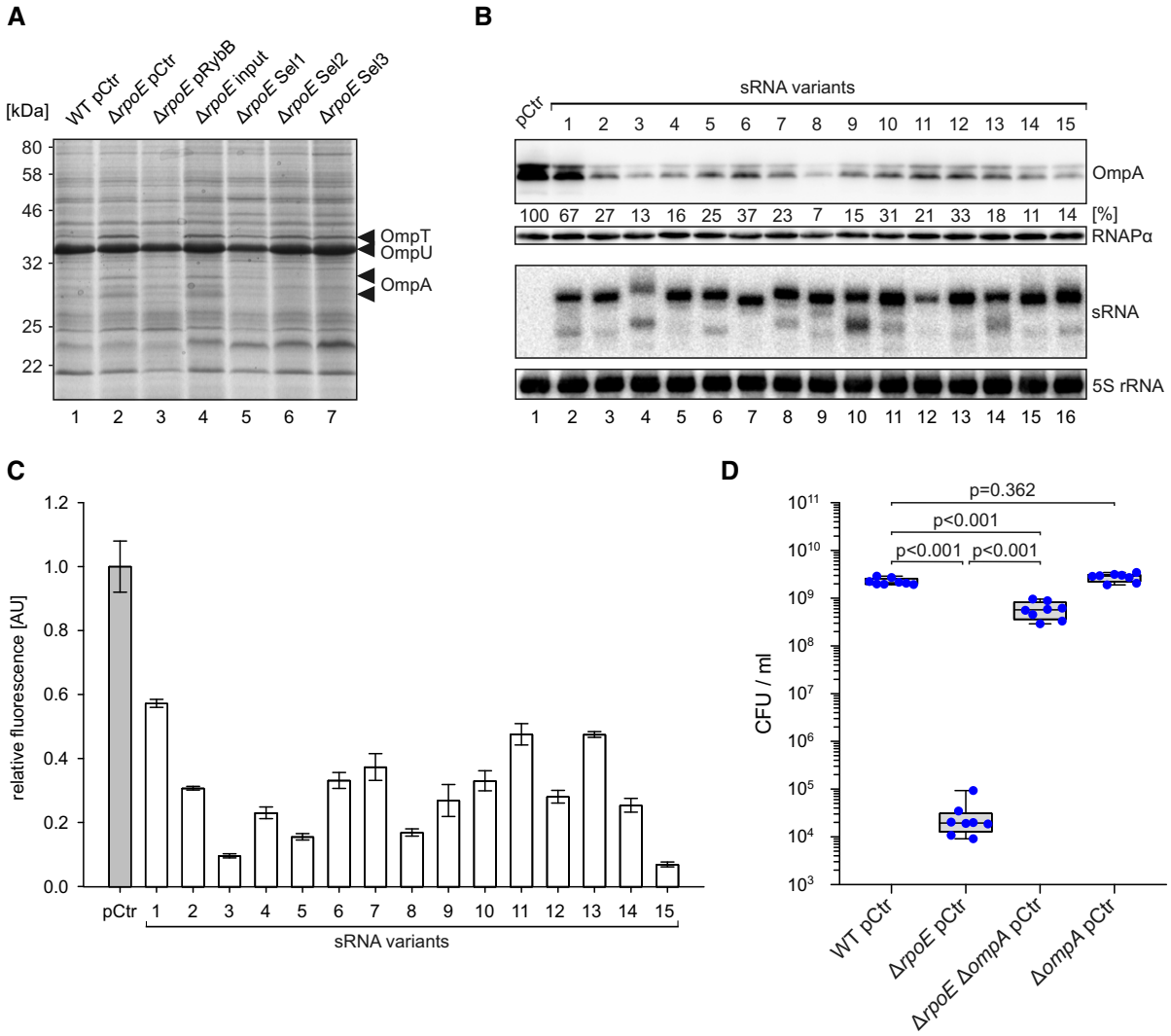


Figure 6. Enriched sRNA variants mediate ethanol resistance by OmpA repression.

A *Vibrio cholerae* wild-type and Δ*rpoE* strains carrying an empty vector control (pCtr), pRybB, or the sRNA library before (input) or after consecutive ethanol selection experiments (Sel1, Sel2, and Sel3) were cultivated in LB medium to OD_{600} of 2.0. Membrane fractions were identified by SDS–PAGE. The indicated bands were identified by mass spectrometry.

B *Vibrio cholerae* Δ*urrA* Δ*micV* cells expressing the *ompA*::3xFLAG gene and carrying an empty vector control (pCtr), or plasmids producing the 15 most highly enriched sRNA variants (sRNA variants 1–15) were grown in LB medium to an OD_{600} of 2.0. RNA and protein samples were collected and tested for sRNA and OmpA::3xFLAG expression on Northern and Western blots, respectively (with 5S rRNA and RNAPα as loading controls).

C *Vibrio cholerae* Δ*urrA* Δ*micV* strains carrying the *ompA*::*gfp* fusion and an empty vector control (pCtr) or the enriched sRNA expression plasmids were grown in M9 minimal medium, and GFP fluorescence was measured. Fluorescence of the control strains was set to 1.

D *Vibrio cholerae* wild-type, Δ*rpoE*, Δ*ompA*, or Δ*rpoE* Δ*ompA* strains were grown in LB medium to OD_{600} of 0.2 and treated with ethanol (3.5% final conc.). After 5 h of treatment, serial dilutions were prepared, recovered on agar plates, and CFU/ml were determined.

Data information: In (C), data are presented as mean ± SD, $n = 3$. In (D), the box plots indicate the median, 75th and 25th percentiles (boxes), and 90th and 10th percentiles (whiskers), $n = 8$. Statistical significance was determined using one-way ANOVA and post hoc Holm–Sidak test.
Source data are available online for this figure.

ompU) remained unchanged (Fig EV4A), suggesting that OmpA repression could be key for ethanol resistance in *V. cholerae*. To explore this possible link, we focused on the 15 most abundant sRNA variants obtained from our final round of selection (Sel3, Figs 5B and EV4B). These top 15 sRNA variants constituted ~54% of all detected sequence variants in the final selection (Fig EV4B) and were strongly enriched during the selection process (Fig EV4C). To confirm the regulatory capacity of these sRNA variants, we isolated all 15 plasmids, transformed them into independent *V. cholerae* Δ*rpoE* cells, and tested for ethanol resistance. In all 15 cases, the presence of the sRNA-expressing plasmid promoted survival (Fig EV4D). In contrast, a plasmid expressing only the *rybB* sRNA scaffold failed to restore ethanol resistance (Fig EV4D).

Next, we investigated the effect of the top 15 sRNAs on OmpA production. To this end, we added a 3XFLAG epitope to the chromosomal *ompA* locus of *V. cholerae* and transformed this strain with

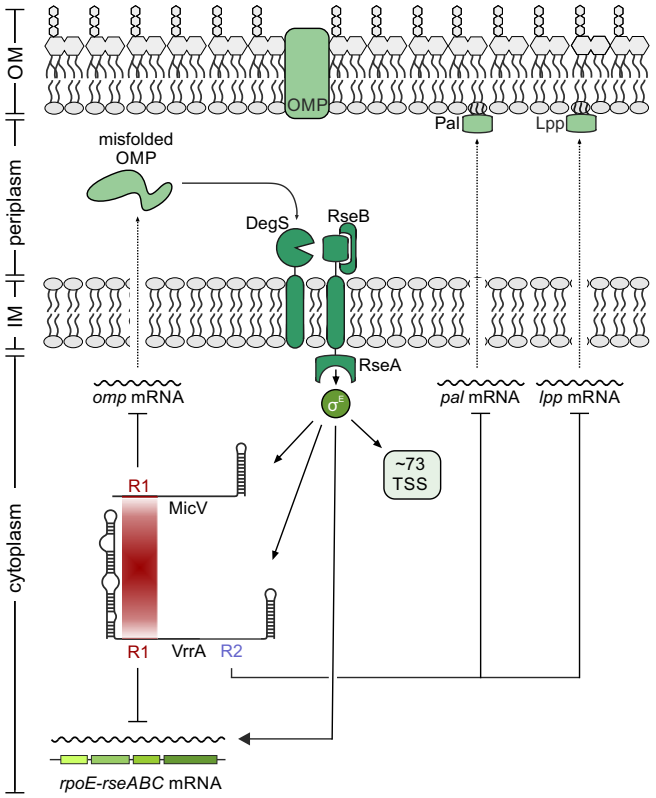


Figure 7. Conserved seed sequences control envelope homeostasis in *V. cholerae*.

Misfolded OMPs activate an intra-membrane proteolysis cascade resulting in the release of σ^E from its anti-σ factor RseA. Free σ^E activates the the expression of at least 73 transcripts in *V. cholerae*, including the *rpoE-rseABC* operon and the MicV and VrrA sRNAs. MicV and VrrA employ the conserved base-pairing region R1 to repress *omp* mRNAs, restoring membrane homeostasis, and the *rpoE-rseABC* operon. VrrA specifically downregulates *pal* and *lpp*, encoding two major lipoproteins, via the base-pairing region R2.

each of the 15 sRNA-expressing plasmids. We cultivated these cells to stationary phase (OD_{600} of 2.0) and monitored OmpA levels by Western blot (Fig 6B). For all 15 sRNA variants, we discovered significantly reduced OmpA production. However, the efficiency of the sRNA variants differed considerably with sRNA variant #1 providing only modest inhibition (~1.5-fold) and variant #8 showing the strongest repression (~14-fold), when compared to the control. All sRNA variants could be detected by Northern blot analysis indicating that the *rybB* 3′ end provides a stable sRNA scaffold (Fig 6B). To corroborate these results with a potential post-transcriptional regulatory mechanism exerted by the sRNA variants, we generated a *ompA::gfp* translational reporter whose transcription is driven by the constitutive P_{TetO} promoter. Co-transformation of this reporter with the plasmids expressing the sRNA variants into *V. cholerae* followed by GFP measurements revealed that all 15 sRNAs inhibited OmpA::GFP production (Fig 6C). Overall, the degree of OmpA protein repression observed in *V. cholerae* (Fig 6B) matched the results of the *ompA::gfp* reporter (Fig 6C), suggesting that OmpA repression by the sRNA variants occurs predominantly at the post-transcriptional level.

Our results indicated elevated OmpA levels as a possible cause of the increased ethanol sensitivity of *V. cholerae* Δ*rpoE* cells (Fig 6A–C). To test this possibility, we deleted the *ompA* gene in wild-type and Δ*rpoE V. cholerae* and assayed ethanol resistance. In line with our previous observation (Fig 4D), *V. cholerae* lacking *rpoE* displayed strongly reduced ethanol resistance, when compared to wild-type cells (Fig 6D). In contrast, cells deficient for *rpoE* and *ompA* exhibited ~10,000-fold improved survival when challenged with ethanol, whereas *V. cholerae* wild-type and Δ*ompA* strains showed highly similar survival numbers (Fig 6D). Taken together, our screen using synthetic sRNAs pinpointed OmpA repression as a key factor for ethanol tolerance in *V. cholerae* and provided evidence that regulation at the post-transcriptional level is a crucial for this phenotype.

Discussion

A main form of transcriptional regulation in bacteria occurs through the exchange of the primary sigma factor subunit of RNA polymerase with alternative sigma factors, which direct the complex to specific promoter sequences. In sharp contrast to σ^{70}, which recognizes the majority of promoters in enterobacterial cells, promoter recognition by extracytoplasmic function σ factors (ECFs) is highly stringent, which restricts the number of target promoters, allowing ECFs to mediate very specific responses (Campagne *et al*, 2015). In *E. coli*, σ^E has been reported to control 89 unique transcription units, which typically function to safeguard the synthesis and homeostasis of the outer membrane and its protein components (Rhodius *et al*, 2006). Here, we identified 73 potential σ^E-controlled TSSs in *V. cholerae* (Appendix Table S1), one of which is responsible for driving *micV* expression (Fig 1).

Three key signals have been suggested to modulate the activity of σ^E. First, σ^E responds to misfolded OMPs activating DegS-mediated cleavage of the anti-sigma factor, RseA, which results in the release of σ^E into the cytoplasm (Mecsas *et al*, 1993). Second, periplasmic lipopolysaccharide intermediates can disassemble the RseA-RseB complex and facilitate proteolytic degradation of RseA (Lima *et al*, 2013). Third, σ^E activity is also activated by limited nutrient availability, which is caused by the production of the alarmone ppGpp and its cofactor DksA (Costanzo *et al*, 2008). In all three cases, activation of σ^E results in the transcription of the *rpoE-rseA-rseB-rseC* operon; however, only conditions supporting continuous degradation of RseA will amplify the response (Chaba *et al*, 2007). Our data suggest the existence of an additional autoregulatory loop controlling the *rpoE-rseA-rseB-rseC* operon. In contrast to the activating function of σ^E on the *rpoE* promoter, MicV and VrrA both base-pair with and reduce the production of *rpoE* (Figs 2C and 7, and Appendix Figs S2A and S3A). σ^E-dependent sRNAs were previously reported to limit σ^E activation; however, in this case the underlying mechanism was associated with the inhibitory effect of these sRNAs on OMP production (Papenfort *et al*, 2006; Thompson *et al*, 2007). Direct repression of *rpoE* by MicV and VrrA could add an additional layer of autorepression, which is independent from the status of OMP synthesis and assembly. Recently, a global screen for base-pairing interactions of Hfq-binding sRNAs suggested that the MicL sRNA, which is also controlled by σ^E (Guo *et al*, 2014), binds to the *rpoE* mRNA in *E. coli* (Melamed *et al*, 2016).

Autorepression of their own transcriptional activator has now been reported for numerous sRNAs (Brosse & Guillier, 2018), and it is interesting to speculate that sRNA-mediated repression of *rpoE* could be a conserved feature of σ^E regulons in various bacterial species. Of note, base-pairing of these sRNAs with the *rpoE* mRNA is not only likely to decrease σ^E levels, but could also inhibit the production of the anti-sigma factor RseA, as well as RseB and RseC, which could further modulate the overall output of the response.

Our work also provides relevant insights into how sRNAs evolve in the context of microbial stress response systems. For example, the *micL* gene is located in the 3′ end of the *cutC* gene of *E. coli*, and although *V. cholerae* also encodes a *cutC* homolog (*vc0730*), we did not detect significant transcription from this locus in *V. cholerae* (Papenfort *et al*, 2015). However, MicL has been shown to repress Lpp synthesis in *E. coli* (Guo *et al*, 2014), and in *V. cholerae*, this function is carried out by VrrA (Figs 2E and 7). Pal, an outer membrane component relevant for cell division and outer membrane integrity (Gerding *et al*, 2007), is repressed by MicA in *E. coli* (Gogol *et al*, 2011) and reduced by VrrA in *V. cholerae*. Finally, OmpA is inhibited by MicA and RybB in *E. coli* and *Salmonella* (Rasmussen *et al*, 2005; Udekwu *et al*, 2005; Papenfort *et al*, 2010) and repressed by VrrA and MicV in *V. cholerae* (Figs 2C and 7). These data suggest that σ^E-dependent sRNAs act as functional analogs, similar to what has been proposed for the widespread group of RyhB-like sRNAs controlling bacterial iron homeostasis (Salvail & Masse, 2012). It also indicates that the establishment of an sRNA–target mRNA interaction is a dynamic process that is driven by the physiological constraints of the overarching physiological pathways (Updegrove *et al*, 2015). In other words, sRNA-mediated repression of major OMPs, such as Lpp, Pal, and OmpA, might be crucial for a fully functional σ^E response and regulation can be achieved by various different base-pairing interactions. Indeed, sequence comparison of the *lpp*, *pal*, and *ompA* base-pairing sites in *V. cholerae* (Fig 3 and Appendix Fig S3) and *E. coli* (Gogol *et al*, 2011; Guo *et al*, 2014) showed that these are not conserved among the two organisms.

Another question pertinent to the evolution of sRNAs and their targets is why certain mRNAs are controlled by two sRNAs, while others only require one. This is particularly interesting for sRNAs which are activated by the same transcription factor, such as RybB and MicA, or MicV and VrrA. Studies in *E. coli* have shown that RybB and MicA share repression of *lamB*, *ompA*, *ompW*, *tsx*, *htrG*, and *yfeX* (Gogol *et al*, 2011), whereas MicV and VrrA both regulate *ompT*, *vca0951*, *rpoE*, *vc1563*, *dsbD*, *vc1485*, *ompA*, and *bamD* (Fig 2C). In fact, the total number of potentially co-regulated MicV/VrrA targets is significantly higher (23; Fig 2B) given that several of these targets are organized in larger operons (Appendix Fig S2). One possible explanation could be that these mRNAs accumulate to high copy numbers in the cell and that rapid repression requires the action of two sRNA regulators. In addition, differences in sRNA stabilities (Fig EV1D) and potency in target regulation (*e.g.*, due to the accessibility of relevant base-pairing sequences) could add to the picture. However, following the decay of the *ompT* mRNA upon σ^E activation, we observed that either VrrA or MicV sufficiently reduced cellular *ompT* levels (Fig 3G). Despite this redundancy in regulation, certain targets are more efficiently regulated by one of the sRNAs. For example, mutation of *micV* resulted in significantly higher OmpT levels, which remained unchanged in cells lacking *vrrA* (Appendix Fig S1B). A potential division of labor among VrrA and MicV is also supported by our phenotypic observations. Laboratory selection experiments suggested that repression of OmpA is key for ethanol resistance of *rpoE*-deficient cells (Fig 6), and although both MicV and VrrA repress *ompA* (Fig 2C), only cells lacking *vrrA* or both *vrrA* and *micV* display a significant reduction in survival upon ethanol exposure (Fig 1F). Therefore, it is likely that VrrA is most relevant in ethanol-stressed cells, while MicV is the more dominant regulator under standard growth conditions, as evident from increased σ^E activity in Δ*micV* relative to wild-type and Δ*vrrA* cells (Appendix Fig S1C). Therefore, one might speculate that MicV and VrrA act as part of the global σ^E regulon to provide protection against specific stress conditions, *e.g.*, ethanol stress.

Detailed analyses of the sequences involved in base-pairing of VrrA and MicV revealed that both sRNAs share a highly conserved seed domain, which is also present in the RybB sRNA of *E. coli* and *Salmonella* (Fig 4A). The same sequence motif was also recovered in our laboratory selection experiments (Fig 5C); the exact *rybB* sequence was also among the top 15 sRNA candidates, which we tested for repression of *ompA* (#11; Fig EV5A). Inspection of the nucleotide distribution of the variable sequence in these 15 highly selected sRNAs revealed a preference for guanine and cytosine residues at the 5′ end of the sequence (Fig EV5B), which could facilitate stable seed pairing with target mRNAs (Gorski *et al*, 2017). Remarkably, all of the 15 sRNA variants selected from > 250,000 initial sequence variants inhibited OmpA production through direct base-pairing with the mRNA (Fig 6C). *In silico* prediction of the corresponding base-pairing sequences suggests that all 15 sRNAs act by blocking access of 30S ribosomes to the *ompA* mRNA (Fig EV5C), and it is also noteworthy that the majority of these sRNAs (12/15) are predicted to interact with a sequence immediately downstream of the *ompA* start codon. In fact, mutation of codons 2–5 in chromosomal *ompA* (while leaving the amino acid sequence unchanged) in Δ*rpoE* cells abrogated rescue of ethanol sensitivity by ten of the selected sRNA variants (Fig EV5C and D; sRNA variants #2 and #14 are still able to base-pair with the mutated *ompA* variant, while variants #1, #3, and #8 base-pair outside the *ompA* coding sequence). The same mutation in *ompA* is also predicted to abolish base-pairing of MicV and VrrA, and consequently, we discovered that a *V. cholerae* strain carrying this mutation displayed ~10-fold reduced ethanol resistance, when compared to the parental wild-type strain (Fig EV5E). This effect is comparable to the decreased ethanol resistance observed for the Δ*vrrA* Δ*micV* strain (Fig 1F).

Our laboratory selection experiment identified repression of OmpA as the single key factor for ethanol resistance, at least in *rpoE*-deficient cells. It would be interesting to test whether other membrane-damaging agents, such as antimicrobial peptides or related antibiotics, would result in the selection of other sRNA variants with altered target specificities. In general, we believe that our strategy of using synthetic sRNA libraries to screen complex microbial phenotypes could become a powerful genetic tool to circumvent the tedious and cost-intensive generation of gene deletion libraries.

Materials and Methods

Bacterial strains and growth conditions

All strains used in this study are listed in Appendix Table S3. Details for strain construction are provided in the

Appendix Supplementary Material and Methods section. *V. cholerae* and *E. coli* cells were grown under aerobic conditions (200 rpm, 37°C) in either LB or M9 minimal medium containing 0.4% glucose and 0.4% casamino acids (final conc.). For stationary phase cultures, samples were collected with respect to the time points when the cells reached an $OD_{600} > 2.0$, *i.e.*, 6 h and 18 h after cells reached an OD_{600} reading of 2.0. Where appropriate, media were supplemented with antibiotics at the following concentrations: 100 μg/ml ampicillin; 20 μg/ml chloramphenicol; 50 μg/ml kanamycin; 50 U/ml polymyxin B; and 5,000 μg/ml streptomycin.

Plasmids and DNA oligonucleotides

A complete list of plasmids and DNA oligonucleotides used in this study is provided in Appendix Tables S4 and S5, respectively. Details on plasmid construction are provided in the Appendix Supplementary Material and Methods section.

RNA isolation and Northern blot analysis

Total RNA was prepared and blotted as described previously (Papenfort *et al*, 2017). Membranes (GE Healthcare Amersham) were hybridized with [^{32}P]-labeled DNA oligonucleotides at 42°C or 63°C when using riboprobes. Riboprobes were generated using the MAXIscript™ T7 Transcription Kit (Thermo Fisher Scientific), according to the manufacturer's instructions. Signals were visualized using a Typhoon Phosphorimager (GE Healthcare) and quantified using GelQuant (BioChemLabSolutions).

Quantitative real-time PCR

Experiments were performed as previously described (Papenfort *et al*, 2017). Briefly, total RNA was isolated using the SV Total RNA Isolation System (Promega), according to the manufacturer's instructions. qRT–PCR was performed using the Luna Universal One-Step RT-qPCR Kit (New England BioLabs) and the MyiQ™ Single-Color Real-Time PCR Detection System (Bio-Rad). *recA* was used as a reference gene.

Transcript stability experiments

Stability of sRNAs was determined as described previously (Papenfort *et al*, 2015). Briefly, biological triplicates of *V. cholerae* wild-type (KPS-0014) and Δ*hfq* (KPS-0054) strains were grown to OD_{600} of 1.0 and transcription was terminated by addition of 250 μg/ml rifampicin. Transcript levels were probed and quantified using Northern blot analysis.

Hfq co-immunoprecipitation

Hfq co-immunoprecipitations were performed as previously described (Chao *et al*, 2012). Briefly, *V. cholerae* wild-type (KPS-0014) and *hfq::3xFLAG* tagged strains (KPS-0995) were grown in LB medium to OD_{600} of 2.0. Lysates corresponding to 50 OD_{600} units were subjected to immunoprecipitation, using monoclonal anti-FLAG antibody (Sigma, #F1804) and Protein G Sepharose (Sigma, #P6649).

RNA-Seq analysis

Biological triplicates of *V. cholerae* Δ*vrrA* Δ*micV* strains harboring the pBAD1K-Ctr, pBAD1K-*vrrA*, or pBAD1K-*micV* plasmids were grown to early stationary phase ($OD_{600} = 1.5$) in LB medium. sRNA expression was induced by addition of L-arabinose (0.2% final conc.). After 10 min of induction, cells were harvested by addition of 0.2 volumes of stop mix (95% ethanol, 5% (v/v) phenol) and snap-frozen in liquid nitrogen. Total RNA was isolated and digested with TURBO DNase (Thermo Fisher Scientific). Ribosomal RNA was depleted using Ribo-Zero kits (Epicentre) for Gram-negative bacteria, and RNA integrity was confirmed using a Bioanalyzer (Agilent). Directional cDNA libraries were prepared using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (NEB, #E7760). The libraries were sequenced using a HiSeq 1500 System in single-read mode for 100 cycles. The read files in FASTQ format were imported into CLC Genomics Workbench v11 (Qiagen) and trimmed for quality and 3′ adaptors. Reads were mapped to the *V. cholerae* reference genome (NCBI accession numbers: NC_002505.1 and NC_002506.1) using the "RNA-Seq Analysis" tool with standard parameters. Reads mapping to annotated coding sequences were counted, normalized (CPM), and transformed (log_2). Differential expression between the conditions was tested using the "Empirical Analysis of DGE" command. Genes with a fold change ≥ 3.0 and a FDR-adjusted *P*-value ≤ 1E-8 were defined as differentially expressed.

Western blot analysis

Experiments were performed as previously described (Papenfort *et al*, 2017). If not stated otherwise, 0.05 OD/lane were separated using SDS–PAGE, stained with "Coomassie blue-silver," or transferred to PVDF membranes for Western blot analysis. 3XFLAG-tagged fusions were detected using anti-FLAG antibody (Sigma, #F1804). RNAPα served as a loading control and was detected using anti-RNAPα antibody (BioLegend, #WP003).

Preparation of membrane protein fractions

Preparation of membrane protein fractions was performed as described previously with minor modifications (Thein *et al*, 2010). Briefly, bacteria were grown to an OD_{600} of 2.0, harvested by centrifugation, and washed in buffer 1 (0.2 M Tris–HCl pH 8, 1 M sucrose, 1 mM EDTA, and 1 mg/ml lysozyme). Cells were centrifuged (200,000 *g*, 4°C, 45 min), and the resulting pellet was resuspended in buffer 2 (10 mM Tris–HCl pH 7.5, 5 mM EDTA, 0.2 mM DTT, and 0.5 mg/ml DNase). Cells were opened using a Bead Ruptor (OMNI International; 6 passes, 30-s ON, 30-s OFF, 40% amplitude, 4°C) and centrifuged to pellet unbroken cells (15,700 *g*, 4°C, 15 min). The resulting supernatants were subjected to ultra-centrifugation (300,000 *g*, 4°C, 3 h) to obtain membrane fractions.

Fluorescence measurements

Fluorescence assays to measure GFP expression were performed as described previously (Corcoran *et al*, 2012). *Vibrio cholerae* strains expressing translational GFP-based reporter fusions were grown

overnight in M9 minimal medium and resuspended in PBS. Fluorescence intensity was quantified using a Spark 10 M plate reader (Tecan). *Vibrio cholerae* strains carrying mKate2 transcriptional reporters were grown in M9 minimal medium, samples were collected at the indicated time points, and mKate2 fluorescence was measured using a Spark 10 M plate reader (Tecan). Control samples not expressing fluorescent proteins were used to subtract background fluorescence.

Generation of a synthetic sRNA library

To construct the synthetic sRNA library, a 210 bp P_L-*rybB* fragment was synthesized *in vitro* (GeneArt) with random nucleotides at positions 1–9 of *rybB*. The fragment was re-amplified with KPO-1491/1492 and cloned into the pMD30 backbone with XbaI and XhoI. Ligated plasmids were transformed into *E. coli* S17 by electroporation and plated on selection agar. Single colonies were harvested by washing the cells off the plates with sterile PBS. Two million clones were pooled to obtain eightfold coverage. The library was conjugated into *V. cholerae* Δ*rpoE lacZ::kanR* to allow for selection on kanamycin and chloramphenicol. Again, single colonies were pooled to obtain the full library, which was subsequently used as input for the selection experiments. Complexity of the obtained library was determined using high-throughput sequencing of the isolated plasmids.

sRNA library sequencing and analysis

To assess the complexity of the initial and selected RybB libraries, the plasmids were re-isolated and digested with XbaI and XhoI. The obtained P_L-*rybB* fragment was purified from agarose gels and used as input for library generation using the NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB, #E7645) and sequenced using an Illumina MiSeq**.** The read files in FASTQ format were imported into CLC Genomics Workbench v11 (Qiagen) and trimmed to remove P_L promoter and *rybB* backbone sequences to obtain reads containing only the nine randomized nucleotides. Abundance of the individual sequences was determined using the custom python script FrequencyAnalyzer, accessible on GitHub (https://github.com/Loxos/srna-tool-kit-python). To normalize for different sequencing depths when comparing library complexity, 800,000 reads were sampled from each replicate and the number of different sequences was counted in each sample.

Ethanol stress assays

Vibrio cholerae strains were grown to exponential phase (OD_{600} of 0.2) in LB medium and challenged with ethanol (3.5% final conc.). Following 5 h of incubation, serial dilutions were prepared and recovered on agar plates to determine CFU/ml. For laboratory selection experiments, the initial Δ*rpoE* sRNA library and control strains (WT pCtr, Δ*rpoE* pCtr, and Δ*rpoE* pRybB) were grown in LB medium to exponential phase (OD_{600} of 0.2) and challenged with ethanol (3.5% final conc.). Following 6 h of incubation, cells were recovered on agar plates to test for survival. At least 1 million single clones were pooled to generate the enriched sRNA libraries, which were used as input for the next round of selection following the same protocol. High-throughput sequencing of the isolated plasmids after each selection step was used to determine library complexity and distribution of the sRNA variants.

Quantification and statistical analysis

Statistical parameters for the respective experiment are indicated in the corresponding figure legends. Details for the performed statistical tests are provided in the corresponding Source data files. Statistical analysis of CFU recovered during ethanol stress assays was performed as follows: The data were log10-transformed and tested for normality and equal variance using Kolmogorov–Smirnov and Brown–Forsythe tests, respectively. The data were tested for significant differences using one-way ANOVA and post hoc Holm–Sidak test or *t*-test. Significance levels are reported in the corresponding figure legends and Source data files. Statistical analysis was performed using SigmaStat v04 (Systat). No blinding or randomization was used in the experiments. No estimation of statistical power was used before performing the experiments, and no data were excluded from analysis.

Data and software availability

The datasets and computer code produced in this study are available in the following databases:

- RNA-Seq and NGS data: Gene Expression Omnibus (GEO) GSE125224 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE125224).
- Variant analysis computer scripts: GitHub (https://github.com/Loxos/srna-tool-kit-python).
- Motif search computer script: Zenodo (https://zenodo.org/record/2543422).

Expanded View for this article is available online.

Acknowledgements
We thank Helmut Blum, Stefan Krebs, and Andreas Brachmann for help with the RNA-sequencing experiments and Andreas Starick for excellent technical support. We also thank Dietrich H. Nies and Kirsten Jung for providing strains, Michaela Huber for help with the Hfq co-immunoprecipitation experiments, Martin Lehmann for support with the mass spectrometry analyses, and Daniel Hoyos for supplying a custom python script. We thank Jörg Vogel and Gisela Storz for comments on the manuscript and all members of the Papenfort laboratory for insightful discussions and suggestions. This work was supported by the German Research Foundation (DFG—GRK2062/1 and Exc114-2), the Human Frontier Science Program (CDA00024/2016-C), and the European Research Council (StG-758212). R.H. and K.P. acknowledge support by the Joachim Herz Foundation and Young Scholars' Programme of the Bavarian Academy of Sciences and Humanities, respectively.

Author contributions
NP, MH, RH, and KP designed the experiments; NP, MH, and RH performed the experiments; NP, MH, RH, KUF, and KP analyzed data; NP, MH, RH, and KP wrote the manuscript.

Conflict of interest
The authors declare that they have no conflict of interest.

References

Balbontin R, Fiorini F, Figueroa-Bossi N, Casadesus J, Bossi L (2010) Recognition of heptameric seed sequence underlies multi-target regulation by RybB small RNA in *Salmonella enterica*. *Mol Microbiol* 78: 380 – 394

Bouvier M, Sharma CM, Mika F, Nierhaus KH, Vogel J (2008) Small RNA binding to 5′ mRNA coding region inhibits translational initiation. *Mol Cell* 32: 827 – 837

Brosse A, Guillier M (2018) Bacterial small RNAs in mixed regulatory networks. *Microbiol Spectr* 6: RWR-0014-2017

Campagne S, Allain FH, Vorholt JA (2015) Extra Cytoplasmic Function sigma factors, recent structural insights into promoter recognition and regulation. *Curr Opin Struct Biol* 30: 71 – 78

Chaba R, Grigorova IL, Flynn JM, Baker TA, Gross CA (2007) Design principles of the proteolytic cascade governing the sigmaE-mediated envelope stress response in Escherichia coli: keys to graded, buffered, and rapid signal transduction. *Genes Dev* 21: 124 – 136

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

Chao MC, Zhu S, Kimura S, Davis BM, Schadt EE, Fang G, Waldor MK (2015) A cytosine methyltransferase modulates the cell envelope stress response in the cholera pathogen [corrected]. *PLoS Genet* 11: e1005666

Chatterjee E, Chowdhury R (2013) Reduced virulence of the *Vibrio cholerae* fadD mutant is due to induction of the extracytoplasmic stress response. *Infect Immun* 81: 3935 – 3941

Corcoran CP, Podkaminski D, Papenfort K, Urban JH, Hinton JC, Vogel J (2012) Superfolder GFP reporters validate diverse new mRNA targets of the classic porin regulator, MicF RNA. *Mol Microbiol* 84: 428 – 445

Costanzo A, Nicoloff H, Barchinger SE, Banta AB, Gourse RL, Ades SE (2008) ppGpp and DksA likely regulate the activity of the extracytoplasmic stress factor sigmaE in Escherichia coli by both direct and indirect mechanisms. *Mol Microbiol* 67: 619 – 632

Cruz JA, Westhof E (2009) The dynamic landscapes of RNA architecture. *Cell* 136: 604 – 609

Freudl R, Schwarz H, Stierhof YD, Gamon K, Hindennach I, Henning U (1986) An outer membrane protein (OmpA) of *Escherichia coli* K-12 undergoes a conformational change during export. *J Biol Chem* 261: 11355 – 11361

Gerding MA, Ogata Y, Pecora ND, Niki H, de Boer PA (2007) The trans-envelope Tol-Pal complex is part of the cell division machinery and required for proper outer-membrane invagination during cell constriction in *E. coli*. *Mol Microbiol* 63: 1008 – 1025

Gogol EB, Rhodius VA, Papenfort K, Vogel J, Gross CA (2011) Small RNAs endow a transcriptional activator with essential repressor functions for single-tier control of a global stress regulon. *Proc Natl Acad Sci USA* 108: 12875 – 12880

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

Guo MS, Updegrove TB, Gogol EB, Shabalina SA, Gross CA, Storz G (2014) MicL, a new sigmaE-dependent sRNA, combats envelope stress by repressing synthesis of Lpp, the major outer membrane lipoprotein. *Genes Dev* 28: 1620 – 1634

Herzog R, Peschek N, Fröhlich KS, Schumacher K, Papenfort K (2019) Three autoinducer molecules act in concert to control virulence gene expression in *Vibrio cholerae*. *Nucleic Acids Res* 47: 3171 – 3183

Hor J, Gorski SA, Vogel J (2018) Bacterial RNA biology on a genome scale. *Mol Cell* 70: 785 – 799

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

Kavita K, de Mets F, Gottesman S (2018) New aspects of RNA-based regulation by Hfq and its partner sRNAs. *Curr Opin Microbiol* 42: 53 – 61

Kovacikova G, Skorupski K (2002) The alternative sigma factor sigma(E) plays an important role in intestinal survival and virulence in *Vibrio cholerae*. *Infect Immun* 70: 5355 – 5362

Lima S, Guo MS, Chaba R, Gross CA, Sauer RT (2013) Dual molecular signals mediate the bacterial response to outer-membrane stress. *Science* 340: 837 – 841

Mecsas J, Rouviere PE, Erickson JW, Donohue TJ, Gross CA (1993) The activity of sigma E, an *Escherichia coli* heat-inducible sigma-factor, is modulated by expression of outer membrane proteins. *Genes Dev* 7: 2618 – 2628

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

Papenfort K, Pfeiffer V, Mika F, Lucchini S, Hinton JC, 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, Bouvier M, Mika F, Sharma CM, Vogel J (2010) Evidence for an autonomous 5′ target recognition domain in an Hfq-associated small RNA. *Proc Natl Acad Sci USA* 107: 20435 – 20440

Papenfort K, Podkaminski D, Hinton JC, Vogel J (2012) The ancestral SgrS RNA discriminates horizontally acquired Salmonella mRNAs through a single G-U wobble pair. *Proc Natl Acad Sci USA* 109: E757 – E764

Papenfort K, Forstner KU, Cong JP, Sharma CM, Bassler BL (2015) Differential RNA-seq of *Vibrio cholerae* identifies the VqmR small RNA as a regulator of biofilm formation. *Proc Natl Acad Sci USA* 112: E766 – E775

Papenfort K, Silpe JE, Schramma KR, Cong JP, Seyedsayamdost MR, Bassler BL (2017) A Vibrio cholerae autoinducer-receptor pair that controls biofilm formation. *Nat Chem Biol* 13: 551 – 557

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

Rehmsmeier M, Steffen P, Hochsmann M, Giegerich R (2004) Fast and effective prediction of microRNA/target duplexes. *RNA* 10: 1507 – 1517

Rhodius VA, Suh WC, Nonaka G, West J, Gross CA (2006) Conserved and variable functions of the sigmaE stress response in related genomes. *PLoS Biol* 4: e2

Sabharwal D, Song T, Papenfort K, Wai SN (2015) The VrrA sRNA controls a stationary phase survival factor Vrp of *Vibrio cholerae*. *RNA Biol* 12: 186 – 196

Salvail H, Masse E (2012) Regulating iron storage and metabolism with RNA: an overview of posttranscriptional controls of intracellular iron homeostasis. *Wiley Interdiscip Rev RNA* 3: 26 – 36

Sauer E, Schmidt S, Weichenrieder O (2012) Small RNA binding to the lateral surface of Hfq hexamers and structural rearrangements upon mRNA target recognition. *Proc Natl Acad Sci USA* 109: 9396 – 9401

Sineva E, Savkina M, Ades SE (2017) Themes and variations in gene regulation by extracytoplasmic function (ECF) sigma factors. *Curr Opin Microbiol* 36: 128 – 137

Song T, Mika F, Lindmark B, Liu Z, Schild S, Bishop A, Zhu J, Camilli A, Johansson J, Vogel J *et al* (2008) A new *Vibrio cholerae* sRNA modulates

colonization and affects release of outer membrane vesicles. *Mol Microbiol* 70: 100 – 111

Song T, Sabharwal D, Wai SN (2010) VrrA mediates Hfq-dependent regulation of OmpT synthesis in *Vibrio cholerae*. *J Mol Biol* 400: 682 – 688

Song T, Sabharwal D, Gurung JM, Cheng AT, Sjostrom AE, Yildiz FH, Uhlin BE, Wai SN (2014) Vibrio cholerae utilizes direct sRNA regulation in expression of a biofilm matrix protein. *PLoS One* 9: e101280

Sorek R, Cossart P (2010) Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity. *Nat Rev Genet* 11: 9 – 16

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

Thein M, Sauer G, Paramasivam N, Grin I, Linke D (2010) Efficient subfractionation of gram-negative bacteria for proteomics studies. *J Proteome Res* 9: 6135 – 6147

Thompson KM, Rhodius VA, Gottesman S (2007) SigmaE regulates and is regulated by a small RNA in *Escherichia coli*. *J Bacteriol* 189: 4243 – 4256

Udekwu KI, Darfeuille F, Vogel J, Reimegard J, Holmqvist E, Wagner EG (2005) Hfq-dependent regulation of OmpA synthesis is mediated by an antisense RNA. *Genes Dev* 19: 2355 – 2366

Updegrove TB, Shabalina SA, Storz G (2015) How do base-pairing small RNAs evolve? *FEMS Microbiol Rev* 39: 379 – 391

Expanded View Figures

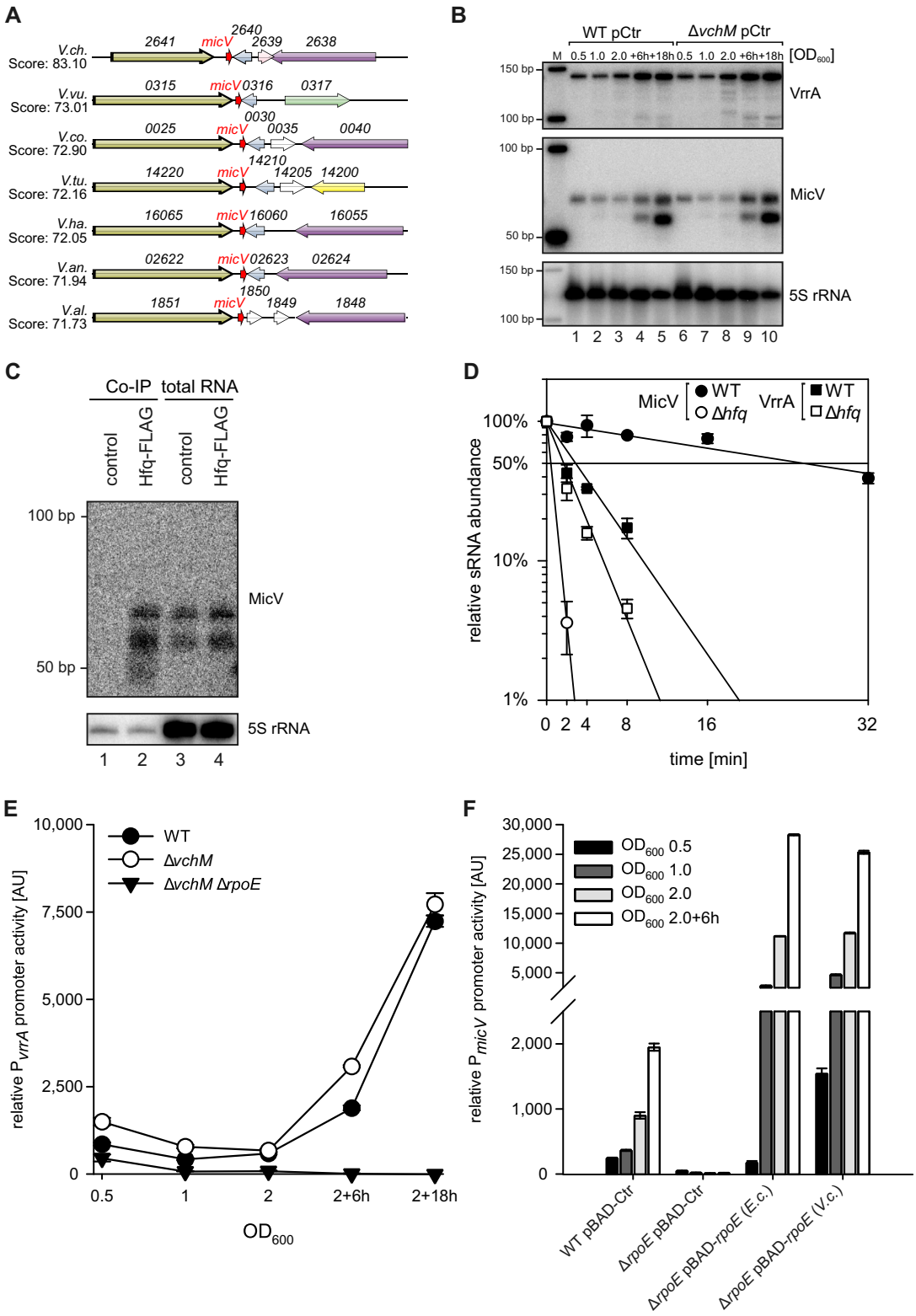


Figure EV1.

Figure EV1. Genomic context and conserved transcriptional control of σ^E-dependent sRNAs in *V. cholerae* (related to Fig 1).

A Gene synteny analysis between the genomic loci encoding *micV* in various *Vibrio* strains. Homologous genes are indicated by the same colors.

B *Vibrio cholerae* wild-type and Δ*vchM* strains carrying empty vector control plasmids (pCtr) were grown in LB medium. At the indicated time points, RNA samples were collected and tested for *micV* and *vrrA* expression by Northern blot analysis. A size marker is provided on the left (M), and 5S rRNA was used as loading control.

C *Vibrio cholerae* wild-type (control) and *hfq*::*hfq-3xFLAG* (Hfq-FLAG) strains were grown to stationary phase (OD_{600} of 2), lysed, and subjected to immunoprecipitation using the anti-FLAG antibody. RNA samples of lysate (total RNA) and co-immunoprecipitated fractions were analyzed on Northern blots. 5S rRNA served as loading control.

D *Vibrio cholerae* wild-type and Δ*hfq* strains were cultivated in LB medium to an OD_{600} of 1.0. Cells were treated with rifampicin to terminate transcription. Total RNA samples were collected at the indicated time points, and MicV or VrrA transcript levels were monitored on Northern blots.

E *Vibrio cholerae* wild-type, Δ*vchM*, and Δ*vchM* Δ*rpoE* strains harboring P*vrrA*::*mKate2* plasmids were grown in M9 minimal medium. Samples were collected at various stages of growth and analyzed for fluorescence.

F *Escherichia coli* BW25113 wild-type and Δ*rpoE* strains carrying P*micV*::*gfp* plasmids and either empty vector control (pBAD-Ctr) or plasmids expressing *rpoE* of *E. coli* (pBAD-*rpoE* (*E.c*)) or of *V. cholerae* (pBAD-*rpoE* (*V.c*)) were grown in LB medium, supplemented with L-arabinose (0.2% final conc.). Samples were collected at various stages of growth and analyzed for fluorescence.

Data information: In (D–F), data are presented as mean ± SD, $n = 3$.

Source data are available online for this figure.

Figure EV2. VrrA harbors two conserved base-pairing regions to regulate mRNA targets (related to Fig 3).

A Alignment of the *vrrA* sequences of several *Vibrio* species. The boxes indicate the conserved seed regions R1 and R2. Mutations used in (B, D, E) are indicated.

B *Vibrio cholerae* Δ*vrrA* Δ*micV* strains carrying pMicV, pMicV M1, pVrrA, pVrrA M1, pVrrA M2, or an empty vector control (pCtr) were grown to OD_{600} of 1.0 in LB medium. RNA samples were collected and monitored for *micV* and *vrrA* expression by Northern blot analysis. 5S rRNA served as loading control.

C *Vibrio cholerae* Δ*rpoE*, Δ*vrrA* Δ*rpoE*, Δ*micV* Δ*rpoE*, or Δ*vrrA* Δ*micV* Δ*rpoE* strains carrying pBAD-*rpoE* plasmids or an empty vector control (pCtr) were grown to early stationary phase (OD_{600} of 1.5), and *rpoE* expression was induced by treatment with L-arabinose (0.2% final conc.). RNA samples were collected at the indicated time points and monitored for *micV* and *vrrA* expression by Northern blot analysis. 5S rRNA served as loading control.

D Predicted base-pairing of VrrA with the 5′UTR of *ompT*.

E *Vibrio cholerae* Δ*vrrA* Δ*micV* strains carrying *ompT*::*gfp* or *ompT* M1*::*gfp* fusions and an empty vector control (pCtr) or *vrrA* expression plasmids (pVrrA, pVrrA M1, or pVrrA M2) were grown in M9 minimal medium. GFP fluorescence was measured, and fluorescence of the control strains was set to 1.

Data information: In (E), data are presented as mean ± SD, $n = 3$.

Source data are available online for this figure.

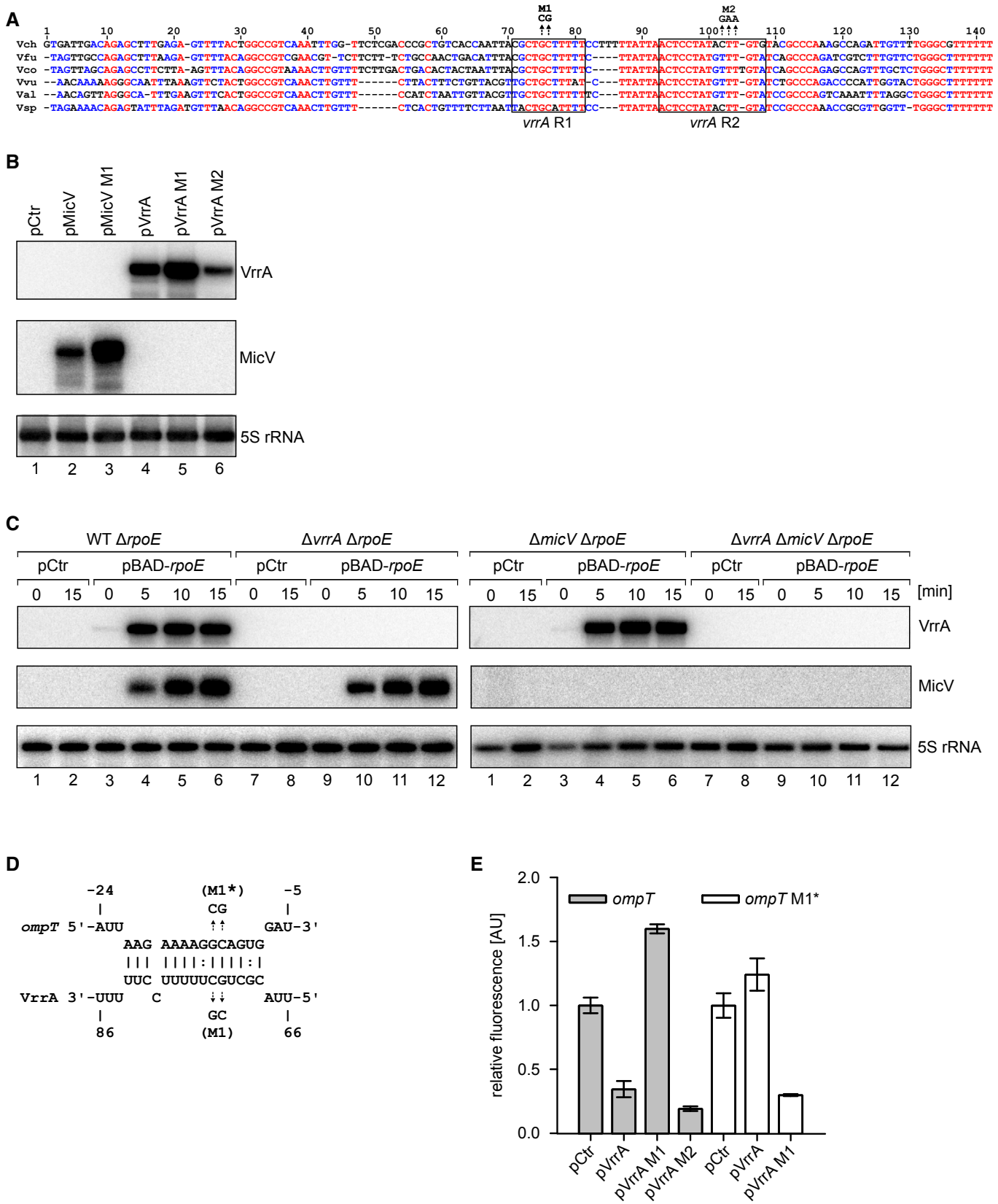

A
M1
CG
M2
GAA
Vch
Vfu
Vco
Vvu
Val
Vsp
vrrA R1
vrrA R2
B
pCtr
pMicV
pMicV M1
pVrrA
pVrrA M1
pVrrA M2
VrrA
MicV
5S rRNA
1 2 3 4 5 6
C
WT ΔrpoE
ΔvrrA ΔrpoE
ΔmicV ΔrpoE
ΔvrrA ΔmicV ΔrpoE
pCtr
pBAD-rpoE
0 15 0 5 10 15
[min]
VrrA
MicV
5S rRNA
1 2 3 4 5 6 7 8 9 10 11 12
D
-24
(M1*)
-5
ompT 5'-AUU
CG
GAU-3'
AAG AAAAGGCAGUG
UUC UUUUUCGUCGC
VrrA 3'-UUU C
GC
AUU-5'
86
(M1)
66
E
ompT
ompT M1*
relative fluorescence [AU]
2.0
1.5
1.0
0.5
0.0
pCtr
pVrrA
pVrrA M1
pVrrA M2
pCtr
pVrrA
pVrrA M1

Figure EV2.

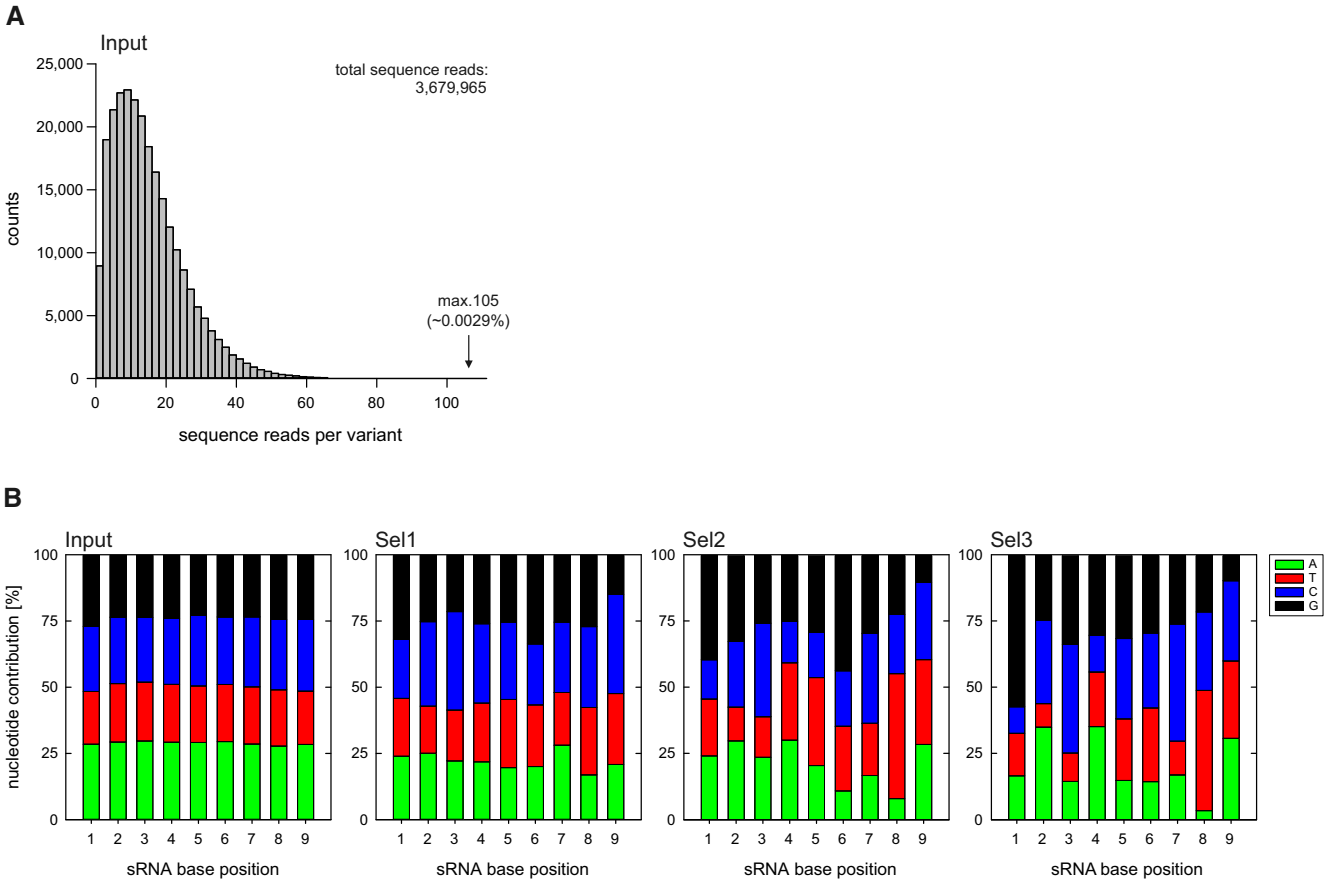


Figure EV3. Synthetic sRNA library composition and nucleotide contributions (related to Fig 5).

A, B A synthetic sRNA library based on a RybB scaffold with nine randomized nucleotides at the 5′ end was cloned into plasmid backbones and transferred into *V. cholerae* Δ*rpoE*. The resulting clones were pooled and treated with ethanol (3.5% final conc.) for 6 h. After treatment, the surviving cells were recovered on agar plates, pooled, and subjected to consecutive rounds of ethanol treatment for a total of three selections. After each selection, plasmid contents of surviving cells were analyzed by high-throughput sequencing. (A) Density histogram depicting the sequence read counts of obtained sRNA variants before ethanol treatment (Input). (B) Nucleotide contributions at the randomized positions in the synthetic sRNA libraries, before ethanol treatment (Input) and after consecutive ethanol treatments (Sel1, 2, 3). A = adenine, T = thymine, C = cytosine, G = guanine.

Data information: In (B), data are presented as mean, $n = 2$.
Source data are available online for this figure.

Figure EV4. Synthetic sRNA variants are enriched in laboratory selection experiments and mediate ethanol resistance (related to Fig 6).

A *Vibrio cholerae* Δ*rpoE* strains carrying the sRNA library before (input) or after consecutive ethanol selection experiments (Sel1, Sel2, and Sel3) were cultivated in LB medium to OD_{600} of 2.0. RNA samples were collected and analyzed for *omp* mRNA levels using qRT–PCR.

B Pie chart indicating the distribution of synthetic sRNA variants after three consecutive ethanol treatments (Sel3). The dashed red line indicates the fraction of the 15 most abundant sequence variants.

C The frequency of the 15 most abundant (top 15) sRNA variants was determined before ethanol treatment (Input) and after consecutive ethanol treatments (Sel1, 2, 3).

D *Vibrio cholerae* wild-type and Δ*rpoE* strains carrying an empty vector control (pCtr), synthetic sRNA expression plasmids (psRNA1-15), *rybB* expression plasmids (pRybB), or expression plasmids containing a *rybB* variant with deletion of nine nucleotides at the 5′ end (pRybBΔ9) were grown to OD_{600} of 0.2. Cells were treated with ethanol (3.5% final conc.) for 5 h. After treatment, the strains were serially diluted (1:10 steps) and spotted onto agar plates.

Data information: In (A, C), data are presented as mean ± SD, $n = 2$.
Source data are available online for this figure.

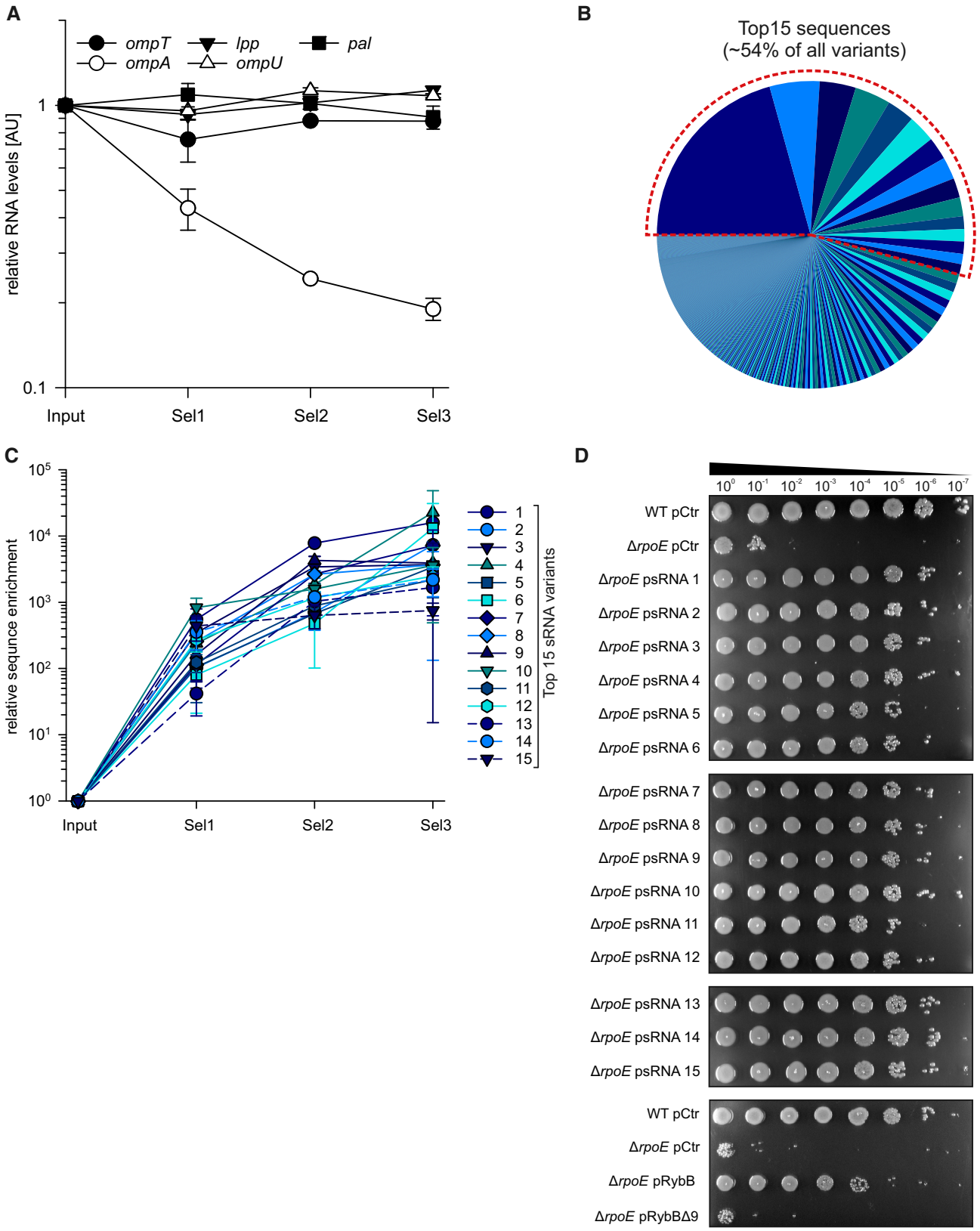

A
ompT
ompA
lpp
ompU
pal
relative RNA levels [AU]
1
0.1
Input
Sel1
Sel2
Sel3
B
Top15 sequences
(~54% of all variants)
C
relative sequnce enrichment
Top 15 sRNA variants
D
WT pCtr
ΔrpoE pCtr
ΔrpoE psRNA 1
ΔrpoE psRNA 2
ΔrpoE psRNA 3
ΔrpoE psRNA 4
ΔrpoE psRNA 5
ΔrpoE psRNA 6
ΔrpoE psRNA 7
ΔrpoE psRNA 8
ΔrpoE psRNA 9
ΔrpoE psRNA 10
ΔrpoE psRNA 11
ΔrpoE psRNA 12
ΔrpoE psRNA 13
ΔrpoE psRNA 14
ΔrpoE psRNA 15
WT pCtr
ΔrpoE pCtr
ΔrpoE pRybB
ΔrpoE pRybBΔ9

Figure EV4.

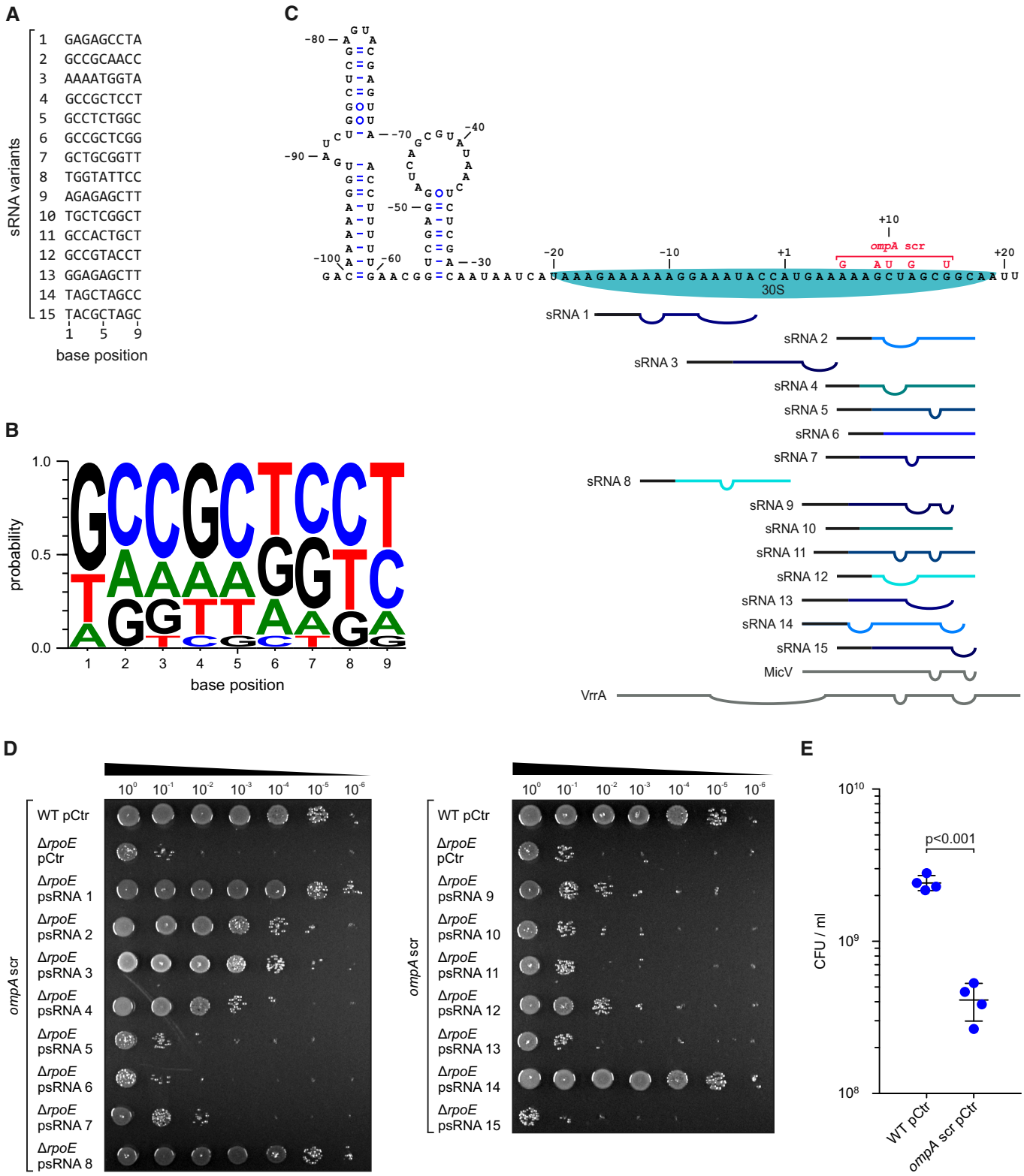


Figure EV5.

Figure EV5. Base-pairing of enriched sRNA variants to *ompA* mRNA is sufficient to mediate ethanol resistance (related to Fig 6).

A Sequence alignment of the 15 most abundant (top 15) sRNA variants.

B Consensus motif for the top 15 sRNA variants.

C Secondary structure model of *ompA* mRNA including the predicted base-pairing interactions of the top 15 sRNA variants. Straight lines indicate pairing bases, and bulges indicate non-pairing bases. Pairing bases corresponding to the variable region of the variants are depicted in color, and pairing bases corresponding to the backbone are depicted in black. MicV and VrrA are shown in gray. Numbers indicate the position on the *ompA* mRNA relative to the AUG start codon. The predicted position of the 30S ribosomal subunit and the *ompA* scr mutation are indicated.

D *Vibrio cholerae* wild-type *ompA* scr and Δ*rpoE ompA* scr strains carrying an empty vector control (pCtr) or synthetic sRNA expression plasmids (psRNA1-15) were grown to OD_{600} of 0.2. Cells were treated with ethanol (3.5% final conc.) for 5 h. After treatment, the strains were serially diluted (1:10 steps) and spotted onto agar plates.

E *Vibrio cholerae* wild-type and *ompA* scr mutant strains carrying empty vector controls (pCtr) were grown to OD_{600} of 0.2 and challenged with ethanol (3.5% final conc.). After 5 h of treatment, serial dilutions were prepared, recovered on agar plates, and CFU/ml were determined.

Data information: In (E), data are presented as mean ± SD, $n = 4$. Statistical significance was determined using a two-tailed, unpaired Student's *t*-test.
Source data are available online for this figure.

4 Concluding discussion

4.1 Studies of the RNA metabolism in a major pathogen

4.1.1 Identification of RNase E sites

The endoribonuclease RNase E plays a major role in post-transcriptional regulation by contributing to sRNA synthesis and turnover as well as target regulation [177, 295]. To study its relevance in *V. cholerae*, we identified RNase E cleavage sites (CS) on a genome-wide level by applying the TIER-seq method (transiently inactivating an endoribonuclease followed by RNA-seq [60]). This approach is based on the temperature-sensitive *rne-3071* mutation first described in *E. coli* [13], which we successfully transferred to the corresponding *V. cholerae* gene based on the strong conservation of the *rne* genes between *E. coli* and *V. cholerae* (chapter 2, Figs. 1A and 1-S1). We detected ~25.000 RNase E cleavage sites by comparing the 5' end profiles of transcripts from the wild-type and rne^{TS} mutant strains at the non-permissive temperature (chapter 2, Figs. 1 and 1-S2). This number is comparable to the TIER-seq results from *Salmonella* [60] and the α-proteobacterium *Rhodobacter sphaeroides* [105]. However, it has to be noted that the detected cleavage patterns can only provide a snapshot of the RNase E activity during the respective growth phase and conditions chosen for the experiment (late exponential phase in rich medium for our study). Especially many stress conditions such as starvation, hypoxia or temperature shock have been shown to globally affect mRNA turnover rates [173, 8, 328, 337]. This could be a result of differential RNase E cleavages, and it would be interesting to repeat the TIER-seq experiment under such conditions. Nevertheless, the present data provides valuable insights into the importance of RNase E for the RNA metabolism in *V. cholerae*.

Approximately 70% of all genes contained at least one cleavage site, supporting the global role of RNase E in bulk RNA turnover [145, 27]. While the majority of CS was found in coding sequences, we also detected 2.348 cleavage events (9.4%) in 5' UTRs. These might contribute to post-transcriptional control as it has been shown for the *cfa* 5' UTR in *Salmonella*, which gets protected from RNase E by binding of the sRNA RydC [108]. Another CS was found in the *rne* 5' UTR itself, which has been reported to be the subject of negative autoregulation through RNase E-mediated cleavage [154]. Although less frequent than in 5' UTRs, 1.475 cleavage sites (5.9%) were discovered in 3' UTRs. Some of these correspond at the same time to the 5' ends of sRNAs that are encoded within the respective 3' UTRs, indicating that these sRNAs are released from their mRNA transcripts through endonucleolytic processing [215, 60, 1]. In total, we detected matching RNase E sites for 15 different 3' UTR-derived sRNAs (chapter 2, Tab. S2). However, this likely

underestimates the true number of processed sRNAs from 3' UTRs, as the detection of an RNase E site by TIER-seq is also dependent on the stability of the respective RNA fragment. In another study recently published by our group, we showed that the FarS sRNA (formerly called Vcr076) is processed from the *fabB* 3' UTR by RNase E [144]. However, we did not detect the corresponding CS in the *fabB* mRNA in our TIER-seq analysis. This can be attributed to the high stability of FarS (half-life > 32 min): as we could not observe a decrease in mature FarS abundance during one hour of RNase E inactivation, our TIER-seq analysis was not able to capture the corresponding cleavage site. The same could be true for other 3' UTR-derived sRNAs, which would then be missed in our analysis. Furthermore, it is likely that not all existing sRNAs have yet been discovered in *V. cholerae*. One part of the aforementioned study on FarS was an Hfq-immunoprecipitation experiment, which revealed 30 new sRNA candidates [144]. About half of them are derived from 3' UTRs and for six of these potential new sRNAs (Vcr203, Vcr213, Vcr216, Vcr219, Vcr222, Vcr228) we could detect matching RNase E sites in our TIER-seq analysis.

RNase E is not only involved in the biogenesis of sRNAs, but also in their maturation and turnover. Our TIER-seq analysis captured the reported processing of the MicX sRNA into two smaller fragments, which we had used to test the functionality of our rne^{TS} mutant ([81] and chapter 2, Fig. 1A). Additionally, we detected internal CS in 46 sRNAs, with numbers ranging from 1 to 103 cleavage sites per sRNA. Remarkably different cleavage patterns could be observed for the three homologous CsrB/C/D sRNAs [180], which showed 65/6/103 cleavage sites, respectively. While the CsrB/C/D sRNAs were initially considered to be fully redundant transcripts [180], these deviating cleavage patterns support recent findings that the three Csr sRNAs differ in their regulation, stability and function [50]. In contrast, the four Qrr sRNAs share a conserved cleavage site at the same position in their second stem loop, which is involved in the regulation of almost all Qrr targets [304, 102]. Similarly, cleavage sites inside seed regions were found for MicX [81], RyhB [210], TfoR [353], VadR (Herzog *et al.*, 2020, in revision) and VqmR [251]. These could potentially contribute to coupled degradation after target regulation [197] or inactivation of the sRNAs in the absence of their targets [22].

RNase E is known to target its substrates through one of two processing pathways: in the 5' end-dependent pathway, a monophosphorylated RNA is recognized by a special binding pocket on the surface of RNase E, which stimulates enzymatic activity at a downstream cleavage site [189, 282]. The resulting 5' fragment is rapidly degraded by exonucleases, while the 3' fragment provides a new target for RNase E, leading to step-wise degradation of the full mRNA [145]. In contrast, the direct entry pathway is independent of the nature of the 5' end. Cleavage occurs efficiently at an internal site without 5' end binding, while secondary structures in the target might be important [74, 24]. These distinct pathways could explain the differential cleavage patterns that were observed for the two 3' UTR-derived sRNAs that were studied in detail in chapter 2, OppZ and CarZ: both are processed from their respective operons by RNase E, but cleavage frequency within the operons varies strongly. While the ~5.8 kb *oppABCDF-OppZ* operon shows 42 cleavage sites (on average 1 CS every 138 nt), only two cleavage sites can be found in the ~4.5 kb *carAB-*

carZ operon, one of them at the sRNA's 5' end. It is tempting to speculate that OppZ might be produced through step-wise processing of the *opp* operon in the 5' end-dependent pathway, while CarZ could be produced by a specific internal cleavage of the *carAB-carZ* operon through the direct entry pathway. However, this hypothesis is challenged by the appearance of multiple longer cleavage intermediates for both sRNAs when RNase E is inactivated (chapter 2, Fig. 1-S4), which rather point towards a similar step-wise processing of the *oppF* and *carB* mRNAs. To distinguish between these two pathways, one could employ several mutants that are impaired in different aspects of RNA turnover: (i) Targets of the 5' end-dependent pathway are often strongly stabilized upon inactivation of *rppH*, which encodes the pyrophosphohydrolase involved in conversion of the 5' triphosphate to monophosphate [88, 189]. (ii) Monophosphate sensing can be abolished by mutations in the 5' sensor domain, allowing only processing of substrates recognized via the direct entry pathway [113]. (iii) Conversely, RNase E mutations that block recognition of structured RNA regions affect the direct entry mode without interfering with 5' sensing [24]. Thus, studying the effects of an *rppH* mutant or of RNase E mutants defective in either of the two modes of target recognition could help to identify the relevant processing pathways for OppZ and CarZ.

Finally, we also detected substantial numbers of cleavage sites in intergenic regions (4.4%) and antisense to annotated genes (asRNA; 7.9%). RNase E activity on IGRs suggests the existence of non-annotated transcripts, which would presumably be non-coding due to the lack of encoded open reading frames. For instance, multiple CS were found within the new sRNA candidates proposed by [144], indicating that the identification of additional sRNAs could be guided by the inspection of IGRs with abundant RNase E sites. The amount of cleavage sites on the antisense strand is in line with reports on widespread antisense transcription from bacterial genomes, which has been assigned both specific and global regulatory roles [306, 274, 251, 114]. A detailed study of multiple individual antisense transcripts showed frequent processing by RNase E [331] and the abundance of CS in asRNA detected by our TIER-seq analysis is comparable to the data from *Salmonella* [60].

4.1.2 3' ends of transcripts and their sensitivity to bicyclomycin

In the publication presented in chapter 2, we showed that the autoregulatory sRNA OppZ induces premature transcription termination within the *opp* operon. To identify the steady-state 3' end of the terminated transcript, we applied Term-seq, a 3' end-specific RNA-seq protocol [76], to our wild-type strain and an *oppB* start codon mutant (chapter 2, Fig. 5C). We also sequenced RNA samples from cultures treated with the antibiotic bicyclomycin (BCM), which specifically inhibits the termination factor Rho by interfering with its ATPase activity [357, 318]. Comparison with mock-treated samples allowed us to determine the sensitivity of RNA 3' ends to bicyclomycin and thus the involvement of Rho in the respective transcription termination events on a genome-wide scale. The following section will focus on the general results obtained from the Term-seq data and BCM dependency, while the results specific to OppZ and CarZ will be discussed in the context of

their functional characterization in section 4.2.

To validate the results obtained by Term-seq, I analyzed the abundance of Term-seq reads at known termination sites in *V. cholerae*. In contrast to the thoroughly studied model organisms *E. coli* and *Salmonella*, the genome of *V. cholerae* is less well annotated and lacks annotations for full transcripts including their untranslated regions [132]. Thus, the sites of transcription termination are defined best for small transcripts that have been studied individually, mostly sRNAs. For all of the 24 sRNAs characterized in *V. cholerae*, the 3' end detected by Term-seq correlated well with the annotated termination site. Moreover, coverage at the 3' end of the transcript was very strongly increased compared to the upstream part. Together, these results indicate that our Term-seq experiment readily captured the steady-state 3' ends of transcripts. It has to be noted that most sRNAs terminate at Rho-independent (intrinsic) terminators and therefore have clearly defined 3' ends [67]. Although Rho-dependent termination has been reported to be more loosely controlled and to terminate in a diffuse manner, most Rho-dependent transcripts still exhibit distinct 3' ends [213, 285]. These can form for instance when the termination site coincides with a stable RNA structure or when structurally unprotected ends are subsequently trimmed by exonucleases [78]. In the case of such trimming, the steady-state 3' ends derived from Rho-dependent terminators do not necessarily match the actual site of transcription termination, but can also be located further upstream.

It was also noteworthy that despite the 3' end-specific protocol for library preparation, many sequencing reads still mapped to the full length of the transcript. The extent of such intragenically mapped reads was comparable to the data from the original Term-seq publication [76]. These reads do not necessarily indicate a lack of technical quality during library preparation, but rather correspond to ongoing transcription and/or degradation processes. In the context of our bicyclomycin treatment, we used these intragenic reads to globally compare the expression levels with or without the antibiotic. While this method cannot produce fully accurate results due to the bias for transcript 3' ends, it still allowed us to have a preliminary glimpse on the importance of Rho for the *V. choleare* transcriptome. For a full analysis of Rho-dependent transcription termination, we would need to prepare standard RNA-seq libraries from our BCM treatment experiment to obtain unbiased read coverage of the transcriptome [76, 78]. The clearest observation in our BCM data was the strong increase in antisense transcription: in non-treated samples, ~6.9% of all reads mapped antisense to annotated coding sequences, whereas this fraction increased fourfold to ~28.1% upon BCM treatment (Fig. 4.1A). This supports the observations from *E. coli* and *B. subtilis* that Rho and NusG are responsible for silencing genome-wide antisense transcription [264, 36]. Moreover, we could observe the action of Rho at some individual transcripts, such as its own messenger RNA. Rho is known to autoregulate its expression by terminating transcription within the 5' UTR of the *rho* gene [202]. We did indeed detect abundant RNA 3' termini in the *rho* 5' UTR, whose levels decreased upon Rho inhibition through BCM (Fig. 4.1B). Simultaneously, read coverage over the *rho* CDS increased strongly as a result of BCM-induced read-through from the upstream terminator. On a more global scale, our analysis revealed 150 repressed and 549

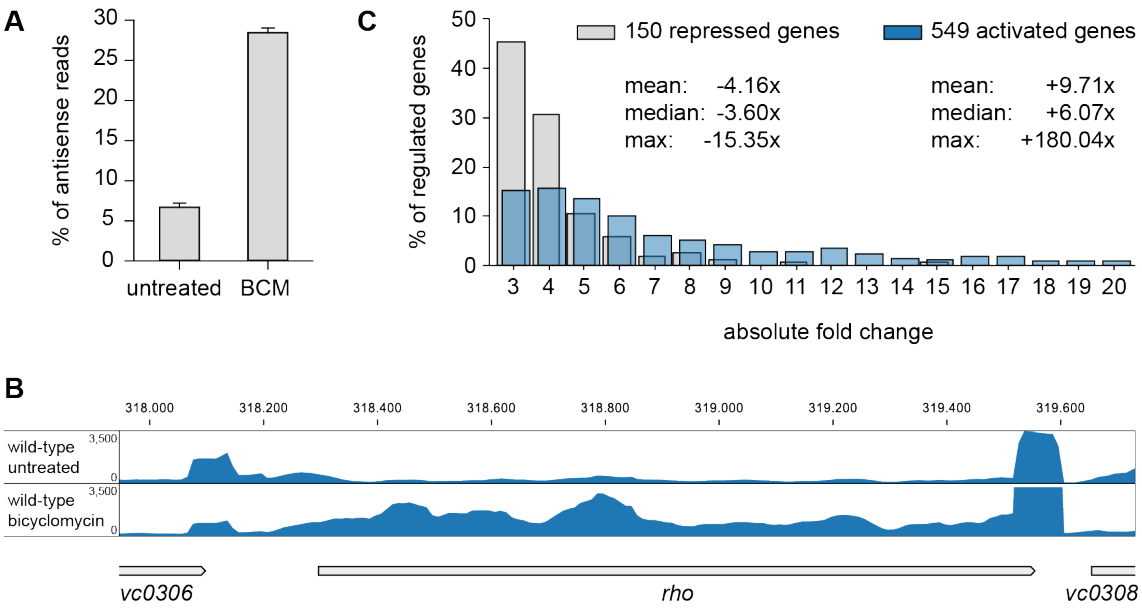


Figure 4.1: **Effect of bicyclomycin on the *V. cholerae* transcriptome.** (A) Percentage of reads mapped in antisense orientation to annotated coding sequences. (B) Read coverage of the *rho* gene. (C) Relative distribution of fold changes for 699 differentially regulated genes in wild-type background after BCM treatment (absolute fold change $\geq$3, FDR p-value $\leq$ 0.05)

activated genes in the BCM treated cells (Fig. 4.1C). On average, the observed gene repression upon Rho inhibition was weaker than gene activation, which is in accordance with the role of Rho as a transcriptional repressor [213].

4.1.3 Connections to other large-scale data sets

The large data sets generated during this work on RNase E sites, RNA 3' termini and bicyclomycin sensitivity provide a valuable resource to the *Vibrio* community. Still, there is more potential in the combination of these results with other data sets that have been generated in previous studies or still have to be obtained. As already mentioned, the *V. cholerae* genome still lacks a comprehensive annotation of transcripts with their 5' and 3' untranslated regions [132]. A previous study from our group identified transcriptional start sites in *V. cholerae* by differential RNA-seq and also suggested substantial corrections of misannotated ORFs [251]. These RNA 5' ends could be combined with the Term-seq data from this work to obtain full mRNA annotations a genome-wide level. Correctly annotated transcripts would also increase the accuracy of future deep sequencing experiments: as current analyses can only be based on read counts mapping to the annotated CDS, any differential regulation in untranslated regions remains undetected during the analysis of regulated genes.

Additional in-depth studies are also necessary to understand the role of Rho in *V. cholerae*. As discussed above, standard RNA-seq without enrichment of RNA 3' ends would be necessary for the bicyclomycin-treated samples to analyze the transcriptomic response to the antibiotic and to determine BCM-dependent read-through at the identified terminators. Following the analyses performed in the Sorek lab [76, 78], RNA termini could be attributed to intrinsic or Rho-dependent

termination events on a genome-wide scale. Moreover, exoribonuclease mutants could be used to define the exact sites of transcription termination, especially for Rho-dependent terminators [78]. Additional application of bioinformatic tools for terminator prediction could help to precisely locate terminator structures [233, 232, 91].

Finally, both the data on RNase E sites and on RNA termini could facilitate the characterization of sRNAs and their targets. While the majority of the detected RNase E sites are likely the result of bulk mRNA turnover [145, 27], some sites could also be targeted by sRNAs to stabilize the respective transcripts [108, 295]. Especially for activated sRNA targets, it might be a promising strategy to compare the site of base-pairing on the mRNA to the presence of any RNase E sites that could be blocked by sRNA binding. Similarly, transcription terminators are predominantly found at the ends of genes and operons to define transcript borders, but many terminators within transcriptional units also have regulatory functions [285]. sRNAs can contribute to this regulation by inducing or inhibiting Rho-dependent termination by uncovering or blocking *rut* sites, respectively ([301, 67, 314, 41] and section 4.2.2). Thus, if for instance a terminator sequence can be found within the 5' UTR of an sRNA target, it might be regulated by an sRNA and should be screened for a potential involvement in sRNA-based control of transcription termination.

4.2 Autoregulatory sRNAs from 3' UTRs

4.2.1 Dual output: interplay of mRNA and sRNA functions from the same transcript

With increasing numbers of sRNA genes being identified in numerous bacterial species, it has become clear that they are not only located within IGRs, but rather frequently overlap with mRNA genes and operons [1]. Conversely, some sRNAs also encode small peptides, thereby functioning as mRNAs themselves [116]. These observations gave rise to the concept of "parallel transcriptional output", which was suggested for both eukaryotic and prokaryotic RNA already two decades ago [203, 340]. Still we do not fully understand the underlying principles and multiple questions remain open for many of the discovered examples: (i) Are overlapping genes transcribed into a single transcript or are multiple independent transcripts produced? (ii) Are the mRNA and sRNA functions of a single transcript mutually exclusive? (iii) If so, what decides which of the two functions is exerted? Is it a timely coordinated process with one function preceding the other, or can a single transcript only ever carry out a single function during its lifetime?

The diverse answers that can be given to these questions for different examples reflect the large variety of bacterial sRNA loci and functions. Some sRNAs that overlap with mRNAs, especially with 3' UTRs, are transcribed from their own sRNA promoters and are therefore independently expressed from the corresponding mRNAs [81, 61, 126]. The two types of transcripts then share some sequence and their transcriptional terminator but are typically functionally unrelated. Thus, mRNA and sRNA functions are uncoupled and exerted by two independent transcripts. In contrast, the sRNAs derived from riboswitches or transcriptional attenuators constitute a mutually exclusive

transcriptional output to their mRNAs [187, 209]. In its OFF state, the riboswitch/attenuator induces premature transcription termination in a way that only the sRNA is produced and can regulate its targets in *trans*. However in the ON state, transcription is continued into the coding region and the sRNA is "trapped" inactively as 5' UTR of the mRNA. Such mutually exclusive expression is again different from peptide-encoding sRNAs, which are at least theoretically able to exert both base pairing and protein coding simultaneously through a single transcript, as the respective regions on the sRNA typically do not overlap [40, 342].

The two sRNAs presented in the study in chapter 2, OppZ and CarZ, belong to yet another class of sRNAs, which are derived from mRNA 3' UTRs by endonucleolytic processing. These sRNAs are transcriptionally coupled to their corresponding mRNAs and are often functionally related [166, 214, 62, 208, 344, 144]. They have to be processed into short regulators to be functionally active ([61, 215] and chapter 2, Fig. 3D). As discussed in section 4.1.1, both OppZ and CarZ are likely produced by step-wise processing of their parental mRNA transcripts through RNase E. Thus, the mRNA is probably inactivated upon production of the functional regulator [61, 62]. Theoretically, 3' UTR-derived sRNAs could also be cleaved from an intact mRNA through the direct entry pathway of RNase E, which would separate both transcripts without destroying either of them. Although such direct entry processing has been shown for SdhX [216], maintaining the intact mRNA despite sRNA processing is rather unlikely: most sRNA genes either overlap to a significant extent with the upstream gene or the sRNA 5' end can be found in close proximity to the stop codon [340, 251, 61, 60]. Cleavage to free the sRNA would therefore truncate the mRNA and result in stalled ribosomes [137]. Additionally, many mRNAs lack a protective stem-loop structure upstream of the cleavage position and are thus susceptible to rapid 3'-to-5' digestion by exonucleases [145, 27].

From an experimental point of view, it has to be noted that the mature OppZ and CarZ species generated *in vivo* will carry a monophosphate at their 5' end due to their processing from the corresponding mRNAs. In contrast, for the plasmid-based sRNA over-expressions used in chapter 2, the sRNAs were transcribed from artificial promoters as primary transcripts starting at their RNase E sites. While this generated sRNAs with the same nucleotide sequence as the processed native variants, such over-expressed sRNAs had a triphosphate group at their 5' ends. The nature of an sRNA's 5' end can affect its regulatory capacity, as monophosphorylated sRNAs can e.g. allosterically stimulate RNase E activity [22]. But over-expressing an sRNA in its processed form without simultaneously over-expressing the corresponding mRNA genes is technically challenging. For our assessment of OppZ processing in wild-type and the Δ*hfq* background (chapter 2, Fig. 2-S1), we constructed a plasmid-based precursor including the 3' end of *oppF* to allow OppZ processing, which would generate monophosphorylated OppZ. However, processing of this truncated *oppZ*-OppF fragment was less efficient than native processing of the full *opp* mRNA and led to the accumulation of the full-length precursor transcript and distinct cleavage intermediates, whereas mature OppZ levels were lower than for direct sRNA over-expression. As the regulatory function of these longer fragments carrying OppZ at their 3' end was unclear, we decided to use only the

sRNA genes for sRNA over-expression despite the caveat of generating triphosphorylated sRNA.

The observation of mRNA inactivation upon sRNA synthesis indicates that mRNA and sRNA functions would be mutually exclusive for transcripts with processed 3' UTR-derived sRNA. However, transcripts may fulfill both mRNA and sRNA functions sequentially, as protein production may occur before the mRNA is turned over and the sRNA is released. In this model, mRNAs may undergo several rounds of translation before giving rise to sRNA production. This should cause a delay in sRNA synthesis compared to the initiation of transcription, as not all mRNA molecules are immediately processed to release the sRNA. Such a delay has been observed for some sRNAs including OppZ (chapter 2, Fig. 8A, and [344]). The mRNA half-live would act as an intrinsic "timer" to define the point of switching from protein production to regulator synthesis and this timing could be tuned by changes in translation efficiency, accessibility of RNase E sites or other factors governing mRNA half-live [44, 87, 282]. Short-term increases or decreases of mRNA turnover should therefore cause a transient rise or drop in the sRNA synthesis rate, respectively, while eventually production of the the two RNA species from the same transcript is inherently fixed to occur at a 1:1 ratio. OppZ and CarZ add another level of complexity to this connection through their feedback capabilities, which will be discussed in section 4.2.3. Despite the equimolarity of mRNA and sRNA synthesis at the transcript level, the ratio of protein and sRNA levels can be different. While this ratio obviously depends on protein and sRNA stabilities, also the half-live of the mRNA is an important factor: higher mRNA stability means more time for protein production from the same transcript before its turnover, increasing the amount of protein produced relative to the sRNA. Tuning mRNA half-live might therefore be an important factor for transcripts where both protein and sRNA are components of the same regulatory pathway and balancing of their abundances is necessary [62, 344, 144]. Additionally, it has to be noted that the 1:1 ratio of mRNA and sRNA production is not necessarily reflected by their steady-state levels in the cell. On the contrary, many processed sRNA showed higher abundance than their corresponding mRNA [61], which is likely caused by their higher stability compared to mRNAs [34, 340].

4.2.2 Divergent regulatory properties of OppZ and CarZ

Molecular mechanisms of target regulation

A variety of molecular mechanisms has been described for sRNA-based regulation and many sRNAs employ more than one mechanism to control the expression of their targets [343, 161]. The most common mode of blocking target translation is often followed by mRNA degradation, although the latter is not always crucial for target regulation [226, 280]. Many experimental approaches to identify sRNA targets, such as sRNA pulse expression experiments, screen for reduced mRNA levels in the presence of the sRNA, which are commonly attributed to fast turnover of the untranslated mRNA [308, 307, 106]. However, although OppZ pulse expression readily reduced *oppBCDF* transcript levels (chapter 2, Figs. 3A and 3-S1C), it did not induce *oppB* turnover (chapter 2, Fig. 3-S1D). Additionally, plasmid-based OppZ over-expression reduced the amount of OppZ produced from the native *oppZ* locus (chapter 2, Fig. 4B). These findings could only be explained by pre-

mature transcription termination as a consequence of inhibited translation, which we validated by observing strongly reduced OppZ levels in an *oppB* start codon mutant (chapter 2, Fig. 4C). Bacteria use two distinct pathways for transcription termination: either intrinsic termination, which depends on structural RNA elements consisting of a strong hairpin followed by an U stretch, or Rho-dependent termination, where the termination factor Rho binds to *rut* sites on the nascent transcript and dislodges RNAP [213, 285]. Due to their sequence-encoded features, intrinsic terminators can be computationally predicted with high accuracy [233] and the only such terminator in the *opp* operon was found at the 3' end of *oppZ*. In contrast, *rut* sites are more loosely defined and their bioinformatic detection is challenging [232, 91]. Instead, we used the Rho inhibitor bicyclomycin to treat the *oppB* start codon mutant and observed a strong increase of *oppBCDF* and OppZ levels, thereby validating the involvement of Rho in the repression of *oppBCDF* (chapter 2, Figs. 5A-B). We obtained similar results when *oppB* translation was inhibited by plasmid-based over-expression of OppZ (chapter 2, Figs. 5D-E). So far, Rho-dependent transcription termination induced by sRNA binding has only been shown for *Salmonella* ChiX and *E. coli* Spot 42 [42, 345]. While the mechanism for increased Rho-dependent termination at the *galTK* junction though Spot 42 is still unknown [345], ChiX represses the distal gene in the *chiPQ* operon by a similar mechanism to the one employed by OppZ: it blocks *chiP* translation and exposes an otherwise hidden *rut* site in *chiP* [42].

Our data on CarZ shares the main findings with the data presented for OppZ: plasmid-based CarZ over-expression represses not only CarA and CarB, but also reduces synthesis of CarZ from the native *carZ* locus in the *carB* 3' UTR (chapter 2, Fig. 7E). This is likely a consequence of blocked *carAB* translation, as stop codon insertions that abolish *carA* or *carB* translation also lead to strongly reduced CarZ levels (chapter 2, Fig. 7F). We therefore concluded that CarZ-dependent regulation follows the same mechanism involving transcription termination as the OppZ regulation. According to OppZ, we proposed the unmasking of *rut* sites within the upstream *carAB* genes as a result of inhibited translation initiation by CarZ. However, in contrast to OppZ, we were unable to directly validate the involvement of Rho by treatment with BCM, as these experiments yielded inconsistent and partly conflicting results regarding CarZ synthesis. We attribute this phenotype to an unexpected interference with the gene encoded downstream of CarZ, *vc2388*, which is transcribed in antisense direction to *carZ*. According to our BCM Term-seq data, *vc2388* transcription is highly Rho-dependent and results in strong antisense transcription at the *carZ* locus upon BCM treatment (Fig. 4.2). This impedes an independent analysis of the BCM effect on CarZ synthesis, as the antisense transcription from *vc2388* likely interferes with correct transcription of the *carAB-carZ* locus [322, 47]. To fully characterize the importance of Rho for CarZ regulation, we would need to isolate the *carAB-carZ* locus from *vc2388* transcription. Nevertheless, we still propose that CarZ regulation by transcription termination is mediated through Rho due to the high similarity to the OppZ pathway and the lack of Rho-independent terminators within the *carAB* sequence.

Despite their shared regulatory pathway of transcription termination, OppZ and CarZ differ in

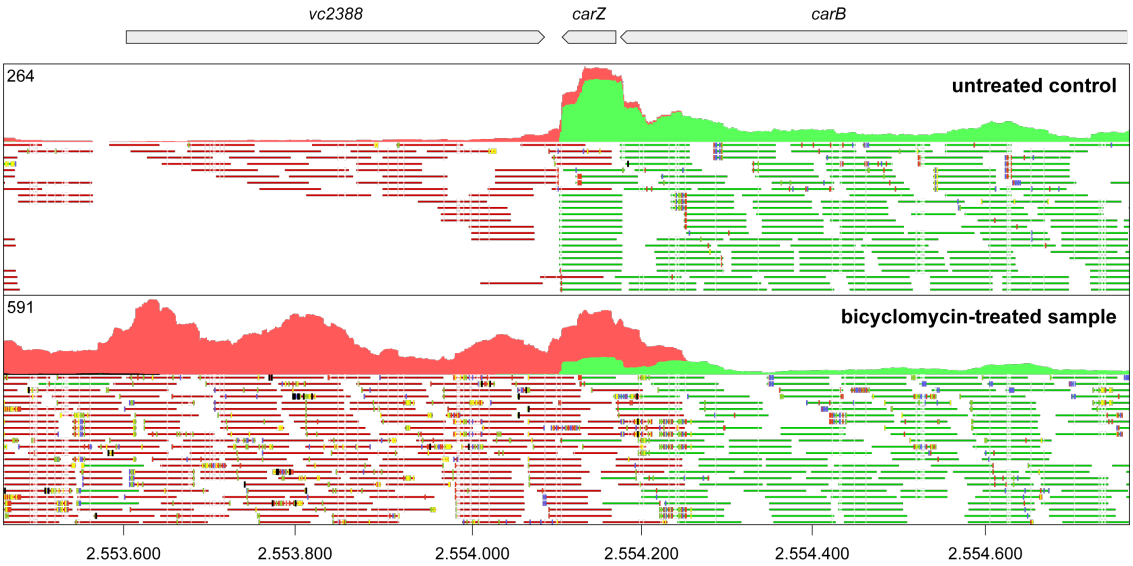


Figure 4.2: **Bicyclomycin-dependent transcription of *vc2388*.** *V. cholerae* was treated with BCM as described in chapter 2, Fig. 5, and Term-seq coverage is shown for one representative biological replicate. Sequencing reads mapping to the forward or reverse strand of the genome are shown in red or green, respectively. The location of *vc2388*, *carZ* and *carB* is indicated above.

their effects on target transcript stability. As discussed above, reduced *oppBCDF* transcript levels result from Rho-dependent termination and repression can be relieved by BCM treatment, as OppZ does not destabilize the *oppB* mRNA (chapter 2, Figs. 5E and 3-S1D). In contrast, over-expression of CarZ reduces the half-lives of *carA* and *carB,* thereby rendering the regulation irreversible through transcript degradation (chapter 2, Figs. 7-S2). This mechanistic variance between the two sRNAs likely reflects the difference in localization of their target genes: as OppZ repression does not affect *oppA* translation, ribosomes on *oppA* protect the *oppABCDF* mRNA from 5' end-dependent turnover by RNase E. In contrast, CarZ inhibits translation of both cistrons, thereby rendering the full *carAB* mRNA susceptible to nucleolytic attack. It can be assumed that BCM treatment would only partially restore *carAB* levels upon CarZ over-expression, as presumably both premature transcription termination and mRNA turnover contribute to the tight regulation observed for *carAB*.

The example of CarZ suggests that sRNA-induced transcription termination through Rho could be more wide-spread in the sRNA world than currently appreciated: without the observation that the sRNA influences its own production from the transcript's 3' end, we would have attributed the decrease in *carAB* transcript levels solely to the increased degradation rates upon CarZ binding. But as discussed in section 4.2.1, faster *carAB* turnover should reduce the delay of CarZ synthesis compared to the onset of transcription, but not decrease absolute CarZ levels. Thus, we only discovered the underlying transcription termination mechanism because of the repression of the 3' UTR-derived sRNA. As most target mRNAs do not carry such sRNAs in their 3' UTRs, RNA levels are typically not compared for multiple positions along the transcript. Thus, transcription termination in addition to accelerated mRNA decay might be an often overlooked regulatory mechanism of sRNAs. It is indeed frequently observed that sRNAs repress multiple genes in an

operon by binding only to the first cistron [307, 251, 261]. While this is commonly attributed to the loss of translational coupling and accelerated degradation of the full messenger transcript, it could just as well involve Rho-dependent transcription termination [41]. To uncouple these effects, pulse expression experiments of an sRNA in question could be repeated in the presence of BCM to detect targets that are no longer regulated at the level of mRNA abundance when Rho is inhibited.

Strength of target repression

OppZ and CarZ share a mechanistically similar regulation of their target genes, nevertheless, they differ in their regulatory capacity: plasmid-borne OppZ over-expression affected OppB protein levels less efficiently than CarZ over-expression affected CarA and CarB levels, both as GFP reporter fusions (chapter 2, Figs. 3C vs. 7-S1B) and as epitope-tagged variants expressed from the native genomic loci (chapter 2, Figs. 4B vs. 7E). The difference was even more pronounced when the sRNAs were not over-expressed from an independent locus, but rather processed from the mRNAs as in their native biogenesis pathway. We generated reporter fusions of *oppB* or *carA* to *gfp* and added the respective sRNA gene to the 3' end of *gfp* to monitor sRNA production and mRNA regulation at the same time (chapter 2, Figs. 3D and 7C). While both sRNAs were efficiently cleaved from the long transcripts, OppB-GFP was repressed ~2-fold by OppZ, whereas CarA-GFP levels decreased ~13-fold when CarZ was present. This quantitative difference in the regulatory capacity of OppZ and CarZ likely results from the interplay of the regulatory mechanisms employed by the two sRNAs as discussed above: while OppZ uses two regulatory mechanisms (translational inhibition and transcription termination), CarZ additionally causes transcript degradation. The individual contributions of the different molecular mechanisms are hard to untangle due to the closely intertwined sRNA production from and regulation of the same locus. For example, blocking mRNA degradation by inactivation of RNase E could normally help to separate the effects of inhibited translation initiation and subsequent mRNA degradation on overall protein repression [226, 125, 107]. But in the case of OppZ and CarZ, RNase E inactivation would at the same time inhibit sRNA synthesis, thereby reducing the levels of active regulator and likely weakening translational repression (chapter 2, Fig. 1-S4). *In vitro* experiments might help to study the different effects individually: toeprinting assays could be used to compare the capability of OppZ and CarZ to block formation of the translation initiation complex, while *in vitro* transcription assays of the *oppB* and *carAB* mRNAs in the presence of Rho and the sRNAs could help to determine the strength of Rho-dependent transcription termination within the two operons [308, 301, 153]. Conversely, inhibiting the action of Rho through BCM or through the identification and mutation of all *rut* sites in *oppB* and *carAB* could shed light on the contribution of transcription termination on target regulation [42]. However, analysis of the *carAB-carZ* locus during BCM treatment is hampered by the interference with antisense transcription from the downstream gene (as discussed in section 4.2.3) and fully excluding any Rho-dependent termination by mutations might be challenging. Additionally, Rho inhibition by either means would simultaneously inactivate the negative feedback on sRNA synthesis and should thereby increase sRNA levels (similar but

opposite to the side-effect of RNase E inactivation).

The regulatory capacity of an sRNA seems to depend as well, at least in part, on its binding strength to the target [131, 31, 262]. I thus analyzed this effect for OppZ and CarZ by calculating the free energy of the RNA duplexes depicted in chapter 2, Figs. 3B and 7B with two different bioinformatic tools. First, I determined the hybridization energy using the RNAhybrid algorithm [277] and obtained similar values (-22.9 kcal/mol for *oppB*-OppZ vs. -20.0 kcal/mol for *carA*-CarZ). However, when comparing the seed regions of both sRNAs to their predicted secondary structures, a remarkable difference is the accessibility of the CarZ seed in a single-stranded region between two stem loops (chapter 2, Fig. 7B), while the first stem loop of OppZ has to be opened to allow *oppB* binding (chapter 2, Figs. 3B and 3-S3). To take this effect into account, I then analyzed the binding strengths with IntaRNA, which also includes the energy for unfolding intramolecular structures [49]. While IntaRNA generally calculated weaker binding energies than RNAhybrid, the difference was much stronger for *oppB*-OppZ (-8.7 kcal/mol with IntaRNA vs. -22.9 kcal/mol with RNAhybrid) due to the energy needed for melting the hairpin than for *carA*-CarZ (-13.4 kcal/mol with IntaRNA vs. -20.0 kcal/mol with RNAhybrid). To test if the weaker duplex strength of *oppB*-OppZ is responsible for the corresponding variation in target regulation, we mutated a single nucleotide of OppZ (C6U) in the *oppB-gfp-oppZ* fusion. This mutation would be expected to increase OppB repression by two means: (i) it improves pairing of OppZ to *oppB* by changing the unpaired C-A in the duplex to an U-A pair (compare chapter 2, Fig. 3B) and (ii) it decreases the strength of the intramolecular OppZ hairpin including the seed region by mutating a C-G pair to an U-G pair (compare chapter 2, Fig. 3-S3). Accordingly, the free energy of *oppB* binding calculated with both tools decreased strongly (RNAhybrid: -28.7 kcal/mol for the C6U mutant vs. -22.9 kcal/mol for wild-type OppZ; IntaRNA: -17.4 kcal/mol for the C6U mutant vs. -8.7 kcal/mol for wild-type OppZ). However, the extent of the OppB-GFP repression did not differ between wild-type and mutated OppZ, despite the large difference in binding strength (data not shown). Thus, at least for OppZ, the hybridization energy to *oppB* does not determine the target repression strength. Finally, it has to be noted that *in vivo*, binding of the OppZ and CarZ to their mRNA targets occurs co-transcriptionally. This might affect folding of the mRNAs and render their sRNA-binding sites more or less accessible than in the steady state structure predictions that were used for the above calculations.

Fate of the sRNA after target binding

The strength of target regulation by an sRNA does not only depend on its regulatory mechanism or the target binding strength, but also on the consequences of target binding for the sRNA itself. The fate of an sRNA after successful base pairing with a target is an important characteristic of any sRNA-dependent target regulation. Three possible outcomes have been described for repressing regulators: (i) sRNA and target are both turned over in a process called coupled degradation [197], (ii) the sRNA is recycled after target degradation, thereby establishing a catalytic reaction [243], or (iii) neither sRNA nor target are degraded, but they sequester each other in an inactive form [102].

These modes of sRNA action differ fundamentally in their associated regulatory properties and kinetics. Coupled degradation results in a stoichiometric regulation, where the relative production rates of sRNA and mRNA determine target expression levels in a threshold-linear response model [183]. A catalytic mode of action resembles transcription factor-based control and is thought to contribute to very tight regulation of strongly expressed targets [130, 102]. Finally, the biological consequence of sequestration is similar to that of coupled degradation as it permanently inactivates sRNA and target, although it does not lead to a detectable reduction of RNA levels [102]. For autoregulatory sRNAs like OppZ and CarZ, the production rates of sRNA and mRNA cannot be tuned individually, as regulator and target are made as one transcript from the same promoter. Thus, the observed strong difference in target regulation strength could stem from different sRNA recycling efficiencies. If CarZ acted by catalytic degradation, sRNA and mRNA levels would become uncoupled despite their connected synthesis, as CarZ could induce *carAB* turnover without being degraded itself. The recycled CarZ molecules would accumulate and contribute to strong *carAB* repression. In contrast, OppZ does not induce *oppB* degradation and thus follows neither the coupled degradation nor the catalytic mode of action. Instead, OppZ might be sequestered upon *oppB* binding, which would inhibit OppZ accumulation after *oppB* repression. Indeed, when the mildly autoregulated *oppB-gfp-oppZ* fusion from chapter 2, Fig. 3D is combined with OppZ over-expression from a second plasmid, the abundance of OppB-GFP is strongly reduced (data not shown). This indicates that the level of free OppZ is a limiting factor for *oppB* repression.

The molecular features determining the fate of an sRNA after target regulation are not yet understood, although there are hints on the importance of RNA duplex binding strength and the mode of binding to Hfq [102, 300]. The Qrr sRNAs employ all of the three described modes of actions to repress three central actors of the QS pathway with distinct kinetics. When comparing the binding strength of the targets, the authors proposed a very tight sRNA-mRNA binding to be the reason for the Qrr sequestration capability of one target [102]. However as discussed above, the *oppB*-OppZ duplex is not very strong and a mutation to increase the binding strength did not affect *oppB* regulation. Thus, if OppZ is indeed sequestered upon *oppB* binding, this outcome is likely mediated by other factors than strong RNA duplex formation.

4.2.3 Feedback on sRNA synthesis through transcription termination

Sequestration or co-degradation of an sRNA upon target binding are important means to control the cellular abundance of the regulator. However, in contrast to other sRNAs, regulation of their respective mRNA targets does not only affect OppZ and CarZ through potential degradation after target binding. As both sRNAs regulate genes encoded upstream in their own operons, inducing transcription termination within these target genes establishes a negative feedback loop on sRNA expression. This adds a novel layer of complexity to the dual output concept discussed for 3' UTR-derived sRNAs in general in section 4.2.1. As illustrated by our model (chapter 2, Fig. 9), during the initial transcription of the *opp* operon, OppZ production is delayed but not reduced compared to mRNA synthesis. However, once the first sRNA molecules have been produced, they

repress the subsequent synthesis of additional OppZ transcripts through a negative feedback circuit: they uncouple transcription and translation by blocking *oppB* translation initiation, rendering the ribosome-devoid *oppB* mRNA more susceptible to being targeted by Rho. This terminates transcription prematurely and reduces OppZ production. Considering the strikingly different half-lives of *oppB* (~50 s; chapter 2, Fig. 3-S1D) and OppZ (> 32 min; chapter 2, Fig. 4-S1A), the feedback regulation through Rho might be necessary to prevent overshooting sRNA synthesis and mRNA repression. We propose that autoregulation of CarZ generally follows the same principles as the OppZ feedback. The main difference is that CarZ additionally destabilizes the *carAB* mRNA, thereby reducing the time for protein production before mRNA turnover and establishing a tighter repression. At the same time, this should decrease the delay of sRNA synthesis compared to the initiation of transcription, a prediction that still awaits experimental validation.

In summary, we consider expression of both sRNAs to be intrinsically limited through Rho-dependent feedback regulation of their own synthesis. However, the distribution of *rut* sites seems to differ between the two operons. According to our stop codon mutants, the *oppB* gene may exhibit multiple *rut* sites with the most important one being located close to the 115th codon, whereas blocked translation of *oppC*, *oppD* or *oppF* did not affect OppZ synthesis (chapter 2, Fig. 6). The locations of the *rut* sites in *oppB* that we inferred from our genetic mutations do not correlate well with the dominant 3' end determined by Term-Seq of the *oppB* start codon mutant, which was found in close proximity to the *oppAB* intergenic region and thus ~300-450 nt upstream of the presumed terminator. This discrepancy between the site of transcription termination and the observed steady-state 3' end of the transcript can be attributed to the activity of 3' to 5' exonucleases, which probably rapidly digest the unprotected part of *oppB* [145, 78, 27]. This decay is likely halted at a stable stem loop structure within the *oppAB* intergenic region, which probably also protects *oppA* from OppZ-dependent regulation [66]. The limited dependency of OppZ production on translation of the *oppCDF* genes is in contrast to the CarZ situation, where abolishing translation of either *carA* or *carB* strongly reduced CarZ levels, indicating that both genes carry at least one *rut* site (chapter 2, Fig. 7F). Alternatively, the two genes could be translationally coupled and blocking *carA* translation would only indirectly induce transcription termination by concurrently blocking *carB* translation and exposing a *rut* site in *carB*. Deletion of *carB* in the *carA* stop codon background could help to determine the presence of a potential terminator sequence within *carA*. Similar to OppA, CarA production is not affected by blocked translation of *carB* and subsequent transcription termination (chapter 2, Fig. 7F). Indeed, inspection of the *carAB* sequence also revealed a stable stem loop structure within the 5' CDS of *carB*, which could protect the upstream gene from exonucleolytic decay.

4.2.4 Biological relevance of autoregulatory sRNAs

Negative feedback regulation of their corresponding mRNAs

Negative autoregulation is one of the most common network motifs in biology [6]. The smallest negative feedback loop consists of just one TF that binds and represses its own promoter like the

tet repressor [288]. Feedback circuits involving sRNAs can act directly on the mRNA of the TF that regulates sRNA expression (such as the OmrA/B sRNAs repressing the mRNA of their TF OmpR [125, 48]) or mediate a more indirect feedback by relieving a stress signal that induces the activity of the TF (like the σ^E-dependent sRNAs that counteract envelope stress [252, 261]).

In contrast, the autoregulation by OppZ and CarZ presented in chapter 2 constitutes a novel type of regulatory feedback on the RNA level. It is not dependent on auxiliary transcription factors and is thus the first reported example of an RNA-only negative autoregulatory circuit. The strength of the feedback repression is proportional to the overall transcriptional output, which might be particularly relevant when the complex transcriptional regulation of the two operons is considered. The *oppABCDF* operon encodes a high affinity ABC transporter for the uptake of oligopeptides, which is also involved in the control of virulence, quorum sensing and competence in diverse bacteria [112, 111, 319]. Thus, its transcription is typically controlled by several different factors such as the leucine responsive protein Lrp and the iron regulator Fur in *E. coli* and *Salmonella* or the global nutritional regulators CodY and ScoC in Gram-positive bacteria [54, 69, 32]. Similarly, the *carAB* genes encoding carbamoyl phosphate synthetase are subject to intricate transcriptional control involving several transcriptional regulators from the arginine, pyrimidines and purines pathways [64]. With regard to this complexity, it seems unfavorable to establish a negative autoregulation loop through an additional DNA-binding transcriptional repressor that would be co-expressed with the operon: the repressor would need to compete with multiple other regulators for promoter binding and the efficiency of autoregulation might be impaired under certain conditions where the other transcriptional regulators are highly active. In contrast to a protein-based feedback loop, expression of the autoregulatory sRNA from the 3' UTR correlates to the absolute promoter activity that results as sum of all transcriptional inputs. This establishes a negative feedback regulation at the RNA level that acts downstream and independent of any transcriptional control.

Regulation by OppZ adds yet another layer of complexity due to its discoordinate nature (chapter 2, Figs. 3C and 4B). Stoichiometry of the proteins encoded by the *oppABCDF* operon follows the typical distribution for ABC transporters with higher abundance of the periplasmic substrate binding protein (OppA) compared to the membrane-bound, structural components of the transporter (OppBCDF) [136, 200]. While a recent publication suggested some internal transcription initiation in the *oppAB* intergenic region of *E. coli* [201], we did not observe any protein expression when the promoter upstream of *oppA* was deleted (chapter 2, Fig. 2C). Thus we conclude that the full *opp* mRNA is transcribed as a single unit and post-transcriptional repression of the *oppBCDF* part by OppZ contributes to the discoordinate expression of the operon. Uncoupling of transcriptionally linked genes has been observed for several sRNAs [221, 90, 255, 86, 344] and this additional flexibility may provide an advantage of OppZ-based regulation over solely transcriptional control of the *opp* operon. Nevertheless, OppZ is probably not the only relevant factor for the discoordinate expression of the *opp* operon, as deletion of *oppZ* only mildly alleviated the

steady-state levels of OppB, whose abundance was still relatively low compared to OppA (chapter 2, Figs. 8-S1). Bacteria have evolved various molecular mechanisms to uncouple the expression of genes linked in an operon such as internal transcription start sites, leaky transcription termination within the operon, varying half-lives of the different cistrons, or different translation efficiencies [2, 273, 77]. Not all of these mechanisms seem to be at play in the *opp* regulation, as the *oppA* mRNA is even slightly less stable than *oppB* (26 vs. 35 s; data not shown) and there is no internal promoter ([251] and chapter 2, Fig. 2C). There is a prominent 3' terminus within the long *oppAB* intercistronic region that becomes less abundant upon bicyclomycin treatment and could represent an intra-operonic Rho-dependent terminator (chapter 2, Fig. 5C). But this structure does not contribute to transcription termination in *E. coli* [136] and mutation of the hairpin in the post-transcriptional reporter plasmid from chapter 2, Fig. 3C had no effect on mKate or GFP fluorescence levels (data not shown). When comparing post-transcriptional *gfp* fusions for *oppA* (starting at the *oppA* TSS) or *oppB* (starting at the *oppAB* IGR), the *oppA-gfp* fusion yields much higher fluorescence than the *oppB-gfp* fusion, suggesting a variation in translation efficiency (data not shown). This is supported by a ribosome profiling study in *E. coli*, which determined a ~4-fold higher translation efficiency for *oppA* compared to *oppB* [184].

In contrast to the relatively mild influence of the *oppZ* knockout on steady-state OppB levels, this effect became much more prominent during the dynamic onset of *oppABCDF* expression, which we studied by replacing the native *opp* promoter by the arabinose-inducible pBAD promoter to artificially control the *opp* expression. Upon transcriptional induction of the operon, OppB accumulated more strongly and over a longer time period in the Δ*oppZ* background (chapter 2, Fig. 8). In contrast, OppB levels in the wild-type strain rapidly reached a plateau at relatively low protein abundance. Accelerated response time and reduced total protein production have been reported to be inherent features of negative autoregulation by transcription factors [288] and we conclude that the RNA-based autoregulation by OppZ shares this behavior. Moreover, we hypothesize that it may help to balance the gene regulatory output in fluctuating environments: the costly production of the structural transporter components OppBCDF would be limited during short-term availability of peptides, while the substrate binding protein OppA would not be not part of the feedback control and its levels could be rapidly increased to capture all available peptides. Additionally, the long half-live of OppZ (> 32 min, chapter 2, Fig. 4-S1A) could provide a memory function by repressing *de novo* OppBCDF synthesis following too shortly after a previous induction of *opp* expression.

Integration of OppZ and CarZ into regulatory networks

With increasing numbers of characterized sRNAs, it has become evident that sRNA-based regulation is often embedded into larger regulatory networks consisting of both transcriptional and post-transcriptional regulators [195, 236]. The ability of many *trans*-acting sRNAs to regulate multiple targets allows them to function as central hubs of their respective regulons and allows for

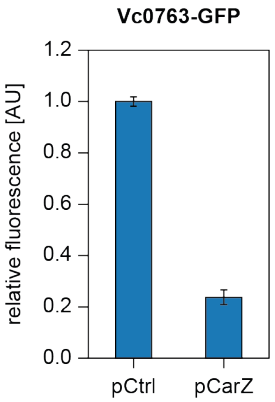


Figure 4.3: **Post-transcriptional repression of *vc0763* by CarZ.** A translational *vc0763::gfp* reporter was constructed by fusing the 5' UTR and the first 20 codons of *vc0763* to the *sfgfp* gene in the pXG10 vector [75]. The reporter was combined with a control plasmid or a CarZ over-expression plasmid and fluorescence was measured as described in chapter 2, Fig. 7-S1B.

target hierarchy and cross-talk [256, 29, 266, 39]. While we did not detect any other targets for OppZ than the *oppBCDF* mRNA (chapter 2, Fig. 3A), bioinformatic target predictions suggested additional targets for CarZ. Preliminary results obtained in our group validated repression of at least one other gene by CarZ (*vc0763* coding for the GTPase EngA involved in ribosome assembly [35]; Fig. 4.3). Depending on the nature of this additional target-sRNA interaction, it may influence the outcome of the *carA*-CarZ regulation by competition between the two mRNAs for their shared regulator: if *vc0763* binding induces for example CarZ degradation or sequestration, increased *vc0763* transcription would titrate CarZ away from the *carAB* mRNA, thereby relieving their CarZ-dependent repression. A similar pattern has been observed for the ChiX-dependent regulation of chitosugar uptake, where ChiX represses both the inner membrane transporter *chbBCA* and the outer membrane chitoporin *chiP*. Chitosugar-induced transcriptional activation of *chbBCA* depletes the cellular ChiX pool through coupled degradation and thereby de-represses *chiP*, allowing for coordinated transport of chitosugars over both membranes [243, 104, 266]. Accordingly, the levels of *vc0763* or potentially other CarZ targets could modulate the strength of the *carAB* repression. Moreover, the existence of a least one *trans*-encoded CarZ target in addition to *carAB* contributes to the necessity of the 3' UTR-derived sRNA instead of a transcriptional feedback regulation of *carAB*: processing of the sRNA does not only serve to control its own operon, but it also links the expression of multiple genes at the post-transcriptional level.

Cross-talk between multiple transcripts is not limited to the competition of mRNAs for a shared sRNA regulator. Instead, sRNA sponges can directly bind other sRNAs and modulate their expression level and regulatory capacity [214, 175, 103]. For instance, the target derived sponge SroC base pairs to GcvB and induces its degradation, thereby alleviating the repression of its parental mRNA and other GcvB targets [214]. The discovery of such sRNA sponges was greatly accelerated by the development of technologies for the study of RNA-RNA interactions such as RIL-seq [208, 207]. When we recently applied RIL-seq to *V. cholerae* in our lab, we detected the sRNA Vcr222 [144] as a potential sponge for OppZ (unpublished data). It is processed from the 3' UTR of *astD* (*vc2616*), an aldehyde dehydrogenase involved in the arginine catabolic pathway [299]. The interaction between Vcr222 and OppZ still has to be validated experimentally and its potential functional consequences are currently unclear. But the involvement of *oppABCDF* in peptide transport and of *astD* in amino acid degradation indicates a possible functional connection

within the nitrogen metabolism network [278]. Additionally, the sRNA GcvB, which targets many genes involved in amino acid transport and metabolism, has been reported to repress *oppA* in *E. coli* and *Salmonella* [335, 305, 270] and a potential GcvB binding site was detected in the 5' UTR of *V. cholerae oppA* (data not shown). But GcvB in *V. cholerae* was solely identified by homology to the enterobacterial sRNAs [305, 270] and awaits further characterization in this organism. It is thus unclear if and how GcvB-dependent regulation of the *opp* operon would extend to OppZ or influence its feedback regulation. Nevertheless, the potential connections to GcvB and the sponge Vcr222 suggest that OppZ may be part of an intertwined regulatory network, despite its seemingly isolated target spectrum consisting of only its own operon [267].

4.3 Synthetic sRNAs for targeted gene regulation

Over the past decade, regulatory RNAs in prokaryotes have emerged as versatile tools for the field of synthetic biology and metabolic engineering [272, 263, 14, 178]. Base-pairing sRNAs are especially useful due to the programmability of their molecular interactions based on Watson-Crick base pairings and their compact, modular structure. They are considered as fast and economic regulators, as they do not require a translation step and control their targets at the post-transcriptional level [206, 148, 160]. Metabolic engineering approaches that involve sRNAs typically rely on deletion or over-expression of native sRNAs, but some of them also employ synthetic sRNAs [110, 71]. Various strategies have been applied to the generation of such synthetic regulators, typically based on a native sRNA scaffold for structure, Hfq binding and transcription termination, which is equipped with an artificial seed region to redirect its target specificity. This seed is often (semi-)rationally designed to match the designated mRNA target sequence [193, 231, 174], but it can also be selected by screening a library of randomly generated seed regions [310, 309]. The present work discusses both design strategies: while we suggest the 3' UTR-derived sRNAs OppZ and CarZ as starting points for rational sRNA engineering, we constructed and employed a large randomized sRNA library for the unbiased study of complex microbial phenotypes.

4.3.1 OppZ and CarZ as blue-prints for RNA-based autoregulation

Negative autoregulation has been reported for almost 40% of all transcriptional regulators in *E. coli* [311, 288] and is also widely used in synthetic gene networks to reduce noise, increase robustness and accelerate the response time of the system [6, 4]. In contrast to the TF-based autoregulation known so far, OppZ and CarZ add an RNA-based feedback loop to their own operons (chapter 2, Fig. 9) and for OppZ, we could show its contribution to faster dynamics of OppB production (chapter 2, Fig. 8). A critical step for OppZ functionality is its processing by RNase E to release the mature regulator from the polycistronic mRNA. Although RNase E sites are only loosely defined on the sequence level ([60, 105] and chapter 2, Fig. 1D), we were able to transfer the sRNA together with its processing site to the 3' UTR of a heterologous *gfp* transcript, indicating that the native sequence context of the *opp* operon is dispensable for functional sRNA maturation

(chapter 2, Fig. 3-S3). Adding the targeted *oppB* sequence to the 5' UTR established feedback regulation of the the *gfp* transcript, an effect that we could repeat for the *carA*-CarZ system (chapter 2, Figs. 3C and 7C). In combination with previous reports on rational re-targeting of natural sRNA scaffolds [193, 231, 174], this suggests that OppZ and CarZ could serve as blueprints for sRNA-based feedback regulation. Such synthetic sRNAs could be inserted into the 3' UTR of any gene without altering the transcriptional control of the gene or regulatory sequences in its 5' UTR. Instead, the sRNA would be equipped with a rationally designed seed sequence targeting the translation initiation region to establish feedback repression. The programmability of the RNA duplex would help to modify feedback strength by adjusting the base pairing strength [131, 231]. This system would require only minimal genetic engineering and could be introduced in addition to any existing regulatory connections, Moreover, synthesis of the regulator would be intrinsically coupled to transcription of the target gene without the need for additional regulatory input controlling sRNA synthesis [164]. Thus, we propose that engineered 3' UTR-derived sRNAs could be used as modular components to add negative autoregulation to any given gene without interfering with established regulatory networks.

4.3.2 A complex library of synthetic sRNAs

A major part of this work was the generation of a synthetic sRNA library. It was motivated by previous reports from our group and others on the modular architecture of sRNAs [265, 249, 152, 9, 108, 262] and is designed for the unbiased study of the underlying regulatory principles of sRNA regulation. Additionally, it can also be used as a tool to decipher complex microbial phenotypes or for the development of synthetic gene regulation. While the following section discusses some general considerations for the construction and application of the library, section 4.4 focuses on the specific use of the library in the context of the *V. cholerae* envelope stress response.

RybB as a scaffold to carry a randomized seed region

When searching for an appropriate scaffold for the sRNA library, RybB from *Salmonella* was chosen as an ideal candidate for several reasons: (i) Its seed sequence is exposed as a single-stranded region at the very 5' end and therefore well accessible for manipulation without changing the overall sRNA structure [21, 249]. This is in contrast to e.g. the seed of OppZ, which is located in the loop region of the first hairpin and requires melting of the stem for target regulation (see section 4.2.2), or to the GcvB seed placed between two hairpins, where sequence changes could lead to rearrangements in the secondary structure [305]. With 79 nt, RybB is relatively short and natively does not fold into secondary structures other than the terminator hairpin [43, 249]. The sequence between the randomized seed and the terminator is relatively AU-rich and part of the Hfq binding site [21], which makes it unlikely to form secondary structures with the randomized seed. (ii) No additional seed regions have been reported for RybB. The existence of a second, non-randomized seed in the sRNA scaffold could establish target regulation independent the actual sRNA variant and thereby bypass sRNA selection based on the randomized seed. (iii) Binding to Hfq and the requirement of

Hfq for base-pairing are well established for RybB [252, 297, 296, 356]. As Hfq-binding sRNAs are to date the most widespread class of known sRNAs [343], we expected an Hfq-dependent library to provide the most valuable insights. (iv) Finally, the *rybB* gene itself is not present in *V. cholerae* [252], which should avoid problems arising through homologous recombination events.

The length of the randomized seed was considered as a trade-off between target binding capacity and the number of resulting permutations. Nine nucleotides are more than the reported minimal seed length of 6-7 nt [162, 21], thereby allowing some flexibility in the positions used for mRNA binding. Still, achieving sufficient coverage for the resulting 262,144 possible variants is within the range of transformation efficiency obtained by standard laboratory methods [94, 151]. Expression of the sRNA library from an inducible promoter on a broad host-range plasmid was chosen to allow (i) easy transfer of the library into different strain backgrounds and subsequent isolation from pooled cells by standard methods, (ii) selection for a single plasmid per cell by an antibiotic resistance marker and (iii) sRNA expression only under the studied condition to maintain variants that would have detrimental effects on bacterial fitness under other conditions.

A regulatory context allowing sRNA selection

The general idea of the sRNA library is to provide *V. cholerae* or another organism of choice with a large pool of potential regulators, from which best-suited candidates to fulfill a designated function can be selected. This function could be mediated through mRNA binding and translational repression following the canonical mode of sRNA action, but it could also involve target activation, regulation of transcription termination, or sponging of other sRNA regulators (see section 1.3.2). To identify the ideal synthetic sRNA variants, they need to be enriched compared to all neutral or detrimental sRNAs. Therefore, every individual cell should carry only a single sRNA variant and a selective pressure is needed to allow those cells with the best variants to become the most abundant. Subsequent analysis of the changed sRNA pool reveals the most strongly enriched sRNA variants, which can then be studied in more detail.

The necessary selective pressure can be achieved by either natural selection (cells with matching sRNA candidates gain a fitness advantage under a given condition) or by an artificially established selection (cells with the desired regulatory outcome are technically sorted based on e.g. fluorescence intensity of a reporter). The strength of the selection will determine the degree of sRNA enrichment: if beneficial sRNA variants provide a strong fitness advantage, they will be rapidly selected over neutral or detrimental variants. Additionally, multiple rounds of selection can be exerted sequentially to further increase the abundance of beneficial sRNAs.

Limitations to the sRNA selection approach

In principle, the complexity of the randomized seed region should allow targeting of any given mRNA without restrictions at the sequence level. However, not all targets may be suitable for sRNA-mediated control depending on additional requirements. For instance, due to the Hfq dependence of the RybB-based library, mRNA targets might need to bind Hfq as well, as described

for Spot 42 targets [31]. This could impede the discovery of a matching regulator especially when the selection context is very narrow, for instance in a technical sorting of cells capable of regulating a specific target-reporter gene fusion. When the selection is defined more broadly as a biological process involving multiple regulatory factors, it is likely that the selected sRNAs will control several of these components. Still, the library might be unable to regulate a specific factor, thereby masking its involvement in the respective process.

Furthermore, it is likely that target repression would be easier to achieve than activation. Target repression by the canonical mode of blocking translation is the predominant outcome of known regulations [343] and the window in which sRNAs can pair to directly suppress translation by ribosome occlusion is relatively broad [43]. In contrast, translational activation by anti-antisense binding requires a self-inhibitory structure as prerequisite, which is not readily found in all mRNAs [191, 128, 269]. Alternatively the synthetic sRNAs could inhibit mRNA turnover by blocking cleavage sites for RNase E, which should be present in basically all mRNAs [316, 108, 60]. This could lead to target activation, as long as mRNA stabilization contributes sufficiently to increased protein levels [275, 108].

Similarly to non-targetable mRNAs, probably not all sRNA variants will be functional regulators. Specific seed permutations could fold into inhibitory secondary structures, thereby masking the seed region itself, blocking Hfq binding or interfering with correct folding and function of the terminator hairpin. This could lead to regulator instability or inactivity, reducing the number of regulation-competent variants as starting point for the selection. Additionally, some variants might be toxic for *E. coli*, the host used for library cloning, or *V. cholerae* itself. We chose an IPTG-inducible sRNA promoter to avoid potential toxic effects, but we still observed weak sRNA expression in the absence of IPTG, presumably due to promoter leakiness, which might also deplete some sRNA variants from the library. However, we were able to obtain at least 253,570 variants (97%) in the *V. cholerae* Δ*rpoE* background (see chapter 3), indicating that only very few sRNA variants have been lost during library construction and transfer.

4.4 Unbiased selection of sRNAs counter-acting envelope stress

4.4.1 Regulatory context for the laboratory selection experiments

In the first part of the publication presented in chapter 3, my colleague Nikolai Peschek identified a new σ^E-dependent sRNA, MicV, which is involved in the envelope stress response in *V. cholerae* together with the previously described sRNA VrrA [324, 325]. He characterized both sRNAs and showed that they act to relieve envelope stress by regulating unique and shared targets including the mRNAs for major OMPs and lipoproteins (chapter 3, Fig. 2B and C). As a proxy for envelope-damaging conditions, cells were exposed to ethanol to induce the σ^E regulon [65] and to study the relevance of σ^E and both sRNAs under these conditions. He showed that cell survival upon ethanol treatment is dramatically decreased in a Δ*rpoE* background (in agreement with [170]), while over-

expression of MicV, VrrA or the homologous RybB from *E. coli* restored survival almost back to wild-type levels (chapter 3, Fig. 4D).

This phenotype provided an ideal background to employ the synthetic sRNA library: a strong selection context, in which cell survival can be increased by several orders of magnitude through sRNA over-expression. It allowed us to generally study the characteristics of sRNAs selected from a large pool of potential regulators and, more specifically, to elucidate the biological determinants underlying the observed cell survival rates.

4.4.2 Detailed study of the 15 most strongly enriched sRNA variants

We performed three consecutive rounds of selection by subjecting $\Delta rpoE$ cells carrying the sRNA library to ethanol stress, plating surviving cells, and determining survival rates and sRNA content before using the pooled cells for the next selection round (chapter 3, Fig. 5). After the third selection, cell survival had increased by four orders of magnitude and was close to wild-type levels. Simultaneously, the number of detected sRNA variants had decreased to only ~5% of the initial complexity, while the 15 most enriched sRNA variants constituted ~54% of the overall library content (chapter 3, Fig. EV4B). Within the scope of the presented study, we focused on these top 15 sRNAs for further characterization. Strikingly, all of them bound to the *ompA* mRNA encoding an outer membrane porin at almost identical regions in the *ompA* coding sequence (except for three variants, which bound slightly upstream at the RBS) (chapter 3, Fig. EV5C). Binding was not restricted to the randomized seed regions, but rather extended into the first few nucleotides of the RybB backbone. The *ompA* translation initiation region contains four poly-A stretches of 3 to 6 consecutive A's, which probably facilitated the extended pairing to the U-rich 5' part of the RybB backbone. While all of the top 15 sRNA variants repressed OmpA protein production post-transcriptionally, the strength of repression varied from ~1.5-fold to ~14-fold (chapter 3, Figs. 6B and C). Several reasons for these differences are evaluated below, although the lack of experimental data limits some considerations to a theoretical discussion.

sRNA expression levels

Expression of all sRNA variants was confirmed via Northern blot. The expression levels differed ~2.5-fold between the most and least abundant variants (#10 and #3, respectively; chapter 3, Fig. 6B). Most likely, this can be attributed to different stabilities of the sRNA variants. These may result from of variations in their intrinsic half-lives due to the individual seed sequences or from different target binding capacities or alternative targets (see below), which would influence base pairing-dependent sRNA turnover. Either way, absolute sRNA levels do not allow conclusions about their target repression strength: when the extend of *ompA* repression is plotted against the sRNA expression level (normalized to the weakest variant, #3), no correlation is detectable (Fig. 4.4A). Of note, processing into shorter 3' fragments can be observed for almost all variants (12/15) to varying extent, indicating again that they undergo decay with different efficiencies. Still, these processed bands are unlikely to contribute to *ompA* repression, as they should lack the 5' encoded seed region. Thus, they

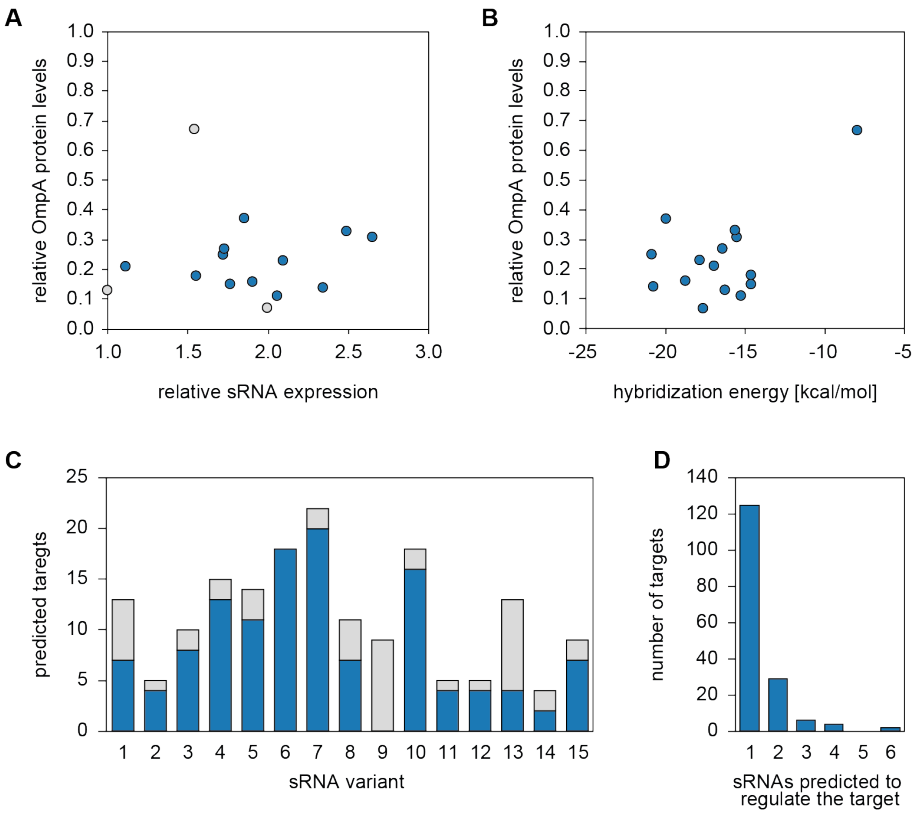


Figure 4.4: **Characteristics of the 15 most strongly enriched sRNA variants upon ethanol stress.** (A) Relative OmpA protein levels in sRNA over-expressing strains compared to a control strain (as shown in chapter 3, Fig. 6B) are plotted against the expression levels of the respective sRNA variants (normalized to the weakest expressed sRNA #3). Grey and blue dots represent sRNAs pairing to the *ompA* RBS and coding sequence, respectively. (B) Relative OmpA levels as in (A) are plotted against the hybridization energy of the sRNA-mRNA duplexes (depicted in chapter 3, Fig. EV5C). (C) Number of targets predicted for the 15 sRNA variants by TargetRNA2. Blue and grey bars represent targets uniquely predicted for the respective sRNA or also predicted for other sRNAs, respectively. (D) Number of targets that were predicted for only a single sRNA variant or for up to 6 different sRNA variants.

are not taken into account for the relative sRNA expression levels shown in Fig. 4.4A.

Position of base pairing

Remarkably, 12 out of 15 variants pair to an almost identical region on the *ompA* mRNA spanning the first 6 codons, which is also the exact same part where MicV binds (chapter 3, Fig. EV5C). Only three variants pair further upstream, at least partly covering the *ompA* RBS. Base pairing to the most 5' coding region has been shown to block ribosome entry similarly to direct pairing at the Shine-Dalgarno sequence, although it is not clear whether the exact pairing position has an influence on the regulation strength [43, 149]. At least for the top 15 sRNA variants, this seems not to be the case, as the three variants pairing at the RBS include both the strongest and the weakest repressor (#8 and #1, respectively; see red dots in Fig. 4.4A).

Strength of base pairing

Previous studies have shown a linear correlation between the strength of target regulation and

the free energy of duplex formation [131, 262], while others doubt that thermodynamics are solely responsible for variations in target regulation strength [31, 39]. For *ompA* repression by the top 15 sRNA variants, no dependency on the strength of the sRNA-mRNA duplex was observed (Fig. 4.4B). Almost all base pairings showed similar free energy values between -15 and -21 kcal/mol. The only exception is variant #1, for which the only predicted base pairing sites exhibits a very small change in free energy (-8 kcal/mol). At the same time, the variant shows the weakest *ompA* repression. However, experimental mutational data would be necessary to determine (i) if the predicted base pairing is indeed the true site of RNA-RNA duplex formation and (ii) if the mild OmpA repression is due to the weak binding and could e.g. be increased by strengthening the duplex.

Mechanistic properties of the individual variants

A variety of other factors potentially influencing the regulatory outcome of *ompA* binding could not be studied experimentally within the publication presented in chapter 3. These include especially the mechanistic details of the respective sRNA variants. It is for instance unclear if blocked *ompA* translation initiation is followed by mRNA decay for some or all sRNA variants. If so, it is also unknown if the sRNAs would act by coupled degradation and be turned over together with the *ompA* mRNA, as it has been shown for other sRNAs [197, 102]. If some variants were instead recycled after *ompA* repression, they could establish a tighter target regulation despite lower sRNA abundance. For instance, variant #3 is the most weakly expressed sRNA, but still shows the third strongest *ompA* repression (chapter 3, Fig. 6B).

Potential additional targets

As multi-targeting is a common feature of native sRNAs [21, 298, 39], it would not be surprising if the synthetic sRNA variants controlled multiple targets, too. Such additional targets would most likely affect *ompA* repression indirectly, for example by competing for the shared regulator. However, as the top 15 sRNAs were obtained through a strong selective pressure, they can be expected to represent the variants which are suited best to survive the ethanol treatment. Hence any additional targets should have beneficial effects, either directly be constituting other factors important for ethanol resistance, or indirectly by modulating *ompA* regulation to an ideal strength. Within the publication presented in chapter 3, we could not perform additional target identification experiments for all sRNA variants. However, to estimate the potential multi-targeting properties of the selected sRNAs, I performed biocomputational target predictions for the top 15 variants using TargetRNA2, a tool which can work independently of phylogenetic conservation of the sRNA or its targets [165]. As the previously determined sRNA pairings to *ompA* extended into the RybB backbone, the first eight nucleotides of the backbone were included into the prediction. The interaction site on the target mRNA was constrained to 80 nt upstream and 20 nt downstream of the start codon, while the minimal length of the required seed region was set to six consecutive nucleotides [162]. The tool correctly identified the interactions with *ompA* for 40% of the sRNA variants

(6/15) and predicted between six and 25 additional targets per sRNA (Fig. 4.4C). Of the 166 targets predicted in total for any of the top 15 sRNAs, about one quarter (41/166) was predicted for more than one sRNA (Fig. 4.4D) with *ompA* being the most commonly shared target. The fraction of predicted targets that were specific to the respective sRNA ranged from none (0/9; sRNA #9) to all targets (18/18; sRNA #6) (Fig. 4.4C). Targets predicted up to four times included the Fe-S cluster regulator *iscR* (involved in pathogenesis of *V. vulnificus* [185] and in the oxidative stress response of *Pseudomonas aeruginosa* [287]), the tmRNA-binding protein *smpB* [223], several metabolic proteins and some uncharacterized genes. However, no obvious functional connections to OmpA and ethanol stress could be discovered and a gene ontology enrichment analysis [211] of all predicted targets produced no significant results.

4.4.3 Further analysis of enriched or depleted variants

In silico analysis of the top 50 sRNA variants

Our results from the top 15 sRNA variants indicated that *ompA* is the primary target of the selected library and base pairing occurs predominantly in the 5' CDS. To evaluate whether these findings also apply to less strongly enriched variants, I extended the analysis to the top 50 sRNAs, which collectively account for 73% of the complete sRNA library after the third selection round (see appendix 6.3 for sRNA sequences). Indeed, base pairing predictions using the full *ompA* 5' UTR and the first 15 codons as input for the RNAhybrid tool [277] suggested *ompA* binding for all of the additional 35 sRNA variants (Fig 4.5A). Again, the majority of the sRNAs (20/35) were predicted to bind to the first few codons of the *ompA* CDS, but binding around the RBS (14/35 = 40%) was more common than in the top 15 variants (3/15 = 20%). Base pairing seems to predominantly occur at these two regions either upstream or downstream of the start codon, although a few pairings are predicted at the very 5' end of the mRNA (Fig. 4.5A). The strength of the predicted duplex shows no correlation with the rank of the respective sRNA (sorted by sRNA frequency, Fig. 4.5B) and neither does the position of base pairing (indicated by the dot color in Fig. 4.5B), suggesting that other factors determine the variation in sRNA enrichment. For a few variants, more than one potential binding site was predicted on *ompA*, which are then individually counted for Fig. 4.5A and represented by multiple dots for the same sRNA in Fig 4.5B.

The consensus sequence generated from the top 50 sRNA genes (Fig. 4.5C) differs slightly from the one of the top 15 sRNAs (chapter 3, Fig. EV5B), reflecting the more diverse binding sites on the *ompA* mRNA. Half of the top 15 variants pair to *ompA* at the very same position, binding the same GCGGC nucleotide stretch with their very 5' end and thereby strongly influencing the top 15 consensus towards the complementary GCCGC sequence at sRNA positions 1 to 5. In contrast, base pairing of the top 50 variants is initiated at 19 different positions, which reduces the influence of a single sequence stretch on the overall consensus motif. There is generally still a preference of G and C over A, probably due to the stronger base pairing capacity of G-C pairs, while the high frequency of T stretches most likely reflects the presence of multiple poly-A sequences within the

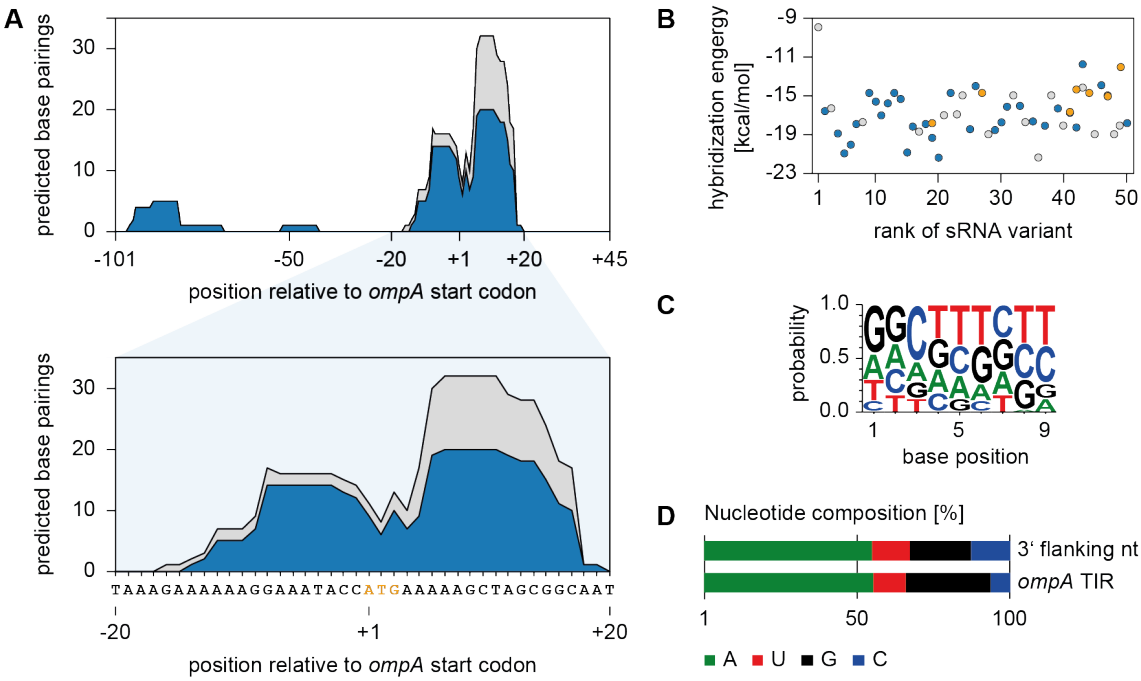


Figure 4.5: **Characteristics of the 50 most strongly enriched sRNA variants upon ethanol stress.** (A) Predicted binding sites of sRNA variants 1-15 (grey) and 16-50 (blue) on the *ompA* mRNA. Numbers indicate positions relative to the A of the *ompA* start codon. Top: Full *ompA* mRNA from the TSS to the 15th codon. Bottom: Zoom on the *ompA* TIR, the start codon is highlighted in yellow. (B) The hybridization energy of the predicted sRNA-mRNA duplexes is plotted against the rank of the respective variant in the selected library sorted by sRNA frequency. Dot color represents the position of the predicted pairing at the *ompA* RBS (grey), coding sequence (blue) or upstream in the 5' UTR (yellow). (C) Consensus motif of the top 50 sRNA variants. (D) Nucleotide content at the 3' flanking position of the predicted sRNA-mRNA duplexes and within the *ompA* TIR.

ompA translation initiation region.

It has been reported previously for the native *Salmonella* RybB that mRNAs often exhibit a flanking adenosine 3' to the base pairing site [249], reminiscent of the flanking A in eukaryotic miRNA targets [25]. Examining the *ompA*-sRNA pairings for the top 50 variants showed that a flanking A indeed appears significantly more often than any of the other three nucleotides (Fig. 4.5D). However, almost all predicted pairings are in the *ompA* translation initiation region, which is generally rich of adenosines. When the 3' flanking positions of the duplexes are compared to the overall nucleotide content of the *ompA* region in which they occur, no enrichment of a particular nucleotide is detectable. Still it cannot be ruled out that the *ompA* TIR might be especially prone to sRNA-based regulation because of the abundance of potential flanking A's. Due to the observed limitation on a narrow binding region on a single mRNA target, the ethanol stress-selected sRNA library cannot contribute substantially to the question of the necessity of a 3' A for sRNA binding. In contrast, the 24 MicV and VrrA base pairings reported in this publication correspond to 16 different targets (Fig. S3 in chapter 3) and show a 3' flanking A for 14 base pairings. This indicates that 3' adenosines are probably not strictly necessary for sRNA pairing, but are nevertheless over-represented compared to other nucleotides.

Derivatives of enriched sRNA variants

Base pairing by *trans*-encoded sRNAs is often imperfect involving bulges and unpaired nucleotides [257, 31, 251, 39]. In this regard it is interesting to note that only some synthetic sRNA variants are specifically enriched, whereas very similar ones are not selected, although they might pair to *ompA* equally well or even better. To study this effect along a representative example, I analyzed the relative sRNA enrichment of sRNA #2 and any of the 27 variants that differ in one of the nine randomized positions (Fig. 4.6A). The original variant is ranked on position 2 when sRNAs are sorted by frequency and is enriched ~7.300-fold compared to the initial, unselected library. In contrast, none of the variants that differ in only one nucleotide is enriched more than 15-fold. To compare the strength of *ompA* binding, I predicted the base pairings for two selected variants that base pair either better (green) or worse (red) than the original #2 variant (blue) (Fig. 4.6B) and calculated their hybridization energy (Fig. 4.6C). Although the calculated strength of pairing to *ompA* differs almost two-fold between the green variant (-22.9 kcal/mol) and the red variant (-12.5 kcal/mol), both are very similarly enriched to an only mild extent (3.3-fold and 3.4-fold, respectively).

Without experimental data on the regulatory properties of such closely related variants, any considerations of potential reasons for their different enrichment can only be speculative. They could differ in their stability and expression levels or in the extent of OmpA protein repression. However, at least for the top 15 sRNA variants, neither sRNA levels, nor base pairing strength or OmpA protein levels could explain their differential enrichment (Fig. 4.4). Seed mutations might also establish or interrupt pairing to unknown additional targets, thereby rewiring a possible regulatory network that is needed in a very specific state to cause high enrichment of the corresponding sRNA variant. Additionally, stochastic events during the first selection round might lead to a bottleneck effect and randomly enrich one potential regulator over another, although both would have had comparably beneficial effects on cell survival. This initial differential enrichment would then be amplified during the next two selection rounds, giving rise to the observed strong differences in frequency of very similar sRNA variants.

4.4.4 A single mRNA target responsible for stress relieve

The results presented above and in chapter 3 identified *ompA* as the key target for ethanol resistance in *V. cholerae*. In wild-type cells, ethanol stress induces MicV and VrrA through the σ^E response, which in turn repress *ompA*. In *rpoE*-deficient cells, sRNA transcription and *ompA* regulation are abolished and cell survival upon ethanol exposure is strongly reduced. The crucial role of *ompA* repression is emphasized by the fact that after an unbiased selection from ~250.000 sRNA variants, all of the 50 most abundant variants are predicted to base pair to *ompA* (Fig. 4.5A) and the top 15 variants have been experimentally shown to reduce OmpA protein levels (chapter 3, Fig. 6B). Bacterial ethanol tolerance is an important factor in both industrial and medical settings, where it is either a desired feature during microbial production processes or a problem for successful disinfection and food safety [92, 156, 159, 169]. Still the molecular mechanisms of ethanol toxicity

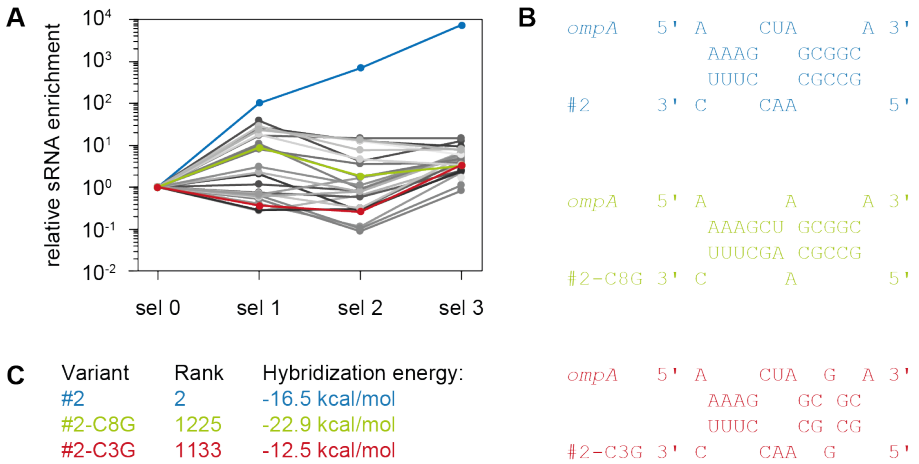


Figure 4.6: **Single nucleotide variants of a highly enriched sRNA.** (A) Relative sRNA enrichment of 27 variants of sRNA #2. Variants were obtained by individually replacing the nine positions of the randomized seed by any of the three other nucleotides. (B) Base pairing to the *ompA* coding sequence of the sRNA #2 (blue), a variant with the C at position 8 replaced by G (C8G; green) or a variant with the C at position 3 replaced by G (C3G; red). (C) Library rank by sRNA frequency for the three sRNAs from (B) and hybridization energy for their respective base pairings to *ompA*.

remain incompletely understood. Most intensively studied is the ethanol-induced increase in cell envelope permeability [150, 302]. Moreover, ethanol has been shown to target the central cellular processes of transcription and translation in *E. coli*, causing aberrant transcription termination, ribosome stalling and translational errors [127]. This could be connected to the SmpB protein of the tmRNA system [223] being predicted as target for multiple synthetic sRNA variants (see section 4.4.2). How reduced OmpA levels can mechanistically protect from ethanol toxicity is yet to be investigated. It might be involved in ethanol uptake, as an *E. coli* Δ*ompA* mutant showed slightly decreased intracellular ethanol concentrations [355]. Alternatively, OmpA itself might be altered upon ethanol exposure, e.g. becoming permeable to other toxic compounds.

While sRNA-mediated *opmA* repression strongly contributed to ethanol resistance in our experimental setting, several observations indicated it might not be the only relevant factor: (i) Deleting the *ompA* gene in the Δ*rpoE* background, thereby bypassing the need for the σ^E-dependent sRNAs, increased cell survival ~10.000-fold. Nevertheless, survival was still ~4-fold reduced in the Δ*rpoE* Δ*ompA* double mutant compared to the wild-type or the Δ*ompA* single mutant, indicating that additional σ^E targets might be involved in ethanol resistance. (ii) Furthermore, deletion of *micV* and *vrrA* in a wild-type background reduced ethanol resistance only ~10-fold and a similar decrease was observed for the *ompA* scr mutant, which can no longer base pair to the sRNAs. Thus, the effect of blocking sRNA-mediated *ompA* repression is much less severe when the remaining σ^E response is functionally intact. Together with the potential role of OmpA in the import of ethanol (or another toxic compound), this suggests the following hypothesis: in wild-type cells, uptake of ethanol or another substance is prevented by *ompA* repression through MicV and VrrA. Inhibiting this regulation reduces cell survival, but the effects of toxicity can still be counterbalanced by other members of the σ^E regulon. In contrast, *rpoE*-deficient cells strongly suffer upon ethanol exposure,

as they can neither reduce the uptake nor mitigate the damage. Deleting *ompA* or suppressing *ompA* translation by direct sRNA over-expression strongly decreases intracellular concentrations of ethanol or another compound, thereby reducing but not completely abolishing the need for the other protective σ^E functions.

4.4.5 Specialized sequence domains within sRNAs

Despite originating from different bacterial species, MicV and VrrA from *V. cholerae* share a conserved seed domain with RybB from *E. coli* and *Salmonella.* RybB is also able to protect the *V. cholerae* Δ*rpoE* mutant from ethanol stress (chapter 3, Figs. 4A and 4D). This prompted us to search for the conserved seed sequence motif in our selected sRNA library and we detected both the original RybB variant as the 11th most strongly enriched sRNA variant and a ~140-fold enrichment of the shared seed sequence motif (chapter 3, Figs. 5C and EV5A). These results indicate that functionally analogous sRNAs might have evolved shared seed domains despite being phylogenetically unrelated and encoded in different species. Such a connection would be reminiscent of the families of protein domains that can be inferred bioinformatically from protein structure predictions [155, 241]. A highly similar seed region has also been reported for the DapZ and GcvB sRNAs, which both recognize the same binding sites on their shared targets *oppA* and *dppA* [305, 61]. These observations suggest a functional classification of sRNAs by their seed regions .

On the other hand, the conserved MicV-like seed sequence is not present in MicA and MicL, the other two σ^E-dependent sRNAs in *E. coli* [252, 126]. Both repress similar targets as MicV and VrrA, but neither the sRNAs themselves nor their target interactions are conserved at the sequence level [118, 126]. Similarly, functional homologs of the iron-responsive RyhB sRNA have been described in many different species [198, 80, 109, 167, 115]. They exert comparable functions in iron homeostasis and regulate some conserved targets, but the regulators themselves share only little or no sequence conservation [292, 63].

Taken together, these observations suggest that the concept of "seed families" might be limited to act in one direction: detection of a seed region that is very similar to that of an existing sRNA can indicate a functional connection, but the absence of such a conserved seed does not exclude functional analogy.

4.5 Summary and Outlook

The present work expands the so far limited knowledge on the sRNA repertoire of *V. cholerae* by the characterization of several candidate sRNAs. Additionally, it provides valuable resources to the field of research on *Vibrio* bacteria, such as the Term-seq data on RNA 3' termini, the analysis of potential σ^E-dependent promoters, and the genome-wide identification of RNase E sites. At the same time, our findings point towards future directions of research. Comparison of the RNase E activity in *V. cholerae* to the previously determined cleavage patterns in *Salmonella* could identify conservation of cleavage sites, which would likely serve specific regulatory functions instead of just

contributing to bulk RNA turnover. The data on RNA 3' ends could be combined with the previous data on transcriptional start sites to obtain full transcript annotations of the *V. cholerae* genome, which would increase the accuracy of transcriptomic analyses in this organism.

The post-transcriptional autoregulation exerted by OppZ and CarZ is the first report of a solely RNA-based feedback circuit that is not dependent on the action of additional transcription factors. It represents a novel implementation of the concept of negative autoregulation, a well-studied network motif which is found ubiquitously throughout biology. Many questions arise from the initial characterization of the sRNA-dependent autoregulation: How does it affect the propagation of noise and phenotypic heterogeneity of the genes under control? How does it compare to TF-based autoregulation with regard to the dynamics and robustness of gene expression? What are the minimal molecular determinants to transfer the autoregulatory sRNA as an orthogonal module to heterologous genes? Is Rho-dependent feedback within the target gene only employed to increase the strength of target repression by using a second mechanism in addition to translational inhibition? Or is it necessary to avoid overshooting sRNA expression, as it limits production of the regulator after successful target regulation? Answers to these question will contribute to both fundamental research on regulatory networks and its application in the context of synthetic biology and engineered feedback loops.

Our work on the sRNAs involved in the envelope stress response of *V. cholerae* identified the OmpA porin as a critical factor during ethanol stress, but the molecular basis for its importance is yet unclear. The synthetic sRNA library that was used to identify the *ompA* repression proved to be a versatile tool to study sRNA-target interactions and their underlying molecular features. Its application to other microbial phenotypes might yield a more diverse target spectrum of the enriched sRNA variants and thus allow a systematic comparison of different sRNA-target pairs and their contribution to the studied phenotype. To facilitate the identification of mRNA targets of the sRNA library, the RIL-seq protocol could be modified to specifically amplify only those sRNA-mRNA chimera that include the synthetic library scaffold. This would allow the simultaneous detection of enriched sRNA variants and their targets without prior knowledge on potentially involved regulatory pathways. It would circumvent the limitations in bioinformatic target prediction for the synthetic sRNAs that naturally lack conservation and it would greatly reduce the time and costs compared to target identification experiments for all individual sRNA variants.

5 References for Chapters 1 and 4

[1] Adams PP and Storz G (2020). Prevalence of small base-pairing RNAs derived from diverse genomic loci. *Biochim Biophys Acta - Gene Regul Mech* 1863.7, p. 194524. PMID: 32147527.

[2] Adhya S (2003). Suboperonic regulatory signals. *Sci STKE* 2003.185, pe22. PMID: 12783981.

[3] Afonyushkin T, Vecerek B, Moll I, Bläsi U, and Kaberdin VR (2005). Both RNase E and RNase III control the stability of *sodB* mRNA upon translational inhibition by the small regulatory RNA RyhB. *Nucleic Acids Res* 33.5, pp. 1678–89. PMID: 15781494.

[4] Afroz T and Beisel CL (2013). Understanding and exploiting feedback in synthetic biology. *Chem Eng Sci* 103, pp. 79–90.

[5] Aït-Bara S, Carpousis AJ, and Quentin Y (2015). RNase E in the γ-Proteobacteria: conservation of intrinsically disordered noncatalytic region and molecular evolution of microdomains. *Mol Genet Genomics* 290.3, pp. 847–862. PMID: 25432321.

[6] Alon U (2007). Network motifs: Theory and experimental approaches. *Nat Rev Genet* 8.6, pp. 450–461. PMID: 17510665.

[7] Altuvia S, Weinstein-Fischer D, Zhang A, Postow L, and Storz G (1997). A small, stable RNA induced by oxidative stress: Role as a pleiotropic regulator and antimutator. *Cell* 90.1, pp. 43–53. PMID: 9230301.

[8] Anderson KL, Roberts C, Disz T, Vonstein V, Hwang K, Overbeek R, Olson PD, Projan SJ, and Dunman PM (2006). Characterization of the *Staphylococcus aureus* heat shock, cold shock, stringent, and SOS responses and their effects on log-phase mRNA turnover. *J Bacteriol* 188.19, pp. 6739–6756. PMID: 16980476.

[9] Andrade JM, Pobre V, and Arraiano CM (2013). Small RNA Modules Confer Different Stabilities and Interact Differently with Multiple Targets. *PLoS One* 8.1, e52866. PMID: 23349691.

[10] Andrade JM, Pobre V, Matos AM, and Arraiano CM (2012). The crucial role of PNPase in the degradation of small RNAs that are not associated with Hfq. *RNA* 18.4, pp. 844–855. PMID: 22355164.

[11] Andreassen PR, Pettersen JS, Szczerba M, Valentin-Hansen P, Møller-Jensen J, and Jørgensen MG (2018). sRNA-dependent control of curli biosynthesis in *Escherichia coli*: McaS directs endonucleolytic cleavage of *csgD* mRNA. *Nucleic Acids Res* 46.13, pp. 6746–6760. PMID: 29905843.

[12] Andrews SC, Robinson AK, and Rodríguez-Quiñones F (2003). Bacterial iron homeostasis. *FEMS Microbiol Rev* 27.2-3, pp. 215–237. PMID: 12829269.

[13] Apirion D and Lassar AB (1978). A conditional lethal mutant of *Escherichia coli* which affects the processing of ribosomal RNA. *J Biol Chem* 253.5, pp. 1738–42. PMID: 342528.

[14] Apura P, Domingues S, Viegas SC, and Arraiano CM (2019). Reprogramming bacteria with RNA regulators. *Biochem Soc Trans* 47.5, pp. 1279–1289. PMID: 31642890.

[15] Arbel-Goren R, Tal A, Friedlander T, Meshner S, Costantino N, Court DL, and Stavans J (2013). Effects of post-transcriptional regulation on phenotypic noise in *Escherichia coli*. *Nucleic Acids Res* 41.9, pp. 4825–34. PMID: 23519613.

[16] Argaman L and Altuvia S (2000). *fhlA* repression by OxyS RNA: Kissing complex formation at two sites results in a stable antisense-target RNA complex. *J Mol Biol* 300.5, pp. 1101–1112. PMID: 10903857.

[17] Argaman L, Hershberg R, Vogel J, Bejerano G, Wagner EGH, Margalit H, and Altuvia S (2001). Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*. *Curr Biol* 11.12, pp. 941–950. PMID: 11448770.

[18] Arnold TE, Yu J, and Belasco JG (1998). mRNA stabilization by the *ompA* 5' untranslated region: Two protective elements hinder distinct pathways for mRNA degradation. *RNA* 4.3, pp. 319–330. PMID: 9510333.

[19] Attaiech L, Boughammoura A, Brochier-Armanet C, Allatif O, Peillard-Fiorente F, Edwards RA, Omar AR, MacMillan AM, Glover M, and Charpentier X (2016). Silencing of natural transformation by an RNA chaperone and a multitarget small RNA. *Proc Natl Acad Sci USA* 113.31, pp. 8813–8818. PMID: 27432973.

[20] Azam MS and Vanderpool CK (2020). Translation inhibition from a distance: the small RNA SgrS silences a ribosomal protein S1-dependent enhancer. *Mol Microbiol* 114.3, pp. 391–408. PMID: 32291821.

[21] Balbontín R, Fiorini F, Figueroa-Bossi N, Casadesús J, and Bossi L (2010). Recognition of heptameric seed sequence underlies multi-target regulation by RybB small RNA in *Salmonella enterica*. *Mol Microbiol* 78.2, pp. 380–394. PMID: 20979336.

[22] Bandyra KJ, Said N, Pfeiffer V, Górna MW, Vogel J, and Luisi BF (2012). The Seed Region of a Small RNA Drives the Controlled Destruction of the Target mRNA by the Endoribonuclease RNase E. *Mol Cell* 47.6, pp. 943–953. PMID: 22902561.

[23] Bandyra KJ, Sinha D, Syrjanen J, Luisi BF, and De Lay NR (2016). The ribonuclease polynucleotide phosphorylase can interact with small regulatory RNAs in both protective and degradative modes. *RNA* 22.3, pp. 360–372. PMID: 26759452.

[24] Bandyra KJ, Wandzik JM, and Luisi BF (2018). Substrate Recognition and Autoinhibition in the Central Ribonuclease RNase E. *Mol Cell* 72.2, 275–285.e4. PMID: 30270108.

[25] Bartel DP (2009). MicroRNAs: Target Recognition and Regulatory Functions. *Cell* 136.2, pp. 215–233. PMID: 19167326.

[26] Battesti A, Majdalani N, and Gottesman S (2011). The RpoS-Mediated General Stress Response in *Escherichia coli*. *Annu Rev Microbiol* 65.1, pp. 189–213. PMID: 21639793.

[27] Bechhofer DH and Deutscher MP (2019). Bacterial ribonucleases and their roles in RNA metabolism. *Crit Rev Biochem Mol Biol* 54.3, pp. 242–300. PMID: 31464530.

[28] Becskel A and Serrano L (2000). Engineering stability in gene networks by autoregulation. *Nature* 405.6786, pp. 590–593. PMID: 10850721.

[29] Beisel CL and Storz G (2010). Base pairing small RNAs and their roles in global regulatory networks. *FEMS Microbiol Rev* 34.5, pp. 866–882. PMID: 20662934.

[30] Beisel CL and Storz G (2011). The Base-Pairing RNA Spot 42 Participates in a Multioutput Feedforward Loop to Help Enact Catabolite Repression in *Escherichia coli*. *Mol Cell* 41.3, pp. 286–297. PMID: 21292161.

[31] Beisel CL, Updegrove TB, Janson BJ, and Storz G (2012). Multiple factors dictate target selection by Hfq-binding small RNAs. *EMBO J* 31.8, pp. 1961–1974. PMID: 22388518.

[32] Belitsky BR, Barbieri G, Albertini AM, Ferrari E, Strauch MA, and Sonenshein AL (2015). Interactive regulation by the *Bacillus subtilis* global regulators CodY and ScoC. *Mol Microbiol* 97.4, pp. 698–716. PMID: 25966844.

[33] Benito Y, Kolb FA, Romby P, Lina G, Etienne J, and Vandenesch F (2000). Probing the structure of RNAIII, the *Staphylococcus aureus agr* regulatory RNA, and identification of

the RNA domain involved in repression of protein A expression. *RNA* 6.5, pp. 668–79. PMID: 10836788.

[34] Bernstein JA, Khodursky AB, Lin PH, Lin-Chao S, and Cohen SN (2002). Global analysis of mRNA decay and abundance in *Escherichia coli* at single-gene resolution using two-color fluorescent DNA microarrays. *Proc Natl Acad Sci USA* 99.15, pp. 9697–9702. PMID: 12119387.

[35] Bharat A, Jiang M, Sullivan SM, Maddock JR, and Brown ED (2006). Cooperative and critical roles for both G domains in the GTPase activity and cellular function of ribosome-associated *Escherichia coli* EngA. *J Bacteriol* 188.22, pp. 7992–7996. PMID: 16963571.

[36] Bidnenko V, Nicolas P, Grylak-Mielnicka A, Delumeau O, Auger S, Aucouturier A, Guerin C, Francis R, Bardowski J, Aymerich S, and Bidnenko E (2017). Termination factor Rho: From the control of pervasive transcription to cell fate determination in *Bacillus subtilis*. *PLoS Genet* 13.7, e1006909. PMID: 28723971.

[37] Biesen T van and Frost LS (1994). The FinO protein of IncF plasmids binds FinP antisense RNA and its target, *traJ* mRNA, and promotes duplex formation. *Mol Microbiol* 14.3, pp. 427–436. PMID: 7533880.

[38] Bilusic I, Popitsch N, Rescheneder P, Schroeder R, and Lybecker M (2014). Revisiting the coding potential of the *Escherichia coli* genome through Hfq co-immunoprecipitation. *RNA Biol* 11.5. PMID: 24922322.

[39] Bobrovskyy M, Azam MS, Frandsen JK, Zhang J, Poddar A, Ma X, Henkin TM, Ha T, and Vanderpool CK (2019). Determinants of target prioritization and regulatory hierarchy for the bacterial small RNA SgrS. *Mol Microbiol* 112.4, pp. 1199–1218. PMID: 31340077.

[40] Boisset S, Geissmann T, Huntzinger E, Fechter P, Bendridi N, Possedko M, Chevalier C, Helfer AC, Benito Y, Jacquier A, Gaspin C, Vandenesch F, and Romby P (2007). *Staphylococcus aureus* RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. *Genes Dev* 21.11, pp. 1353–1366. PMID: 17545468.

[41] Bossi L, Figueroa-Bossi N, Bouloc P, and Boudvillain M (2020). Regulatory interplay between small RNAs and transcription termination factor Rho. *Biochim Biophys Acta - Gene Regul Mech* 1863.7, p. 194546. PMID: 32217107.

[42] Bossi L, Schwartz A, Guillemardet B, Boudvillain M, and Figueroa-Bossi N (2012). A role for Rho-dependent polarity in gene regulation by a noncoding small RNA. *Genes Dev* 26.16, pp. 1864–73. PMID: 22895254.

[43] Bouvier M, Sharma CM, Mika F, Nierhaus KH, and Vogel J (2008). Small RNA Binding to 5' mRNA Coding Region Inhibits Translational Initiation. *Mol Cell* 32.6, pp. 827–837. PMID: 19111662.

[44] Braun F, Le Derout J, and Régnier P (1998). Ribosomes inhibit an RNase E cleavage which induces the decay of the *rpsO* mRNA of *Escherichia coli*. *EMBO J* 17.16, pp. 4790–4797. PMID: 9707438.

[45] Breaker RR (2018). Riboswitches and translation control. *Cold Spring Harb Perspect Biol* 10.11, a032797. PMID: 29844057.

[46] Bronesky D, Wu Z, Marzi S, Walter P, Geissmann T, Moreau K, Vandenesch F, Caldelari I, and Romby P (2016). *Staphylococcus aureus* RNAIII and Its Regulon Link Quorum Sensing, Stress Responses, Metabolic Adaptation, and Regulation of Virulence Gene Expression. *Annu Rev Microbiol* 70.1, pp. 299–316. PMID: 27482744.

[47] Brophy JA and Voigt CA (2016). Antisense transcription as a tool to tune gene expression. *Mol Syst Biol* 12.1, p. 854. PMID: 26769567.

[48] Brosse A, Korobeinikova A, Gottesman S, and Guillier M (2016). Unexpected properties of sRNA promoters allow feedback control via regulation of a two-component system. *Nucleic Acids Res* 44.20, pp. 9650–9666. PMID: 27439713.

[49] Busch A, Richter AS, and Backofen R (2008). IntaRNA: Efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. *Bioinformatics* 24.24, pp. 2849–2856. PMID: 18940824.

[50] Butz HA, Mey AR, Ciosek AL, and Payne SM (2019). *Vibrio cholerae* CsrA directly regulates *varA* to increase expression of the three nonredundant Csr small RNAs. *MBio* 10, e01042–19. PMID: 31164471.

[51] Cai L, Friedman N, and Xie XS (2006). Stochastic protein expression in individual cells at the single molecule level. *Nature* 440.7082, pp. 358–362. PMID: 16541077.

[52] Callaghan AJ, Aurikko JP, Ilag LL, Günter Grossmann J, Chandran V, Kühnel K, Poljak L, Carpousis AJ, Robinson CV, Symmons MF, and Luisi BF (2004). Studies of the RNA degradosome-organizing domain of the *Escherichia coli* ribonuclease RNase E. *J Mol Biol* 340.5, pp. 965–979. PMID: 15236960.

[53] Callaghan AJ, Marcaida MJ, Stead JA, McDowall KJ, Scott WG, and Luisi BF (2005). Structure of *Escherichia coli* RNase E catalytic domain and implications for RNA turnover. *Nature* 437.7062, pp. 1187–1191. PMID: 16237448.

[54] Calvo JM and Matthews RG (1994). The leucine-responsive regulatory protein, a global regulator of metabolism in *Escherichia coli*. *Microbiol Rev* 58.3, pp. 466–490. PMID: 7968922.

[55] Camas FM, Blázquez J, and Poyatos JF (2006). Autogenous and nonautogenous control of response in a genetic network. *Proc Natl Acad Sci USA* 103.34, pp. 12718–12723. PMID: 16908855.

[56] Cameron TA and De Lay NR (2016). The phosphorolytic exoribonucleases polynucleotide phosphorylase and RNase PH stabilize sRNAs and facilitate regulation of their mRNA targets. *J Bacteriol* 198.24, pp. 3309–3317. PMID: 27698082.

[57] Cameron TA, Matz LM, Sinha D, and De Lay NR (2019). Polynucleotide phosphorylase promotes the stability and function of Hfq-binding sRNAs by degrading target mRNA-derived fragments. *Nucleic Acids Res* 47.16, pp. 8821–8837. PMID: 31329973.

[58] Caron MP, Lafontaine DA, and Massé E (2010). Small RNA-mediated regulation at the level of transcript stability. *RNA Biol* 7.2, pp. 140–144. PMID: 20220305.

[59] Cech TR and Steitz JA (2014). The noncoding RNA revolution - Trashing old rules to forge new ones. *Cell* 157.1, pp. 77–94. PMID: 24679528.

[60] Chao Y, Li L, Girodat D, Förstner KU, Said N, Corcoran C, Smiga M, Papenfort K, Reinhardt R, Wieden HJ, Luisi BF, and Vogel J (2017). In Vivo Cleavage Map Illuminates the Central Role of RNase E in Coding and Non-coding RNA Pathways. *Mol Cell* 65.1, pp. 39–51. PMID: 28061332.

[61] Chao Y, Papenfort K, Reinhardt R, Sharma CM, and Vogel J (2012). An atlas of Hfq-bound transcripts reveals 3' UTRs as a genomic reservoir of regulatory small RNAs. *EMBO J* 31.20, pp. 4005–4019. PMID: 22922465.

[62] 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* 61.3, pp. 352–363. PMID: 26805574.

[63] Chareyre S and Mandin P (2018). Bacterial Iron Homeostasis Regulation by sRNAs. *Microbiol Spectr* 6.2, RWR–0010–2017. PMID: 29573257.

[64] Charlier D, Nguyen Le Minh P, and Roovers M (2018). Regulation of carbamoylphosphate synthesis in *Escherichia coli*: an amazing metabolite at the crossroad of arginine and pyrimidine biosynthesis. *Amino Acids* 50.12, pp. 1647–1661. PMID: 30238253.

[65] Chatterjee E and Chowdhury R (2013). Reduced virulence of the *Vibrio cholerae fadD* mutant is due to induction of the extracytoplasmic stress response. *Infect Immun* 81.10, pp. 3935–3941. PMID: 23918781.

[66] Chen CYA, Beatty JT, Cohen SN, and Belasco JG (1988). An intercistronic stem-loop structure functions as an mRNA decay terminator necessary but insufficient for *puf* mRNA stability. *Cell* 52.4, pp. 609–619. PMID: 2449287.

[67] Chen J, Morita T, and Gottesman S (2019). Regulation of transcription termination of small RNAs and by small RNAs: Molecular mechanisms and biological functions. *Front Cell Infect Microbiol* 9, p. 201. PMID: 31249814.

[68] Chen S, Lesnik EA, Hall TA, Sampath R, Griffey RH, Ecker DJ, and Blyn LB (2002). A bioinformatics based approach to discover small RNA genes in the *Escherichia coli* genome. *BioSystems* 65.2-3, pp. 157–177. PMID: 12069726.

[69] Chen Z, Lewis KA, Shultzaberger RK, Lyakhov IG, Zheng M, Doan B, Storz G, and Schneider TD (2007). Discovery of Fur binding site clusters in *Escherichia coli* by information theory models. *Nucleic Acids Res* 35.20, pp. 6762–6777. PMID: 17921503.

[70] Chevalier C, Boisset S, Romilly C, Masquida B, Fechter P, Geissmann T, Vandenesch F, and Romby P (2010). *Staphylococcus aureus* RNAIII Binds to Two Distant Regions of *coa* mRNA to Arrest Translation and Promote mRNA Degradation. *PLoS Pathog* 6.3, e1000809. PMID: 20300607.

[71] Cho SH, Haning K, and Contreras LM (2015). Strain engineering via regulatory noncoding RNAs: Not a one-blueprint-fits-all. *Curr Opin Chem Eng* 10, pp. 25–34.

[72] Choi E, Han Y, Cho YJ, Nam D, and Lee EJ (2017). A *trans*-acting leader RNA from a *Salmonella* virulence gene. *Proc Natl Acad Sci USA* 114.38, pp. 10232–10237. PMID: 28874555.

[73] Cisse II, Kim H, and Ha T (2012). A rule of seven in Watson-Crick base-pairing of mismatched sequences. *Nat Struct Mol Biol* 19.6, pp. 623–627. PMID: 22580558.

[74] Clarke JE, Kime L, Romero A D, and McDowall KJ (2014). Direct entry by RNase E is a major pathway for the degradation and processing of RNA in *Escherichia coli*. *Nucleic Acids Res* 42.18, pp. 11733–51. PMID: 25237058.

[75] Corcoran CP, Podkaminski D, Papenfort K, Urban JH, Hinton JC, and Vogel J (2012). Superfolder GFP reporters validate diverse new mRNA targets of the classic porin regulator, MicF RNA. *Mol Microbiol* 84.3, pp. 428–445. PMID: 22458297.

[76] Dar D, Shamir M, Mellin JR, Koutero M, Stern-Ginossar N, Cossart P, and Sorek R (2016). Term-seq reveals abundant ribo-regulation of antibiotics resistance in bacteria. *Science* 352.6282, aad9822. PMID: 27120414.

[77] Dar D and Sorek R (2018a). Extensive reshaping of bacterial operons by programmed mRNA decay. *PLOS Genet* 14.4, e1007354. PMID: 29668692.

[78] Dar D and Sorek R (2018b). High-resolution RNA 3'-ends mapping of bacterial Rho-dependent transcripts. *Nucleic Acids Res* 46.13, pp. 6797–6805. PMID: 29669055.

[79] Darfeuille F, Unoson C, Vogel J, and Wagner EGH (2007). An Antisense RNA Inhibits Translation by Competing with Standby Ribosomes. *Mol Cell* 26.3, pp. 381–392. PMID: 17499044.

[80] Davis BM, Quinones M, Pratt J, Ding Y, and Waldor MK (2005). Characterization of the small untranslated RNA RyhB and its regulon in *Vibrio cholerae*. *J Bacteriol* 187.12, pp. 4005–4014. PMID: 15937163.

[81] Davis BM and Waldor MK (2007). RNase E-dependent processing stabilizes MicX, a *Vibrio cholerae* sRNA. *Mol Microbiol* 65.2, pp. 373–85. PMID: 17590231.

[82] Davis BM and Waldor MK (2009). High-throughput sequencing reveals suppressors of *Vibrio cholerae rpoE* mutations: One fewer porin is enough. *Nucleic Acids Res* 37.17, pp. 5757–5767. PMID: 19620211.

[83] De Felice M, Guardiola J, Lamberti A, and Iaccarino M (1973). *Escherichia coli* K-12 Mutants Altered in the Transport Systems for Oligo- and Dipeptides. *J Bacteriol* 116.2, pp. 751–756. PMID: 4126826.

[84] De Lay NR and Gottesman S (2011). Role of polynucleotide phosphorylase in sRNA function in *Escherichia coli*. *RNA* 17.6, pp. 1172–1189. PMID: 21527671.

[85] De Lay NR and Gottesman S (2012). A complex network of small non-coding RNAs regulate motility in *Escherichia coli*. *Mol Microbiol* 86.3, pp. 524–538. PMID: 22925049.

[86] De Mets F, Van Melderen L, and Gottesman S (2019). Regulation of acetate metabolism and coordination with the TCA cycle via a processed small RNA. *Proc Natl Acad Sci USA* 116.3, pp. 1043–1052. PMID: 30591570.

[87] Deana A and Belasco JG (2005). Lost in translation: The influence of ribosomes on bacterial mRNA decay. *Genes Dev* 19.21, pp. 2526–2533. PMID: 16264189.

[88] Deana A, Celesnik H, and Belasco JG (2008). The bacterial enzyme RppH triggers messenger RNA degradation by 5' pyrophosphate removal. *Nature* 451.7176, pp. 355–358. PMID: 18202662.

[89] Desnoyers G and Massé E (2012). Noncanonical repression of translation initiation through small RNA recruitment of the RNA chaperone Hfq. *Genes Dev* 26.7, pp. 726–739. PMID: 22474262.

[90] Desnoyers G, Morissette A, Prévost K, and Massé E (2009). Small RNA-induced differential degradation of the polycistronic mRNA *iscRSUA*. *EMBO J* 28.11, pp. 1551–1561. PMID: 19407815.

[91] Di Salvo M, Puccio S, Peano C, Lacour S, and Alifano P (2019). RhoTermPredict: An algorithm for predicting Rho-dependent transcription terminators based on *Escherichia coli*, *Bacillus subtilis* and *Salmonella enterica* databases. *BMC Bioinformatics* 20, p. 117. PMID: 30845912.

[92] Dien BS, Cotta MA, and Jeffries TW (2003). Bacteria engineered for fuel ethanol production: Current status. *Appl Microbiol Biotechnol* 63.3, pp. 258–266. PMID: 13680206.

[93] Dimastrogiovanni D, Fröhlich KS, Bandyra KJ, Bruce HA, Hohensee S, Vogel J, and Luisi BF (2014). Recognition of the small regulatory RNA RydC by the bacterial Hfq protein. *Elife* 3, e05375. PMID: 25551292.

[94] Dower WJ, Miller JF, and Ragsdale CW (1988). High efficiency transformation of *Escherichia coli* by high voltage electroporation. *Nucleic Acids Res* 16.13, pp. 6127–6145. PMID: 3041370.

[95] Dubrac S and Touati D (2000). Fur positive regulation of iron superoxide dismutase in *Escherichia coli*: Functional analysis of the *sodB* promoter. *J Bacteriol* 182.13, pp. 3802–3808. PMID: 10850997.

[96] Ellis JC and Brown JW (2009). The RNase P family. *RNA Biol* 6.4, pp. 362–9. PMID: 19738420.

[97] Ellis MJ, Trussler RS, Charles O, and Haniford DB (2017). A transposon-derived small RNA regulates gene expression in *Salmonella Typhimurium*. *Nucleic Acids Res* 45.9, pp. 5470–5486. PMID: 28335027.

[98] Esakova O and Krasilnikov AS (2010). Of proteins and RNA: The RNase P/MRP family. *RNA* 16.9, pp. 1725–1747. PMID: 20627997.

[99] Escolar L, Pérez-Martín J, and De Lorenzo V (1999). Opening the iron box: Transcriptional metalloregulation by the Fur protein. *J Bacteriol* 181.20, pp. 6223–6229. PMID: 10515908.

[100] Even S, Pellegrini O, Zig L, Labas V, Vinh J, Bréchemmier-Baey D, and Putzer H (2005). Ribonucleases J1 and J2: Two novel endoribonucleases in *Bacillus subtilis* with functional homology to *Escherichia coli* RNase E. *Nucleic Acids Res* 33.7, pp. 2141–2152. PMID: 15831787.

[101] Fender A, Elf J, Hampel K, Zimmermann B, and Wagner EGH (2010). RNAs actively cycle on the Sm-like protein Hfq. *Genes Dev* 24.23, pp. 2621–2626. PMID: 21123649.

[102] Feng L, Rutherford ST, Papenfort K, Bagert JD, Van Kessel JC, Tirrell DA, Wingreen NS, and Bassler BL (2015). A Qrr noncoding RNA deploys four different regulatory mechanisms to optimize quorum-sensing dynamics. *Cell* 160.1-2, pp. 228–240. PMID: 25579683.

[103] Figueroa-Bossi N and Bossi L (2018). Sponges and Predators in the Small RNA World. *Microbiol Spectr* 6.4, RWR-0021–2018. PMID: 30003868.

[104] 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 Dev* 23.17, pp. 2004–2015. PMID: 19638370.

[105] Förstner KU, Reuscher CM, Haberzettl K, Weber L, and Klug G (2018). RNase E cleavage shapes the transcriptome of *Rhodobacter sphaeroides* and strongly impacts phototrophic growth. *Life Sci Alliance* 1.4, e201800080. PMID: 30456366.

[106] Fröhlich KS, Haneke K, Papenfort K, and Vogel J (2016). The target spectrum of SdsR small RNA in *Salmonella*. *Nucleic Acids Res* 44.21, pp. 10406–10422. PMID: 27407104.

[107] Fröhlich KS, Papenfort K, Berger AA, and Vogel J (2012). A conserved RpoS-dependent small RNA controls the synthesis of major porin OmpD. *Nucleic Acids Res* 40.8, pp. 3623–40. PMID: 22180532.

[108] Fröhlich KS, Papenfort K, Fekete A, and Vogel J (2013). A small RNA activates CFA synthase by isoform-specific mRNA stabilization. *EMBO J* 32.22, pp. 2963–2979. PMID: 24141880.

[109] Gaballa A, Antelmann H, Aguilar C, Khakh SK, Song KB, Smaldone GT, and Helmann JD (2008). The *Bacillus subtilis* iron-sparing response is mediated by a Fur-regulated small RNA and three small, basic proteins. *Proc Natl Acad Sci USA* 105.33, pp. 11927–11932. PMID: 18697947.

[110] Gaida SM, Al-Hinai MA, Indurthi DC, Nicolaou SA, and Papoutsakis ET (2013). Synthetic tolerance: Three noncoding small RNAs, DsrA, ArcZ and RprA, acting supra-additively against acid stress. *Nucleic Acids Res* 41.18, pp. 8726–8737. PMID: 23892399.

[111] Garai P, Chandra K, and Chakravortty D (2017). Bacterial peptide transporters: Messengers of nutrition to virulence. *Virulence* 8.3, pp. 297–309. PMID: 27589415.

[112] Gardan R, Besset C, Guillot A, Gitton C, and Monnet V (2009). The oligopeptide transport system is essential for the development of natural competence in *Streptococcus thermophilus* strain LMD-9. *J Bacteriol* 191.14, pp. 4647–4655. PMID: 19447907.

[113] Garrey SM and Mackie GA (2011). Roles of the 5'-phosphate sensor domain in RNase E. *Mol Microbiol* 80.6, pp. 1613–1624. PMID: 21518390.

[114] Georg J and Hess WR (2018). Widespread Antisense Transcription in Prokaryotes. *Microbiol Spectr* 6.4, RWR-0029–2018. PMID: 30003872.

[115] Georg J, Kostova G, Vuorijoki L, Schön V, Kadowaki T, Huokko T, Baumgartner D, Müller M, Klähn S, Allahverdiyeva Y, Hihara Y, Futschik ME, Aro EM, and Hess WR (2017). Acclimation of Oxygenic Photosynthesis to Iron Starvation Is Controlled by the sRNA IsaR1. *Curr Biol* 27.10, 1425–1436.e7. PMID: 28479323.

[116] Gimpel M and Brantl S (2017). Dual-function small regulatory RNAs in bacteria. *Mol Microbiol* 103.3, pp. 387–397. PMID: 27750368.

[117] Girard ME, Gopalkrishnan S, Grace ED, Halliday JA, Gourse RL, and Herman C (2018). DksA and ppGpp regulate the σ^S stress response by activating promoters for the small RNA DsrA and the anti-adapter protein IraP. *J Bacteriol* 200.2, e00463–17. PMID: 29061665.

[118] Gogol EB, Rhodius VA, Papenfort K, Vogel J, and Gross CA (2011). Small RNAs endow a transcriptional activator with essential repressor functions for single-tier control of a global stress regulon. *Proc Natl Acad Sci USA* 108.31, pp. 12875–12880. PMID: 21768388.

[119] Göpel Y, Papenfort K, Reichenbach B, Vogel J, and Görke B (2013). Targeted decay of a regulatory small RNA by an adaptor protein for RNase E and counteraction by an anti-adaptor RNA. *Genes Dev* 27.5, pp. 552–564. PMID: 23475961.

[120] Görke B and Stülke J (2008). Carbon catabolite repression in bacteria: Many ways to make the most out of nutrients. *Nat Rev Microbiol* 6.8, pp. 613–624. PMID: 18628769.

[121] Gorski SA, Vogel J, and Doudna JA (2017). RNA-based recognition and targeting: Sowing the seeds of specificity. *Nat Rev Mol Cell Biol* 18.4, pp. 215–228. PMID: 28196981.

[122] Gottesman S (2019). Trouble is coming: Signaling pathways that regulate general stress responses in bacteria. *J Biol Chem* 294.31, pp. 11685–11700. PMID: 31197038.

[123] Grantcharova N, Peters V, Monteiro C, Zakikhany K, and Römling U (2010). Bistable expression of CsgD in biofilm development of *Salmonella enterica* serovar Typhimurium. *J Bacteriol* 192.2, pp. 456–466. PMID: 19897646.

[124] Guillier M and Gottesman S (2006). Remodelling of the *Escherichia coli* outer membrane by two small regulatory RNAs. *Mol Microbiol* 59.1, pp. 231–247. PMID: 16359331.

[125] Guillier M and Gottesman S (2008). The 5' end of two redundant sRNAs is involved in the regulation of multiple targets, including their own regulator. *Nucleic Acids Res* 36.21, pp. 6781–94. PMID: 18953042.

[126] Guo MS, Updegrove TB, Gogol EB, Shabalina SA, Gross CA, and Storz G (2014). MicL, a new σ^E-dependent sRNA, combats envelope stress by repressing synthesis of Lpp, the major outer membrane lipoprotein. *Genes Dev* 28.14, pp. 1620–1634. PMID: 25030700.

[127] Haft RJ, Keating DH, Schwaegler T, Schwalbach MS, Vinokur J, Tremaine M, Peters JM, Kotlajich MV, Pohlmann EL, Ong IM, Grass JA, Kiley PJ, and Landick R (2014). Correcting direct effects of ethanol on translation and transcription machinery confers ethanol tolerance in bacteria. *Proc Natl Acad Sci USA* 111.25, E2576–E2585. PMID: 24927582.

[128] Hammer BK and Bassler BL (2007). Regulatory small RNAs circumvent the conventional quorum sensing pathway in pandemic *Vibrio cholerae*. *Proc Natl Acad Sci USA* 104.27, pp. 11145–11149. PMID: 17556542.

[129] Hao Y, Updegrove TB, Livingston NN, and Storz G (2016). Protection against deleterious nitrogen compounds: role of σ^S-dependent small RNAs encoded adjacent to *sdiA*. *Nucleic Acids Res* 44.14, pp. 6935–48. PMID: 27166377.

[130] Hao Y, Xu L, and Shi H (2011). Theoretical analysis of catalytic-sRNA-mediated gene silencing. *J Mol Biol* 406.1, pp. 195–204. PMID: 21145897.

[131] Hao Y, Zhang Z, Erickson DW, Huang M, Huang Y, Li J, Hwa T, and Shi H (2011). Quantifying the sequence-function relation in gene silencing by bacterial small RNAs. *Proc Natl Acad Sci USA* 108.30, pp. 12473–12478. PMID: 21742981.

[132] Heidelberg JF, Elsen JA, Nelson WC, Clayton RA, Gwinn ML, Dodson RJ, Haft DH, Hickey EK, Peterson JD, Umayam L, Gill SR, Nelson KE, Read TD, Tettelin H, Richardson D, Ermolaeva MD, Vamathevan J, Bass S, Halving Q, Dragol I, Sellers P, McDonald L, Utterback T, Fleishmann RD, Nierman WC, White O, Saizberg SL, Smith HO, Colwell RR, Mekalanos JJ, Venter CJ, and Fraser CM (2000). DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature* 406.6795, pp. 477–483. PMID: 10952301.

[133] Heidrich N, Bauriedl S, Barquist L, Li L, Schoen C, and Vogel J (2017). The primary transcriptome of *Neisseria meningitidis* and its interaction with the RNA chaperone Hfq. *Nucleic Acids Res* 45.10, pp. 6147–6167. PMID: 28334889.

[134] Helmann JD (2019). Where to begin? Sigma factors and the selectivity of transcription initiation in bacteria. *Mol Microbiol* 112.2, pp. 335–347. PMID: 31119812.

[135] Herzog R, Peschek N, Fröhlich KS, Schumacher K, and Papenfort K (2019). Three autoinducer molecules act in concert to control virulence gene expression in *Vibrio cholerae*. *Nucleic Acids Res* 47.6, pp. 3171–3183. PMID: 30649554.

[136] Hiles ID, Gallagher MP, Jamieson DJ, and Higgins CF (1987). Molecular characterization of the oligopeptide permease of *Salmonella typhimurium*. *J Mol Biol* 195.1, pp. 125–142. PMID: 2821267.

[137] Himeno H, Nameki N, Kurita D, Muto A, and Abo T (2015). Ribosome rescue systems in bacteria. *Biochimie* 114, pp. 102–112. PMID: 25446863.

[138] Hoe CH, Raabe CA, Rozhdestvensky TS, and Tang TH (2013). Bacterial sRNAs: Regulation in stress. *Int J Med Microbiol* 303.5, pp. 217–229. PMID: 23660175.

[139] Hoekzema M, Romilly C, Holmqvist E, and Wagner EGH (2019). Hfq-dependent mRNA unfolding promotes sRNA-based inhibition of translation. *EMBO J* 38.7, e101199. PMID: 30833291.

[140] 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* 70.5, 971–982.e6. PMID: 29804828.

[141] Holmqvist E, Reimegård J, Sterk M, Grantcharova N, Römling U, and Wagner EGH (2010). Two antisense RNAs target the transcriptional regulator CsgD to inhibit curli synthesis. *EMBO J* 29.11, pp. 1840–50. PMID: 20407422.

[142] Holmqvist E, Unoson C, Reimegård J, and Wagner EGH (2012). A mixed double negative feedback loop between the sRNA MicF and the global regulator Lrp. *Mol Microbiol* 84.3, pp. 414–427. PMID: 22324810.

[143] Holmqvist E and Wagner EGH (2017). Impact of bacterial sRNAs in stress responses. *Biochem Soc Trans* 45.6, pp. 1203–1212. PMID: 29101308.

[144] Huber M, Fröhlich KS, Radmer J, and Papenfort K (2020). Switching fatty acid metabolism by an RNA-controlled feed forward loop. *Proc Natl Acad Sci USA* 117.14, pp. 8044–8054. PMID: 32193348.

[145] Hui MP, Foley PL, and Belasco JG (2014). Messenger RNA Degradation in Bacterial Cells. *Annu Rev Genet* 48.1, pp. 537–559. PMID: 25292357.

[146] Huntzinger E, Boisset S, Saveanu C, Benito Y, Geissmann T, Namane A, Lina G, Etienne J, Ehresmann B, Ehresmann C, Jacquier A, Vandenesch F, and Romby P (2005). *Staphylococcus aureus* RNAIII and the endoribonuclease III coordinately regulate *spa* gene expression. *EMBO J* 24.4, pp. 824–835. PMID: 15678100.

[147] Hussein R and Lim HN (2011). Disruption of small RNA signaling caused by competition for Hfq. *Proc Natl Acad Sci USA* 108.3, pp. 1110–1115. PMID: 21189298.

[148] Hussein R and Lim HN (2012). Direct comparison of small RNA and transcription factor signaling. *Nucleic Acids Res* 40.15, pp. 7269–79. PMID: 22618873.

[149] Hüttenhofer A and Noller H (1994). Footprinting mRNA-ribosome complexes with chemical probes. *EMBO J* 13.16, pp. 3892–3901. PMID: 8070416.

[150] Ingram LO (1989). Ethanol tolerance in bacteria. *Crit Rev Biotechnol* 9.4, pp. 305–319. PMID: 2178781.

[151] Inoue H, Nojima H, and Okayama H (1990). High efficiency transformation of *Escherichia coli* with plasmids. *Gene* 96.1, pp. 23–28. PMID: 2265755.

[152] Ishikawa H, Otaka H, Maki K, Morita T, and Aiba H (2012). The functional Hfq-binding module of bacterial sRNAs consists of a double or single hairpin preceded by a U-rich sequence and followed by a 3' poly(U) tail. *RNA* 18.5, pp. 1062–1074. PMID: 22454537.

[153] Jagodnik J, Brosse A, Le Lam TN, Chiaruttini C, and Guillier M (2017). Mechanistic study of base-pairing small regulatory RNAs in bacteria. *Methods* 115, pp. 67–76. PMID: 27693881.

[154] Jain C and Belasco JG (1995). RNase E autoregulates its synthesis by controlling the degradation rate of its own mRNA in *Escherichia coli*: Unusual sensitivity of the *rne* transcript to RNase E activity. *Genes Dev* 9.1, pp. 84–96. PMID: 7530223.

[155] Janin J and Wodak SJ (1983). Structural domains in proteins and their role in the dynamics of protein function. *Prog Biophys Mol Biol* 42, pp. 21–78. PMID: 6353481.

[156] Jarboe LR, Grabar TB, Yomano LP, Shanmugan KT, and Ingram LO (2007). Development of ethanologenic bacteria. *Adv Biochem Eng Biotechnol* 108, pp. 237–261. PMID: 17665158.

[157] Johansen J, Rasmussen AA, Overgaard M, and Valentin-Hansen P (2006). Conserved Small Non-coding RNAs that belong to the σ^{E} Regulon: Role in Down-regulation of Outer Membrane Proteins. *J Mol Biol* 364.1, pp. 1–8. PMID: 17007876.

[158] Jørgensen MG, Nielsen JS, Boysen A, Franch T, Møller-Jensen J, and Valentin-Hansen P (2012). Small regulatory RNAs control the multi-cellular adhesive lifestyle of *Escherichia coli*. *Mol Microbiol* 84.1, pp. 36–50. PMID: 22250746.

[159] Kampf G and Kramer A (2004). Epidemiologic background of hand hygiene and evaluation of the most important agents for scrubs and rubs. *Clin Microbiol Rev* 17.4, pp. 863–893. PMID: 15489352.

[160] Kang Z, Zhang C, Zhang J, Jin P, Zhang J, Du G, and Chen J (2014). Small RNA regulators in bacteria: Powerful tools for metabolic engineering and synthetic biology. *Appl Microbiol Biotechnol* 98.8, pp. 3413–3424. PMID: 24519458.

[161] Kavita K, De Mets F, and Gottesman S (2018). New aspects of RNA-based regulation by Hfq and its partner sRNAs. *Curr Opin Microbiol* 42, pp. 53–61. PMID: 29125938.

[162] Kawamoto H, Koide Y, Morita T, and Aiba H (2006). Base-pairing requirement for RNA silencing by a bacterial small RNA and acceleration of duplex formation by Hfq. *Mol Microbiol* 61.4, pp. 1013–1022. PMID: 16859494.

[163] Kawano M, Reynolds AA, Miranda-Rios J, and Storz G (2005). Detection of 5'- and 3'-UTR-derived small RNAs and *cis*-encoded antisense RNAs in *Escherichia coli*. *Nucleic Acids Res* 17.33, pp. 1040–50. PMID: 15718303.

[164] Kelly CL, Harris AW, Steel H, Hancock EJ, Heap JT, and Papachristodoulou A (2018). Synthetic negative feedback circuits using engineered small RNAs. *Nucleic Acids Res* 46.18, pp. 9875–9889. PMID: 30212900.

[165] Kery MB, Feldman M, Livny J, and Tjaden BC (2014). TargetRNA2: identifying targets of small regulatory RNAs in bacteria. *Nucleic Acids Res* 42.Web Server issue, W124–9. PMID: 24753424.

[166] Kim HM, Shin JH, Cho YB, and Roe JH (2014). Inverse regulation of Fe- and Ni-containing SOD genes by a Fur family regulator Nur through small RNA processed from 3'UTR of the *sodF* mRNA. *Nucleic Acids Res* 42.3, pp. 2003–14. PMID: 24234448.

[167] Kim JN (2016). Roles of two RyhB paralogs in the physiology of *Salmonella enterica*. *Microbiol Res* 186-187, pp. 146–152. PMID: 27242152.

[168] Kim W and Lee Y (2020). Mechanism for coordinate regulation of *rpoS* by sRNA-sRNA interaction in *Escherichia coli*. *RNA Biol* 17.2, pp. 176–187. PMID: 31552789.

[169] Köhler AT, Rodloff AC, Labahn M, Reinhardt M, Truyen U, and Speck S (2019). Evaluation of disinfectant efficacy against multidrug-resistant bacteria: A comprehensive analysis of different methods. *Am J Infect Control* 47.10, pp. 1181–1187. PMID: 31060869.

[170] Kovacikova G and Skorupski K (2002). The alternative sigma factor σ^{E} plays an important role in intestinal survival and virulence in *Vibrio cholerae*. *Infect Immun* 70.10, pp. 5355–5362. PMID: 12228259.

[171] Kudla G, Granneman S, Hahn D, Beggs JD, and Tollervey D (2011). Cross-linking, ligation, and sequencing of hybrids reveals RNA-RNA interactions in yeast. *Proc Natl Acad Sci USA* 108.24, pp. 10010–10015. PMID: 21610164.

[172] Künne T, Swarts DC, and Brouns SJ (2014). Planting the seed: Target recognition of short guide RNAs. *Trends Microbiol* 22.2, pp. 74–83. PMID: 24440013.

[173] Kuzj AE, Medberry PS, and Schottel JL (1998). Stationary phase, amino acid limitation and recovery from stationary phase modulate the stability and translation of chloramphenicol acetyltransferase mRNA and total mRNA in *Escherichia coli*. *Microbiology* 144.3, pp. 739–750. PMID: 9534243.

[174] Lahiry A, Stimple SD, Wood DW, and Lease RA (2017). Retargeting a Dual-Acting sRNA for Multiple mRNA Transcript Regulation. *ACS Synth Biol* 6.4, pp. 648–658. PMID: 28067500.

[175] Lalaouna D, Carrier MC, Semsey S, Brouard JS, Wang J, Wade JT, 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* 58.3, pp. 393–405. PMID: 25891076.

[176] Lalaouna D, Eyraud A, Devinck A, Prévost K, and Massé E (2019). GcvB small RNA uses two distinct seed regions to regulate an extensive targetome. *Mol Microbiol* 111.2, pp. 473–486. PMID: 30447071.

[177] Lalaouna D, Simoneau-Roy M, Lafontaine DA, and Massé E (2013). Regulatory RNAs and target mRNA decay in prokaryotes. *Biochim Biophys Acta - Gene Regul Mech* 1829.6-7, pp. 742–747. PMID: 23500183.

[178] Leistra AN, Curtis NC, and Contreras LM (2019). Regulatory non-coding sRNAs in bacterial metabolic pathway engineering. *Metab Eng* 52, pp. 190–214. PMID: 30513348.

[179] Lejars M and Hajnsdorf E (2020). The world of asRNAs in Gram-negative and Gram-positive bacteria. *Biochim Biophys Acta - Gene Regul Mech* 1863.2, p. 194489. PMID: 31935527.

[180] Lenz DH, Miller MB, Zhu J, Kulkarnl RV, and Bassler BL (2005). CsrA and three redundant small RNAs regulate quorum sensing in *Vibrio cholerae*. *Mol Microbiol* 58.4, pp. 1186–1202. PMID: 16262799.

[181] Lenz DH, Mok KC, Lilley BN, Kulkarni RV, Wingreen NS, and Bassler BL (2004). The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in *Vibrio harveyi* and *Vibrio cholerae*. *Cell* 118.1, pp. 69–82. PMID: 15242645.

[182] Levine E and Hwa T (2008). Small RNAs establish gene expression thresholds. *Curr Opin Microbiol* 11.6, pp. 574–579. PMID: 18935980.

[183] Levine E, Zhang Z, Kuhlman T, and Hwa T (2007). Quantitative characteristics of gene regulation by small RNA. *PLoS Biol* 5.9, pp. 1998–2010. PMID: 17713988.

[184] Li GW, Burkhardt D, Gross C, and Weissman JS (2014). Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cell* 157.3, pp. 624–635. PMID: 24766808.

[185] Lim JG and Choi SH (2014). IscR is a global regulator essential for pathogenesis of *Vibrio vulnificus* and induced by host cells. *Infect Immun* 82.2, pp. 569–578. PMID: 24478072.

[186] Livny J and Waldor MK (2007). Identification of small RNAs in diverse bacterial species. *Curr Opin Microbiol* 10.2, pp. 96–101. PMID: 17383222.

[187] Loh E, Dussurget O, Gripenland J, Vaitkevicius K, Tiensuu T, Mandin P, Repoila F, Buchrieser C, Cossart P, and Johansson J (2009). A *trans*-Acting Riboswitch Controls Expression of the Virulence Regulator PrfA in *Listeria monocytogenes*. *Cell* 139.4, pp. 770–779. PMID: 19914169.

[188] Loh E, Righetti F, Eichner H, Twittenhoff C, and Narberhaus F (2018). RNA Thermometers in Bacterial Pathogens. *Microbiol Spectr* 6.2, RWR-0012–2017. PMID: 29623874.

[189] Luciano DJ, Vasilyev N, Richards J, Serganov A, and Belasco JG (2017). A Novel RNA Phosphorylation State Enables 5' End-Dependent Degradation in *Escherichia coli*. *Mol Cell* 67.1, 44–54.e6. PMID: 28673541.

[190] MacKenzie KD, Wang Y, Shivak DJ, Wong CS, Hoffman LJ, Lam S, Kröger C, Cameron AD, Townsend HG, Köster W, and White AP (2015). Bistable expression of CsgD in *Salmonella enterica* serovar Typhimurium connects virulence to persistence. *Infect Immun* 83.6, pp. 2312–2326. PMID: 25824832.

[191] Majdalani N, Cunning C, Sledjeski D, Elliott T, and Gottesman S (1998). DsrA RNA regulates translation of *rpoS* message by an anti-antisense mechanism, independent of its action as an antisilencer of transcription. *Proc Natl Acad Sci USA* 95.21, pp. 12462–12467. PMID: 9770508.

[192] Majdalani N, Hernandez D, and Gottesman S (2002). Regulation and mode of action of the second small RNA activator of *rpoS* translation, RprA. *Mol Microbiol* 46.3, pp. 813–826. PMID: 12410838.

[193] Man S, Cheng R, Miao C, Gong Q, Gu Y, Lu X, Han F, and Yu W (2011). Artificial *trans*-encoded small non-coding RNAs specifically silence the selected gene expression in bacteria. *Nucleic Acids Res* 39.8, e50. PMID: 21296758.

[194] Mandin P and Gottesman S (2010). Integrating anaerobic/aerobic sensing and the general stress response through the ArcZ small RNA. *EMBO J* 29.18, pp. 3094–3107. PMID: 20683441.

[195] Mandin P and Guillier M (2013). Expanding control in bacteria: Interplay between small RNAs and transcriptional regulators to control gene expression. *Curr Opin Microbiol* 16.2, pp. 125–132. PMID: 23415757.

[196] Mangan S and Alon U (2003). Structure and function of the feed-forward loop network motif. *Proc Natl Acad Sci USA* 100.21, pp. 11980–11985. PMID: 14530388.

[197] Massé E, Escorcia FE, and Gottesman S (2003). Coupled degradation of a small regulatory RNA and its mRNA targets in *Escherichia coli*. *Genes Dev* 17.19, pp. 2374–2383. PMID: 12975324.

[198] Massé E and Gottesman S (2002). A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*. *Proc Natl Acad Sci USA* 99.7, pp. 4620–4625. PMID: 11917098.

[199] Massé E, Vanderpool CK, and Gottesman S (2005). Effect of RyhB small RNA on global iron use in *Escherichia coli*. *J Bacteriol* 187.20, pp. 6962–6971. PMID: 16199566.

[200] Masuda T, Saito N, Tomita M, and Ishihama Y (2009). Unbiased quantitation of *Escherichia coli* membrane proteome using phase transfer surfactants. *Mol Cell Proteomics* 8.12, pp. 2770–2777. PMID: 19767571.

[201] Masulis IS, Sukharycheva NA, Kiselev SS, Andreeva ZS, and Ozoline ON (2020). Between computational predictions and high-throughput transcriptional profiling: in depth expression analysis of the OppB transmembrane subunit of *Escherichia coli* OppABCDF oligopeptide transporter. *Res Microbiol* 171.2, pp. 55–63. PMID: 31704256.

[202] Matsumoto Y, Shigesada K, Hirano M, and Imai M (1986). Autogenous regulation of the gene for transcription termination factor *rho* in *Escherichia coli*: Localization and function of its attenuators. *J Bacteriol* 166.3, pp. 945–958. PMID: 2423505.

[203] Mattick JS and Gagen MJ (2001). The evolution of controlled multitasked gene networks: The role of introns and other noncoding RNAs in the development of complex organisms. *Mol Biol Evol* 18.9, pp. 1611–1630. PMID: 11504843.

[204] McBroom AJ and Kuehn MJ (2007). Release of outer membrane vesicles by Gram-negative bacteria is a novel envelope stress response. *Mol Microbiol* 63.2, pp. 545–558. PMID: 17163978.

[205] McCullen CA, Benhammou JN, Majdalani N, and Gottesman S (2010). Mechanism of positive regulation by DsrA and RprA small noncoding RNAs: Pairing increases translation and protects *rpoS* mRNA from degradation. *J Bacteriol* 192.21, pp. 5559–5571. PMID: 20802038.

[206] Mehta P, Goyal S, and Wingreen NS (2008). A quantitative comparison of sRNA-based and protein-based gene regulation. *Mol Syst Biol* 4, p. 221. PMID: 18854820.

[207] Melamed S, Adams PP, Zhang A, Zhang H, and Storz G (2020). RNA-RNA Interactomes of ProQ and Hfq Reveal Overlapping and Competing Roles. *Mol Cell* 77.2, 411–425.e7. PMID: 31761494.

[208] Melamed S, Peer A, Faigenbaum-Romm R, Gatt YE, Reiss N, Bar A, Altuvia Y, Argaman L, and Margalit H (2016). Global Mapping of Small RNA-Target Interactions in Bacteria. *Mol Cell* 63.5, pp. 884–897. PMID: 27588604.

[209] Melior H, Li S, Madhugiri R, Stötzel M, Azarderakhsh S, Barth-Weber S, Baumgardt K, Ziebuhr J, and Evguenieva-Hackenberg E (2019). Transcription attenuation-derived small RNA rnTrpL regulates tryptophan biosynthesis gene expression in *trans*. *Nucleic Acids Res* 47.12, pp. 6396–6410. PMID: 30993322.

[210] Mey AR, Craig SA, and Payne SM (2005). Characterization of *Vibrio cholerae* RyhB: The RyhB regulon and role of RyhB in biofilm formation. *Infect Immun* 73.9, pp. 5706–5719. PMID: 16113288.

[211] Mi H, Muruganujan A, Ebert D, Huang X, and Thomas PD (2019). PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res* 47.D1, pp. D419–D426. PMID: 30407594.

[212] Mitchell AM and Silhavy TJ (2019). Envelope stress responses: balancing damage repair and toxicity. *Nat Rev Microbiol* 17.7, pp. 417–428. PMID: 31150012.

[213] Mitra P, Ghosh G, Hafeezunnisa M, and Sen R (2017). Rho Protein: Roles and Mechanisms. *Annu Rev Microbiol* 71.1, pp. 687–709. PMID: 28731845.

[214] Miyakoshi M, Chao Y, and Vogel J (2015a). Cross talk between ABC transporter mRNAs via a target mRNA-derived sponge of the GcvB small RNA. *EMBO J* 34.11, pp. 1478–1492. PMID: 25630703.

[215] Miyakoshi M, Chao Y, and Vogel J (2015b). Regulatory small RNAs from the 3' regions of bacterial mRNAs. *Curr Opin Microbiol* 24, pp. 132–139. PMID: 25677420.

[216] 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 Res* 47.4, pp. 2075–2088. PMID: 30541135.

[217] Mizuno T, Chou MY, and Inouye M (1984). A unique mechanism regulating gene expression: Translational inhibition by a complementary RNA transcript (micRNA). *Proc Natl Acad Sci USA* 81.7 I, pp. 1966–1970. PMID: 6201848.

[218] Modi SR, Camacho DM, Kohanski MA, Walker GC, and Collins JJ (2011). Functional characterization of bacterial sRNAs using a network biology approach. *Proc Natl Acad Sci USA* 108.37, pp. 15522–15527. PMID: 21876160.

[219] Mohanraju P, Makarova KS, Zetsche B, Zhang F, Koonin EV, and Van Der Oost J (2016). Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. *Science* 353.6299, aad5147. PMID: 27493190.

[220] Moll I, Afonyushkin T, Vytvytska O, Kaberdin VR, and Bläsi U (2003). Coincident Hfq binding and RNase E cleavage sites on mRNA and small regulatory RNAs. *RNA* 9.11, pp. 1308–1314. PMID: 14561880.

[221] Møller T, Franch T, Udesen C, Gerdes K, and Valentin-Hansen P (2002). Spot 42 RNA mediates discoordinate expression of the *Escherichia coli* galactose operon. *Genes Dev* 16.13, pp. 1696–1706. PMID: 12101127.

[222] Moon K and Gottesman S (2009). A PhoQ/P-regulated small RNA regulates sensitivity of *Escherichia coli* to antimicrobial peptides. *Mol Microbiol* 74.6, pp. 1314–1330. PMID: 19889087.

[223] Moore SD and Sauer RT (2007). The tmRNA System for Translational Surveillance and Ribosome Rescue. *Annu Rev Biochem* 76.1, pp. 101–124. PMID: 17291191.

[224] Morfeldt E, Taylor D, Gabain A von, and Arvidson S (1995). Activation of alpha-toxin translation in *Staphylococcus aureus* by the *trans*-encoded antisense RNA, RNAIII. *EMBO J* 14.18, pp. 4569–4577. PMID: 7556100.

[225] Morita T, Maki K, and Aiba H (2005). RNase E-based ribonucleoprotein complexes: Mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs. *Genes Dev* 19.18, pp. 2176–2186. PMID: 16166379.

[226] Morita T, Mochizuki Y, and Aiba H (2006). Translational repression is sufficient for gene silencing by bacterial small noncoding RNAs in the absence of mRNA destruction. *Proc Natl Acad Sci USA* 103.13, pp. 4858–4863. PMID: 16549791.

[227] Morita T, Nishino R, and Aiba H (2017). Role of the terminator hairpin in the biogenesis of functional Hfq-binding sRNAs. *RNA* 23.9, pp. 1419–1431. PMID: 28606943.

[228] Morita T, Ueda M, Kubo K, and Aiba H (2015). Insights into transcription termination of Hfq-binding sRNAs of *Escherichia coli* and characterization of readthrough products. *RNA* 21.8, pp. 1490–1501. PMID: 26106215.

[229] Mückstein U, Tafer H, Hackermüller J, Bernhart SH, Stadler PF, and Hofacker IL (2006). Thermodynamics of RNA-RNA binding. *Bioinformatics* 22.10, pp. 1177–82. PMID: 16446276.

[230] Mutalik VK, Nonaka G, Ades SE, Rhodius VA, and Gross CA (2009). Promoter strength properties of the complete σ^{E} regulon of *Escherichia coli* and *Salmonella enterica*. *J Bacteriol* 191.23, pp. 7279–7287. PMID: 19783623.

[231] Na D, Yoo SM, Chung H, Park H, Park JH, and Lee SY (2013). Metabolic engineering of *Escherichia coli* using synthetic small regulatory RNAs. *Nat Biotechnol* 31.2, pp. 170–174. PMID: 23334451.

[232] Nadiras C, Eveno E, Schwartz A, Figueroa-Bossi N, and Boudvillain M (2018). A multivariate prediction model for Rho-dependent termination of transcription. *Nucleic Acids Res* 46.16, pp. 8245–8260. PMID: 29931073.

[233] Naville M, Ghuillot-Gaudeffroy A, Marchais A, and Gautheret D (2011). ARNold: A web tool for the prediction of Rho-independent transcription terminators. *RNA Biol* 8.1, pp. 11–13. PMID: 21282983.

[234] Nikaido H (2003). Molecular Basis of Bacterial Outer Membrane Permeability Revisited. *Microbiol Mol Biol Rev* 67.4, pp. 593–656. PMID: 14665678.

[235] Nitzan M, Fechter P, Peer A, Altuvia Y, Bronesky D, Vandenesch F, Romby P, Biham O, and Margalit H (2015). A defense-offense multi-layered regulatory switch in a pathogenic bacterium. *Nucleic Acids Res* 43.3, pp. 1357–69. PMID: 25628364.

[236] Nitzan M, Rehani R, and Margalit H (2017). Integration of Bacterial Small RNAs in Regulatory Networks. *Annu Rev Biophys* 46.1, pp. 131–148. PMID: 28532217.

[237] Novick R, Ross H, Projan SJ, Kornblum J, Kreiswirth B, and Moghazeh S (1993). Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule. *EMBO J* 12.10, pp. 3967–3975. PMID: 7691599.

[238] Oglesby-Sherrouse AG and Murphy ER (2013). Iron-responsive bacterial small RNAs: Variations on a theme. *Metallomics* 5.4, pp. 276–286. PMID: 23340911.

[239] Olejniczak M and Storz G (2017). ProQ/FinO-domain proteins: another ubiquitous family of RNA matchmakers? *Mol Microbiol* 104.6, pp. 905–915. PMID: 28370625.

[240] Opdyke JA, Fozo EM, Hemm MR, and Storz G (2011). RNase III participates in GadY-dependent cleavage of the *gadX-gadW* mRNA. *J Mol Biol* 406.1, pp. 29–43. PMID: 21147125.

[241] Orengo CA, Jones DT, and Thornton JM (1994). Protein superfamilies and domain superfolds. *Nature* 372.6507, pp. 631–634. PMID: 7990952.

[242] Otaka H, Ishikawa H, Morita T, and Aiba H (2011). PolyU tail of Rho-independent terminator of bacterial small RNAs is essential for Hfq action. *Proc Natl Acad Sci USA* 108.32, pp. 13059–64. PMID: 21788484.

[243] Overgaard M, Johansen J, Møller-Jensen J, and Valentin-Hansen P (2009). Switching off small RNA regulation with trap-mRNA. *Mol Microbiol* 73.5, pp. 790–800. PMID: 19682266.

[244] Ow MC and Kushner SR (2002). Initiation of tRNA maturation by RNase E is essential for cell viability in *Escherichia coli*. *Genes Dev* 16.9, pp. 1102–1115. PMID: 12000793.

[245] 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* 12.5, pp. 509–513. PMID: 25760244.

[246] Panja S, Schu DJ, and Woodson SA (2013). Conserved arginines on the rim of Hfq catalyze base pair formation and exchange. *Nucleic Acids Res* 41.15, pp. 7536–7546. PMID: 23771143.

[247] Panja S and Woodson SA (2012). Hfq proximity and orientation controls RNA annealing. *Nucleic Acids Res* 40.17, pp. 8690–8697. PMID: 22761405.

[248] Papenfort K and Bassler BL (2016). Quorum sensing signal-response systems in Gram-negative bacteria. *Nat Rev Microbiol* 14.9, pp. 576–588. PMID: 27510864.

[249] Papenfort K, Bouvier M, Mika F, Sharma CM, and Vogel J (2010). Evidence for an autonomous 5' target recognition domain in an Hfq-associated small RNA. *Proc Natl Acad Sci USA* 107.47, pp. 20435–20440. PMID: 21059903.

[250] Papenfort K, Espinosa E, Casadesús J, and Vogel J (2015). Small RNA-based feedforward loop with AND-gate logic regulates extrachromosomal DNA transfer in *Salmonella*. *Proc Natl Acad Sci USA* 112.34, E4772–81. PMID: 26307765.

[251] Papenfort K, Förstner KU, Cong JP, Sharma CM, and Bassler BL (2015). Differential RNA-seq of *Vibrio cholerae* identifies the VqmR small RNA as a regulator of biofilm formation. *Proc Natl Acad Sci USA* 112.7, E766–E775. PMID: 25646441.

[252] Papenfort K, Pfeiffer V, Mika F, Lucchini S, Hinton JC, and Vogel J (2006). σ^{E}-dependent small RNAs of *Salmonella* respond to membrane stress by accelerating global *omp* mRNA decay. *Mol Microbiol* 62.6, pp. 1674–1688. PMID: 17427289.

[253] Papenfort K, Podkaminski D, Hinton JC, and Vogel J (2012). The ancestral SgrS RNA discriminates horizontally acquired *Salmonella* mRNAs through a single G-U wobble pair. *Proc Natl Acad Sci USA* 109.13, E757–64. PMID: 22383560.

[254] Papenfort K, Said N, Welsink T, Lucchini S, Hinton JC, and Vogel J (2009). Specific and pleiotropic patterns of mRNA regulation by ArcZ, a conserved, Hfq-dependent small RNA. *Mol Microbiol* 74.1, pp. 139–158. PMID: 19732340.

[255] Papenfort K, Sun Y, Miyakoshi M, Vanderpool CK, and Vogel J (2013). Small RNA-mediated activation of sugar phosphatase mRNA regulates glucose homeostasis. *Cell* 153.2, pp. 426–437. PMID: 23582330.

[256] Papenfort K and Vogel J (2009). Multiple target regulation by small noncoding RNAs rewires gene expression at the post-transcriptional level. *Res Microbiol* 160.4, pp. 278–287. PMID: 19366629.

[257] Peer A and Margalit H (2011). Accessibility and evolutionary conservation mark bacterial small-RNA target-binding regions. *J Bacteriol* 193.7, pp. 1690–1701. PMID: 21278294.

[258] Peer A and Margalit H (2014). Evolutionary patterns of *Escherichia coli* small RNAs and their regulatory interactions. *RNA* 20.7, pp. 994–1003. PMID: 24865611.

[259] Peng Y, Soper TJ, and Woodson SA (2014). Positional effects of AAN motifs in *rpoS* regulation by sRNAs and Hfq. *J Mol Biol* 426.2, pp. 275–285. PMID: 24051417.

[260] Pérez-Reytor D, Plaza N, Espejo RT, Navarrete P, Bastías R, and Garcia K (2017). Role of non-coding regulatory RNA in the virulence of human pathogenic *Vibrios*. *Front Microbiol* 7, p. 2160. PMID: 28123382.

[261] Peschek N, Hoyos M, Herzog R, Förstner KU, and Papenfort K (2019). A conserved RNA seed-pairing domain directs small RNA-mediated stress resistance in enterobacteria. *EMBO J* 38.16, e101650. PMID: 31313835.

[262] Peterman N, Lavi-Itzkovitz A, and Levine E (2014). Large-scale mapping of sequence-function relations in small regulatory RNAs reveals plasticity and modularity. *Nucleic Acids Res* 42.19, pp. 12177–88. PMID: 25262352.

[263] Peters G, Coussement P, Maertens J, Lammertyn J, and De Mey M (2015). Putting RNA to work: Translating RNA fundamentals into biotechnological engineering practice. *Biotechnol Adv* 33.8, pp. 1829–1844. PMID: 26514597.

[264] Peters JM, Mooney RA, Grass JA, Jessen ED, Tran F, and Landick R (2012). Rho and NusG suppress pervasive antisense transcription in *Escherichia coli*. *Genes Dev* 26.23, pp. 2621–2633. PMID: 23207917.

[265] Pfeiffer V, Papenfort K, Lucchini S, Hinton JC, and Vogel J (2009). Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. *Nat Struct Mol Biol* 16.8, pp. 840–846. PMID: 19620966.

[266] Plumbridge J, Bossi L, Oberto J, Wade JT, and Figueroa-Bossi N (2014). Interplay of transcriptional and small RNA-dependent control mechanisms regulates chitosugar uptake in *Escherichia coli* and *Salmonella*. *Mol Microbiol* 92.4, pp. 648–658. PMID: 24593230.

[267] Prasse D and Schmitz RA (2018). Small RNAs Involved in Regulation of Nitrogen Metabolism. *Microbiol Spectr* 6.4, RWR-0018–2018. PMID: 30027888.

[268] Prévost K, Desnoyers G, Jacques JF, Lavoie F, and Massé E (2011). Small RNA-induced mRNA degradation achieved through both translation block and activated cleavage. *Genes Dev* 25.4, pp. 385–396. PMID: 21289064.

[269] Prévost K, Salvail H, Desnoyers G, Jacques JF, Phaneuf É, and Massé E (2007). The small RNA RyhB activates the translation of *shiA* mRNA encoding a permease of shikimate, a compound involved in siderophore synthesis. *Mol Microbiol* 64.5, pp. 1260–1273. PMID: 17542919.

[270] Pulvermacher SC, Stauffer LT, and Stauffer GV (2008). The role of the small regulatory RNA GcvB in GcvB/mRNA posttranscriptional regulation of *oppA* and *dppA* in *Escherichia coli*. *FEMS Microbiol Lett* 281.1, pp. 42–50. PMID: 18312576.

[271] Pulvermacher SC, Stauffer LT, and Stauffer GV (2009). Role of the sRNA GcvB in regulation of *cycA* in *Escherichia coli*. *Microbiology* 155.1, pp. 106–114. PMID: 19118351.

[272] Qi LS and Arkin AP (2014). A versatile framework for microbial engineering using synthetic non-coding RNAs. *Nat Rev Microbiol* 12.5, pp. 341–354. PMID: 24736794.

[273] Quax TE, Wolf YI, Koehorst JJ, Wurtzel O, Van der Oost R, Ran W, Blombach F, Makarova KS, Brouns SJ, Forster AC, Wagner EGH, Sorek R, Koonin EV, and Van der Oost J (2013). Differential translation tunes uneven production of operon-encoded proteins. *Cell Rep* 4.5, pp. 938–944. PMID: 24012761.

[274] Raghavan R, Sloan DB, and Ochman H (2012). Antisense transcription is pervasive but rarely conserved in enteric bacteria. *MBio* 3.4, e00156–12. PMID: 22872780.

[275] Ramirez-Peña E, Treviño J, Liu Z, Perez N, and Sumby P (2010). The group A *Streptococcus* small regulatory RNA FasX enhances streptokinase activity by increasing the stability of the *ska* mRNA transcript. *Mol Microbiol* 78.6, pp. 1332–1347. PMID: 21143309.

[276] Rasmussen AA, Eriksen M, Gilany K, Udesen C, Franch T, Petersen C, and Valentin-Hansen P (2005). Regulation of *ompA* mRNA stability: The role of a small regulatory RNA in growth phase-dependent control. *Mol Microbiol* 58.5, pp. 1421–1429. PMID: 16313626.

[277] Rehmsmeier M, Steffen P, Höchsmann M, and Giegerich R (2004). Fast and effective prediction of microRNA/target duplexes. *RNA* 10.10, pp. 1507–1517. PMID: 15383676.

[278] Reitzer L (2005). Catabolism of Amino Acids and Related Compounds. *EcoSal Plus* 1.2. PMID: 26443507.

[279] Rhodius VA, Suh WC, Nonaka G, West J, and Gross CA (2006). Conserved and variable functions of the σ^{E} stress response in related genomes. *PLoS Biol* 4.1, pp. 0043–0059. PMID: 16336047.

[280] Rice JB and Vanderpool CK (2011). The small RNA SgrS controls sugar-phosphate accumulation by regulating multiple PTS genes. *Nucleic Acids Res* 39.9, pp. 3806–19. PMID: 21245045.

[281] Rice JB, Balasubramanian D, and Vanderpool CK (2012). Small RNA binding-site multiplicity involved in translational regulation of a polycistronic mRNA. *Proc Natl Acad Sci USA* 109.40, E2691–E2698. PMID: 22988087.

[282] Richards J and Belasco JG (2019). Obstacles to Scanning by RNase E Govern Bacterial mRNA Lifetimes by Hindering Access to Distal Cleavage Sites. *Mol Cell* 74.2, 284–295.e5. PMID: 30852060.

[283] Richter AS and Backofen R (2012). Accessibility and conservation: General features of bacterial small RNA-mRNA interactions? *RNA Biol* 9.7, pp. 954–965. PMID: 22767260.

[284] Rivas E, Klein RJ, Jones TA, and Eddy SR (2001). Computational identification of noncoding RNAs in *Escherichia coli* by comparative genomics. *Curr Biol* 11.17, pp. 1369–1373. PMID: 11553332.

[285] Roberts JW (2019). Mechanisms of Bacterial Transcription Termination. *J Mol Biol* 431.20, pp. 4030–4039. PMID: 30978344.

[286] Romeo T, Vakulskas CA, and Babitzke P (2013). Post-transcriptional regulation on a global scale: Form and function of Csr/Rsm systems. *Environ Microbiol* 15.2, pp. 313–324. PMID: 22672726.

[287] Romsang A, Duang-Nkern J, Leesukon P, Saninjuk K, Vattanaviboon P, and Mongkolsuk S (2014). The Iron-Sulphur Cluster Biosynthesis Regulator IscR Contributes to Iron Homeostasis and Resistance to Oxidants in *Pseudomonas aeruginosa*. *PLoS One* 9.1, e86763. PMID: 24466226.

[288] Rosenfeld N, Elowitz MB, and Alon U (2002). Negative autoregulation speeds the response times of transcription networks. *J Mol Biol* 323.5, pp. 785–793. PMID: 12417193.

[289] Rutherford ST, Valastyan JS, Taillefumier T, Wingreen NS, and Bassler BL (2015). Comprehensive analysis reveals how single nucleotides contribute to noncoding RNA function in bacterial quorum sensing. *Proc Natl Acad Sci USA* 112.44, E6038–47. PMID: 26483489.

[290] Rutherford ST, Van Kessel JC, Shao Y, and Bassler BL (2011). AphA and LuxR/HapR reciprocally control quorum sensing in *Vibrios*. *Genes Dev* 25.4, pp. 397–408. PMID: 21325136.

[291] Salvail H, Lanthier-Bourbonnais P, Sobota JM, Caza M, Benjamin JAM, Mendieta MES, Lépine F, Dozois CM, Imlay J, and Massé E (2010). A small RNA promotes siderophore production through transcriptional and metabolic remodeling. *Proc Natl Acad Sci USA* 107.34, pp. 15223–15228. PMID: 20696910.

[292] Salvail H and Massé E (2012). Regulating iron storage and metabolism with RNA: An overview of posttranscriptional controls of intracellular iron homeostasis. *Wiley Interdiscip Rev RNA* 3.1, pp. 26–36. PMID: 21793218.

[293] Santiago-Frangos A and Woodson SA (2018). Hfq chaperone brings speed dating to bacterial sRNA. *Wiley Interdiscip Rev RNA* 9.4, e1475. PMID: 29633565.

[294] Santos RF dos, Arraiano CM, and Andrade JM (2019). New molecular interactions broaden the functions of the RNA chaperone Hfq. *Curr Genet* 65.6, pp. 1313–1319. PMID: 31104083.

[295] Saramago M, Bárria C, Santos RF dos, Silva IJ, Pobre V, Domingues S, Andrade JM, Viegas SC, and Arraiano CM (2014). The role of RNases in the regulation of small RNAs. *Curr Opin Microbiol* 18, pp. 105–115. PMID: 24704578.

[296] Sauer E, Schmidt S, and Weichenrieder O (2012). Small RNA binding to the lateral surface of Hfq hexamers and structural rearrangements upon mRNA target recognition. *Proc Natl Acad Sci USA* 109.24, pp. 9396–9401. PMID: 22645344.

[297] Sauer E and Weichenrieder O (2011). Structural basis for RNA 3'-end recognition by Hfq. *Proc Natl Acad Sci USA* 108.32, pp. 13065–70. PMID: 21737752.

[298] Schmiedel JM, Axmann IM, and Legewie S (2012). Multi-Target Regulation by Small RNAs Synchronizes Gene Expression Thresholds and May Enhance Ultrasensitive Behavior. *PLoS One* 7.8, e42296. PMID: 22927924.

[299] Schneider BL, Kiupakis AK, and Reitzer LJ (1998). Arginine catabolism and the arginine succinyltransferase pathway in *Escherichia coli*. *J Bacteriol* 180.16, pp. 4278–4286. PMID: 9696779.

[300] Schu DJ, Zhang A, Gottesman S, and Storz G (2015). Alternative Hfq-sRNA interaction modes dictate alternative mRNA recognition. *EMBO J* 34.20, pp. 2557–2573. PMID: 26373314.

[301] Sedlyarova N, Shamovsky I, Bharati BK, Epshtein V, Chen J, Gottesman S, Schroeder R, and Nudler E (2016). sRNA-Mediated Control of Transcription Termination in *Escherichia coli*. *Cell* 167.1, 111–121.e13. PMID: 27662085.

[302] Segura A, Molina L, Fillet S, Krell T, Bernal P, Muñoz-Rojas J, and Ramos JL (2012). Solvent tolerance in Gram-negative bacteria. *Curr Opin Biotechnol* 23.3, pp. 415–421. PMID: 22155018.

[303] Seo SW, Kim D, Latif H, O'Brien EJ, Szubin R, and Palsson BO (2014). Deciphering Fur transcriptional regulatory network highlights its complex role beyond iron metabolism in *Escherichia coli*. *Nat Commun* 5, p. 4910. PMID: 25222563.

[304] Shao Y, Feng L, Rutherford ST, Papenfort K, and Bassler BL (2013). Functional determinants of the quorum-sensing non-coding RNAs and their roles in target regulation. *EMBO J* 32.15, pp. 2158–2171. PMID: 23838640.

[305] Sharma CM, Darfeuille F, Plantinga TH, 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* 21.21, pp. 2804–2817. PMID: 17974919.

[306] Sharma CM, Hoffmann S, Darfeuille F, Reignier J, Findeiß S, Sittka A, Chabas S, Reiche K, Hackermüller J, Reinhardt R, Stadler PF, and Vogel J (2010). The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature* 464.7286, pp. 250–255. PMID: 20164839.

[307] Sharma CM, Papenfort K, Pernitzsch SR, Mollenkopf HJ, Hinton JC, and Vogel J (2011). Pervasive post-transcriptional control of genes involved in amino acid metabolism by the Hfq-dependent GcvB small RNA. *Mol Microbiol* 81.5, pp. 1144–1165. PMID: 21696468.

[308] Sharma CM and Vogel J (2009). Experimental approaches for the discovery and characterization of regulatory small RNA. *Curr Opin Microbiol* 12.5, pp. 536–546. PMID: 19758836.

[309] Sharma V, Sakai Y, Smythe KA, and Yokobayashi Y (2013). Knockdown of *recA* gene expression by artificial small RNAs in *Escherichia coli*. *Biochem Biophys Res Commun* 430.1, pp. 256–259. PMID: 23178465.

[310] Sharma V, Yamamura A, and Yokobayashi Y (2012). Engineering artificial small RNAs for Conditional gene silencing in *Escherichia coli*. *ACS Synth Biol* 1.1, pp. 6–13. PMID: 23651005.

[311] Shen-Orr SS, Milo R, Mangan S, and Alon U (2002). Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nat Genet* 31.1, pp. 64–68. PMID: 11967538.

[312] Shimoni Y, Friedlander G, Hetzroni G, Niv G, Altuvia S, Biham O, and Margalit H (2007). Regulation of gene expression by small non-coding RNAs: a quantitative view. *Mol Syst Biol* 3.1, p. 138. PMID: 17893699.

[313] Silhavy TJ, Kahne D, and Walker S (2010). The bacterial cell envelope. *Cold Spring Harb Perspect Biol* 2.5, a000414. PMID: 20452953.

[314] Silva IJ, Barahona S, Eyraud A, Lalaouna D, Figueroa-Bossi N, Massé E, and Arraiano CM (2019). SraL sRNA interaction regulates the terminator by preventing premature transcription termination of *rho* mRNA. *Proc Natl Acad Sci* 116.8, pp. 3042–3051. PMID: 30718400.

[315] Silva IJ, Ortega ÁD, Viegas SC, García-Del Portillo F, and Arraiano CM (2013). An RpoS-dependent sRNA regulates the expression of a chaperone involved in protein folding. *RNA* 19.9, pp. 1253–1265. PMID: 23893734.

[316] Silva IJ, Saramago M, Dressaire C, Domingues S, Viegas SC, and Arraiano CM (2011). Importance and key events of prokaryotic RNA decay: The ultimate fate of an RNA molecule. *Wiley Interdiscip Rev RNA* 2.6, pp. 818–836. PMID: 21976285.

[317] Sittka A, Lucchini S, Papenfort K, Sharma CM, Rolle K, Binnewies TT, Hinton JC, and Vogel J (2008). Deep Sequencing Analysis of Small Noncoding RNA and mRNA Targets of the Global Post-Transcriptional Regulator, Hfq. *PLoS Genet* 4.8, e1000163. PMID: 18725932.

[318] Skordalakes E, Brogan AP, Park BS, Kohn H, and Berger JM (2005). Structural mechanism of inhibition of the Rho transcription termination factor by the antibiotic bicyclomycin. *Structure* 13.1, pp. 99–109. PMID: 15642265.

[319] Slamti L and Lereclus D (2019). The oligopeptide ABC-importers are essential communication channels in Gram-positive bacteria. *Res Microbiol* 170.8, pp. 338–344. PMID: 31376485.

[320] Smirnov A, Förstner KU, 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 USA* 113.41, pp. 11591–11596. PMID: 27671629.

[321] Smirnov A, Wang C, Drewry LL, and Vogel J (2017). Molecular mechanism of mRNA repression in *trans* by a ProQ-dependent small RNA. *EMBO J* 36.8, pp. 1029–1045. PMID: 28336682.

[322] Sneppen K, Dodd IB, Shearwin KE, Palmer AC, Schubert RA, Callen BP, and Egan JB (2005). A mathematical model for transcriptional interference by RNA polymerase traffic in *Escherichia coli*. *J Mol Biol* 346.2, pp. 399–409. PMID: 15670592.

[323] Sobrero P and Valverde C (2012). The bacterial protein Hfq: Much more than a mere RNA-binding factor. *Crit Rev Microbiol* 38.4, pp. 276–299. PMID: 22435753.

[324] Song T, Mika F, Lindmark B, Liu Z, Schild S, Bishop A, Zhu J, Camilli A, Johansson J, Vogel J, and Wai SN (2008). A new *Vibrio cholerae* sRNA modulates colonization and affects release of outer membrane vesicles. *Mol Microbiol* 70.1, pp. 100–111. PMID: 18681937.

[325] Song T, Sabharwal D, Gurung JM, Cheng AT, Sjöström AE, Yildiz FH, Uhlin BE, and Wai SN (2014). *Vibrio cholerae* Utilizes Direct sRNA Regulation in Expression of a Biofilm Matrix Protein. *PLoS One* 9.7, e101280. PMID: 25054332.

[326] Sonnleitner E, Gonzalez N, Sorger-Domenigg T, Heeb S, Richter AS, Backofen R, Williams P, Hüttenhofer A, Haas D, and Bläsi U (2011). The small RNA PhrS stimulates synthesis of the *Pseudomonas aeruginosa* quinolone signal. *Mol Microbiol* 80.4, pp. 868–885. PMID: 21375594.

[327] Svenningsen SL, Tu KC, and Bassler BL (2009). Gene dosage compensation calibrates four regulatory RNAs to control *Vibrio cholerae* quorum sensing. *EMBO J* 28.4, pp. 429–439. PMID: 19165149.

[328] Tamura M, Moore CJ, and Cohen SN (2013). Nutrient dependence of RNase E essentiality in *Escherichia coli*. *J Bacteriol* 195.6, pp. 1133–41. PMID: 23275245.

[329] Tani TH, Khodursky AB, Blumenthal RM, Brown PO, and Matthews RG (2002). Adaptation to famine: A family of stationary-phase genes revealed by microarray analysis. *Proc Natl Acad Sci USA* 99.21, pp. 13471–13476. PMID: 12374860.

[330] Thieffry D, Huerta AM, Pérez-Rueda E, and Collado-Vides J (1998). From specific gene regulation to genomic networks: a global analysis of transcriptional regulation in *Escherichia coli*. *Bioessays* 20.5, pp. 433–40.

[331] Thomason MK, Bischler T, Eisenbart SK, Förstner KU, Zhang A, Herbig A, Nieselt K, Sharma CM, and Storz G (2015). Global transcriptional start site mapping using differential RNA sequencing reveals novel antisense RNAs in *Escherichia coli*. *J Bacteriol* 197.1, pp. 18–28. PMID: 25266388.

[332] Tu KC, Long T, Svenningsen SL, Wingreen NS, and Bassler BL (2010). Negative Feedback Loops Involving Small Regulatory RNAs Precisely Control the *Vibrio harveyi* Quorum-Sensing Response. *Mol Cell* 37.4, pp. 567–579. PMID: 20188674.

[333] Updegrove TB, Shabalina SA, and Storz G (2015). How do base-pairing small RNAs evolve? *FEMS Microbiol Rev* 39.3, pp. 379–91. PMID: 25934120.

[334] Updegrove TB, Zhang A, and Storz G (2016). Hfq: The flexible RNA matchmaker. *Curr Opin Microbiol*. PMID: 26907610.

[335] Urbanowski ML, Stauffer LT, and Stauffer GV (2000). The *gcvB* gene encodes a small untranslated RNA involved in expression of the dipeptide and oligopeptide transport systems in *Escherichia coli*. *Mol Microbiol* 37.4, pp. 856–868. PMID: 10972807.

[336] Vanderpool CK and Gottesman S (2004). Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system. *Mol Microbiol* 54.4, pp. 1076–1089. PMID: 15522088.

[337] Vargas-Blanco DA, Zhou Y, Gregory Zamalloa L, Antonelli T, and Shell SS (2019). mRNA degradation rates are coupled to metabolic status in *Mycobacterium smegmatis*. *MBio* 10.4, e00957–19. PMID: 31266866.

[338] Vecerek B, Moll I, and Bläsi U (2007). Control of Fur synthesis by the non-coding RNA RyhB and iron-responsive decoding. *EMBO J* 26.4, pp. 965–75. PMID: 17268550.

[339] Vincent HA, Henderson CA, Ragan TJ, Garza-Garcia A, Cary PD, Gowers DM, Malfois M, Driscoll PC, Sobott F, and Callaghan AJ (2012). Characterization of *Vibrio cholerae* Hfq provides novel insights into the role of the Hfq C-terminal region. *J Mol Biol* 420.1-2, pp. 56–69. PMID: 22484176.

[340] Vogel J, Bartels V, Tang TH, Churakov G, Slagter-Jäger JG, Hüttenhofer A, and Wagner EGH (2003). RNomics in *Escherichia coli* detects new sRNA species and indicates parallel transcriptional output in bacteria. *Nucleic Acids Res* 31.22, pp. 6435–6443. PMID: 14602901.

[341] Vogel J and Luisi BF (2011). Hfq and its constellation of RNA. *Nat Rev Microbiol* 9.8, pp. 578–589. PMID: 21760622.

[342] Wadler CS and Vanderpool CK (2007). A dual function for a bacterial small RNA: SgrS performs base pairing-dependent regulation and encodes a functional polypeptide. *Proc Natl Acad Sci USA* 104.51, pp. 20454–20459. PMID: 18042713.

[343] Wagner EGH and Romby P (2015). Small RNAs in Bacteria and Archaea: Who They Are, What They Do, and How They Do It. *Adv Genet* 90, pp. 133–208. PMID: 26296935.

[344] 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 Res* 48.4, pp. 2126–2143. PMID: 31863581.

[345] Wang X, Ji SC, Jeon HJ, Lee Y, and Lim HM (2015). Two-level inhibition of *galK* expression by Spot 42: Degradation of mRNA mK2 and enhanced transcription termination before the *galK* gene. *Proc Natl Acad Sci USA* 112.24, pp. 7581–7586. PMID: 26045496.

[346] Wassarman KM (2018). 6S RNA, a Global Regulator of Transcription. *Microbiol Spectr* 6.3, RWR-0019–2018. PMID: 29916345.

[347] Wassarman KM, Repoila F, Rosenow C, Storz G, and Gottesman S (2001). Identification of novel small RNAs using comparative genomics and microarrays. *Genes Dev* 15.13, pp. 1637–1651. PMID: 11445539.

[348] Waters LS and Storz G (2009). Regulatory RNAs in Bacteria. *Cell* 136.4, pp. 615–628. PMID: 19239884.

[349] Waters SA, McAteer SP, Kudla G, Pang I, Deshpande NP, Amos TG, Leong KW, Wilkins MR, Strugnell R, Gally DL, Tollervey D, and Tree JJ (2017). Small RNA interactome of pathogenic *Escherichia coli* revealed through crosslinking of RNase E. *EMBO J* 36.3, pp. 374–387. PMID: 27836995.

[350] Weber L, Thoelken C, Volk M, Remes B, Lechner M, and Klug G (2016). The conserved *dcw* gene cluster of *R. sphaeroides* is preceded by an uncommonly extended 5' leader featuring the sRNA UpsM. *PLoS One* 11.11. PMID: 27802301.

[351] Wright PR, Georg J, Mann M, Sorescu DA, Richter AS, Lott S, Kleinkauf R, Hess WR, and Backofen R (2014). CopraRNA and IntaRNA: predicting small RNA targets, networks and interaction domains. *Nucleic Acids Res* 42.Web Server issue, W119–23. PMID: 24838564.

[352] Wright PR, Richter AS, Papenfort K, Mann M, Vogel J, Hess WR, Backofen R, and Georg J (2013). Comparative genomics boosts target prediction for bacterial small RNAs. *Proc Natl Acad Sci USA* 110.37, E3487–96. PMID: 23980183.

[353] Yamamoto S, Izumiya H, Mitobe J, Morita M, Arakawa E, Ohnishi M, and Watanabe H (2011). Identification of a chitin-induced small RNA that regulates translation of the *tfoX* gene, encoding a positive regulator of natural competence in *Vibrio cholerae*. *J Bacteriol* 193.8, pp. 1953–1965. PMID: 21317321.

[354] Zhang A, Schu DJ, Tjaden BC, Storz G, and Gottesman S (2013). Mutations in interaction surfaces differentially impact *Escherichia coli* Hfq association with small RNAs and their mRNA targets. *J Mol Biol* 425.19, pp. 3678–3697. PMID: 23318956.

[355] Zhang D feng, Ye J zhou, Dai H hou, Lin X min, Li H, and Peng X xian (2018). Identification of ethanol tolerant outer membrane proteome reveals OmpC-dependent mechanism in a manner of EnvZ/OmpR regulation in *Escherichia coli*. *J Proteomics*. PMID: 29518576.

[356] Zuzanna W and Mikolaj O (2016). Hfq assists small RNAs in binding to the coding sequence of *ompD* mRNA and in rearranging its structure. *RNA* 22.7, pp. 979–994. PMID: 27154968.

[357] Zwiefka A, Kohn H, and Widger WR (1993). Transcription Termination Factor Rho: The Site of Bicyclomycin Inhibition in *Escherichia coli*. *Biochemistry* 32.14, pp. 3564–3570. PMID: 8466900.

6 Appendix

6.1 Figure supplements to chapter 2: "Gene autoregulation by 3' UTR-derived bacterial small RNAs"

RESEARCH ARTICLE

Figures and figure supplements

Gene autoregulation by 3′ UTR-derived bacterial small RNAs

Mona Hoyos *et al*

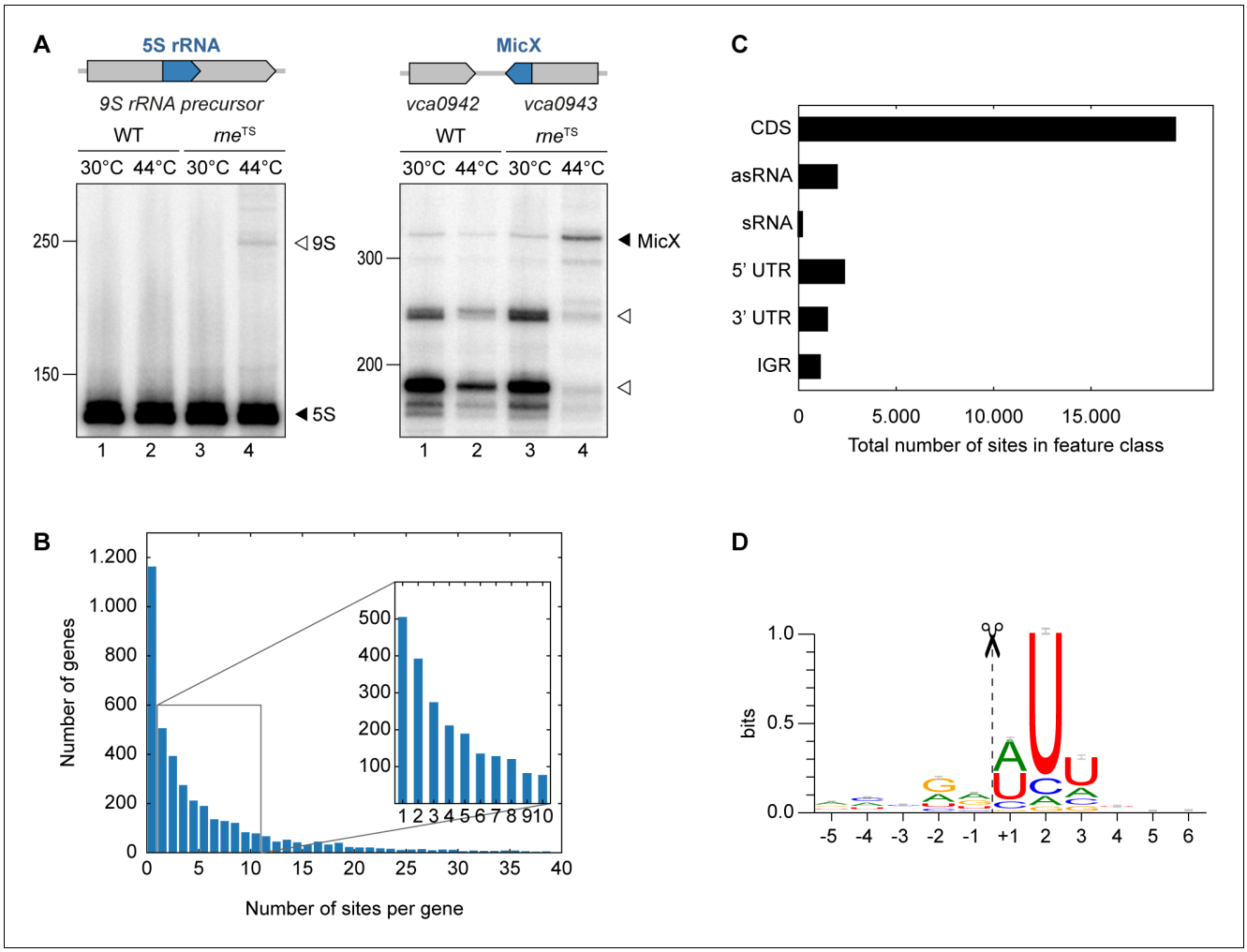


Figure 1. TIER-seq analysis of *V. cholerae*. (**A**) *V. cholerae* wild-type and *rne*^{TS} strains were grown at 30˚C to stationary phase (OD_{600} of 2.0). Cultures were divided in half and continuously grown at either 30˚C or 44˚C for 60 min. Cleavage patterns of 5S rRNA and 3' UTR-derived MicX were analyzed on Northern blots. Closed triangles indicate mature 5S or full-length MicX, open triangles indicate the 9S precursor or MicX processing products. (**B, C, D**) Biological triplicates of *V. cholerae* wild-type and *rne*^{TS} strains were grown at 30˚C to late exponential phase (OD_{600} of 1.0). Cultures were divided in half and continuously grown at either 30˚C or 44˚C for 60 min. Isolated RNA was subjected to RNA-seq and RNase E cleavage sites were determined as described in the materials and methods section. (**B**) Number of cleavage sites detected per gene. (**C**) Classification of RNase E sites by their genomic location. (**D**) The RNase E consensus motif based on all detected cleavage sites. The total height of the error bar is twice the small sample correction.

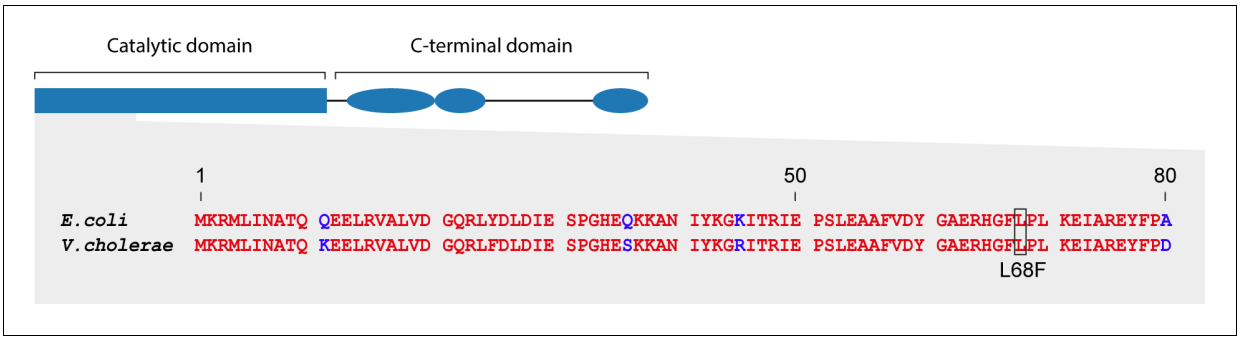


Figure 1—figure supplement 1. Conservation of RNase E between *E. coli* and *V. cholerae*. Sequence alignment of the first 80 N-terminal amino acids of RNase E from *E. coli* and *V. cholerae*. The temperature-sensitive *rne-3071* mutation changing a leucine to phenylalanine at position 68 is indicated.

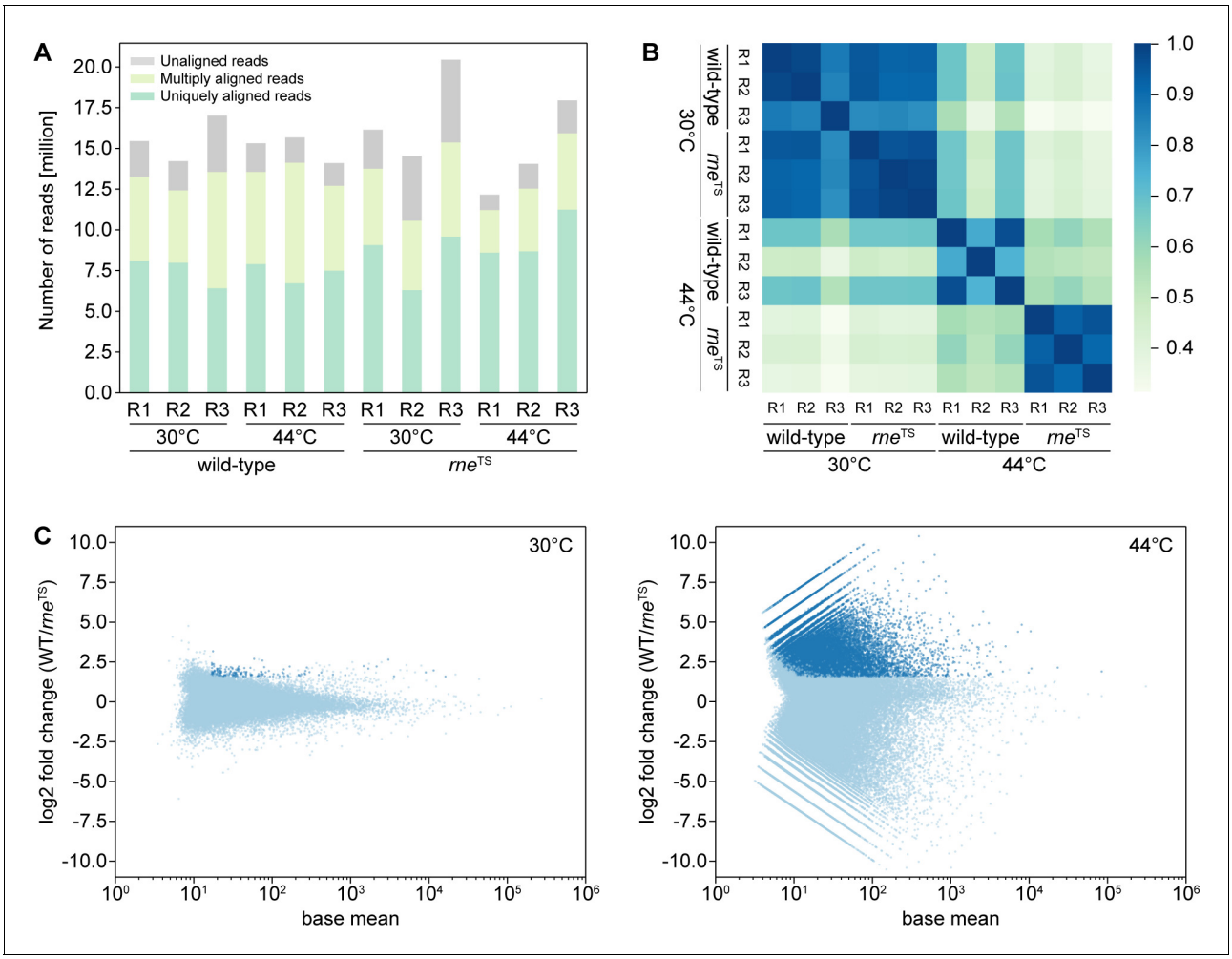


Figure 1—figure supplement 2. TIER-Seq read mapping statistics. TIER-seq was performed as described in ***Figure 1***. (**A**) Total number of raw cDNA reads obtained for all samples, showing the fractions of uniquely aligned reads (dark green), multiply aligned reads (light green) or unaligned reads (grey). R1-R3 indicate the biological triplicates. (**B**) Similarity of 5′ ends profiles of uniquely aligned reads, obtained by comparison of all detected 5′ end positions between the respective cDNA libraries. Colored rectangles show the Pearson correlation coefficient corresponding to the scale bar on the right. (**C**) Global analysis of 5′ profiles at the permissive (30˚C, left) and non-permissive temperature (44˚C, right). Plots show average coverage levels of 5′ read ends and the respective $\log_2$ fold change in wild-type samples compared to *rne*^{TS} samples. Candidate RNase E cleavage sites were determined as positions enriched ≥3 fold in the wild-type (p-value<0.05) and are shown in dark blue.

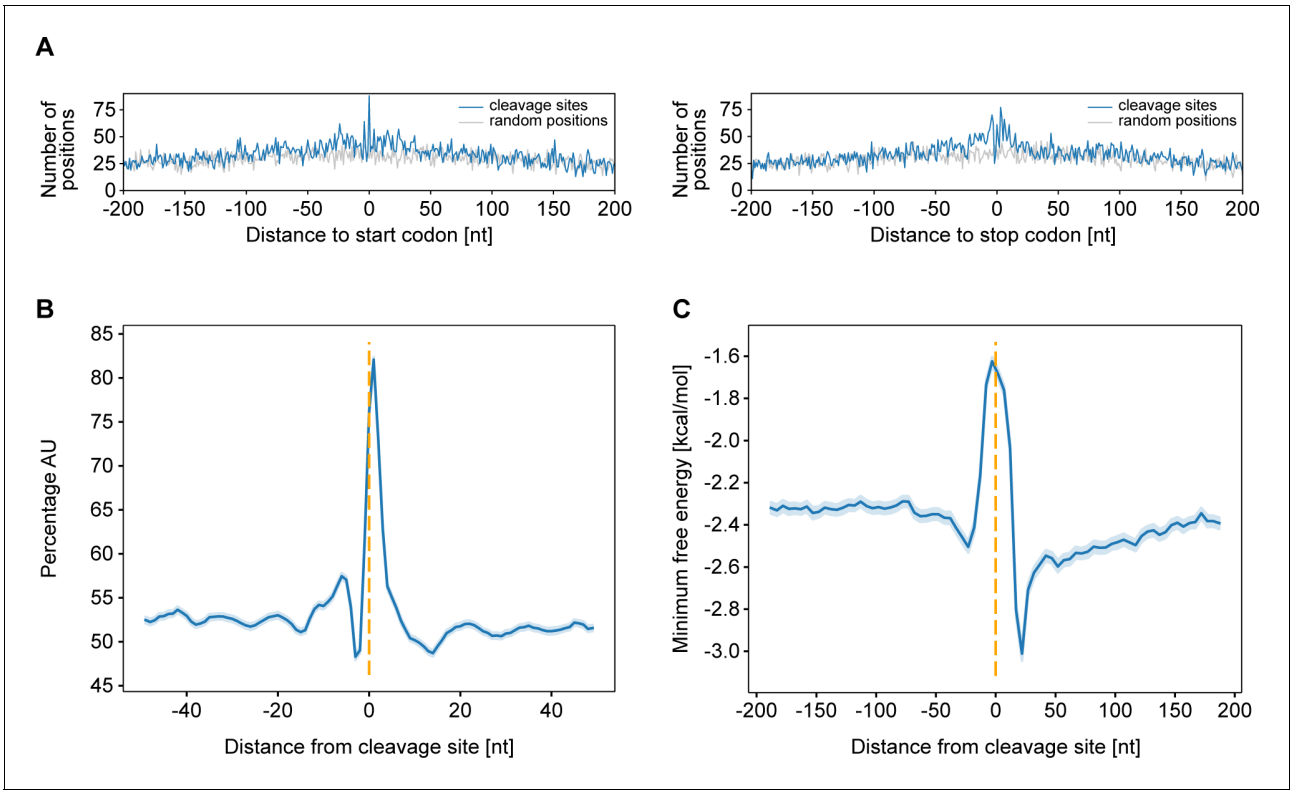


Figure 1—figure supplement 3. Position and characteristics of RNase E cleavage sites. TIER-seq was performed as described in ***Figure 1***. (**A**) Frequency of RNase E sites or the same number of randomly selected genome positions dependent on their relative position to start codons (left) and stop codons (right). (**B**) AU content around the RNase E cleavage sites. The 95% confidence interval is indicated in light blue. (**C**) Degree of RNA structure around RNase E cleavage sites. Minimal folding energy (MFE) was calculated in five nt steps for each 25 nt window. The 95% confidence interval is indicated in light blue.

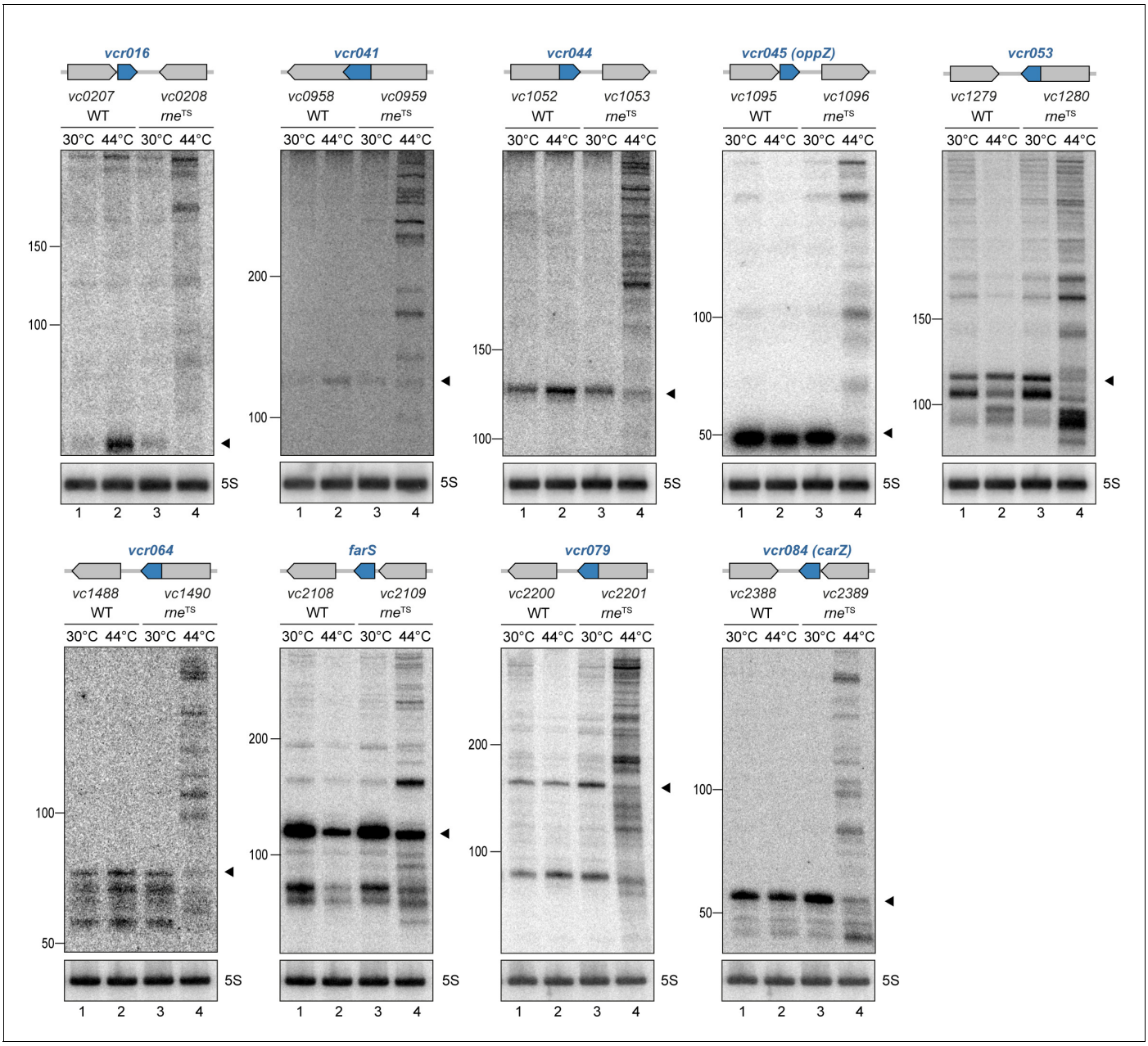


Figure 1—figure supplement 4. RNase E-mediated maturation of sRNAs from 3′ UTRs. *V. cholerae* wild-type and *rne*^{TS} strains were grown at 30˚C to stationary phase (OD_{600} of 2.0). Cultures were divided in half and continuously grown at either 30˚C or 44˚C for 30 min. Cleavage patterns of 3′ UTR-derived sRNAs were monitored on Northern blots. The genomic locations and relative orientations are shown above the gels. Genes are shown in gray; sRNAs are shown in blue. Filled triangles indicate the size of mature sRNAs. 5S rRNA served as loading control.

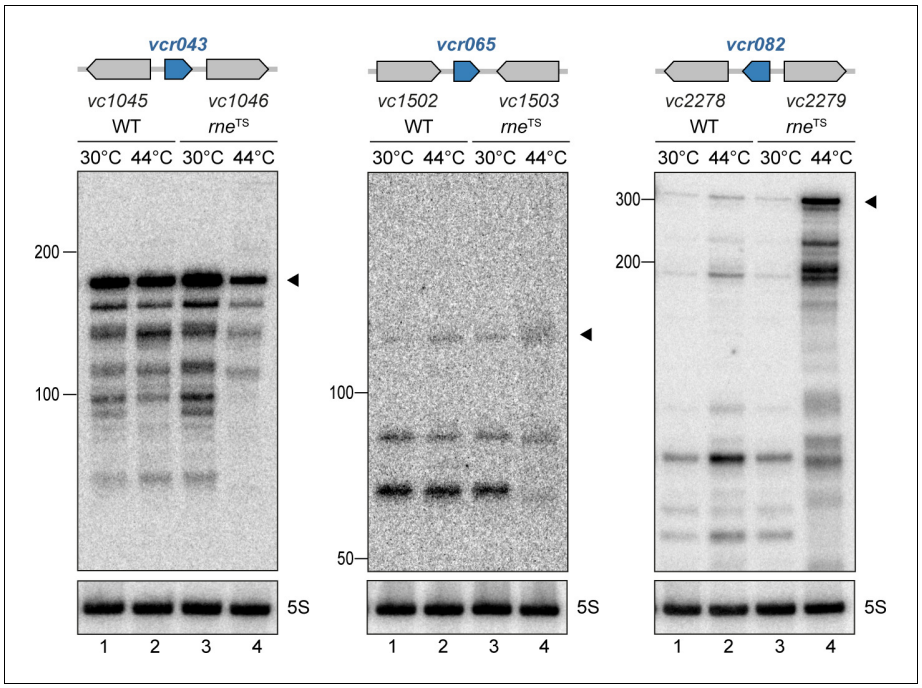


Figure 1—figure supplement 5. RNase E-mediated maturation of sRNAs from IGRs. *V. cholerae* wild-type and *rne*^{TS} strains were grown at 30˚C to stationary phase (OD_{600} of 2.0). Cultures were divided in half and continuously grown at either 30˚C or 44˚C for 30 min. Cleavage patterns of intergenic sRNAs were monitored on Northern blots. The genomic locations and relative orientations are shown above the gels. Genes are shown in gray; sRNAs are shown in blue. Filled triangles indicate the size of unprocessed sRNAs. 5S rRNA served as loading control.

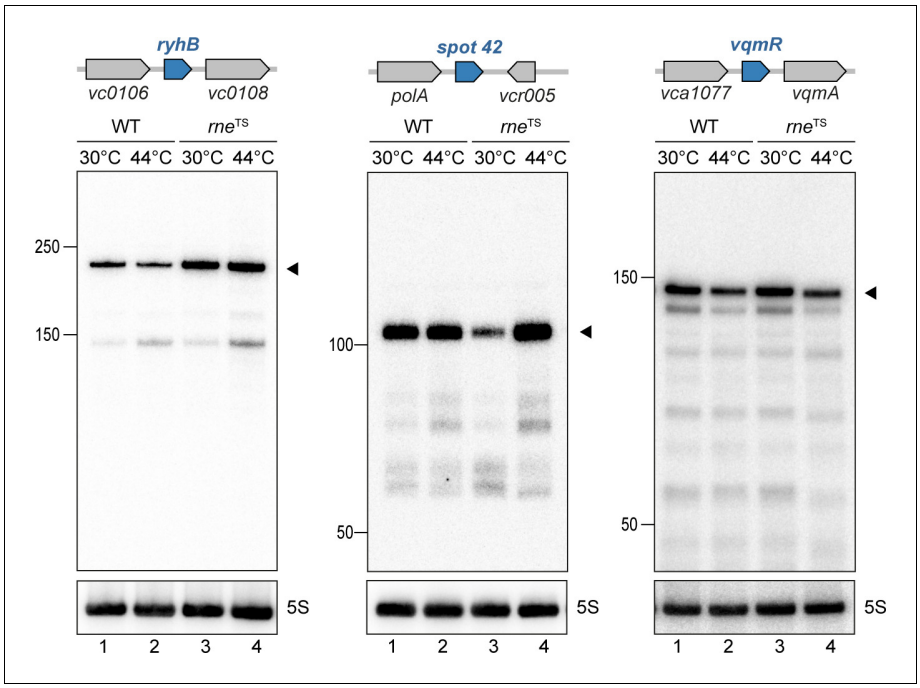


Figure 1—figure supplement 6. Expression of RNase E-independent sRNAs. *V. cholerae* wild-type and *rne*^{TS} strains were grown at 30˚C to early exponential phase (OD_{600} of 0.2; RyhB and Spot 42) or to stationary phase (OD_{600} of 2.0; VqmR). Cultures were divided in half and continuously grown at either 30˚C or 44˚C for 30 min. Cleavage patterns of sRNAs without detectable RNase E cleavage sites were monitored on Northern blots. The genomic locations and relative orientations are shown above the gels. Genes are shown in gray; sRNAs are shown in blue. Filled triangles indicate the size of full-length sRNAs. 5S rRNA served as loading control.

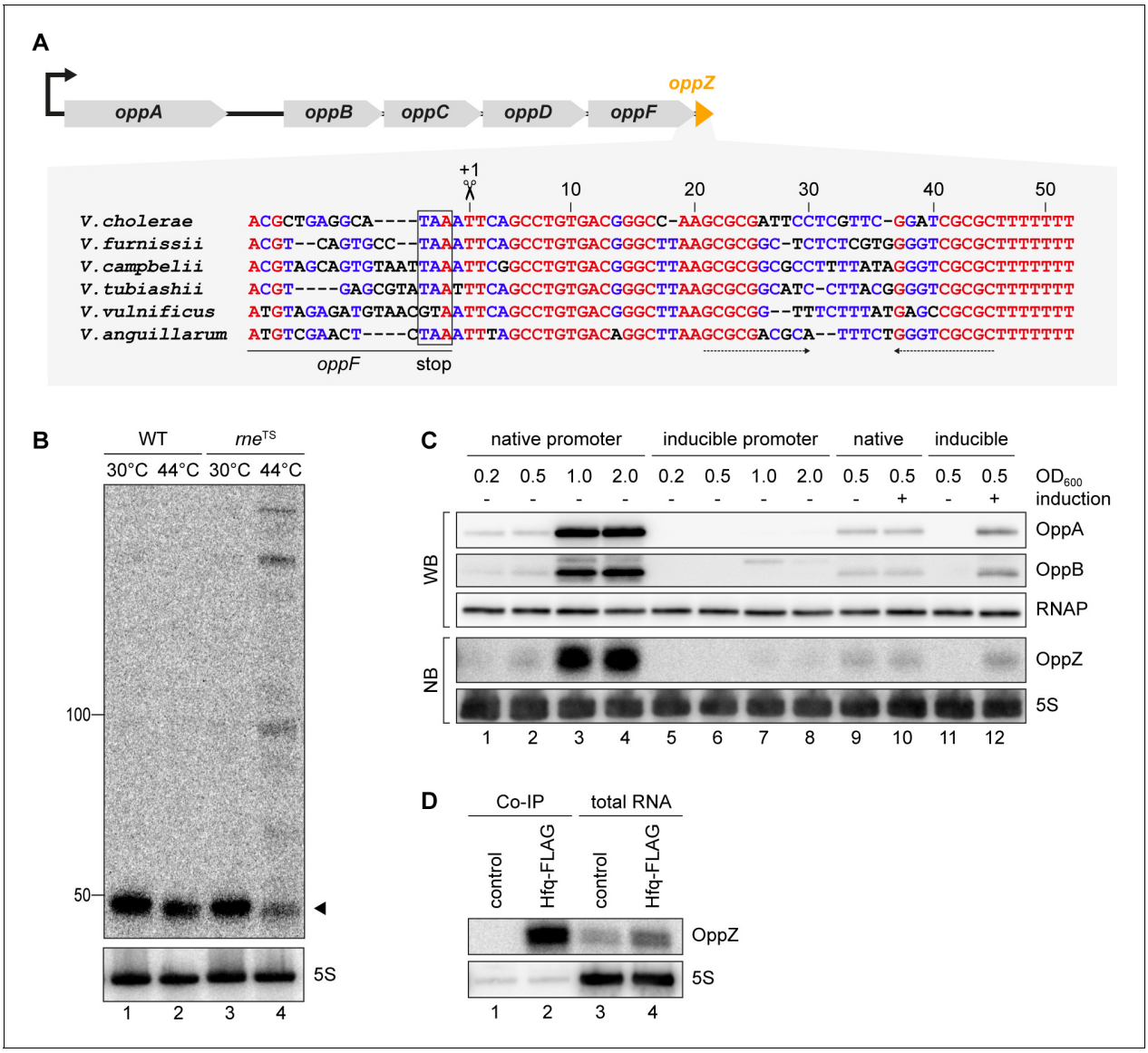


Figure 2. OppZ is produced from the *oppABCDF* 3′ end. (**A**) Top: Genomic organization of *oppABCDF* and *oppZ*. Bottom: Alignment of *oppZ* sequences, including the last codons of *oppF*, from various *Vibrio* species. The *oppF* stop codon, the RNase E cleavage site and the Rho-independent terminator are indicated. (**B**) *V. cholerae* wild-type and *rne*^{TS} strains were grown at 30˚C to stationary phase (OD_{600} of 2.0). Cultures were divided in half and continuously grown at either 30˚C or 44˚C for 30 min. OppZ synthesis was analyzed by Northern blot with 5S rRNA as loading control. The triangle indicates the size of mature OppZ. (**C**) Protein and RNA samples were obtained from *V. cholerae oppA*::3XFLAG *oppB*::3XFLAG strains carrying either the native *oppA* promoter or the inducible pBAD promoter upstream of *oppA*. Samples were collected at the indicated OD_{600} and tested for OppA and OppB production by Western blot and for OppZ expression by Northern blot. RNAP and 5S rRNA served as loading controls for Western and Northern blots, respectively. Lanes 1–8: Growth without L-arabinose. Lanes 9–12: Growth with either H_2O (-) or L-arabinose (+) (0.2% final conc.). (**D**) *V. cholerae* wild-type (control) and *hfq*::3XFLAG (Hfq-FLAG) strains were grown to stationary phase (OD_{600} of 2.0), lysed, and subjected to immunoprecipitation using the anti-FLAG antibody. RNA samples of lysate (total RNA) and co-immunoprecipitated fractions were analyzed on Northern blots. 5S rRNA served as loading control.

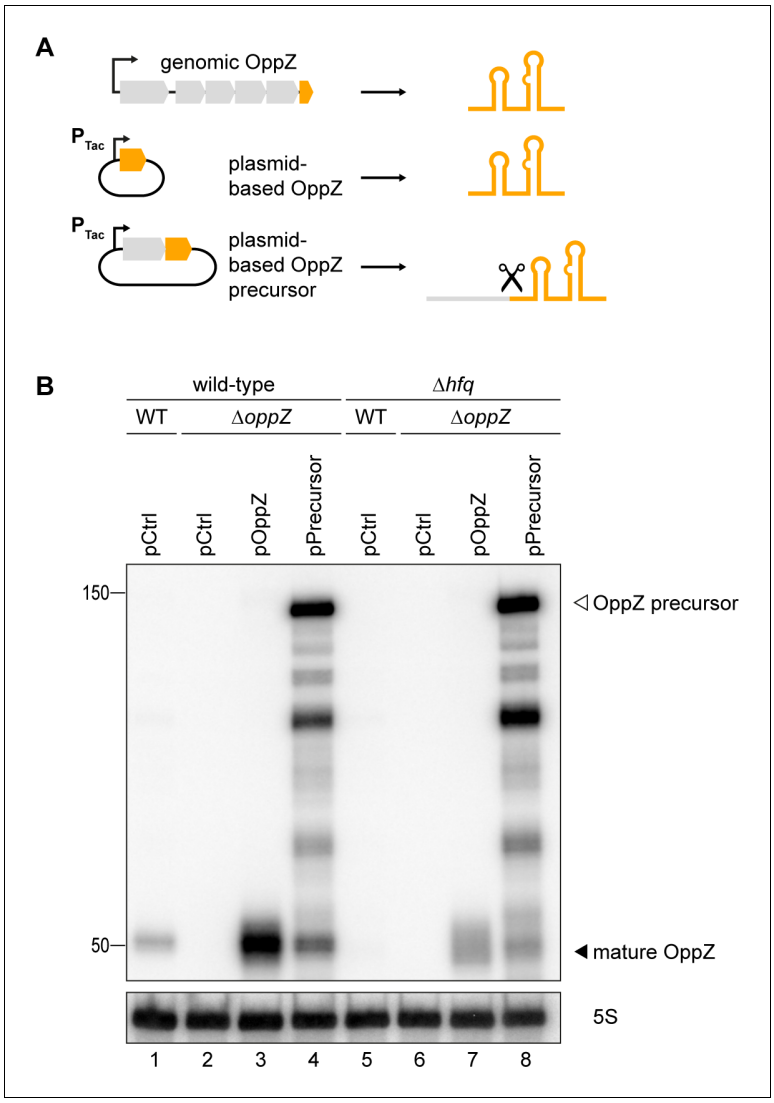


Figure 2—figure supplement 1. Hfq dependence of OppZ processing. (**A**) Schematic description of the analyzed OppZ variants. OppZ was produced natively from the genomic *opp* locus, expressed as mature sRNA from a plasmid (pOppZ) or cleaved from a plasmid-encoded precursor transcript including the 3' end of *oppF* (pPrecursor). Expression of both plasmid-based *oppZ* variants was driven by a constitutive promoter. (**B**) *V. cholerae* wild-type, Δ*oppZ*, Δ*hfq* or Δ*hfq* Δ*oppZ* strains carrying *oppA*::3XFLAG *oppB*::3XFLAG genes and a control plasmid or the indicated OppZ expression plasmid were grown to stationary phase (OD_{600} of 2.0). RNA samples were collected and OppZ processing was analyzed by Northern blot. 5S rRNA served as loading control.

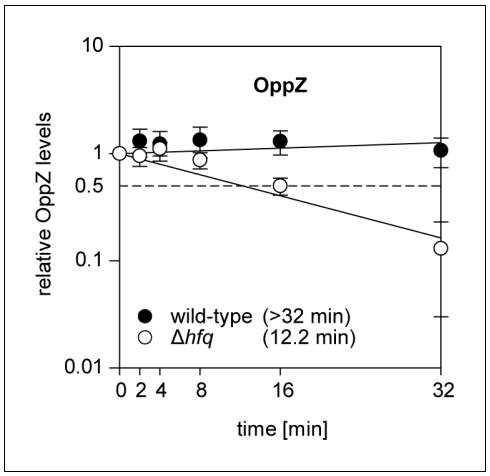


Figure 2—figure supplement 2. Hfq dependence of OppZ stability. *V. cholerae* wild-type and Δ*hfq* strains were grown to early stationary phase (OD_{600} of 1.5) and treated with rifampicin to terminate transcription. RNA samples were obtained at the indicated time points and OppZ transcript levels were monitored by Northern blot and normalized to 5S rRNA levels as loading control. Error bars represent the SD of three biological replicates.

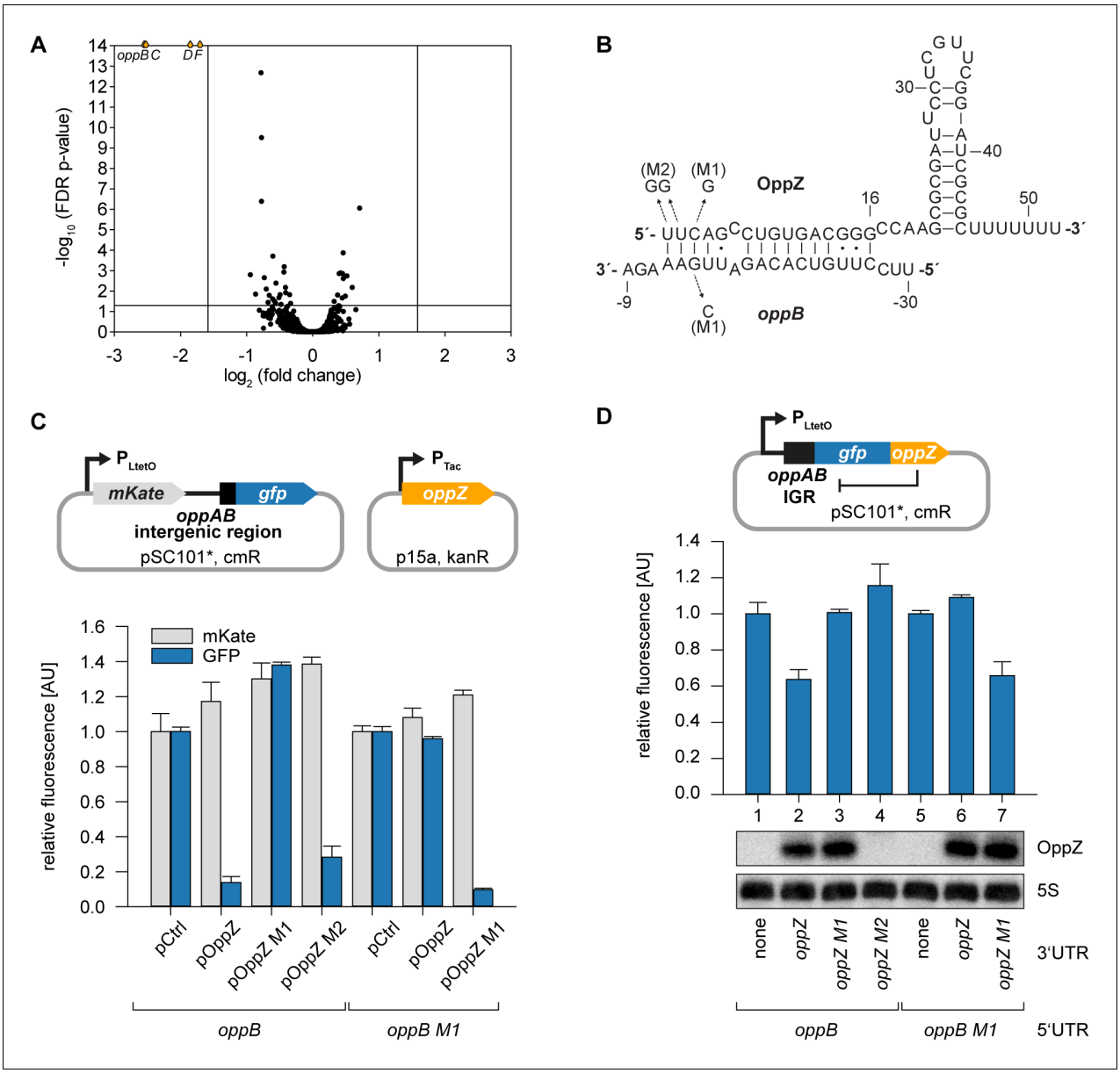


Figure 3. Feedback autoregulation at the suboperonic level. (**A**) Volcano plot of genome-wide transcript changes in response to inducible OppZ over-expression. Lines indicate cut-offs for differentially regulated genes at 3-fold regulation and FDR-adjusted p-value≤0.05. Genes with an FDR-adjusted p-value<10^{-14} are indicated as droplets at the top border of the graph. (**B**) Predicted OppZ secondary structure and base-pairing to *oppB*. Arrows indicate the mutations tested in (**C**) and (**D**). (**C**) *E. coli* strains carrying a translational reporter plasmid with the *oppAB* intergenic region placed between *mKate2* and *gfp* were co-transformed with a control plasmid or the indicated OppZ expression plasmids. Transcription of the reporter and *oppZ* were driven by constitutive promoters. Cells were grown to OD_{600} = 1.0 and fluorophore production was measured. mKate and GFP levels of strains carrying the control plasmid were set to 1. Error bars represent the SD of three biological replicates. (**D**) Single-plasmid regulation was measured by inserting the indicated *oppZ* variant into the 3′ UTR of a translational *oppB::gfp* fusion. Expression was driven from a constitutive promoter. *E. coli* strains carrying the respective plasmids were grown to OD_{600} = 1.0 and GFP production was measured. Fluorophore levels from control fusions without an sRNA gene were set to one and error bars represent the SD of three biological replicates. OppZ expression was tested by Northern blot; 5S rRNA served as loading control.

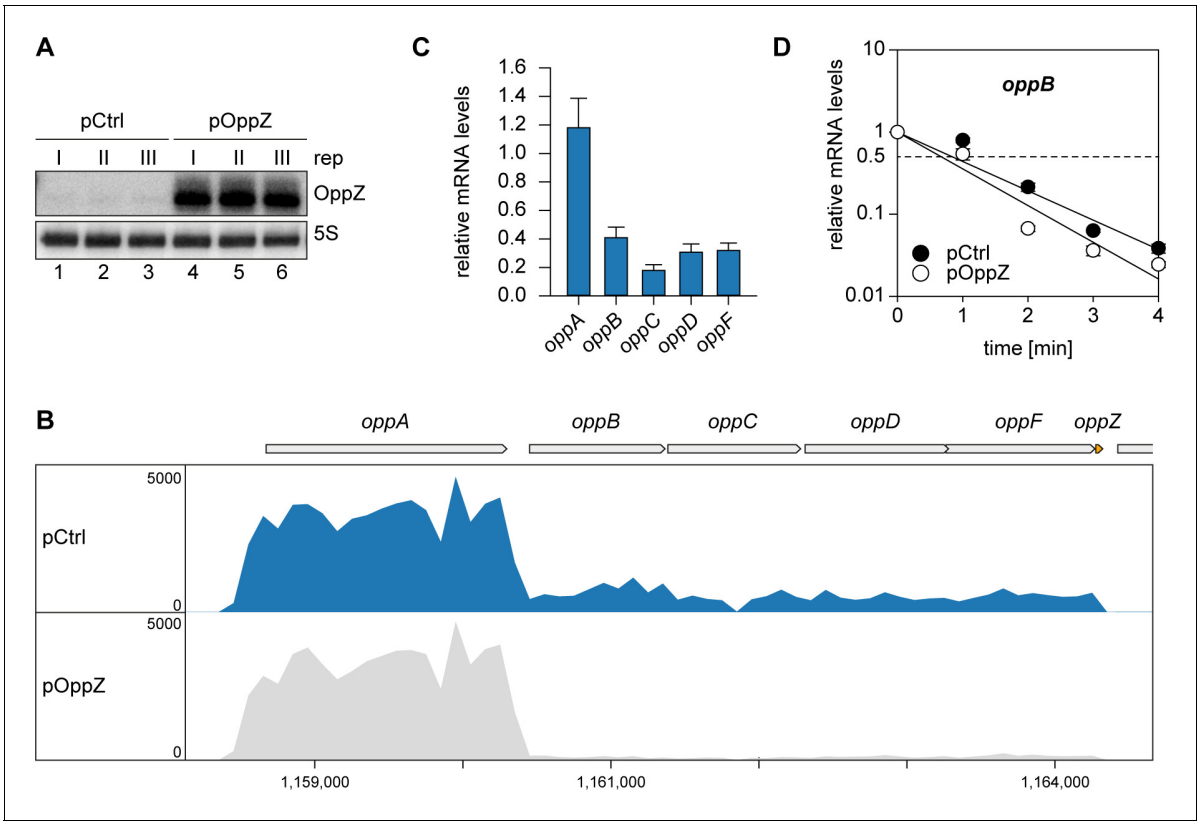


Figure 3—figure supplement 1. Pulse expression of OppZ reduces *oppBCDF* transcript levels. (**A**) *V. cholerae* carrying pBAD1K-*oppZ* (pOppZ) or a control plasmid (pCtrl) were grown in biological triplicates to exponential phase (OD_{600} of 0.5) and *oppZ* expression was induced by L-arabinose (0.2% final conc.). RNA samples were collected after 15 min and analyzed for OppZ levels by Northern blot; 5S rRNA served as loading control. (**B**) Samples from (**A**) were subjected to RNA-seq and average coverage of the *opp* operon is shown for one representative replicate. (**C**) *V. cholerae* Δ*oppZ* carrying pBAD1K-*oppZ* or a control plasmid were grown to late exponential phase (OD_{600} of 1.0) and *oppZ* expression was induced by L-arabinose (0.2% final conc.) for 15 min. mRNA levels of *oppABCDF* were analyzed by qRT-PCR. Bars show mRNA levels upon OppZ induction compared to the control; error bars represent the SD of three biological replicates. (**D**) *V. cholerae* Δ*oppZ* strains carrying either pBAD1K-ctrl (pCtrl) or pBAD1K-*oppZ* (pOppZ) were grown to late exponential phase (OD_{600} of 1.0) and treated with L-arabinose (0.2% final conc.) to induce sRNA expression. After 15 min of induction, rifampicin was added to terminate transcription. RNA samples were obtained at the indicated time points and *oppB* transcript levels were monitored by qRT-PCR. Error bars represent the SD of three biological replicates.

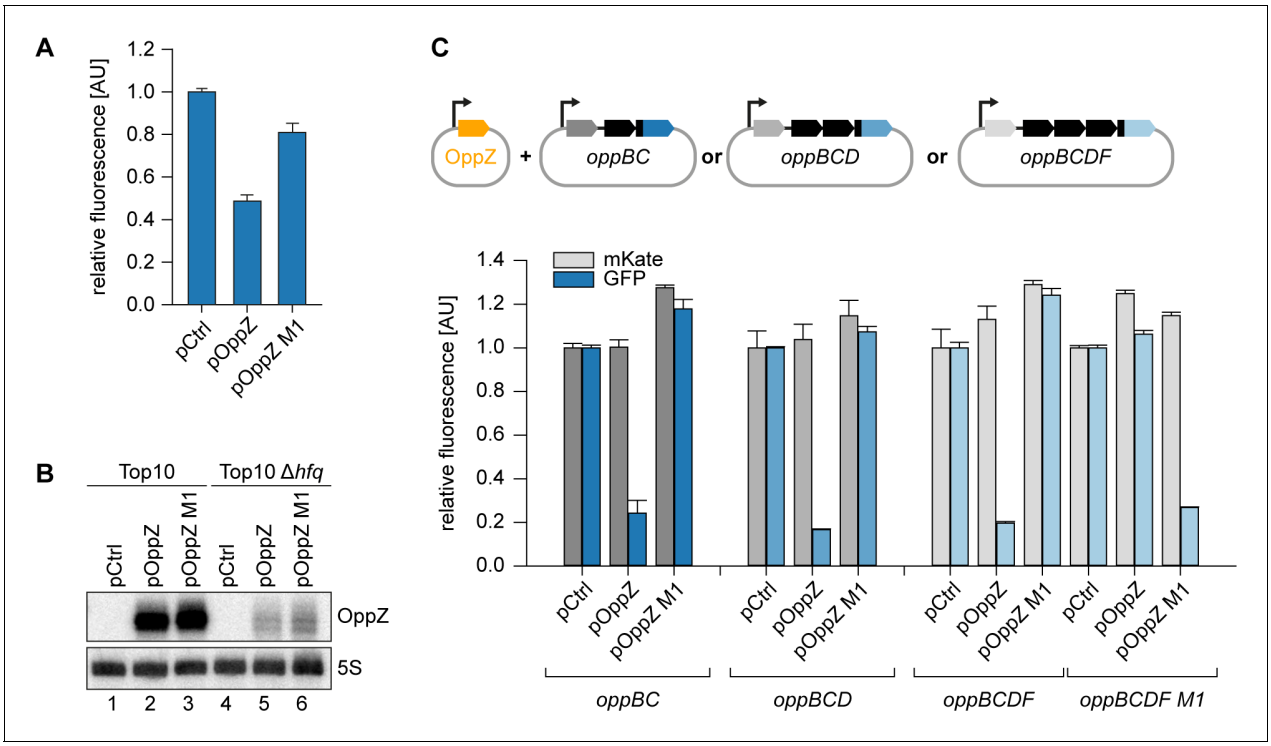


Figure 3—figure supplement 2. Hfq-dependent, post-transcriptional repression of OppBCDF by OppZ. (**A**) *E. coli* Δ*hfq* strains carrying the translational *oppB-gfp* reporter plasmid and either a control plasmid or the indicated OppZ expression plasmids were grown to OD_{600} = 1.0 and fluorophore production was measured. GFP levels of the control strain were set to 1. Error bars represent the SD of three biological replicates. (**B**) *E. coli* wild-type or Δ*hfq* strains carrying the translational *oppB-gfp* reporter plasmid and either a control plasmid or the indicated OppZ expression plasmids were grown to OD_{600} = 1.0. RNA samples were analyzed for OppZ levels by Northern blot; 5S rRNA served as loading control. (**C**) *E. coli* strains carrying translational reporter plasmids with the indicated parts of the *opp* operon placed between *mKate2* and *gfp* were co-transformed with a control plasmid or the respective OppZ expression plasmids. Transcription of the reporter and *oppZ* were driven by constitutive promoters. Cells were grown to OD_{600} = 1.0 and fluorophore production was measured. mKate and GFP levels of strains carrying the control plasmid were set to 1. Error bars represent the SD of three biological replicates.

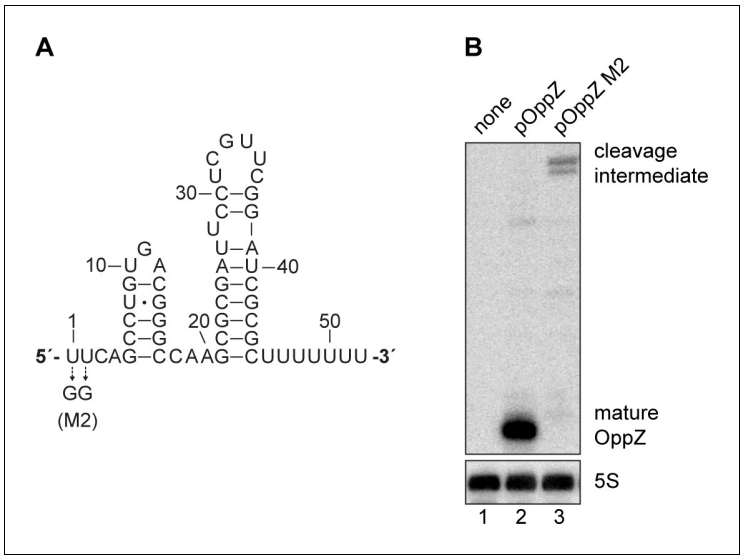


Figure 3—figure supplement 3. Mutational analysis of the RNase E site in *oppZ*. (**A**) Predicted structure of the OppZ sRNA. The M2 mutation blocking cleavage by RNase E is indicated. (**B**) *E. coli* strains carrying the empty pXG10-SF plasmid or derivatives with the indicated *oppZ* gene in the 3′ UTR of *gfp* were grown to OD_{600} = 1.0. RNA samples were analyzed for OppZ processing by Northern blot; 5S rRNA served as loading control.

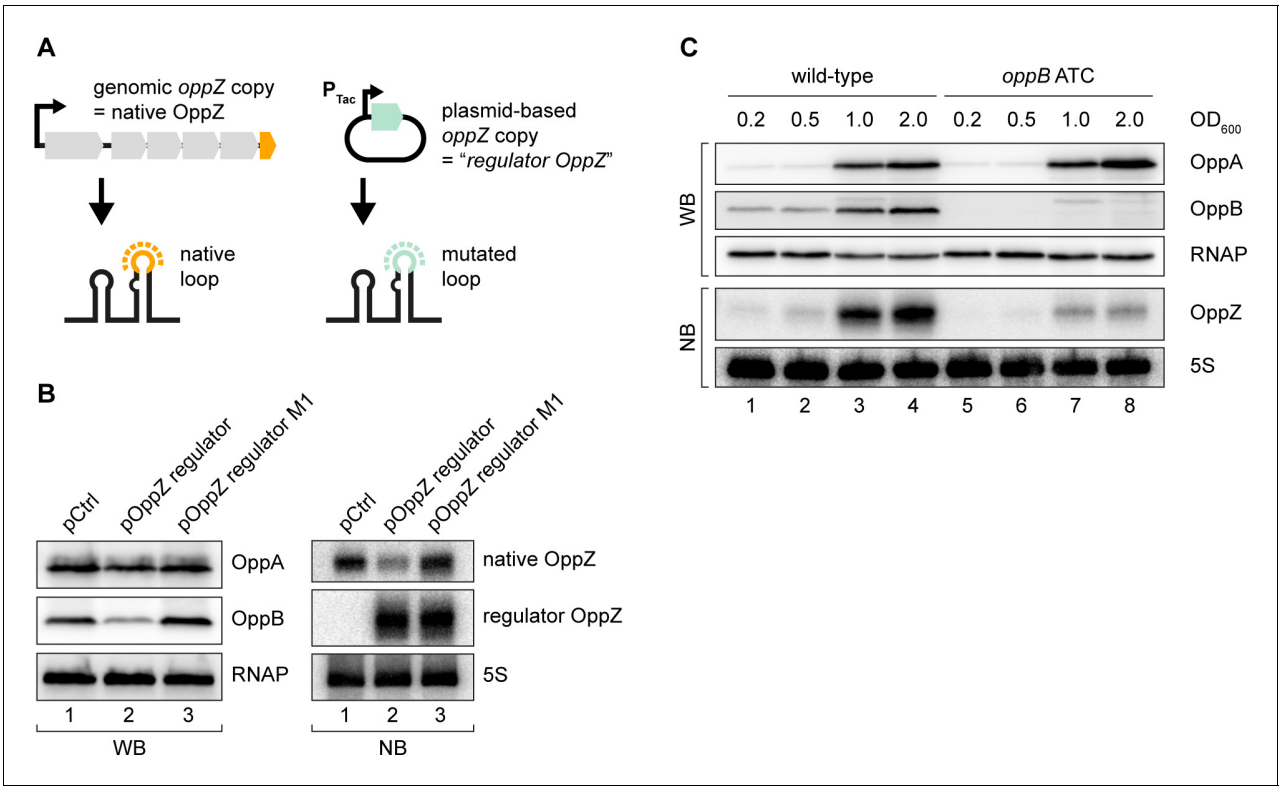


Figure 4. Translational control of OppZ synthesis. (**A**) Schematic of the analyzed OppZ variants containing the native stem loop sequence (produced from the genomic *oppZ* locus) or a mutated stem loop sequence ('*regulator OppZ*' produced from a plasmid-based constitutive promoter). (**B**) *V. cholerae oppA*::3XFLAG *oppB*::3XFLAG carrying a control plasmid (pCMW-1) or a plasmid expressing *regulator OppZ* (pMD194, pMD195) were grown to stationary phase (OD_{600} of 2.0). OppA and OppB production were tested by Western blot and expression of native OppZ and regulator OppZ was monitored on Northern blot using oligonucleotides binding to the respective loop sequence variants. RNAP and 5S rRNA served as loading controls for Western blot and Northern blot, respectively. (**C**) The *oppB* start codon was mutated to ATC in an *oppA*::3XFLAG *oppB*::3XFLAG background. *V. cholerae* strains with wild-type or mutated *oppB* start codon were grown in LB medium. Protein and RNA samples were collected at the indicated OD_{600} and tested for OppA and OppB production by Western blot and for OppZ expression by Northern blot. RNAP and 5S rRNA served as loading controls for Western and Northern blots, respectively.

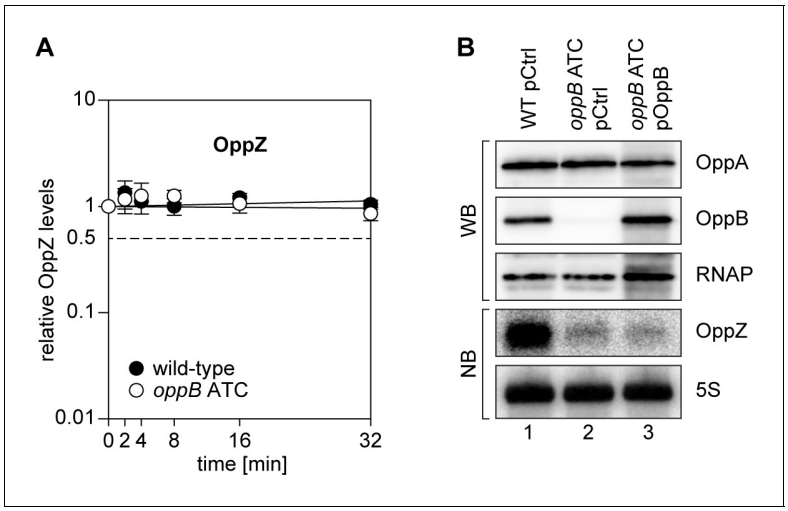


Figure 4—figure supplement 1. Translational control of OppZ synthesis. (**A**) *V. cholerae* wild-type and *oppB* ATC strains were grown to stationary phase (OD_{600} of 2.0) and treated with rifampicin to terminate transcription. RNA samples were obtained at the indicated time points and OppZ transcript levels were monitored by Northern blot and normalized to 5S rRNA levels as loading control. Error bars represent the SD of three biological replicates. (**B**) *V. cholerae* wild-type and *oppB* ATC strains carrying either a control plasmid (pCtrl) or an inducible *oppB* complementation plasmid (pOppB) were grown to late exponential phase (OD_{600} of 1.0) and *oppB* expression was induced by the addition of L-arabinose (0.2% final conc.). Protein and RNA samples were obtained after 60 min and tested for OppA and OppB production by Western blot and for OppZ expression by Northern blot. RNAP and 5S rRNA served as loading controls for Western and Northern blots, respectively.

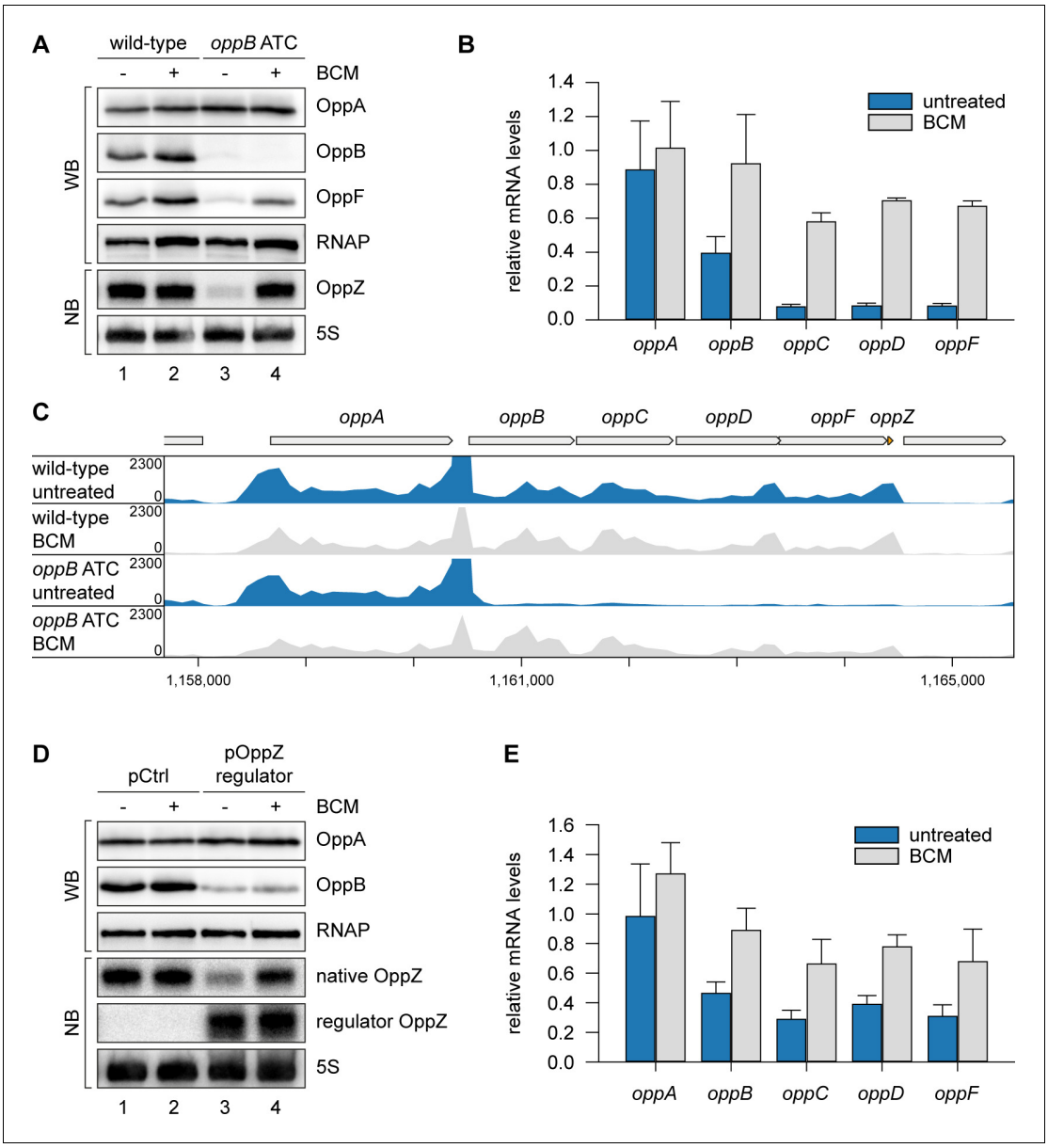


Figure 5. OppZ promotes transcription termination through Rho. (**A**) *V. cholerae oppA*::3XFLAG *oppB*::3XFLAG *oppF*::3XFLAG strains with wild-type or mutated *oppB* start codon were grown to early stationary phase (OD_{600} of 1.5). Cultures were divided in half and treated with either H_2O or BCM (25 µg/ml final conc.) for 2 hr before protein and RNA samples were collected. OppA, OppB and OppF production were tested by Western blot and OppZ expression was monitored by Northern blot. RNAP and 5S rRNA served as loading controls for Western and Northern blots, respectively. (**B**) Biological triplicates of *V. cholerae oppA*::3XFLAG *oppB*::3XFLAG strains with wild-type or mutated *oppB* start codon were treated with BCM as described in (**A**). *oppABCDF* expression in the *oppB* start codon mutant compared to the wild-type control was analyzed by qRT-PCR. Error bars represent the SD of three biological replicates. (**C**) Triplicate samples from (**B**) were subjected to Term-seq and average coverage of the *opp* operon is shown for one representative replicate. The coverage cut-off was set at the maximum coverage of annotated genes. (**D**) *V. cholerae oppA*::3XFLAG *oppB*::3XFLAG strains carrying a control plasmid (pMD397) or a plasmid expressing *regulator OppZ* (pMD398) were treated with BCM as described in (**A**). OppA and OppB production were tested by Western blot and expression of native OppZ and regulator OppZ was monitored on Northern blot using oligonucleotides binding to the respective loop sequence variants. RNAP and 5S rRNA served as loading controls for Western and Northern blots, respectively. (**E**) Levels of *oppABCDF* in the experiment described in (**D**) were analyzed by qRT-PCR. Error bars represent the SD of three biological replicates.

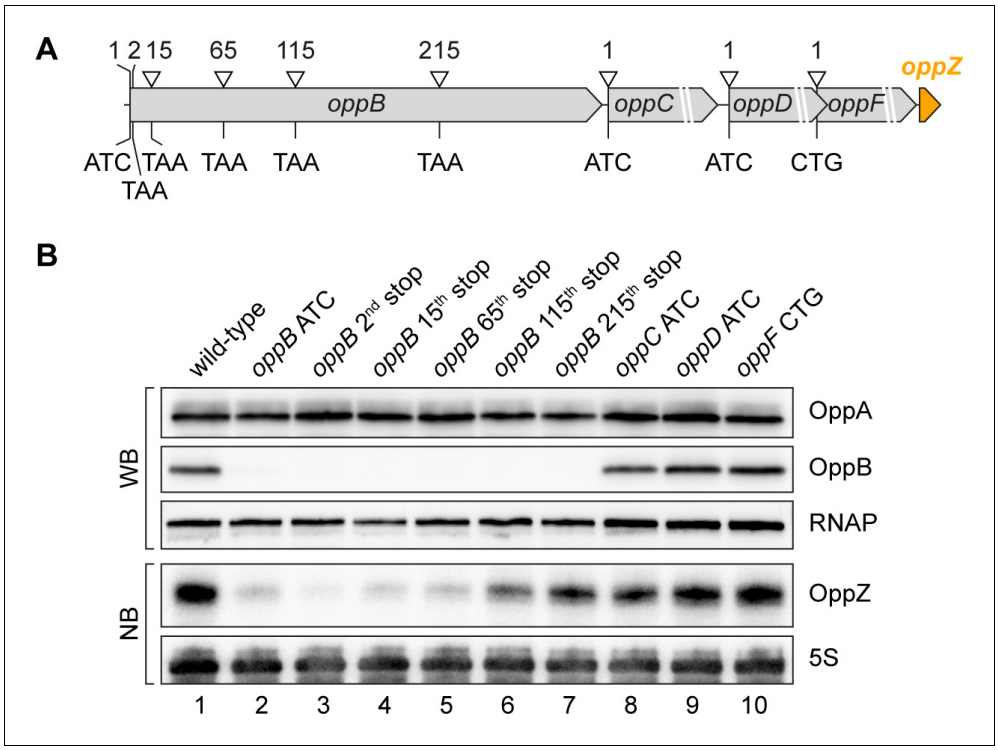


Figure 6. Influence of OppBCDF translation on OppZ expression. (**A**) The depicted mutations were individually inserted into the *opp* locus to inactivate the start codons of *oppB*, *oppC*, *oppD* or *oppF* or to insert STOP codons at the positions 2, 15, 65, 115 or 215 of *oppB*. (**B**) *V. cholerae oppA*::3XFLAG *oppB*::3XFLAG strains with the described *opp* mutations were grown: wild-type (lane 1), the *oppB* start codon mutated (lane 2), a STOP codon inserted at the 2nd, 15th, 65th, 115th or 215th codon of *oppB* (lanes 3–7) or mutated start codons of *oppC*, *oppD* or *oppF* (lanes 8–10). At stationary phase (OD_{600} of 2.0), protein and RNA samples were collected and tested for OppA and OppB production by Western blot and for OppZ expression by Northern blot. RNAP and 5S rRNA served as loading controls for Western and Northern blots, respectively.

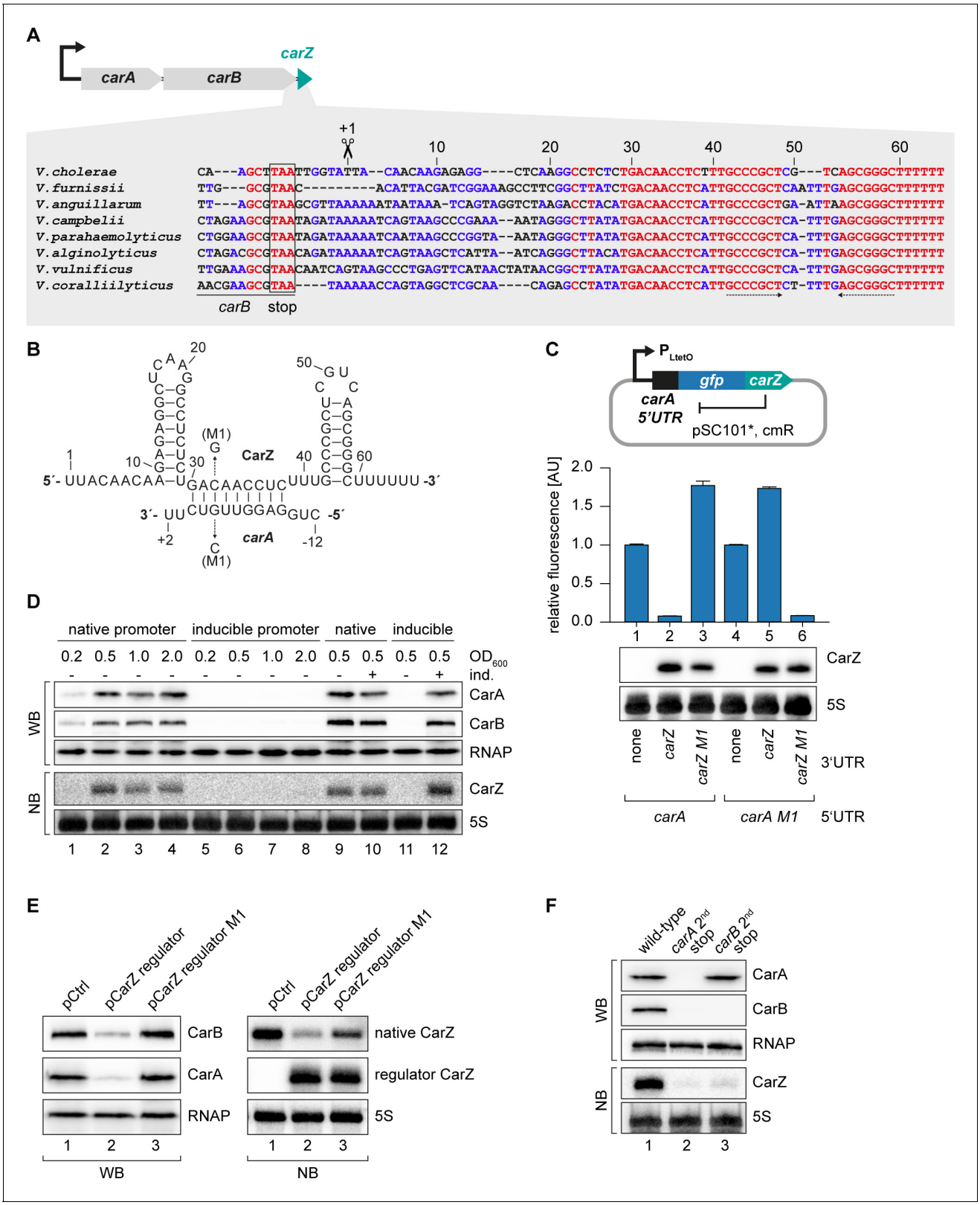


Figure 7. CarZ is another autoregulatory sRNA from *V. cholerae*. (**A**) Top: Genomic context of *carAB* and *carZ*. Bottom: Alignment of *carZ* sequences, including the last codons of *carB*, from various *Vibrio* species. The *carB* stop codon, the RNase E cleavage site and the Rho-independent terminator are

Figure 7 continued on next page

Figure 7 continued

indicated. (**B**) Predicted CarZ secondary structure and base-pairing to *carA*. Arrows indicate the single nucleotide mutations tested in (C). (**C**) Single-plasmid feedback regulation of *carA* by CarZ was measured by inserting the indicated *carZ* variant into the 3' UTR of a translational *carA::gfp* fusion. Expression was driven from a constitutive promoter. *E. coli* strains carrying the respective plasmids were grown to OD_{600} = 1.0 and GFP production was measured. Fluorophore levels from control fusions without an sRNA gene were set to one and error bars represent the SD of three biological replicates. CarZ expression was tested by Northern blot; 5S rRNA served as loading control. (**D**) Protein and RNA samples were obtained from *V. cholerae carA*::3XFLAG *carB*::3XFLAG carrying either the native *carA* promoter or the inducible pBAD promoter upstream of *carA*. Samples were collected at the indicated OD_{600} and tested for CarA and CarB production by Western blot and for CarZ expression by Northern blot. RNAP and 5S rRNA served as loading controls for Western and Northern blots, respectively. Lanes 1–8: Growth without L-arabinose. Lanes 9–12: Growth with either H_2O (-) or L-arabinose (+) (0.2% final conc.). (**E**) *V. cholerae carA*::3XFLAG *carB*::3XFLAG strains carrying a control plasmid or a plasmid expressing a CarZ variant with a mutated stem loop (*regulator CarZ*) were grown to late exponential phase (OD_{600} of 1.0). CarA and CarB production were tested by Western blot and expression of native CarZ or regulator CarZ was monitored on Northern blot using oligonucleotides binding to the respective loop sequence variants. RNAP and 5S rRNA served as loading controls for Western blot and Northern blot, respectively. (**F**) *V. cholerae carA*::3XFLAG *carB*::3XFLAG strains with the following *carA* or *carB* mutations were grown: wild-type (lane 1) or a STOP codon inserted at the 2nd codon of *carA* (lane 2) or *carB* (lane 3), respectively. At late exponential phase (OD_{600} of 1.0), protein and RNA samples were collected and tested for CarA and CarB production by Western blot and for CarZ expression by Northern blot. RNAP and 5S rRNA served as loading controls for Western and Northern blots, respectively.

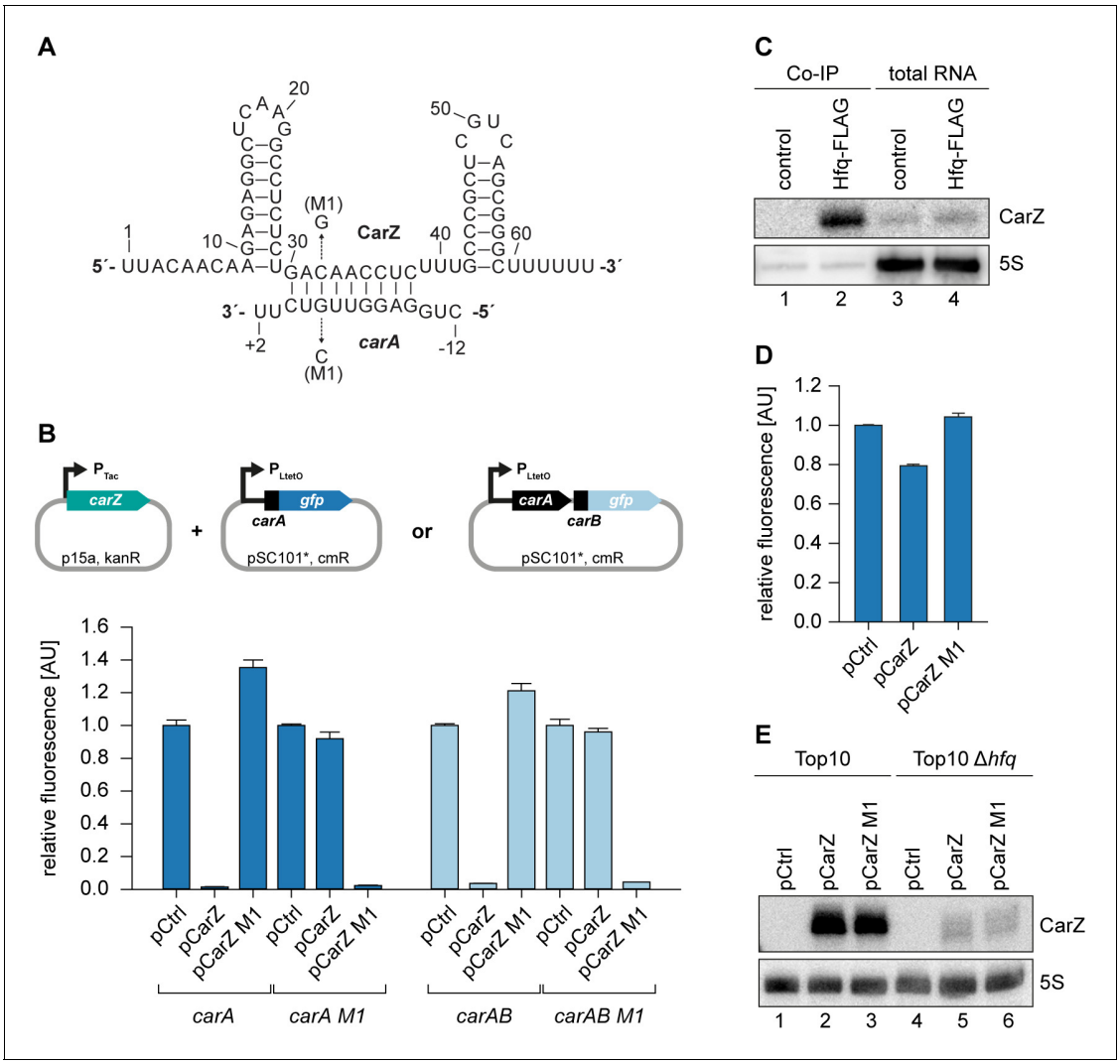


Figure 7—figure supplement 1. Hfq-dependent, post-transcriptional repression of CarA and CarB by CarZ. (**A**) Predicted CarZ secondary structure and base-pairing to *carA*. Arrows indicate the single nucleotide mutations tested in (**B**). (**B**) *E. coli* strains carrying translational reporter plasmids for *carA::gfp* or *carAB::gfp* were co-transformed with a control plasmid or the indicated CarZ expression plasmids. Transcription of the reporter and *carZ* were driven by constitutive promoters. Cells were grown to OD_{600} = 1.0 and fluorophore production was measured. GFP levels of strains carrying the control plasmid were set to 1. Error bars represent the SD of three biological replicates. (**C**) *V. cholerae* wild-type (control) and *hfq*::3XFLAG (Hfq-FLAG) strains were grown to stationary phase (OD_{600} of 2.0), lysed, and subjected to immunoprecipitation using the anti-FLAG antibody. RNA samples of lysate (total RNA) and co-immunoprecipitated fractions were analyzed on Northern blots. 5S rRNA served as loading control. (**D**) *E. coli* Δ*hfq* strains carrying the translational *carA::gfp* reporter plasmid and either a control plasmid or the indicated CarZ expression plasmids were grown to OD_{600} = 1.0 and fluorophore production was measured. GFP levels of the control strain were set to 1. Error bars represent the SD of three biological replicates. (**E**) *E. coli* wild-type or Δ*hfq* strains carrying the translational *carA::gfp* reporter plasmid and either a control plasmid or the indicated CarZ expression plasmids were grown to OD_{600} = 1.0. RNA samples were analyzed for CarZ levels by Northern blot; 5S rRNA served as loading control.

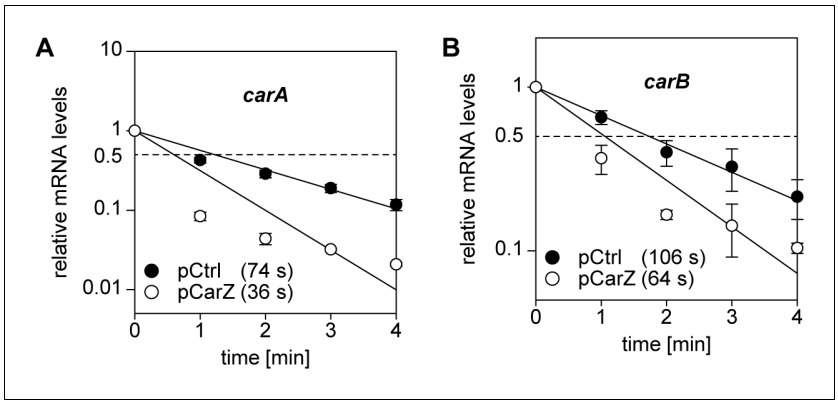


Figure 7—figure supplement 2. CarZ induces *carAB* degradation. (**A, B**) *V. cholerae* Δ*carZ* strains carrying either pBAD1K-ctrl (pCtrl) or pBAD1K-*carZ* (pCarZ) were grown to late exponential phase (OD_{600} of 1.0) and treated with L-arabinose (0.2% final conc.) to induce sRNA expression. After 15 min of induction, rifampicin was added to terminate transcription. RNA samples were obtained at the indicated time points and transcript levels of *carA* (**A**) and *carB* (**B**) were monitored by qRT-PCR. Error bars represent the SD of three biological replicates.

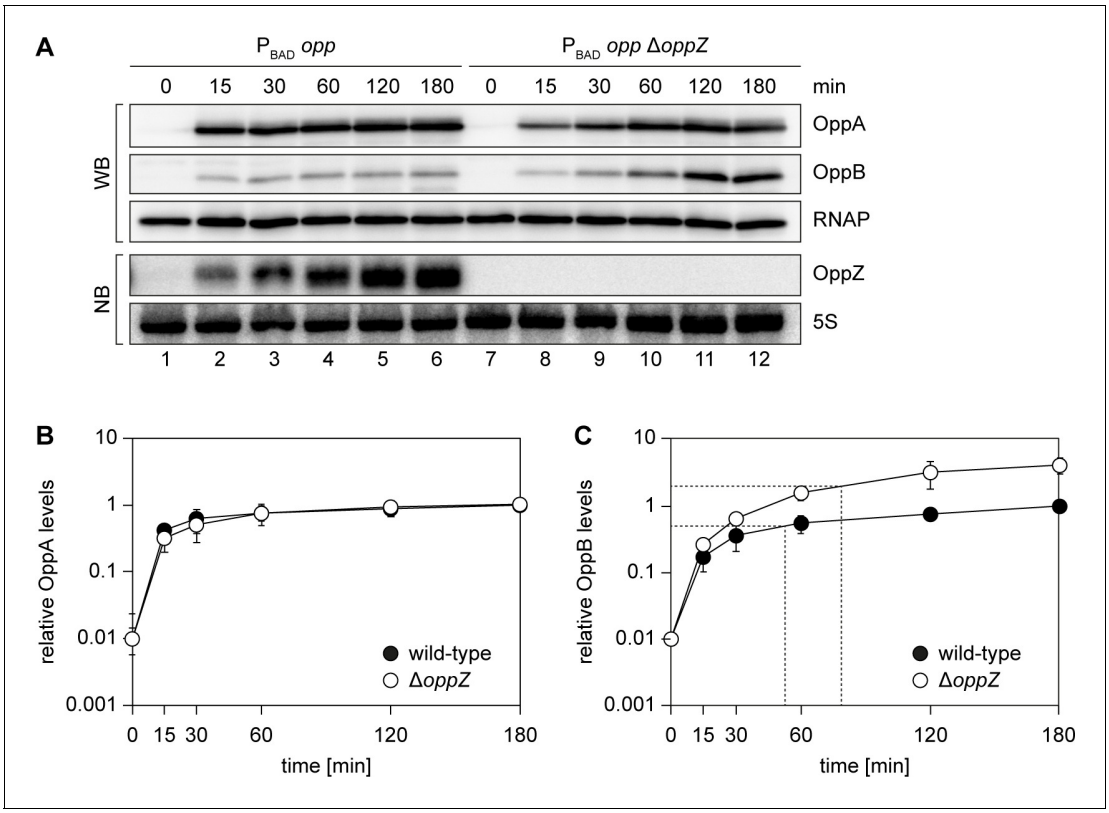


Figure 8. Modified kinetics of gene induction by autoregulatory OppZ. (**A**) Expression of the *opp* operon including the *oppA*::3XFLAG and *oppB*::3XFLAG genes and the native *oppZ* gene (lanes 1–6) or an *oppZ* deletion (lanes 7–12) was induced from the pBAD promoter at late exponential phase (OD_{600} of 1.0) by the addition of L-arabinose (0.2% final conc.). Protein and RNA samples were obtained at the indicated time points and tested for OppA and OppB production by Western blot and for OppZ expression by Northern blot. RNAP and 5S rRNA served as loading controls for Western and Northern blots, respectively. (**B, C**) Quantification of OppA (**B**) or OppB (**C**) levels from the experiment in (**A**); error bars represent the SD of three biological replicates. Data are presented as fold regulation of OppA or OppB in Δ*oppZ* compared to the wild-type. Dashed lines in (**C**) indicate the time points of half-maximum OppB expression.

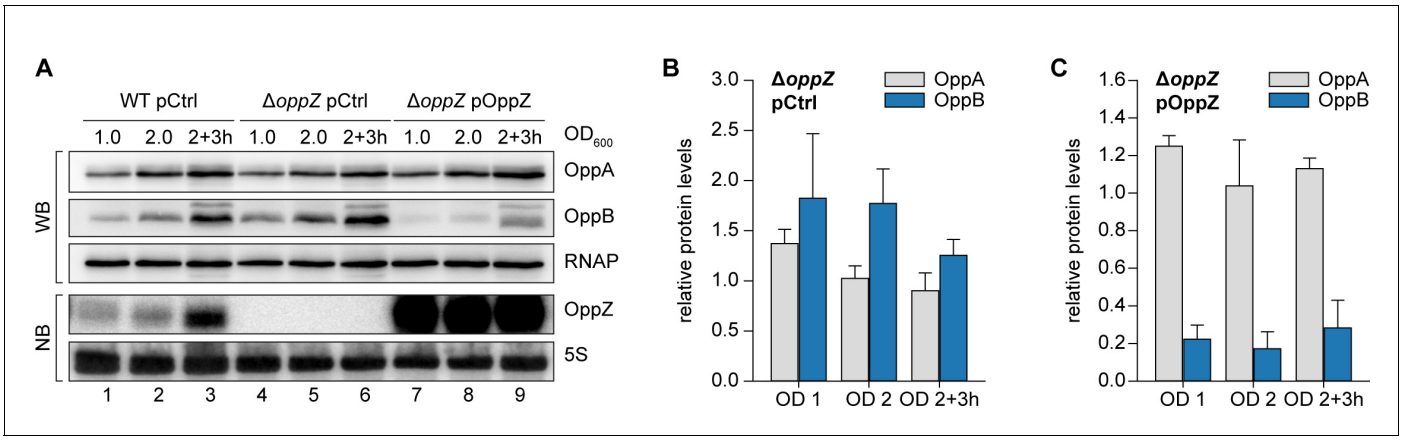


Figure 8—figure supplement 1. OppZ-dependent repression of OppA and OppB protein levels. (**A**) *V. cholerae* wild-type and Δ*oppZ* strains carrying the *oppA*::3XFLAG and *oppB*::3XFLAG genes and either a control plasmid or a constitutive OppZ expression plasmid were grown to obtain protein and RNA samples at the indicated OD_{600}. OppA and OppB production were analyzed by Western blot and OppZ expression was tested by Northern blot. RNAP and 5S rRNA served as loading controls for Western and Northern blots, respectively. (**B**) Quantification of (**A**), bars show fold regulation of OppA and OppB in Δ*oppZ* compared to the wild-type; error bars represent the SD of four biological replicates. (**C**) Quantification of (**A**), bars show fold regulation of OppA and OppB upon OppZ overexpression in the Δ*oppZ* background compared to the wild-type control; error bars represent the SD of four biological replicates.

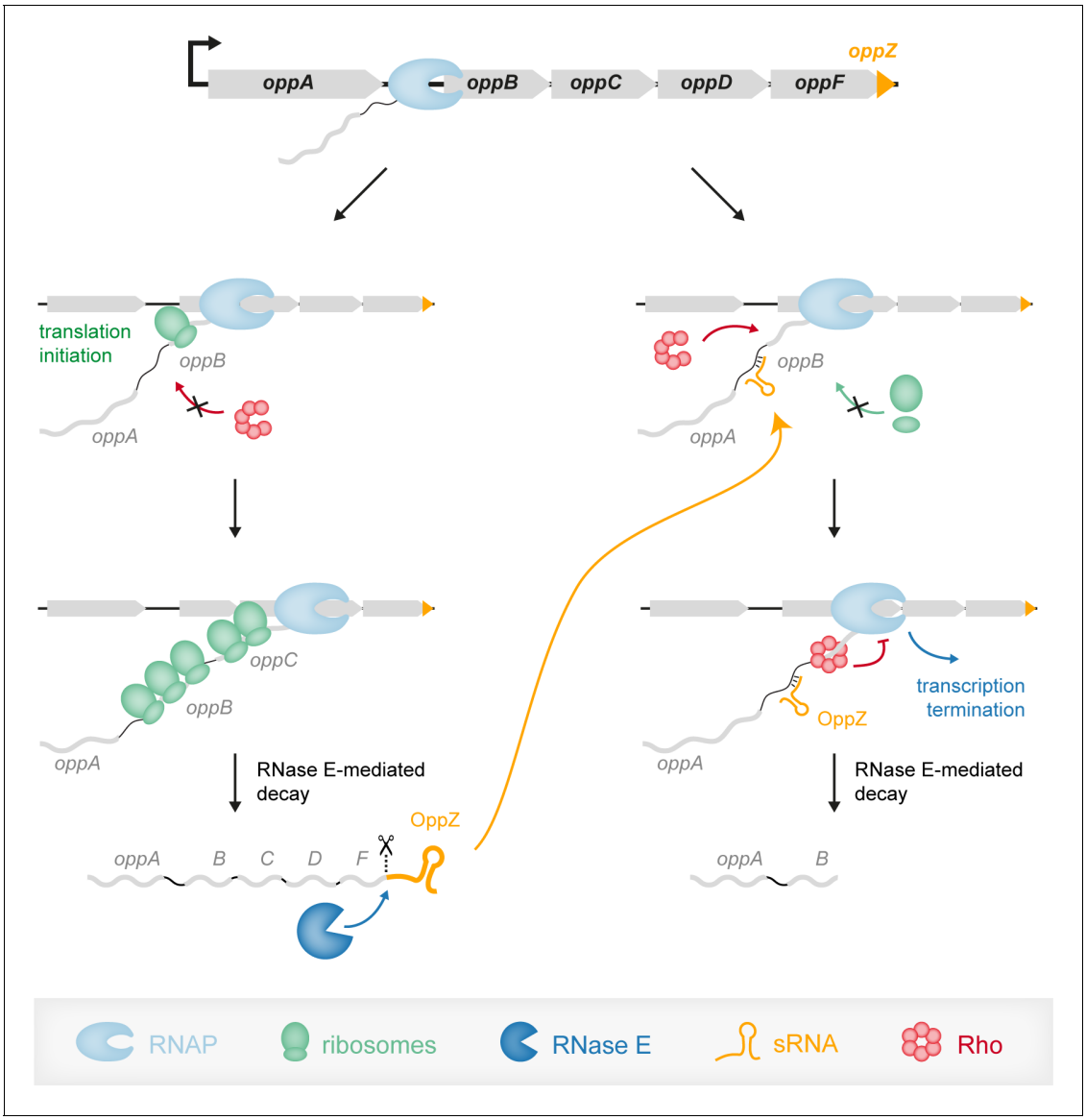


Figure 9. Model of the OppZ-dependent mechanism of *opp* regulation. Transcription of the *oppABCDF* operon initiates upstream of *oppA* and in the absence of OppZ (left) involves all genes of the operon as well as OppZ. In this scenario, all cistrons of the operon are translated. In the presence of OppZ (right), the sRNA blocks translation of *oppB* and the ribosome-free mRNA is recognized by termination factor Rho. Rho catches up with the transcribing RNAP and terminates transcription pre-maturely within *oppB*. Consequently, *oppBCDF* are not translated and OppZ is not produced.

6.2 Appendix to chapter 3: "A conserved RNA seed-pairing domain directs small RNA-mediated stress resistance in enterobacteria"

A conserved seed-pairing domain affords small RNA-mediated stress resistance in enterobacteria

Nikolai Peschek[1, 2], Mona Hoyos[1], Roman Herzog[1],
Konrad U. Förstner[3,4] and Kai Papenfort[1, 2, #]

[1] Faculty of Biology I, Department of Microbiology, Ludwig-Maximilians-University of Munich, 82152 Martinsried, Germany
[2] Munich Center for Integrated Protein Science (CIPSM)
[3] TH Köln - University of Applied Sciences, Institute of Information Science, 50678 Cologne, Germany
[4] ZB MED - Information Centre for Life Sciences, 50931 Cologne, Germany

[#] Corresponding author: kai.papenfort@lmu.de

This supplement contains:

Appendix Figures S1-S3
Appendix Tables S1-S5
Appendix Supplementary Material and Methods
Appendix Supplementary References

TABLE OF CONTENTS

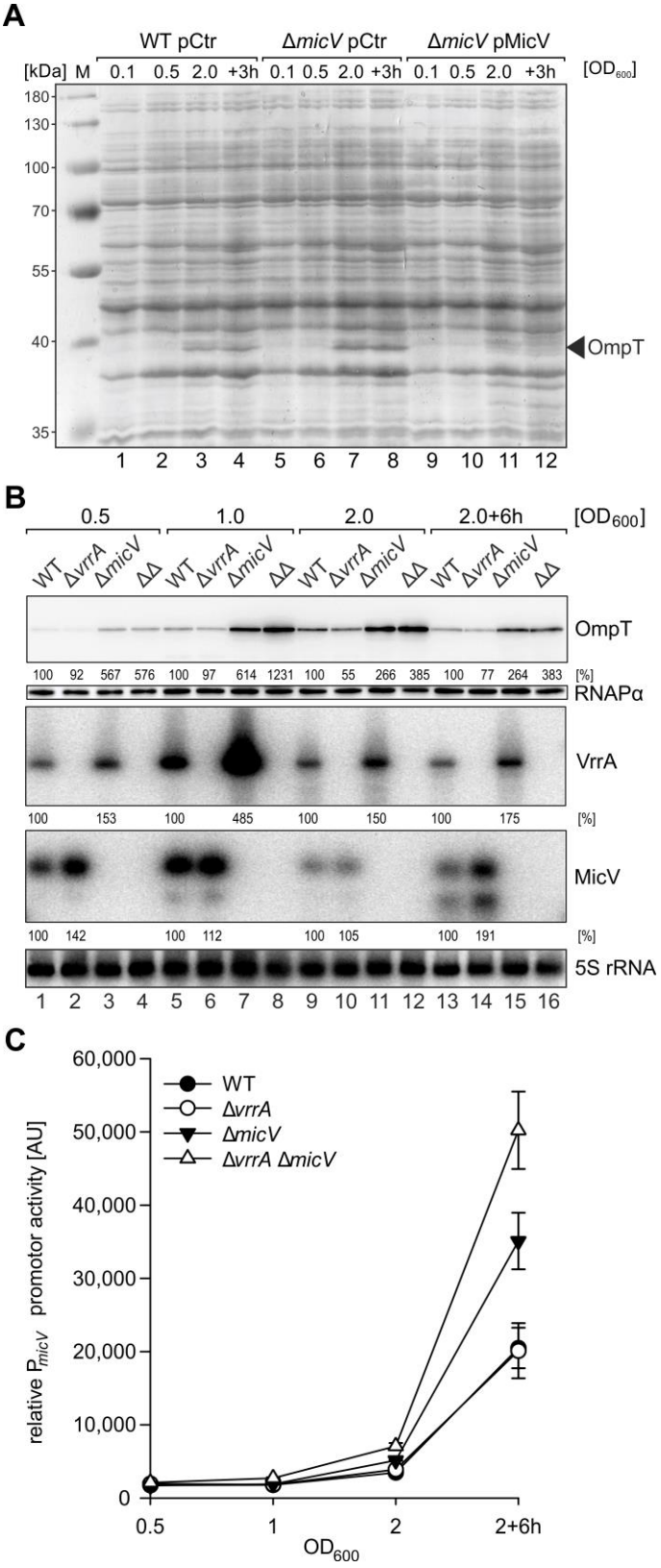
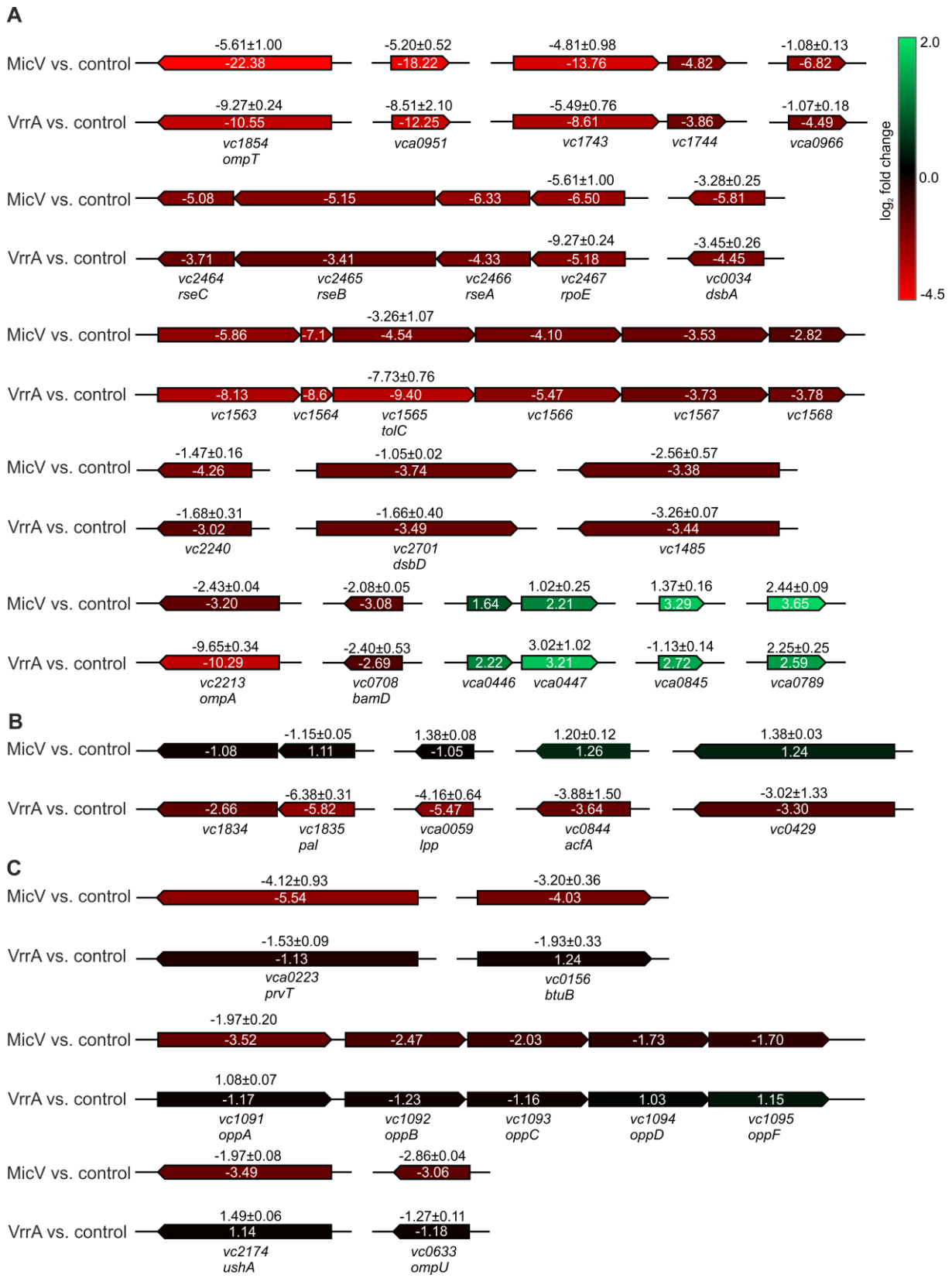

A
WT pCtr
ΔmicV pCtr
ΔmicV pMicV
[kDa] M 0.1 0.5 2.0 +3h 0.1 0.5 2.0 +3h 0.1 0.5 2.0 +3h [OD600]
180
130
100
70
55
40
35
OmpT
1 2 3 4 5 6 7 8 9 10 11 12
B
0.5 1.0 2.0 2.0+6h [OD600]
WT ΔvrrA ΔmicV ΔΔ WT ΔvrrA ΔmicV ΔΔ WT ΔvrrA ΔmicV ΔΔ WT ΔvrrA ΔmicV ΔΔ
OmpT
100 92 567 576 100 97 614 1231 100 55 266 385 100 77 264 383 [%]
RNAPα
VrrA
100 153 100 485 100 150 100 175 [%]
MicV
100 142 100 112 100 105 100 191 [%]
5S rRNA
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
C
WT
ΔvrrA
ΔmicV
ΔvrrA ΔmicV
relative PmicV promotor activity [AU]
60,000
50,000
40,000
30,000
20,000
10,000
0
0.5
1
2
2+6h
OD600

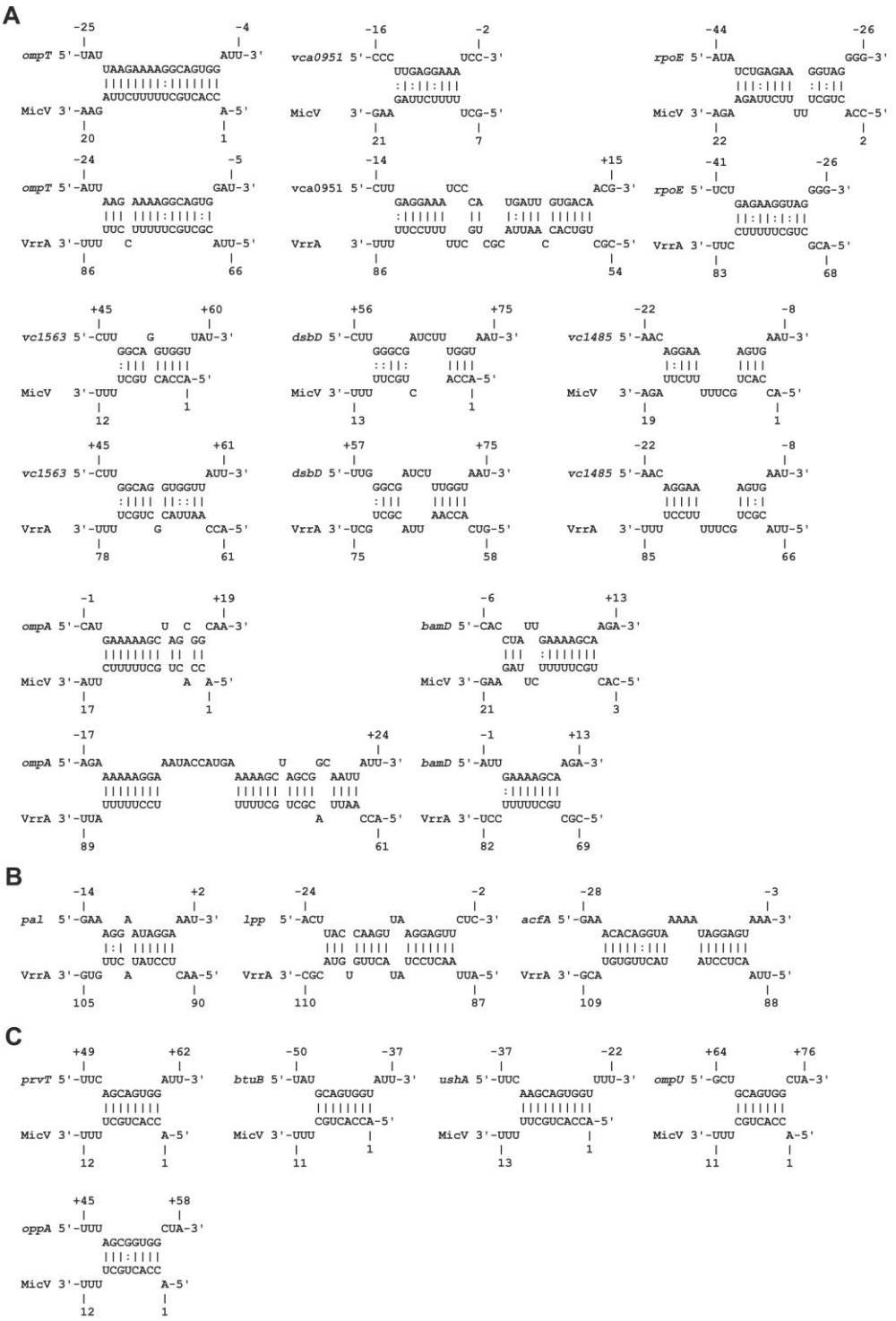
Figure Appendix S1: MicV controls OmpT production and OMP homeostasis

A) *V. cholerae* wild-type and Δ*micV* strains carrying empty vector control (pCtr) or *micV* expression plasmids (pMicV) were grown in LB. At the indicated time points, protein samples were collected and analyzed using SDS-PAGE and Coomassie staining. A molecular weight marker is provided on the left (M). Bands with different intensities were analyzed by mass spectrometry. **B)** *V. cholerae* wild-type, Δ*vrrA*, Δ*micV* or Δ*vrrA* Δ*micV* (ΔΔ) strains carrying the *ompT*::3XFLAG gene were grown in LB and at the indicated stages of growth, RNA and protein samples were collected. RNA samples were monitored for MicV and VrrA expression on Northern blots. Protein samples were investigated for OmpT::3xFLAG production using Western blot analysis. RNAPα served as a loading control for Western blots and 5S rRNA served as loading control for Northern blots. **C)** *V. cholerae* wild-type, Δ*vrrA*, Δ*micV* or Δ*vrrA* Δ*micV* strains carrying P*micV::mKate2* plasmids were cultivated in M9 minimal medium and at the indicated stages of growth, samples were collected and tested for mKate2 fluorescence. Data information: In (C), data are presented as mean ± SD, n = 3.


A
MicV vs. control
-5.61±1.00
-22.38
-5.20±0.52
-18.22
-4.81±0.98
-13.76
-4.82
-1.08±0.13
-6.82
VrrA vs. control
-9.27±0.24
-10.55
vc1854
ompT
-8.51±2.10
-12.25
vca0951
-5.49±0.76
-8.61
vc1743
-3.86
vc1744
-1.07±0.18
-4.49
vca0966
MicV vs. control
-5.08
-5.15
-6.33
-5.61±1.00
-6.50
-3.28±0.25
-5.81
VrrA vs. control
-3.71
vc2464
rseC
-3.41
vc2465
rseB
-4.33
vc2466
rseA
-9.27±0.24
-5.18
vc2467
rpoE
-3.45±0.26
-4.45
vc0034
dsbA
MicV vs. control
-5.86
-7.1
-3.26±1.07
-4.54
-4.10
-3.53
-2.82
VrrA vs. control
-8.13
vc1563
-8.6
vc1564
-7.73±0.76
-9.40
vc1565
tolC
-5.47
vc1566
-3.73
vc1567
-3.78
vc1568
MicV vs. control
-1.47±0.16
-4.26
-1.05±0.02
-3.74
-2.56±0.57
-3.38
VrrA vs. control
-1.68±0.31
-3.02
vc2240
-1.66±0.40
-3.49
vc2701
dsbD
-3.26±0.07
-3.44
vc1485
MicV vs. control
-2.43±0.04
-3.20
-2.08±0.05
-3.08
1.64
1.02±0.25
2.21
1.37±0.16
3.29
2.44±0.09
3.65
VrrA vs. control
-9.65±0.34
-10.29
vc2213
ompA
-2.40±0.53
-2.69
vc0708
bamD
2.22
vca0446
3.02±1.02
3.21
vca0447
-1.13±0.14
2.72
vca0845
2.25±0.25
2.59
vca0789
B
MicV vs. control
-1.08
-1.15±0.05
1.11
1.38±0.08
-1.05
1.20±0.12
1.26
1.38±0.03
1.24
VrrA vs. control
-2.66
vc1834
-6.38±0.31
-5.82
vc1835
pal
-4.16±0.64
-5.47
vca0059
lpp
-3.88±1.50
-3.64
vc0844
acfA
-3.02±1.33
-3.30
vc0429
C
MicV vs. control
-4.12±0.93
-5.54
-3.20±0.36
-4.03
VrrA vs. control
-1.53±0.09
-1.13
vca0223
prvT
-1.93±0.33
1.24
vc0156
btuB
MicV vs. control
-1.97±0.20
-3.52
-2.47
-2.03
-1.73
-1.70
VrrA vs. control
1.08±0.07
-1.17
vc1091
oppA
-1.23
vc1092
oppB
-1.16
vc1093
oppC
1.03
vc1094
oppD
1.15
vc1095
oppF
MicV vs. control
-1.97±0.08
-3.49
-2.86±0.04
-3.06
VrrA vs. control
1.49±0.06
1.14
vc2174
ushA
-1.27±0.11
-1.18
vc0633
ompU
log2 fold change
2.0
0.0
-4.5

Appendix Figure S2: MicV and VrrA mRNA target validation

A, B and C) *V. cholerae* Δ*vrrA* Δ*micV* strains carrying either an empty vector control (control), pBAD-micV (MicV) or pBAD-vrrA (VrrA) were grown in LB to OD_{600}=1.5. RNA samples were collected 10 min after induction with L-arabinose (0.2% final conc.) and analyzed for mRNA levels using RNA-seq or qRT-PCR. Targets determined by RNA-seq are depicted by arrows and are labelled with the fold change (white numbers). Arrows are colored according to the log_2 transformed fold change (right scalebar). For targets validated using qRT-PCR, the resulting fold-change is indicated above the tested mRNA target. The targets were grouped as follows: regulated by both sRNAs (A), regulated only by VrrA (B) or regulated only by MicV (C). Data information: In (A-C), qRT-PCR data are presented as mean ± SD, n = 2.



Appendix Figure S3: Base-pairing interactions of MicV and VrrA with target mRNAs

A, B and C) Prediction of base-pairing interactions of the MicV and VrrA sRNAs with their respective target mRNAs. The numbers indicate the position relative to the AUG start codon (mRNA) or the transcriptional start site (sRNA). The targets were grouped as follows: regulated by both sRNAs (A), regulated only by VrrA (B) or regulated only by MicV (C).

Appendix Table S1: Global identification of σ^E-dependent promoters in *V. cholerae*

Potential promotor sites in *V. cholerae* predicted with MEME on a motif search based on σ^E-dependent promotor sites that are found in maximal distance of 50 nt to a TSS (see Appendix Supplementary Methods). Entries marked as "orphans" are TSS which are not associated with any gene, i.e. are not located 300 nt or less upstream or downstream of an ORF (see Papenfort *et al.,* 2015, Figure 1B).

NC_ 002505						
Motif hit start	**Motif hit end**	**Motif hit seq**	**TSS position**	**TSS strand**	**TSS locus_tag**	**TSS gene product**
13737	13764	GAAATTTAGAGAAAAAAGGAAAGTCTAA	13729	-	orphan	orphan
134057	134084	TAAACGTTCTTCACAACTGGTAATCAAT	134017	-	VC0140	hypothetical protein
139822	139849	TGAACCTGTTTAAGTGATGGTGGTCATA	139855	+	VC0150	RNA polymerase factor sigma-32
142040	142067	CAAATTTATCGGCACTGTAGGTGTCGAT	142104	+	VC0151	soluble pyridine nucleotide transhydrogenase
163858	163885	GGAACTAGAGCAAAAAATGTTTGCCAAA	163852	-	VC0165	hypothetical protein
255476	255503	CGAACTCTATAAAATTTAGCTGTTCTGA	255435	-	VC0249*,#	RfbL protein
282410	282437	CGAACTCAACAATGCCGGTTTTATCTGA	282451	+	VC0276	bifunctional phosphoribosylaminoimidazolecarboxamide formyltransferase%2FIMP cyclohydrolase
369659	369686	TGAACTTCGCTTTGAGAAGATGGTCGAA	369692	+	VC0346	tRNA delta%282%29-isopentenylpyrophosphate transferase
477347	477374	TGAATTTTTCTGGAGTGAAGCGGTCTGA	477308	-	VC0446*	organic solvent tolerance protein
491005	491032	GGCACTTTATCGGTCCACTACAGTCGAA	490998	-	VC0461	hypothetical protein
522911	522938	TGAATTTATTGGTTTTGGTTTTATCTGT	522955	+	orphan	orphan
523404	523431	GGAAATAAACTTGACTTGGCTGATCATA	523444	+	VC0490	hypothetical protein
530444	530471	TGAACTTATAGATAGTTGAACGGCCTAA	530436	-	VC0496	hypothetical protein
573241	573268	TGAACCGTTTGGCGCTTTGGATGCCAAA	573275	+	VC0541	sulfate ABC transporter ATP-binding protein
580501	580528	GGCACCATTGAGAGTTTACGATGTCATA	580551	+	VC0548	carbon storage regulator
586889	586916	CGAACTGATTCACAAAAACAAGGTCATA	586922	+	VC0554	insulinase family protease%2Finsulinase family protease
598379	598406	GGAACCTTTCAGAATCACACTCGTCTAA	598413	+	VC0565	protease DegS
609707	609734	GGAACTTCCACAGATGAAAATCGTCGAA	609700	-	VC0580	hypothetical protein
743157	743184	GAACATAAACACTGATTTTGTGGTCAAA	743191	+	VC0694	hypothetical protein
757821	757848	CGAACTTTTTAAAAAACCGTGAGACTAA	757815	-	VC0708*,#	hypothetical protein
772170	772197	TGAACCTTATATGAAAATTTGTTTGCAA	772227	+	VC0719	DNA-binding response regulator PhoB
803818	803845	CGAACTTTGCCCCGGCTCTGCAATCTGA	803811	-	VC0751	co-chaperone HscB
867987	868014	CGCATTTTCCGGGTTAATTGCTGCCAAA	868043	+	VC0812	helicase-like protein
916669	916696	GCAACCAAACTCAAAATTCACAGTCTCA	916662	-	VC0851*	small protein A
929880	929907	TGAATTAAATCGCTTTCCTGTTGTCAGA	929861	-	orphan	orphan
1063364	1063391	GAACCTTGAATATGTTGAGTCGATCAAA	1063426	+	VC0997	glutaminyl-tRNA synthetase
1114450	1114477	TGAACTCCTTCGCATAATTTCTGTCCTA	1114444	-	VC1045	RNA polymerase sigma factor
1158499	1158526	TGAATTAATTTGCCATAAAAATGTCTTA	1158493	-	orphan	orphan
1169175	1169202	TCAAATTTTTTTTGTTATATTTGTCCAT	1169156	-	VC1098	acetate kinase
1301806	1301833	GGAACTCCATCGAAAACTCGAAGTCTGA	1301799	-	orphan	orphan
1374686	1374713	CGAACATTTTTTGAGTGGTCGTATCAGA	1374748	+	orphan	orphan
1469405	1469432	TCCACTTCTTCCTTATTATGTTATCTAT	1469364	-	VC1376	GGDEF family protein
1591531	1591558	GGAACTTTTGGAAGAATTGCTTGCCAAT	1591525	-	VC1486	ABC transporter ATPase
1591564	1591591	TGAACCAACCAACGATTTAGATATCGAA	1591525	-	VC1486	ABC transporter ATPase
1602295	1602322	GAAATGTTACTGAACAGGTGTTGTCCAA	1602328	+	VC1492	hypothetical protein
1617120	1617147	AAAACCTGTCGCTAATTTCAGTATCTGT	1617163	+	orphan	orphan
1675554	1675581	TGAACTTTCCTTATCATCCTTAGTCTGA	1675587	+	VC1563	pseudo

1744512	1744539	GGAACATCACGCCATTAATCGAATCGAA	1744545	+	VC1623	carboxynorspermidine decarboxylase
1856343	1856370	GGAACTTTTTGCGTGTCCAGTTGACTGA	1856377	+	VC1718	hypothetical protein
1878650	1878677	GGAACTCTTTGCCAAACGCCCAGTCTGA	1878684	+	orphan	***vrrA***
1903913	1903940	TGCACTAATCAGCATATTGTTTATCTGA	1903967	+	VC1764	hypothetical protein
1937499	1937526	CAAACTATTAGCTGTAGTGGTCAACTAA	1937568	+	VC1788	hypothetical protein
2049280	2049307	TAAACTTTCGTTAAAAAACGCGATCTAA	2049274	-	orphan	orphan
2068844	2068871	CGAACCTTTTGAAATTATGCGCATCCTA	2068808	-	VC1918	peptidyl-prolyl cis-trans isomerse D
2093516	2093543	CAAACGTTTGCCTGTGGATGTTATCAAA	2093565	+	VC1942	bifunctional 5%2C10-methylene-tetrahydrofolate dehydrogenase%2F 5%2C10-methylene-tetrahydrofolate cyclohydrolase
2111173	2111200	GAACAGTATGCGCAATTTGGTTGTCAGA	2111167	-	VC1957	hypothetical protein
2140363	2140390	GGAACTTGCGCAGCTACTTGGGGTCGAT	2140356	-	VC1987*,#	outer membrane lipoprotein Slp
2164623	2164650	TGACTTTATCGAGGATTATGGTGTCTGA	2164617	-	orphan	orphan
2196838	2196865	GCAACCAAAGCTGGAATTCACTGTCTGA	2196871	+	VC2040	hypothetical protein
2248788	2248815	GGAATTTCGCACCAAGATAGCGCTCTAA	2248822	+	VC2087	2-oxoglutarate dehydrogenase E1
2302696	2302723	TAAATCGATTGGCAAGTTATTGATCAAA	2302657	-	VC2149	hypothetical protein
2306486	2306513	GGAACCAGCACGCCCAATCGTTGCCCAA	2306480	-	VC2156*	lipoprotein
2524108	2524135	GGAACCCTGAGAGTATTCGCTTGTCAGA	2524101	-	VC2366	ribonuclease activity regulator protein RraA
2613708	2613735	TGAATTTTTAGCGCAATATCTGGTCTTA	2613701	-	VC2437	pseudo
2649846	2649873	TGAACTTTCTCGATAATGCCGAGTCTCT	2649839	-	VC2467*,#	RNA polymerase sigma factor RpoE
2654232	2654259	CACACTATTTTTGTTTAGGTTTTTCTAT	2654270	+	VC2473	hypothetical protein
2709569	2709596	GGAACCTTCACTGCTGGAGATTGCCAAA	2709603	+	VC2524	3-deoxy-D-manno-octulosonate 8-phosphate phosphatase
2812169	2812196	TGAACCTTTTGCTTAGAGCTCTGTCTAT	2812162	-	VC2640	***micV***
2908373	2908400	GGAACTCATTGCCACATTGCCTCTCTAA	2908365	-	VC2734	general secretion pathway protein C
NC_ 002506						
Motif hit start	**Motif hit end**	**Motif hit seq**	**TSS position**	**TSS strand**	**TSS locus_tag**	**TSS gene product**
35140	35167	GGCACTTTCTGCTCCTGCATCAGTCAAA	35174	+	VCA0027	chitinase
67115	67142	CAAATTTTTTCCAGACAAATTTGTCACT	67182	+	VCA0061	DEAD%2FDEAH box helicase
92487	92514	TGAACAAGCTGTCATTCTCTCTATCAAA	92447	-	VCA0080	diguanylate cyclase
214685	214712	GAAACTCATTGACAAAACGAACATCAAA	214667	-	VCA0198	site-specific DNA-methyltransferase
357164	357191	CAAACTATTAGCTGTAGTGGTCAACTAA	357233	+	VCA0370	hypothetical protein
400386	400413	GGCAATTTAATGTCAAAAATTTATCAAA	400463	+	VCA0447	hemagglutinin associated protein
424133	424160	TAAACAAGAAGTCGATGAAGTTGTCGAA	424183	+	VCA0485	MazG domain-containing protein
510200	510227	GAAACTTCACATTGAATGAACTATCATA	510193	-	VCA0572	D-alanyl-alanine synthetase A
513961	513988	CAACCTTATATTGATAAAGGTGAACTAA	513940	-	VCA0575	LysR family transcriptional regulator
524874	524901	GAAACTCAAAGCCTATTTGAGAATCCAA	524866	-	VCA0588	peptide ABC transporter ATP-binding protein
921498	921525	TGCCATTTTCGCTGAAAAACTTGTCCAT	921452	-	VCA0974	methyl-accepting chemotaxis protein
929810	929837	GGAATCCAAAGCCATTTGCTTAGTCCAT	929863	+	VCA0981	hypothetical protein
949359	949386	TAAAATTTATCGATTGAAATTCATCAAA	949317	-	VCA0994	hypothetical protein
957852	957879	TAAACTAACCGCTGATAAACTACTCAGA	957911	+	VCA1004	hypothetical protein

*Listed as σ^E-dependent in *E. coli* K12, according to Ecocyc database (https://ecocyc.org/)

#Listed as σ^E-dependent in *E. coli* K12 (Rhodius *et al.*, 2006)

Appendix Table S2: Genes differentially regulated by either *micV* or *vrrA* pulse expression

Gene	Description[#]	Fold change* *micV* pulse	Fold change* *vrrA* pulse
ompT	outer membrane protein OmpT	-22.38	-10.55
vca0951	hypothetical protein	-18.22	-12.25
vc1743	hypothetical protein	-13.76	-8.61
vca0966	hypothetical protein	-6.82	-4.49
rpoE	RNA polymerase sigma factor RpoE	-6.50	-5.18
rseA	sigma-E factor negative regulatory protein RseA	-6.33	-4.33
vc1563	pseudogene	-5.86	-8.13
dsbA	thiol:disulfide interchange protein DsbA	-5.81	-4.45
rseB	sigma-E factor negative regulatory protein RseB	-5.15	-3.41
rseC	sigma-E factor negative regulatory protein RseC	-5.08	-3.71
vc1744	hypothetical protein	-4.82	-3.86
tolC	outer membrane protein TolC	-4.54	-9.40
vc2240	phenolic acid decarboxylase	-4.26	-3.02
vc1566	putative ABC transport system permease	-4.10	-5.47
dsbD	thiol:disulfide interchange protein DsbD	-3.74	-3.49
vc1567	putative ABC transport system permease	-3.53	-3.73
vc1485	hypothetical protein	-3.38	-3.44
ompA	outer membrane protein OmpA	-3.20	-10.29
bamD	outer membrane protein assembly factor BamD	-3.08	-2.69
vc1568	ABC transporter ATP-binding protein	-2.82	-3.78
vca0447	site-specific DNA-methyltransferase	2.21	3.21
vca0845	hypothetical protein	3.29	2.72
vca0789	putative membrane protein	3.65	2.59
pal	peptidoglycan-associated lipoprotein	1.11	-5.82
lpp	major outer membrane lipoprotein	-1.05	-5.47
acfA	accessory colonization factor AcfA	1.26	-3.64
vc0429	hypothetical protein	1.24	-3.30
prtV	immune inhibitor A, protease	-5.54	-1.31
btuB	vitamin B12 transporter	-4.03	1.24
oppA	oligopeptide transport substrate-bind. protein	-3.52	-1.17
ushA	5'-nucleotidase / UDP-sugar diphosphatase	-3.49	1.14
ompU	outer membrane protein OmpU	-3.06	-1.18

[#]Description is based on the annotation at KEGG (https://www.genome.jp/kegg)

*Fold change is based on transcriptomic analysis of pBAD-derived *micV* or *vrrA* expression using RNA-seq. Genes with a fold-change of at least 3.0-fold in either condition and a FDR adjusted p-value ≤ 1E-8 were considered to be differentially expressed.

Appendix Table S3: Bacterial strains used in this study

Strain	Relevant markers/ genotype	Reference/ source
V. cholerae		
KPS-0014	C6706 Wild-type	(Thelin & Taylor, 1996)
KPS-0054	C6706 Δ*hfq*	(Svenningsen et al., 2009)
KPS-0995	C6706 *hfq::hfq*-3XFlag	This study
KPVC-10072	C6706 Δ*vrrA*	This study
KPVC-10075	C6706 Δ*micV* Δ*vrrA*	This study
KPVC-10076	C6706 Δ*micV*	This study
KPVC-10122	C6706 *ompT::ompT-3xFLAG*	This study
KPVC-10124	C6706 Δ*micV ompT::ompT-3xFLAG*	This study
KPVC-10137	C6706 Δ*vrrA ompT::ompT-3xFLAG*	This study
KPVC-10139	C6706 Δ*micV* Δ*vrrA ompT::ompT-3xFLAG*	This study
KPVC-10814	C6706 Δ*vchM*	This study
KPVC-10822	C6706 Δ*vchM* Δ*rpoE*	This study
KPVC-10824	C6706 Δ*vchM* Δ*rpoE* Δ*vrrA*	This study
KPVC-10826	C6706 Δ*vchM* Δ*rpoE* Δ*micV*	This study
KPVC-10828	C6706 Δ*vchM* Δ*rpoE* Δ*vrrA* Δ*micV*	This study
KPVC-12139	C6706 Δ*ompA*	This study
KPVC-12143	C6706 Δ*vchM* Δ*rpoE* Δ*ompA*	This study
KPVC-12203	C6706 Δ*micV* Δ*vrrA ompA::ompA-3xFlag*	This study
KPVC-12647	C6706 *ompA::ompA* scr	This study
KPVC-12651	C6706 Δ*vchM* Δ*rpoE ompA::ompA* scr	This study
E. coli		
BW25113	*lacI*$^{+}$*rrnB*$_{T14}$ Δ*lacZ*$_{WJ16}$ *hsdR*514 Δ*araBAD*$_{AH33}$ Δ*rhaBAD*$_{LD78}$ *rph-1* Δ*(araB–D)567* Δ*(rhaD–B)568* Δ*lacZ4787*(::*rrnB-3*) *hsdR514 rph-1*	(Datsenko & Wanner, 2000)
TOP10	F- *mcrA* Δ(*mrr-hsdRMS-mcrBC*) φ80*lacZ*ΔM15 Δ*lacX*74 *nupG recA*1 *araD*139 Δ(*ara-leu*)7697 *galE*15 *galK*16 *rpsL*(Str^{R}) *endA1* λ^{-}	Invitrogen
S17λpir	Δ*lacU*169 (Φ*lacZ*ΔM15), *recA*1, *endA*1, *hsdR*17, *thi*-1, *gyrA*96, *relA*1, λ*pir*	(Simon et al., 1983)
ECA101	*E.coli* BW25113 Δ*rpoE*	(Egler et al., 2005)
KPEC-52214	BW25113 *ompA::kan*^{R}	(Baba et al., 2006)
KPEC-52215	BW25113 *ompC::kan*^{R}	(Baba et al., 2006)

Appendix Table S4: Plasmids used in this study

Plasmid trivial name	Plasmid stock name-	Relevant fragment	Comment	Origin, marker	Reference
pBAD1K-Ctr	pMD004		Control plasmid	p15A, Kan^R	Papenfort lab plasmid collection
pBAD1K-*micV*	pNP016	P_{BAD}-*micV*	*micV* expression plasmid	p15A, Kan^R	This study
pBAD1K-rpoE	pNP018	P_{BAD}-*rpoE*	*rpoE* expression plasmid	p15A, Kan^R	This study
pBAD1K-vrrA	pNP022	P_{BAD}-*vrrA*	*vrrA* expression plasmid	p15A, Kan^R	This study
pBAD5A	pKP8-35	empty	Control plasmid	pBR322, Amp^R	(Papenfort et al., 2006)
pBAD5A-*rpoE* (*E.c.)*	pKP142-2	*rpoE* (*E.coli*)	*rpoE* expression plasmid	pBR322, Amp^R	(Papenfort et al., 2010)
pBAD5A-*rpoE* (*V.c.*)	pRH011	*rpoE* (*V.cholerae*)	*rpoE* expression plasmid	pBR322, Amp^R	This study
pCMW-1C	pCtr		Promotorless plasmid for transcriptional reporters	p15A, Cm^R	(Herzog et al., 2019)
pCMW-1C-*mKATE2*	pYH-010	*mKATE2*	Promoterless plasmid for transcriptional reporters	p15A, Cm^R	(Herzog et al., 2019)
pCMW-1C-P*micV*::*mKate2*	pNP074	P*micV*::*mKATE2*	Transcriptional reporter P*micV::mKATE2*	p15A, Cm^R	This study
pCMW-1C-P*vrrA*::*mKate2*	pNP075	P*vrrA*::*mKATE2*	Transcriptional reporter P*vrrA::mKATE2*	p15A, Cm^R	This study
pCMW-1K	pCtr		Control plasmid	p15A, Kan^R	Papenfort lab plasmid collection
pCMW-1K-P*micV*::*gfp*	pNP017	P*micV::gfp*	Transcriptional reporter P*micV::gfp*	p15A, Kan^R	This study
pEVS143-1K		P*tac* promoter	Constitutive over-expression plasmid	p15A, Kan^R	(Dunn et al., 2006)
pEVS143-1C		P*tac* promoter	Constitutive over-expression plasmid	p15A, Cm^R	Papenfort lab plasmid collection
pEVS-*micV*	pNP002	*micV*	*micV* expression plasmid	p15A, kan^R	This study
pEVS-*micV* M1	pRG001	*micV* M1	*micV* M1 expression plasmid	p15A, Kan^R	This study
pEVS-P_L-*rybB*	pMD030	*rybB*	*rybB* expression plasmid	p15A, Cm^R	This study
pEVS-P_L-*rybB*	pMD251	*rybB*	*rybB* expression plasmid	p15A, Kan^R	This study
pEVS-P_L-*rybB*Δ9	pNP088	*rybB*Δ9	*rybB*Δ9 expression plasmid	p15A, Kan^R	This study
pEVS-P_L-*sRNA* var.01	pMD241	*sRNA* var.1	*sRNA* var.1 expression plasmid	p15A, Kan^R	This study
pEVS-P_L-*sRNA* var.02	pMD242	*sRNA* var.2	*sRNA* var.2 expression plasmid	p15A, Kan^R	This study
pEVS-P_L-*sRNA* var.03	pMD243	*sRNA* var.3	*sRNA* var.3 expression plasmid	p15A, Kan^R	This study

pEVS-P_L-*sRNA* var.04	pMD244	*sRNA* var.4	*sRNA* var.4 expression plasmid	p15A, Kan^R	This study
pEVS-P_L-*sRNA* var.05	pMD245	*sRNA* var.5	*sRNA* var.5 expression plasmid	p15A, Kan^R	This study
pEVS-P_L-*sRNA* var.06	pMD246	*sRNA* var.6	*sRNA* var.6 expression plasmid	p15A, Kan^R	This study
pEVS-P_L-*sRNA* var.07	pMD247	*sRNA* var.7	*sRNA* var.7 expression plasmid	p15A, Kan^R	This study
pEVS-P_L-*sRNA* var.08	pMD248	*sRNA* var.8	*sRNA* var.8 expression plasmid	p15A, Kan^R	This study
pEVS-P_L-*sRNA* var.09	pMD249	*sRNA* var.9	*sRNA* var.9 expression plasmid	p15A, Kan^R	This study
pEVS-P_L-*sRNA* var.10	pMD250	*sRNA* var.10	*sRNA* var.10 expression plasmid	p15A, Kan^R	This study
pEVS-P_L-*sRNA* var.11	pMD251	*sRNA* var.11	*sRNA* var.11 expression plasmid	p15A, Kan^R	This study
pEVS-P_L-*sRNA* var.12	pMD252	*sRNA* var.12	*sRNA* var.12 expression plasmid	p15A, Kan^R	This study
pEVS-P_L-*sRNA* var.13	pMD253	*sRNA* var.13	*sRNA* var.13 expression plasmid	p15A, Kan^R	This study
pEVS-P_L-*sRNA* var.14	pMD254	*sRNA* var.14	*sRNA* var.14 expression plasmid	p15A, Kan^R	This study
pEVS-P_L-*sRNA* var.15	pMD255	*sRNA* var.15	*sRNA* var.15 expression plasmid	p15A, Kan^R	This study
pEVS-P_L-*sRNA* library		*sRNA* 9nt variants	*sRNA* library expression plasmid	p15A, Cm^R	This study
pEVS-*rybB*	pRH013	*rybB*	*rybB* expression plasmid	p15A, Kan^R	This study
pEVS-*vrrA*	pRH001	*vrrA*	*vrrA* expression plasmid	p15A, kan^R	This study
pEVS-*vrrA* M1	pRG002	*vrrA* M1	*vrrA* M1 expression plasmid	p15A, Kan^R	This study
pEVS-*vrrA* M2	pRG004	*vrrA* M2	*vrrA* M2 expression plasmid	p15A, Kan^R	This study
pKAS32	pKAS32		suicide plasmid for allelic exchange	R6K, Amp^R	(Skorupski & Taylor, 1996)
pKAS32-*hfq*::3xFLAG	pKP431	*hfq:3xFLAG:*	*hfq::3xFLAG* allelic replacement	R6K, Amp^R	This study
pKAS32-*ompA::3xFLAG*	pNP089	*ompA::3xFLAG*	*ompA::3xFLAG* allelic replacement	R6K, Amp^R	This study
pKAS32-*ompA*	pNP090	*ompA*	*ompA* region	R6K, Amp^R	This study
pKAS32-*ompA* scr	pNP091	*ompA* scr	*ompA* scr allelic replacement	R6K, Amp^R	This study
pKAS32-*ompT::3xFLAG*	pNP021	*ompT::3xFLAG*	*ompT::3xFLAG* allelic replacement	R6K, Amp^R	This study
pKAS32-Δ*micV*	pNP024	up-/downstream flanks of *micV*	suicide plasmid for *micV* knock-out	R6K, Amp^R	This study
pKAS32-Δ*ompA*	pEE001	up-/downstream flanks of *ompA*	suicide plasmid for *ompA* knock-out	R6K, Amp^R	This study

pKAS32-Δ*rpoE*	pNP023	up-/downstream flanks of *rpoE*	suicide plasmid for *rpoE* knock-out	R6K, Amp^R	This study
pKAS32-Δ*vchM*	pNP076	up-/downstream flanks of *vchM*	suicide plasmid for *vchM* knock-out	R6K, Amp^R	This study
pKAS32-Δ*vrrA*	pNP026	up-/downstream flanks of *vrrA*	suicide plasmid for *vrrA* knock-out	R6K, Amp^R	This study
P_L-*rybB*	pFM1-1	P_L-*rybB*	*rybB* expression plasmid	ColE1, Amp^R	(Bouvier et al., 2008)
pXG10-1C	pXG10-1C	'*lacZ::gfp*	template plasmid for translational reporters	pSC101*, Cm^R	Papenfort lab plasmid collection
pXG10-1C-*acfA::gfp*	pNP059	*acfA::gfp*	Translational reporter *acfA::gfp*	pSC101*, Cm^R	This study
pXG10-1C-*bamD::gfp*	pNP082	*bamD::gfp*	Translational reporter *bamD::gfp*	pSC101*, Cm^R	This study
pXG10-1C-*btuB::gfp*	pNP029	*btuB::gfp*	Translational reporter *btuB::gfp*	pSC101*, Cm^R	This study
pXG10-1C-*dsbD::gfp*	pNP079	*dsbD::gfp*	Translational reporter *dsbD::gfp*	pSC101*, Cm^R	This study
pXG10-1C-*lpp* M2	pNP087	*lpp* M2**::gfp*	Translational reporter *lpp* M2**::gfp*	pSC101*, Cm^R	This study
pXG10-1C-*lpp::gfp*	pNP086	*lpp::gfp*	Translational reporter *lpp::gfp*	pSC101*, Cm^R	This study
pXG10-1C-*ompA::gfp*	pNP081	*ompA::gfp*	Translational reporter *ompA::gfp*	pSC101*, Cm^R	This study
pXG10-1C-*ompT* M1	pRG009	*ompT* M1**::gfp*	Translational reporter *ompT* M1**::gfp*	pSC101*, Cm^R	This study
pXG10-1C-*ompT::gfp*	pKP465	*ompT::gfp*	Translational reporter *ompT::gfp*	pSC101*, Cm^R	This study
pXG10-1C-*ompU::gfp*	pNP085	*ompU::gfp*	Translational reporter *ompU::gfp*	pSC101*, Cm^R	This study
pXG10-1C-*oppA::gfp*	pNP084	*oppA::gfp*	Translational reporter *oppA::gfp*	pSC101*, Cm^R	This study
pXG10-1C-*pal::gfp*	pNP062	*pal::gfp*	Translational reporter *pal::gfp*	pSC101*, Cm^R	This study
pXG10-1C-*prvT::gfp*	pNP083	*prvT::gfp*	Translational reporter *prvT::gfp*	pSC101*, Cm^R	This study
pXG10-1C-*rpoE::gfp*	pNP044	*rpoE::gfp*	Translational reporter *rpoE::gfp*	pSC101*, Cm^R	This study
pXG10-1C-*ushA* M1	pRG008	*ushA* M1**::gfp*	Translational reporter *ushA* M1**::gfp*	pSC101*, Cm^R	This study
pXG10-1C-*ushA::gfp*	pNP054	*ushA::gfp*	Translational reporter *ushA::gfp*	pSC101*, Cm^R	This study
pXG10-1C-*vc1485::gfp*	pNP080	*vc1485::gfp*	Translational reporter *vc1485::gfp*	pSC101*, Cm^R	This study
pXG10-1C-*vc1563::gfp*	pNP078	*vc1563::gfp*	Translational reporter *vc1563::gfp*	pSC101*, Cm^R	This study
pXG10-1C-*vca0951::gfp*	pNP077	*vca0951::gfp*	Translational reporter *vca0951::gfp*	pSC101*, Cm^R	This study
pXG10-SF	pXG10SF	'*lacZ::gfp*	template plasmid for translational reporters	PSC101*, Cm^R	(Corcoran et al., 2012)

Appendix Table S5: DNA oligonucleotides used in this study

ID	Sequence (5'→3'); P denotes a monophosphate	Description
KPO-0012	GCAGATCGAACTGGAAGCT	pKP431
KPO-0019	CGAGATTATCGATCTTATTCA	pKP431
KPO-0066	TATCATGATCTTTATAATCACCGTCATGGTCTTTGTAGTCCTCTTCAGACTTC TCTGCTGG	pKP431
KPO-0067	GATCATGATATCGACTACAAAGATGACGATAAATAGTTCTTTGCACAATTATT TAAGGAG	pKP431
KPO-0092	CCACACATTATACGAGCCGA	pNP002, pRH001/013
KPO-0148	GTTTTTGGTACCATCCCAATGATCCACAAAGA	pKP431
KPO-0149	GTTTTTCCTAGGAAACAGTCTCTACCGCTTGG	pKP431
KPO-0196	GGAGAAACAGTAGAGAGTTGCG	pNP016/018/022
KPO-0236	GGGCGTACACAAGTATAGGAGT	VrrA oligoprobe
KPO-0243	TTCGTTTCACTTCTGAGTTCGG	5S rRNA oligoprobe
KPO-0267	TAATAGGCCTAGGATGCATATG	pNP026/076/089, pEE01
KPO-0268	CGTTAACAACCGGTACCTCTA	pNP026/076/089, pEE01
KPO-0282	CACTGACACCCTCATCAGTG	pMD241-255
KPO-0640	P-TTTTACCGCGACACCGTGGC	pNP16
KPO-0820	GGCCTTCTTAGAGTCTTCTAAGAA	MicV oligoprobe
KPO-0999	P-ACCACTGCTTTTTCTTAGAAGAC	pNP002/016
KPO-1000	GTTTTTTCTAGAGGATTAGAACCCGAATTAAACT	pNP002
KPO-1023	GTTTTTTCTAGAGGATCCGGTGATTGATTGAG	pRH001/013, pNP002
KPO-1064	GTTTTTTATGCATGAATCTAATGGCGGTGGTG	pKP465
KPO-1065	GTTTTTTGCTAGCAGCTGCGTTTACAGAGCCT	pKP465
KPO-1082	P-GTGATTGACAGAGCTTTGAGA	pRH001
KPO-1083	GTTTTTTCTAGATCGCCAATGAACCGACTTG	pRH001
KPO-1180	GCTTATTTGGAGATGTTTGAGC	pNP024
KPO-1181	AGAGCTCTAAGCAAAAGGTTCAT	pNP024
KPO-1182	AACCTTTTGCTTAGAGCTCTTGCGTAGCAGAAAGTTTAATTCG	pNP024
KPO-1183	ACCAAATCCCGCTGCTGCAT	pNP024
KPO-1184	GTTTTTGGTACCCTAGCGCGTTTAAACACCTCA	pNP024
KPO-1185	GTTTTTCCTAGGAGATCAAGGACGCATTGCCG	pNP024
KPO-1186	TCAATCGTAAAAGGCTCGACAC	pNP023
KPO-1187	GAAGGTAGGGGAATAACAATATTCCGTAATGACTATGGTGAATAG	pNP023
KPO-1188	ATTGTTATTCCCCTACCTTCTC	pNP023
KPO-1189	TTATCTTCAGTGATCAAATCCAGC	pNP023
KPO-1214	GTTTTTGGTACCTTCATCACCACGGCGGATC	pNP023
KPO-1215	GTTTTTCCTAGGTTATCTTGCAAGGACGTCTGC	pNP023
KPO-1235	GTTTTTTGTCGACTGCTCTTCAGCAAGCTCAAGC	pNP017/074
KPO-1236	GTTTTTTGCATGCGTGGTACAGTAATAGACAGAG	pNP017/074
KPO-1237	GTTTTTGGTACCGGATTAGAACCCGAATTAAACT	pNP016
KPO-1324	AGAGGTACCGGTTGTTAACGGATAATGGTGCAGCTTGGTG	pNP026
KPO-1325	ATGAACCGACTTGAACTATTCAGACTGGGCGTTTG	pNP026
KPO-1326	GTTCAAGTCGGTTCATTGG	pNP026
KPO-1327	TATGCATCCTAGGCCTATTAGGTGTAGATAAAGCAAGTTTC	pNP026
KPO-1397	GATCCGGTGATTGATTGAGC	pNP018/022
KPO-1398	CGCAACTCTCTACTGTTTCTCCTAGGGGAATAACAATAGGAGTG	pNP018
KPO-1399	GCTCAATCAATCACCGGATCACCATAGTCATTACGGAATTTGC	pNP018

KPO-1409	TCGTATAATGTGTGGGCCACTGCTTTTCTTTGATGTC	pRH013
KPO-1410	ACCGGATCCTCTAGAGGTTGAGAGGGTTGCAGGG	pRH013
KPO-1417	TAGAGGTACCGGTTGTTAACGCAAAGAGTTGGAAAACCACCTTC	pNP021
KPO-1418	CCAGTAGATACGAGCACCGA	pNP021
KPO-1419	GATCTCGAACACGTTTATTGAG	pNP021
KPO-1420	CATATGCATCCTAGGCCTATTAGAAGAGCGCTCTCGATTTC	pNP021
KPO-1421	TCGGTGCTCGTATCTACTGGGACTACAAAGACCATGACGGTG	pNP021
KPO-1422	CTCAATAAACGTGTTCGAGATCTTACTATTTATCGTCATCTTTGTAGTC	pNP021
KPO-1423	TCTAGATTAAATCAGAACGCAGAAG	pRH011
KPO-1424	GGAGAAACAGTAGAGAGTTGC	pRH011
KPO-1425	GCAACTCTCTACTGTTTCTCCTAGGGGAATAACAATAGGAGTG	pRH011
KPO-1426	GCGTTCTGATTTAATCTAGAACCATAGTCATTACGGAATTTGC	pRH011
KPO-1478	CGCAACTCTCTACTGTTTCTCCGTGATTGACAGAGCTTTGAGA	pNP022
KPO-1479	GCTCAATCAATCACCGGATCTCGCCAATGAACCGACTTG	pNP022
KPO-1491	CTTTCGTCTTCACCTCGAGAATTGTGAGCGGATAACAATTGAC	synthetic sRNA library
KPO-1492	GATAAAACGAAAGGCCCAGTCTTTCGACTGAGCCTTTCG	synthetic sRNA library
KPO-1505	GTTTTTTTAATACGACTCACTATAGGGAGGGCACTGCGAGTGCTAATAGAG	*ompT* riboprobe
KPO-1506	GGTGACCAAACAAAGAGTTGG	*ompT* riboprobe
KPO-1525	GCGGCCCTCTCACTTCC	pMD241-255
KPO-1529	GGAAGTGAGAGGGCCGCGGCAAAGCCGTTTTTCCATAG	pMD241-255
KPO-1660	GTTTTTATGCATATGACCTATACCGTCCGC	pMD241-255
KPO-1681	GTTTTTATGCATGTTATGCAGTGGTATTGCAC	pNP029
KPO-1682	GTTTTTGCTAGCGGTAAGCAGCGATGCTAGA	pNP029
KPO-1683	GTTTTTATGCATAAGTTTTATCCGCACTCCAAG	pNP054
KPO-1684	GTTTTTGCTAGCAATGGCTGCACTGAGGAC	pNP054
KPO-1702	ATGCATGTGCTCAGTATCTCTATC	pNP081/084/085
KPO-1703	GCTAGCGGATCCGCTGG	pNP081/084/085
KPO-1704	GAGATACTGAGCACATGCATACGAAAATGGCTGAGCCATC	pNP085
KPO-1712	GAGATACTGAGCACATGCATATGATTGCTAATGTGTGCCGCA	pNP084
KPO-1716	GAGATACTGAGCACATGCATGCGGTGAAACCAAGCGTTTAAC	pNP062
KPO-1717	GAGCCAGCGGATCCGCTAGCTGGTAGCGCAATCAGCAGAC	pNP062
KPO-1718	GAGATACTGAGCACATGCATAATAAAATGTGAAACACAGGTAAAAATAG	pNP059
KPO-1719	GAGCCAGCGGATCCGCTAGCCGGTGCTGCATTTGCTGATAAAG	pNP059
KPO-1826	GAGCCAGCGGATCCGCTAGCAGCAAAAAGTAACGTCGCTGAA	pNP081
KPO-1831	GAGATACTGAGCACATGCATGACAAAAAGGTGATCTGGCTC	pNP081
KPO-1840	GTTTTTATGCATGCTCATGCAAGTAGTGGTGTC	pNP044
KPO-1841	GTTTTTGCTAGCCTGAACTCGCTCAATCAACAC	pNP044
KPO-1846	GTGTGTATGGAAGGCCCTAATC	*ushA* qRT-PCR
KPO-1847	CACTCGTAAGCTTGAACAATGTAAG	*ushA* qRT-PCR
KPO-1850	TGCCGGAGAGAAAGACAAATC	*oppA* qRT-PCR
KPO-1851	ACCCATCATCATCACGAAGTAAG	*oppA* qRT-PCR
KPO-1852	CTGAGCAAGAACTGAAAGAACAAG	*pal* qRT-PCR
KPO-1853	AGCTAGCATTGCTTCGTAGTC	*pal* qRT-PCR
KPO-2193	AGAGGTACCGGTTGTTAACGCACTGTCTGATGAACTGATCTTC	pNP076
KPO-2194	TGAAGCATGTAAAAAGGGAGTTAACTGTATCACCATACTACCTC	pNP076
KPO-2195	CTCCCTTTTTACATGCTTCACAG	pNP076
KPO-2196	TATGCATCCTAGGCCTATTACGATCTTTGCGTTGATATTCAGG	pNP076
KPO-2297	GTCTGATGCACTACACGATTCT	*ompT* qRT-PCR
KPO-2298	GCTAGCTCTTGCTTTGCATTATC	*ompT* qRT-PCR

KPO-2311	GACCACTCGTTTTTCTTAGAAGACTCTAAGAAGG	pRG001
KPO-2312	TCTAAGAAAAACGAGTGGTCCCACACATTATACG	pRG001
KPO-2313	CAATTACGCTCGTTTTTCCTTTTTATTAACTCCTATAG	pRG002
KPO-2314	AGGAAAAACGAGCGTAATTGGTGACAGCG	pRG002
KPO-2315	CTATAGAAGTGTACGCCCAAAGCCAGATTG	pRG004
KPO-2316	CTTTGGGCGTACACTTCTATAGGAGTTAATAAAAAG	pRG004
KPO-2378	GGTAACCCAGAAACTACCACTG	*recA* qRT-PCR
KPO-2379	CACCACTTCTTCGCCTTCTT	*recA* qRT-PCR
KPO-2418	CAACGAGTGGTTTTCATCAGTTCAAAGGTATGAC	pRG008
KPO-2419	GAACTGATGAAAACCACTCGTTGAAATGTCGTTG	pRG008
KPO-2426	CCATATTAAGAAAAGCGAGTGGATTAAC	pRG009
KPO-2427	CACTCGCTTTTCTTAATATGGGAATTCC	pRG009
KPO-2503	GTTTTTGCATGCCAAGCGATTAACATCACATTTTCTCG	pNP075
KPO-2504	GTTTTTGTCGACGTCTATTCAGACTGGGCGTTTG	pNP075
KPO-3328	GATGCGGTTGATTGGCTTAAA	*vca0447* qRT-PCR
KPO-3329	CCGTGTAGTCGTACCTATTTGTC	*vca0447* qRT-PCR
KPO-3330	TTCAGGGTAAGGTGGCTTTG	*vca0845* qRT-PCR
KPO-3331	GCGAGCAGCAGACTAAAGAT	*vca0845* qRT-PCR
KPO-3332	GACCGCCTATGTCTTGATGTT	*vca0789* qRT-PCR
KPO-3333	GTGTAGAGCCGATCAAGGTATT	*vca0789* qRT-PCR
KPO-3334	CAACAACGCATGCCCAATAC	*vc1743* qRT-PCR
KPO-3335	GGAGCCATTCGAGCATTTCTA	*vc1743* qRT-PCR
KPO-3336	CCAAGCAAAGATCTGACCAAAG	*vca0966* qRT-PCR
KPO-3337	CGCGTATTTCTTCACGCTTATG	*vca0966* qRT-PCR
KPO-3338	GAAGCCATTCTTGGTGCTAAC	*vca0951* qRT-PCR
KPO-3339	TCTCGTTCATAAGTGCCAGAG	*vca0951* qRT-PCR
KPO-3340	GTTTTTATGCATGTTTTTTTGAACTTTCCTTATCATCC	pNP078
KPO-3341	GTTTTTGCTAGCTGAAGACTCAGGGGTATAAGTG	pNP078
KPO-3346	GTTTTTATGCATTACTATCACCGGTAAATGATTAATC	pNP079
KPO-3347	GTTTTTGCTAGCGTTATTGCCAGCGTTATTACCAA	pNP079
KPO-3348	GTTTTTATGCATGAACTTGAAGCTCTTCCGCAA	pNP080
KPO-3349	GTTTTTGCTAGCCATCACTTGGTAGAGTGCCG	pNP080
KPO-3350	GTTTTTATGCATACATCAAAAACATCCCTTGAGGAA	pNP077
KPO-3351	GTTTTTGCTAGCGTTAGCACCAAGAATGGCTTC	pNP077
KPO-3360	GTTTTTATGCATACTAGTATGGAAAAATACGCCGAC	pNP082
KPO-3361	GTTTTTGCTAGCGCAACCAAATAACAGGGATAACG	pNP082
KPO-3362	GTTTTTATGCATAAATACTTTACATATGGATATGTACTATG	pNP083
KPO-3363	GTTTTTGCTAGCGCCTAAATCAATGGGTGTTTGAG	pNP083
KPO-3364	GTTTTTATGCATGTCCATATTTTAATTTTCGATAAGTATAG	pNP086
KPO-3365	GTTTTTGCTAGCTGCAGTGGTAGCTTCATCAGG	pNP086
KPO-3418	CCAGCGGATCCGCTAGCAACATTCGCTGCAAAAGAGGTG	pNP084
KPO-3419	CCAGCGGATCCGCTAGCAGCGTAAGCGCCAGTAGC	pNP085
KPO-3420	AGCGTCTTCCGTACTTCTACT	*lpp* qRT-PCR
KPO-3421	GCTGACCTGAGTGCTGATTT	*lpp* qRT-PCR
KPO-3422	CAGTGTCACCGAAAGTGTAGAT	*ompU* qRT-PCR
KPO-3423	CTGTTGACGCAATGGGTAATG	*ompU* qRT-PCR
KPO-3424	CTGAACTGGATGACTGGCTTAC	*vc1565* qRT-PCR
KPO-3425	CCACTTTGACTCTGCTGCTTAG	*vc1565* qRT-PCR
KPO-3426	CCAATTCGCTGCCTTTGATTAC	*prvT* qRT-PCR

KPO-3427	CCTGTGTACTGGGTGTCATATTC	*prvT* qRT-PCR
KPO-3428	AAACGGCGCACCATAGAA	*dsbA* qRT-PCR
KPO-3429	CGTAAGCCACCGAAAGATGA	*dsbA* qRT-PCR
KPO-3464	CCTCGTAACTCAAGCCATCAA	*rpoE* qRT-PCR
KPO-3465	ATCGAACCCGGAGAACATTAC	*rpoE* qRT-PCR
KPO-3466	CATCACGGTACGCTTCCATAA	*vc2240* qRT-PCR
KPO-3467	GCATGGTGCCATTTACTTTCC	*vc2240* qRT-PCR
KPO-3468	GTACTCATTGACCGAAGGTGAG	*dsbD* qRT-PCR
KPO-3469	AGGCAGCGGAACAGATAAAG	*dsbD* qRT-PCR
KPO-3470	TTCCATGCACGGGTATATAAGG	*vc1485* qRT-PCR
KPO-3471	GACGTGACAACGTATCGTAGAA	*vc1485* qRT-PCR
KPO-3472	CTGGAATTCAGGGATCACTAGC	*ompA* qRT-PCR
KPO-3473	CAGCTAAAGGTCTAGGCGAAAG	*ompA* qRT-PCR
KPO-3474	GTGCTGACCTTCACCTTCTT	*vc0429* qRT-PCR
KPO-3475	GCTCGACAATCTGCTCTAACT	*vc0429* qRT-PCR
KPO-3478	CGAAGCACAACCTCAAGAAAC	*btuB* qRT-PCR
KPO-3479	TCGATATCTTGGCGGGTAATG	*btuB* qRT-PCR
KPO-3480	CTCCAGCGTTACCCAAACA	*bamD* qRT-PCR
KPO-3481	AGAAATCTGCGGTCGCTAAA	*bamD* qRT-PCR
KPO-3484	TTTAGGCTAACAGCGTCACTT	*acfA* qRT-PCR
KPO-3485	GCAAATGCAGCACCGTATATT	*acfA* qRT-PCR
KPO-3562	CCTTCTTAAGGAGTTCTCTATGAAC	pNP087
KPO-3563	CTTAAGAAGGTAAGTCGGTGTTATTG	pNP087
KPO-4040	AGAGGTACCGGTTGTTAACGCACTGCTAAACCATGACTCAAG	pEE001
KPO-4041	CATCAAGATTCAATCTACAAAGGC	pEE001
KPO-4042	TTGTAGATTGAATCTTGATGAGCCTTTCGGTTATTATTTTGTCAC	pEE001
KPO-4043	TATGCATCCTAGGCCTATTAGTGCAATGATCTTGGGTGATG	pEE001
KPO-4110	AGAGGTACCGGTTGTTAACGGGAAATACCATGAAAAAGCTAGC	pNP089
KPO-4111	TTCAGTAACTTGGTACTGGAATTC	pNP089
KPO-4112	TCCAGTACCAAGTTACTGAAGACTACAAAGACCATGACGGTG	pNP089
KPO-4113	ATCTTGATGATTACCGTAAATTACTATTTATCGTCATCTTTGTAGTC	pNP089
KPO-4114	TTTACGGTAATCATCAAGATTCAATC	pNP089
KPO-4115	TATGCATCCTAGGCCTATTAAACCATGACTCAAGTCCATGC	pNP089
KPO-4308	GTGCTCAGTATCTTGTTATCCGCTC	pNP088
KPO-4309	GATAACAAGATACTGAGCACTTTCTTTGATGTCCCCATTTTGTGGAG	pNP088
KPO-4356	GCTCCACAAAATGGGGAC	rybB-scaffold oligoprobe
KPO-4962	AGAGGTACCGGTTGTTAACGCACTCGATTTTTGTTATCACCAG	pNP090
KPO-4963	TATGCATCCTAGGCCTATTAGTGCAATGATCTTGGGTGATG	pNP090
KPO-5122	ATTATAGCAGCAAATTTCTTCATGGTATTTCCTTTTTTCTTTATG	pNP091
KPO-5123	GAAGAAATTGGCTGCTATAATTTCAGCGACGTTACTTTTTGC	pNP091

APPENDIX SUPPLEMENTARY MATERIALS AND METHODS

Plasmid construction

The plasmids used in this study are listed in Table S4, used DNA oligonucleotides are listed in Table S5. The plasmids pNP074, pNP075 and pNP017 were obtained by amplification of the promotor regions of *micV* and *vrrA* from KPS-0014 chromosomal DNA, using the oligonucleotides KPO-1235/1236 and KPO-2503/2504, respectively. The promotor inserts were digested using SphI and SalI restriction enzymes and ligated into an equally treated pYH010 backbone (pNP074, pNP075) or a pCMW-1K backbone (pNP017). The inserts for the sRNA expression plasmids pRH001, pRH013 and pNP002 were obtained by amplification with KPO-1082/1083, KPO-1409/1410 or KPO-0999/1000, respectively. Fragments were introduced into linearized pEVS plasmid backbones (KPO-0092/1023) using XbaI restriction (pRH001, pNP002) and ligation, or Gibson assembly (pRH013). The plasmid pRH011 was generated via Gibson assembly using linearized pBAD5A backbone (KPO-1423/1424) and a KPO-1425/1426 amplified *rpoE* insert, obtained from KPS-0014 chromosomal DNA. The *micV, vrrA and rpoE* fragments were PCR amplified from KPS-0014 chromosomal DNA using the primer pairs KPO-0999/1237, KPO-1398/1399 or KPO-1478/1479, respectively. KpnI restriction of the *micV* fragment and a KPO-0196/0640 linearized pBAD5K backbone, yielded pNP016. Gibson assembly of the *vrrA* and *rpoE* fragments with KPO-0196/1397 linearized pBAD5K backbone yielded pNP022 and pNP018, respectively. Plasmid pKP431 was cloned by PCR amplification of the *hfq* flanking regions with KPO-0012/0066 and KPO-0019/0067, thereby introducing the 3xFLAG tag with primer overhangs. The resulting fragments were fused via overlap PCR using KPO-0148/0149, and introduced into pKAS32 backbone using KpnI and AvrII restriction sites. The plasmids pNP023 and pNP024 were constructed by amplification of insert fragments using KPO-1186/1187 and KPO-1188/1189, or KPO-1180/1181 and KPO-1182/1183, respectively. The inserts were fused using overlap PCR with the oligonucleotides KPO-1214/125 (pNP023) or KPO-1184/1185 (pNP024), digested with KpnI and AvrII and ligated into an equally digested pKAS32 backbone. The plasmids pNP026, pNP076, pEE001, pNP021 and pNP089 were constructed by Gibson assembly, using a KPO-0267/0268 linearized pKAS32 backbone. The insert fragments were PCR amplified from KPS-0014 chromosomal DNA using the following oligonucleotides: pNP021 (KPO-1324/1325 and KPO-1326/1327), pNP076 (KPO-2193/2194 and KPO-2195/2196), pEE001 (KPO-4040/4041, KPO-4042/4043), pNP021 (KPO-1417/1418, KPO-1419/1420 and KPO-1421/1422 amplified from KPS-0995 chromosomal DNA), pNP089 (KPO-4110/4111, KPO-4114/4115 and KPO-4112/4113 amplified from KPS-0995 chromosomal DNA). GFP fusions were cloned as described previously (Corcoran et al., 2012) and employing previously determined transcriptional start site annotations (Papenfort et al., 2015). Briefly, *acfA* (pNP059), *bamD* (pNP082), *btuB* (pNP029), *dsbD* (pNP079), *lpp* (pNP086), *ompA* (pNP081), *ompT* (pKP465),

ompU (pNP085), *oppA* (pNP084), *pal* (pNP062), *prvT* (pNP083), *rpoE* (pNP044), *ushA* (pNP054), *vc1485* (pNP080), *vc1563* (pNP078), *vca0951* (pNP077) inserts for translational reporters were PCR amplified using the primers indicated in Table S6 and introduced into pXG10-1C backbones using NheI, NsiI restriction sites or Gibson assembly. The pMD030 plasmid was constructed by restriction digest of pFM1-1 with XbaI and XhoI, yielding the P_L-*rybB* fragment and insertion into an equally treated pEVS backbone. pNP088 was obtained by site-directed mutagenesis PCR using KPO-4308/4309 and the parental plasmid pMD251 as a template. pMD241-255 plasmids were obtained by sequencing plasmids derived from EtOH resistant colonies. The Cm^R resistance cassettes were replaced with Kan^R cassettes using linearization with KPO-0282/1529 and amplification of the KanR cassette from the pCMW-1K plasmid using KPO-1160/1525. The plasmid pNP089 was generated by amplification of the insert from KPS-0014 chromosomal DNA using KPO-4962/4963 and Gibson assembly with KPO-0267/0268 linearized pKAS32 backbone. Quickchange PCR using pNP090 as template, and KPO-5122/5123 yielded pNP091. Mutations for compensatory base pair exchanges were introduced using the oligonucleotides listed in Table S5, and the respective parental plasmids as a template.

V. cholerae **strain construction**

A complete list of strains used in this study is provided in Table S3. *V. cholerae* C6706 was used as the wild-type strain in this study. *V. cholerae* mutant strains were generated as described previously (Papenfort et al., 2017). RK2/RP4-based conjugal transfer was used to introduce plasmids into *V. cholerae* from *E. coli* S17λpir plasmid donor strains. Subsequently, transconjugants were selected using appropriate antibiotics, and polymyxin B to specifically inhibit *E. coli* growth.

Identification of σ^E-dependent promoters in *V. cholerae*

To detect σ^E-dependent promoters in *V. cholerae*, the promotor sequence of 60 σ^E-dependent genes of *E. coli* (Mutalik et al., 2009) were used to construct a motif with MEME (Bailey et al., 2015). The motif was searched with FIMO (Bailey et al., 2015) in the genome sequence of *V. cholerae* (Accession NC_002505, NC_002506) accepting only hits with a p-value below 0.0001. The 626 motif matching sites in *V. cholerae* were filtered by proximity to transcription start site (Papenfort *et al.,* 2015) and only those with a maximal distance of 50 nt were reported in Table S1. A Unix shell script that represents the analyses has been deposited at Zenodo (https://doi.org/10.5281/zenodo.2543422)

In silico analyses

Genomic loci encoding *micV* in various *Vibrio* strains were analyzed for gene synteny using SyntTax (Fig. EV1A) (Oberto, 2013). The following strains were used for analysis: *V.ch*., *Vibrio cholerae* (NCBI:txid243277); *V.vu.*, *Vibrio vulnificus* (NCBI:txid914127); *V.co., Vibrio coralliilyticus* (NCBI:txid1384040); *V.tu*., *Vibrio tubiashii* (NCBI:txid1051646); *V.ha*., *Vibrio harveyi* (ATCC:33843); *V.an*., *Vibrio anguillarum* (NCBI:txid882102); *V.al*., *Vibrio alginolyticus* (NCBI:txid1219076). To generate the alignment of *micV* sequences (Fig. 1A), the following strains were used: Vch, *Vibrio cholerae* (NCBI:txid243277); Vfu, *Vibrio furnissii* (NCBI:txid903510); Vvu, *Vibrio vulnificus* (NCBI:txid216895); Van, *Vibrio anguillarum* (NCBI:txid882102); Vsp, *Vibrio splendidus* (NCBI:txid575788); Vex, *Vibrio* sp. Ex25 (NCBI:txid150340); Vpa, *Vibrio parahaemolyticus* (NCBI:txid223926); Vej, *Vibrio* sp. EJY3 (NCBI:txid1116375); Asa, *Aliivibrio salmonicida* (NCBI:txid316275); Afi, *Aliivibrio fischeri* (NCBI:txid312309). To generate the alignment of *vrrA* sequences (Fig. EV2A), the following strains were used: Vch, *Vibrio cholerae* (NCBI:txid243277); Vco, *Vibrio coralliilyticus* (NCBI:txid1384040); Vvu, *Vibrio vulnificus* (NCBI:txid672); Val, *Vibrio alginolyticus* (NCBI:txid1219076); Vsp, *Vibrio splendidus* (NCBI:txid575788).

APPENDIX SUPPLEMENTARY REFERENCES

Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H (2006) Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* 2: 2006 0008

Bailey TL, Johnson J, Grant CE, Noble WS (2015) The MEME Suite. *Nucleic Acids Res* 43: W39-49

Bouvier M, Sharma CM, Mika F, Nierhaus KH, Vogel J (2008) Small RNA binding to 5' mRNA coding region inhibits translational initiation. *Mol Cell* 32: 827-37

Chao Y, Papenfort K, Reinhardt R, Sharma CM, Vogel J (2012) An atlas of Hfq-bound transcripts reveals 3' UTRs as a genomic reservoir of regulatory small RNAs. *EMBO J* 31: 4005-19

Corcoran CP, Podkaminski D, Papenfort K, Urban JH, Hinton JC, Vogel J (2012) Superfolder GFP reporters validate diverse new mRNA targets of the classic porin regulator, MicF RNA. *Mol Microbiol* 84: 428-45

Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci U S A* 97: 6640-5

Dunn AK, Millikan DS, Adin DM, Bose JL, Stabb EV (2006) New rfp- and pES213-derived tools for analyzing symbiotic Vibrio fischeri reveal patterns of infection and lux expression in situ. *Appl Environ Microbiol* 72: 802-10

Egler M, Grosse C, Grass G, Nies DH (2005) Role of the extracytoplasmic function protein family sigma factor RpoE in metal resistance of Escherichia coli. *J Bacteriol* 187: 2297-307

Herzog R, Peschek N, Fröhlich KS, Schumacher K, Papenfort K (2019) Three autoinducer molecules act in concert to control virulence gene expression in Vibrio cholerae. *Nucleic Acids Research*: gky1320-gky1320

Mutalik VK, Nonaka G, Ades SE, Rhodius VA, Gross CA (2009) Promoter strength properties of the complete sigma E regulon of Escherichia coli and Salmonella enterica. *J Bacteriol* 191: 7279-87

Oberto J (2013) SyntTax: a web server linking synteny to prokaryotic taxonomy. *BMC Bioinformatics* 14: 4
Papenfort K, Bouvier M, Mika F, Sharma CM, Vogel J (2010) Evidence for an autonomous 5' target recognition domain in an Hfq-associated small RNA. *Proc Natl Acad Sci U S A* 107: 20435-40
Papenfort K, Forstner KU, Cong JP, Sharma CM, Bassler BL (2015) Differential RNA-seq of Vibrio cholerae identifies the VqmR small RNA as a regulator of biofilm formation. *Proc Natl Acad Sci U S A* 112: E766-75
Papenfort K, Pfeiffer V, Mika F, Lucchini S, Hinton JC, Vogel J (2006) SigmaE-dependent small RNAs of Salmonella respond to membrane stress by accelerating global omp mRNA decay. *Mol Microbiol* 62: 1674-88
Papenfort K, Silpe JE, Schramma KR, Cong JP, Seyedsayamdost MR, Bassler BL (2017) A Vibrio cholerae autoinducer-receptor pair that controls biofilm formation. *Nat Chem Biol* 13: 551-557
Simon R, Priefer U, Pühler A (1983) A Broad Host Range Mobilization System for In Vivo Genetic Engineering: Transposon Mutagenesis in Gram Negative Bacteria. *Bio/Technology* 1: 784
Skorupski K, Taylor RK (1996) Positive selection vectors for allelic exchange. *Gene* 169: 47-52
Svenningsen SL, Tu KC, Bassler BL (2009) Gene dosage compensation calibrates four regulatory RNAs to control Vibrio cholerae quorum sensing. *EMBO J* 28: 429-39
Thelin KH, Taylor RK (1996) Toxin-coregulated pilus, but not mannose-sensitive hemagglutinin, is required for colonization by Vibrio cholerae O1 El Tor biotype and O139 strains. *Infect Immun* 64: 2853-6

6.3 Seed sequences of the top 50 enriched sRNA variants from chapter 3

#	sequence	#	sequence
1	GAGAGCCTA	26	CGCCGAACC
2	GCCGCAACC	27	CGCCAACTT
3	AAAATGGTA	28	TTATGGTGT
4	GCCGCTCCT	29	ACAGCTAGC
5	GCCTCTGGC	30	GCTAGCCGT
6	GCCGCTCGG	31	GTCGCTCCT
7	GCTGCGGTT	32	TACTTGGTA
8	TGGTATTCC	33	GGATAGCTT
9	AGAGAGCTT	34	GGTAATTCC
10	TGCTCGGCT	35	ATGTTAGCT
11	GCCACTGCT	36	GGTATTTCC
12	GCCGTACCT	37	TGGTAGCTT
13	GGAGAGCTT	38	AACTTGGTA
14	TAGCTAGCC	39	GCAATTAGC
15	TACGCTAGC	40	GTATTTCCT
16	TTGCCGCCT	41	GCCACTCGC
17	AGCATGGTG	42	GTTGCTAGG
18	CGCCACTGG	43	TGTTTCCTT
19	CGCCGGAGC	44	GACCTTACC
20	ATCGCTAGC	45	GACATGGTG
21	AGATGGTAT	46	GCCCAACTT
22	CAGCTAGTC	47	TGCTACTCC
23	AGGTATTCC	48	AACATGGTG
24	GACTTGGTA	49	AGCTTGGTA
25	AACGTTAGC	50	GTCGTTAGG

Table 6.1: Seed sequences of the top 50 enriched sRNA variants

Acknowledgements

Many people have accompanied me along the way of my PhD and I would like to express my great appreciation for all their help and support.

First and foremost, many thanks to Kai for always being a very engaged and enthusiastic supervisor during these years. I highly appreciated our scientific discussions, your critical advice and your profound interest in my ideas. It was also a fascinating experience to see a lab grow from the start and to help shaping parts of it.

My sincere gratitude goes to Prof. Dr. Kirsten Jung, who generously agreed to be the second examiner of this thesis, and to PD Dr. Bettina Bölter and Prof. Dr. John Parsch from my thesis committee for their time and interest. Special thanks also goes to Prof. Dr. Marc Bramkamp and Dr. Sabine Schneider for their guidance and support as members of my TAC.

Furthermore, I would like to thank Prof. Dr. Konrad Förstner for our long-term TIER-seq collaboration and my undergraduate students Steffi, Caroline, Kateryna, Alex and Lisa for their valuable contributions to various projects.

Belonging to the GRK2062 was a great part of my PhD experience and I'm very grateful for manifold kinds of support and financial aid. Many thanks to Dr. Beate Hafner for her patience and organizational help and thanks to the whole GRK crowd for many inspiring late night science meetings, especially to Aurore, Karsten, Mattheus, Martin, Marcel and Philipp.

Of course a big thanks goes to all my past and present colleagues from the Papenfort group for sharing parts of this exciting PhD journey, for a nice lab atmosphere, for fun times outside the lab and for some great collaborations. Thanks to Roman and Niko for all the great times together since our iGEM days, thanks to Micha, Kavyaa and Daniel for being awesome office mates, thanks to Kathi for being always interested in our *Vibrio* projects and giving some very helpful postdoc advice, thanks to Andreas for being a great lab manager and thanks to Yannik and Alex for all your positivity and motivation. I am curious to see where all our lives will lead us and I hope our paths will cross again in the future.

Finally, these years wouldn't have been possible without the continuous support from my family and friends. You were always there to cheer me up, to encourage me to push through and to finally celebrate the happy ending. A special thanks goes to my little niece Johanna for spreading so much joy and to my wonderful husband Daniel, who is a never-ending source of optimism and strength. I could not wish for more than to have you by my side.

Curriculum Vitae

Removed for protection of privacy.

Aus Datenschutzgründen entfernt.