Regulatory role of small RNAs and RNA-binding proteins in carbon metabolism and collective behaviour of *Vibrio cholerae*

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

vorgelegt von
Kavyaa Venkat

aus Bengaluru, Indien
München, September 2021

Diese Dissertation wurde angefertigt
unter der Leitung von Prof. Dr. Kai Papenfort
im Bereich von Department Biologie I
an der Ludwig-Maximilians-Universität München

Erstgutachter: Prof. Dr. Kai Papenfort
Zweitgutachter: Prof. Dr. Simon Ringgaard

Tag der Abgabe: 02.09.2021
Tag der mündlichen Prüfung: 03.02.2022

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. I also declare that I have not submitted or defended a dissertation previously without success.

München, 02.09.2021

Kavyaa Venkat

Contents

List of abbreviations

% (v/v)	% (Volume/volume)
% (w/v)	% (Weight/volume)
↑	Upregulated (of genes)
↓	Repressed (of genes)
°C	Degree Celsius
A	Adenosine
aa	Amino acid
AI	Autoinducer
AIP	Autoinducer peptide
ANOVA	Analysis of variance
Ara	L-arabinose
ATP	Adenosine triphosphate
AU	Arbitarary units
bp	Base pair
C	Cytosine
cAMP	Cyclic adenosine monophosphate
CCR	Carbon catabolite repression
cDNA	Complementary DNA
CDS	Coding sequencing
Ci	Curie
CIP	Calf intestinal phosphate
CLIP-seq	Cross-linking and immunoprecipitation followed by sequencing
Cm	Chloramphenicol
CRP	c-AMP receptor protein
CS	Citrate synthase
CT	Cholera toxin
Ctrl	Control
Da	Dalton
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide
DTT	Dithiothreitol
EDTA	Ethylene diamine tetraacetic acid
EMSA	Electrophoretic mobility shift assay
ETS	External transcribed spacers
FDR	False discovery rate
Fur	Ferric uptake regulator
G	Guanosine
gDNA	Genomic DNA
HCD	High cell density
HGT	Horizontal gene transfer
IGR	Intergenic region
IP	Immunoprecipitation
IPTG	Isopropyl beta-D-1-thiogalactopyranoside
iTIS	Internal translation initiation sites
LB	Lennox-broth
LCD	Low cell density
LC-MS	Liquid chromatography tandem mass spectromety
M	Molar or size marker
min	Minute
mRNA	Messenger RNA
N	Any nucleotide (A, T, G or C)
NADH	Nicotinamide adenine dinucleotide with hydrogen
NAG	N-acetly glucosamine
nt	Nucleotide
OD	Optical density
OMP	Outer membrane porin
OOF	Out-of-frame
ORF	Open reading frame
P	Phosphate group
P:C:I	Phenol:chloroform:isoamylalcohol
PAA	Polyacrylamide
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffer saline
PCR	Polymerase chain reaction
PNK	Polynucleotide kinase
PTS	Phosphoenolpyruvate transferase system
qRT-PCR	Quantitative real-time PCR
QS	Quorum sensing
RBP	RNA-binding protein
RBD	RNA-binding domain
RBS	Ribosome binding site
RIP-seq	RNA immnuoprecipitation followed by sequencing
RNA	Ribonucleic acid
RNAP	RNA polymerase
RNase	Ribonuclease
RNA-seq	RNA sequencing
rpm	Revolutions per minute
RRM	RNA recognition domain
rRNA	Ribosomal RNA
s	Seconds
SAM	S-adenosyl methionine
SD	Shine-Dalgarno or standard deviation
SDS	Sodium dodecyl sulfate
SEM	Standard error of means
sfGFP	Superfolder green fluorescent protein
SPA	Sequential peptide affinity
sRNA	Small regulatory RNA
SSC	Saline-sodium citrate
t	Time
T	Thymidine
TA	Toxin / anti-toxin
T3SS	Type III secretion system
TCP	Toxin coregulated pilus
Tet	Tetracyclin
TSS	Transcriptional start site
U	Unit or Uridine
UTP	Uridine triphosphate
UTR	Untranslated region
vol	Volume
VPI	Vibrio pathogenecity island
σ	Sigma (factor)

Summary

The importance of small regulatory RNAs (sRNAs) has been recognized across all domains of life. Originally considered "non-coding RNAs," several bacterial sRNAs have been found to encode functional proteins that are under 50 amino acids long. This group of regulators are called dual-function regulators. To date, only five such regulators have been characterized in bacteria. In the primary study, the first dual-function RNA of *Vibrio cholerae* was discovered and characterized. The pathogen colonizes and infects the upper intestines by producing two key virulence determinants – toxin co-regulated pilus (TCP) and cholera toxin (CT). While all the known sRNAs of *V. cholerae* act directly or indirectly to regulate the production of TCP, the sRNA VqmR is the only known direct repressor of CT production to date. Therefore, a forward genetic screen was employed to score for CT repression. This screen identified another promising candidate called Vcr082. Interestingly, Vcr082 also encodes 29 amino acids long ORF and hence was re-named VcdRP, for *V. cholerae* dual RNA regulator and protein, eponymous to their roles. The dual regulator is controlled by the global transcription factor of carbon utilization, cAMP-CRP. The riboregulatory component is conserved at the 3′ end of the dual regulator. By employing a conserved stretch of four cytosines, VcdR base-pairs with and represses mRNAs that encode for transporters that import PTS sugars. Additionally, VcdR also downregulates the phosphor-carrier proteins PtsH and PtsI that are involved in the phospho-relay during glycolysis. The small protein, VcdP exerts its regulatory role by interacting with and accelerating the activity of citrate synthase enzyme, opening the gateway into the TCA cycle. This way, both VcdR and VcdP act to block sugar uptake and modulate the flux through the TCA cycle, thereby striking a balance to maintain overall carbon metabolism in *V. cholerae*.

The diverse environments that *V. cholerae* inhabits necessitates that the organism rapidly perceives changes in its external environment and appropriately tailors its gene expression paradigm. To achieve this, the bacteria employ quorum sensing (QS) to communicate and coordinate a suitable response. While this mechanism of census taking has been well-documented early on in several marine bacteria, more recent studies have identified additional QS systems in *V. cholerae*. Similarly, while biofilm formation has been extensively studied, the transition into and subsequent dispersal was only documented recently. These incomplete underpinnings thereby prompted further investigation of the QS pathway. Therefore, in the second study, a forward genetic screen in a *V. cholerae* mutant library was employed to score for an altered QS phenotypic transition. This screen identified a novel RNA-binding protein called MbrA (membrane-bound RNA-binding protein A). This protein localizes to the membrane and contains two trans-membrane domains at the N-terminus and a conserved RNA recognition motif-type RNA-binding domain located towards the C-terminus. MbrA is activated by the global transcription factor cAMP-CRP and a subsequent transcriptome analysis revealed its role in the regulation of motility genes and flagellar assembly complex in *V. cholerae*.

"Sometimes you climb out of bed in the morning, and you think, I'm not going to make it, but you laugh inside — remembering all the times you've felt that way."

— Charles Bukowski

Chapter 1

Introduction

Tight control of gene expression is of utmost importance to all living organisms. In particular, pathogenic bacteria have to alter their gene expression patterns based on their surrounding niche. Small regulatory RNAs (sRNAs) are the largest class of post-transcriptional regulators in bacteria and have emerged as key players in the spatio-temporal regulation of gene expression. sRNAs often act in association with RNA-binding proteins (RBPs). Three global bacterial RBPs are Hfq, CsrA and ProQ, however there has been a growing list of predicted RBPs that need to be validated. Elucidating the role of one such putative RBP is the focus of this study. sRNAs are known to regulate a plethora of regulatory processes including sugar uptake, central metabolism, and virulence. Although they are non-coding, increasing evidence has shown that some sRNAs also harbor an open reading frame (ORF) that encode proteins. These are referred to as dual-function regulators. To date, only a handful of these regulators have been characterized in bacteria including only one known dual-regulator called SgrST in Gram-negative bacteria. Therefore, this thesis also focuses on the discovery and characterization of a second Gram-negative regulator and its role in balancing carbon metabolism and virulence.

1.1 The RNA-binding proteins driving sRNA-mediated regulation of genes

The sRNAs are the most abundant class of post-transcriptional regulators in bacteria and are typically 50-500 nucleotides (nt) long. These non-coding molecules are highly structured, comprising numerous stem loops (Waters & Storz, 2009; Fu *et al*, 2019). Base-pairing sRNAs are classified based on their genomic location with respect to the target gene(s) (Waters & Storz, 2009). *cis*-encoded sRNAs are encoded on the antisense strand of their target genes and therefore have extensive sequence complementarity with each other (Brantl, 2002, 2007). In contrast, *trans*-encoded sRNAs only share partial complementarity to their target gene(s) and therefore need RNA-binding proteins (RBPs) to serve as chaperones to facilitate base-pairing (Gottesman & Storz, 2011)

The RBPs are a diverse class of proteins ubiquitously found in all living organisms that control the fate and function of the bound RNA (Andresen & Holmqvist, 2018). The capacity of these proteins to recognize and bind to RNA molecules stems from the recognition of short stretch of sequences present in the RNA-binding domains (RBDs) of RBPs. Some of the well-defined domains include the S1 domain, cold shock domain, K homology domain, RNA recognition motifs (RRMs), amongst others (Holmqvist & Vogel, 2018). RBPs serve as: (a) chaperones for the interaction of mRNAs and sRNAs, (b) modulators of the ribosome binding site (RBS) accessibility or (c) recruiters of ribonucleases to initiate transcript decay (Van Assche *et al*, 2015).

While the number of eukaryotic RBPs identified by global screening approaches have been categorically catalogued, for example at the ENCODE Project Consortium (ENCODE Project Consortium, 2012), our understanding of bacterial RBPs have largely been attributed to serendipitous discoveries (Quendera *et al*, 2020). This lack of knowledge emanates from the absence of two important features in bacteria when compared to eukaryotic genomes – the lack of poly(A) tails on transcripts and inefficient incorporation of crosslinking-enhancing artificial nucleotides (Hör *et al*, 2018). However, this has also led to the development of innovative RNA-centric strategies to overcome these challenges such as Grad-Seq (gradient profiling by sequencing) and GradR (glycerol gradient sedimentation with RNase treatment and mass spectrometry) to study the bacterial RNA-protein landscape (Smirnov *et al*, 2016; Hör *et al*, 2020a, 2020b; Gerovac *et al*, 2020). Three global bacterial RBPs that are currently known are described below (Figure 1).

1.1.1 The versatile roles of Hfq

Hfq was first discovered nearly 60 years ago in *Escherichia coli* as an essential host factor of the RNA bacteriophage Qß. At the time, it was already speculated that this factor may contribute to changes in RNA secondary structure, thus improving the replication efficiency of the phage genome (Franze de Fernandez *et al*, 1968). About a half a decade later, biochemical characterization of Hfq in *E. coli* revealed a strong preference for AU-rich sequences (Hori & Yanazaki, 1974; Carmichael *et al*, 1975). The innate importance of Hfq beyond its role in phage replication became evident in mid-to-late 90s when a mutant of *hfq* in *E. coli* resulted in decreased growth rate, increased cell length and sensitivity to UV-light and impaired stress response, and subsequent mutagenic studies in *Brucella abortus* also implicated its role in diminished virulence (Tsui *et al*, 1994; Robertson & Roop, 1999). The pleiotropic effects of impaired stress response were linked to mutations in the *rpoS* gene, that is essential for stationary phase response, and subsequent studies in *E. coli* and *Salmonella typhimurium* suggested the direct or indirect role of Hfq in the destabilization of several RNAs in a RpoS-dependent manner (Brown & Elliott, 1996; Muffler *et al*, 1996). Phylogenetic studies in the early 2000s revealed orthologs of Hfq in approximately half of all sequenced bacteria as well as the archeon, *Methanothermobacter thermoautotrophicum* (Sun *et al*, 2002). However, one of the first associations between the ability of Hfq to bind sRNAs only unfolded in studies on OxyS and RprA regulation of *rpoS* (Zhang *et al*, 2001; Majdalani *et al*, 2001; Zhang *et al*, 2002). The now observed widespread role of Hfq in facilitating short-imperfect base-pairing interactions between sRNAs and their target mRNAs by disparate mechanisms is no longer limited to the model organisms *E. coli* and *Salmonella*.

While Hfq has been implicated to be prevalent in most prokaryotes and some archaea, a major challenge in understanding Hfq-mediated regulation is that the RNA substrates vary considerably in size, structure, and sequence motif combinations (Bouloc & Repoila, 2016). Notably, the role of Hfq in regulating mRNAs, independent of sRNAs have also been discovered. For example, the *mutS*

mRNA is regulated by the *E. coli* sRNA ArcZ, facilitated by the RNA chaperone Hfq. However, Hfq still interacts with *mutS* in the absence of ArcZ, thus preventing its translation through the formation of an inhibitory hairpin in the translation initiation region (TIR; Chen & Gottesman, 2017). Additionally, there have been reports of sRNAs in *Staphylococcus aureus* and *Bacillus subtilis* that do not rely on Hfq for the modulation of gene expression and turnover of mRNAs ((Heidrich *et al*, 2006, 1; Bohn *et al*, 2007; Heidrich *et al*, 2007). The stability of NrrF sRNA in *Neisseria meningitides* remains unchanged in the absence of Hfq. Moreover, it has been demonstrated that Hfq is expendable for NrrF-mediated regulation of *sodB* that is essential for iron-responsive gene regulation (Mellin *et al*, 2010). Recently, ChvR sRNA of the alpha-proteobacteria *Caulobacter crescentus* has been shown to modulate *chvT* expression independent of Hfq (Fröhlich *et al*, 2018).

1.1.2 CsrA and CsrA-like proteins

CsrA was initially identified as a regulator of carbon storage and glycogen biosynthesis in *E. coli* (Romeo *et al*, 1993). Its counterpart in *Pseudomonas aeruginosa* is RsmA, a repressor of stationary phase metabolites. The CsrA/RsmA family of RBPs are widely found in Gram-negative and Gram-positive bacteria with implicated roles in facilitating biofilm formation, virulence gene expression, carbon utilization and motility (Marden *et al*, 2013; Romeo & Babitzke, 2018).

This class of RBPs primarily affect translation of mRNAs by directly binding to their 5′ untranslated regions (UTRs) by recognizing A/CUGGA motifs (Dubey *et al*, 2005; Duss *et al*, 2014; Holmqvist *et al*, 2016). Typically, CsrA binding occurs in the ribosome bindig site (RBS) sequence or overlaps with the initiation codon. Similar to Hfq, CsrA/RsmA can also regulate transcript stability by controlling the access of ribonucleases (RNases; Yakhnin *et al*, 2013). Interestingly, the activity of the RBP itself is regulated by the two-component systems BarA/UvrY in *E. coli*, GacS/GacA in *P. aeruginosa*, BarA/SirA in *Salmonella* and VarS/VarA in *Vibrio cholerae*. Phosphorylation of the response regulators induces the expression of multiple sRNAs – CsrB/C/D, RsmY/Z. The sRNAs themselves are composed of repetitive sequence elements of the GGA recognition motif and consequently, have a high affinity for the RBPs, thereby act as ''sponges'' to sequester them. This feedback regulation allows to control the levels of active RBPs available in response to changing environmental cues (Timmermans & Van Melderen, 2010; Butz *et al*, 2019). While CsrB/C do not interact with the other most prominent RBP Hfq, CsrA can repress the synthesis of *E. coli* Hfq *in vivo* by binding to a site overlapping the Shine-Dalgarno (SD) sequence (Baker *et al*, 2007). However, the sRNA targetome of both Hfq and CsrA were assumed to be independent of each other, until the findings by Jørgensen *et al*, 2013 revealed the sRNA McaS that acts in both Hfq and CsrA regulons to modulate biofilm formation in *E. coli*. Yet, McaS requires only Hfq, but not CsrA for its stability. With increasing discovery of new regulatory RNAs as well as non-canonical mechanisms of post-transcriptional regulation, the general categorization of sRNAs based on initially identified functions has begun to become hazy, thus not only emphasizing on the extent of their functional importance but also solicits for the need for a systematic inventory of bacterial RNA-protein interactions.

1.1.3 The third pillar of sRNA-mediated regulation is mediated by ProQ

Well-characterized model organisms like *E. coli* and *Salmonella* harbor many sRNAs that do not associate with Hfq or CsrA. Furthermore, epsilonproteobacteria like *Camplylobacter jejuni* and *Helicobacter pylori* lack functional homologs of Hfq, while *Streptococcus pneumoniae* lacks both Hfq and CsrA, thus hinting at the presence of additional RNA-binding proteins that have been overlooked (Quendera *et al*, 2020). This prompted (Smirnov *et al*, 2016) to develop a global high-

throughput method called Grad-Seq to provide an RNA/protein complexome resource for *Salmonella*. Remarkably, this technique unveiled ProQ as a global RBP. ProQ is composed of a N-terminal FinO-like domain, a flexible linker, and a Hfq-like C-terminal domain. However, unlike Hfq and CsrA, ProQ binds to RNA in a sequence-independent manner, but shows a structural preference (Mark Glover *et al*, 2015). While the FinO family of proteins were discovered in the 1980s, ProQ was best known then as a regulator of *E. coli* proline transporter ProP (Stalmach *et al*, 1983; Milner & Wood, 1989). Its role as a global RBP only emerged based on co-immunoprecipitation, CLIP-Seq (cross-linking and immunoprecipitation followed by high throughput sequencing), dual-RNA Seq and RIL-Seq (RNA interaction by ligation and sequencing) analyses that revealed hundreds of ligands associated with ProQ (Smirnov *et al*, 2016; Melamed *et al*, 2016; Holmqvist *et al*, 2018; Westermann *et al*, 2019; Melamed *et al*, 2020). Owing to ProQ's structural preference for binding, it appears to bind to *cis*-encoded RNAs. Yet, the *Salmonella* sRNA RaiZ was the first *trans*-encoded that effectively co-precipitated with both Hfq as well as ProQ, however only relying on the latter for its stability (Chao *et al*, 2012; Smirnov *et al*, 2017). RIP-Seq (RNA immunoprecipitation followed by sequencing) analysis of Hfq in *V. cholerae* revealed OppZ as one of the 82 sRNAs that co-precipitated with the RBP (Huber *et al*, 2020). Interestingly, unpublished data from the lab also indicates OppZ strongly associates with ProQ in stationary phase, thus re-iterating the need to re-visit existing global datasets and categorically compile RNA-protein interaction maps for bacteria.

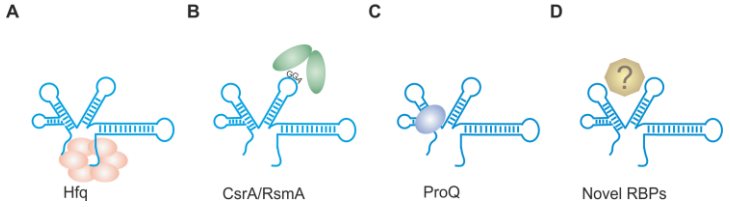

Figure 1. RNA chaperones assisting sRNAs for post-transcriptional regulation. A) Hfq is a hexameric RBP that globally binds to an assemblage of several hundred sRNAs by providing a contact match-making surface to act on their respective target mRNAs. **B)** CsrA and RsmA RBPs of *E. coli* and *P. aeruginosa* bind to GGA motifs on the mRNA, while themselves being regulated by the sequestration via the sRNAs CsrB/C and RsmY/Z, respectively. **C)** ProQ unlike Hfq and CsrA typically has a structure specificity rather than sequence specificity to regulate mRNAs and sRNAs. **D)** Novel RBPs and their mode of regulation remain to be discovered and characterized.

1.1.4 Other bacterial regulators with RNA-binding domains

sRNA-mediated networks are advantageous in rapidly and effectively reprogramming gene expression (Dutta & Srivastava, 2018). Their reliance on RBPs to facilitate their interaction with mRNAs have also been well-documented. However, the lack of functional homologs or the lack of ''typical'' recognition motifs in certain Gram-negative and Gram-positive bacteria has resulted in the surging interest in developing global high-throughput approaches in bacteria to identify new RBPs. Often these techniques have been adapted from prior research in eukaryotes, albeit with modifications incorporated to suit bacterial genomes. As a result, *E. coli* alone has about 180 annotated RBPs (Holmqvist & Vogel, 2018). The roles of a subset of RBPs in Gram-negative bacteria are discussed below.

Glucosamine-6-phosphate (GlcN6P) serves as a precursor for initiating cell envelope biosynthesis in *Enterobacteriaceae* (Göpel *et al*, 2011). The enzyme that catalyzes this step is GlmS. The regulation of *glmS* is mediated by two homologous sRNAs GlmY and GlmZ (Urban & Vogel, 2008). GlmY acts upstream of GlmZ to antagonize GlmZ inactivation, whereas GlmY in concert with Hfq activates *glmS*. The RBP RapZ (short for RNAse adaptor protein for GlmZ) was initially discovered in a phenotypic screen resulting in overproduction of GlmS protein upon its deletion (Kalamorz *et al*, 2007). RapZ binds to GlmZ in the presence of GlcN6P, resulting in the degradation of the sRNA by

RNase recruitment. However, upon depletion of GlcN6P, GlmY accumulates and sequesters RapZ, concomitantly suppressing GlmZ decay, thus ensuring homeostasis of the cell envelope (Khan *et al*, 2020; Durica-Mitic *et al*, 2020).

Analogous to *E. coli* CsrA/RsmA and their sponge CsrB/C and RsmY/Z, *Pseudomonads* encode Crc protein that is sequestered by the sRNAs CrcY/Z. Crc is an RBP that is involved in catabolite repression, and binds RNA by acting as a translational repressor to control the assimilation of carbon sources (Moreno & Rojo, 2008).

The actin-like cytoskeletal protein MreB maintains the rod-shaped morphology of many bacteria. RodZ (formerly YfgA) was identified as an anchoring protein that co-localized with MreB in the inner membrane of *E. coli* (Shiomi *et al*, 2008; Bendezú *et al*, 2009). In addition to being a cell-shape determinant, RodZ of *Shigella* has also been reported to be a bi-functional RBP that promotes virulence and mRNA decay in *Shigella sonnei*. (Mitobe *et al*, 2011).

RIP-Seq analyses of the cold shock proteins CspE and CspR of *Salmonella* revealed a myriad of associated transcripts to the tune of 20% of the bacterial genome. Transcriptome and phenotypic screens underscored the importance of these two RBPs in mediating stress resistance, motility and biofilm formation (Michaux *et al*, 2017).

A diverse family of proteins with RNA-chaperoning properties have emerged over the decades, however, the evolution of these proteins, their mode of regulation, sub-cellular localization, and the biological processes they control remain unanswered. Elucidating the role of one such putative RNA-binding protein of *V. cholerae* is the focus of this thesis.

1.2 Mechanisms and consequences of base-pairing

Since *trans*-encoded sRNAs have limited complementarity with their mRNA targets, single sRNAs can target multiple mRNAs, and single mRNAs can be a target of multiple sRNAs (Beisel & Storz, 2011). A well-established example of an sRNA that regulates many targets is RyhB (Massé & Gottesman, 2002). RyhB is repressed by the transcription factor Fur (ferric uptake regulator) in response to high concentrations of intracellular iron. The pioneering work by the laboratory of Eric Massé has shed light on the role of RyhB in maintaining iron homeostastis by regulating over 50 genes. The post-transcriptional regulation by RyhB thus serves as a model sRNA to outline the general regulatory principles employed by sRNAs to regulate gene expression. Through the various illustrations depicted in Figure 2A-E, this section not only aims to give an overview of general mechanisms employed by sRNAs to modulate gene expression, but importantly, emphasize the effect of one sRNA with many modes of action.

1.2.1 The canonical mRNA repression involving coupled degradation

Base-pairing by sRNAs usually occurs in the translation initiation region (TIR) of the target mRNA, in a region encompassing the RBS and the translation start codon (Rice *et al*, 2012). Hfq-dependent sRNAs like RyhB often compete with ribosomes for the RBS, leading to inhibition of translation initiation via ribosome occlusion. Consequently, the ribonuclease RNase E and the degradosome complex get recruited. This then leads to the coupled degradation of both the mRNA and the sRNA in a stoichiometric manner (Waters & Storz, 2009).

The *sodB* gene is one of three *E. coli* genes that encodes superoxide dismutase (Niederhoffer *et al*, 1990). However, unlike the two other genes, only *sodB* uses iron in its active site. Upon depletion of iron, RyhB represses *sodB* via ribosome occlusion (Figure 2A). Subsequently, the initial cleavage by RNase E occurs at a site that is distant from the base-pairing region. Distal cleavage ensures that translating ribosomes finish translation before degrading the mRNA, avoiding stalled ribosomes and incomplete proteins (Geissmann & Touati, 2004; Massé *et al*, 2005a).

1.2.2 Translational activation of the mRNA

Although negative regulation of mRNAs has been observed more frequently, several different mechanisms by which sRNAs can directly or indirectly promote target gene expression in various bacterial species have been described (Fröhlich & Vogel, 2009). The most common mechanism by which direct base-pairing of an sRNA to an mRNA activates its expression is by an anti-antisense mechanism (Papenfort & Vanderpool, 2015). Some transcripts contain inhibitory hairpin-like structures, that occlude regions essential for translation initiation like the RBS. Consequently, direct base-pairing of an sRNA with such a hairpin relieves the inhibitory structure to facilitate the unmasking of the RBS that would then allow translation of the mRNA.

The *shiA* gene of *E. coli* encodes shikimate permease that is essential for siderophore synthesis (Whipp *et al*, 1998). However, *shiA* is poorly translated because the RBS is blocked by an inhibitory structure in its 5'UTR (untranslated region). Upon iron depletion, RyhB is expressed. The sRNA then base-pairs with the 5'UTR to prevent the formation of this inhibitory structure, thus allow *shiA* to be translated (Figure 2B). This way, RyhB activates siderophore production through shikimate acquisition (Prévost *et al*, 2007).

RyhB also positively acts on another target, although by a mechanism that differs from the *shiA* regulation. The CirA protein serves as a receptor for the antibiotic colicin (Jakes & Finkelstein, 2010). In the absence of RyhB, Hfq binds to *cirA* and prevents translation initiation. However, upon iron depletion, RyhB base-pairs with *cirA* to promote structural changes within the mRNA. Consequently, Hfq gets displaced, thereby allowing translation (Figure 2C; Salvail *et al*, 2013).

1.2.3 Discoordinate regulation of an operon

Many bacterial genes are clustered into operons, with the genes of each operon transcribed on a single, polycistronic mRNA (Sáenz-Lahoya *et al*, 2019). Typically, the translation of the constituent genes (cistrons) in an operon is coordinated, *i.e.,* the ratio of translation products is constant. This mechanism of coordinated operon expression ensures the simultaneous synthesis of the encoded proteins, which are often functionally related (Adhya, 2003). However, certain growth or stress conditions require the increased or decreased expression of individual cistrons, thus resulting in the discoordinate expression of an operon (Mitarai *et al*, 2009).

RyhB mediates the differential control of the expression of genes within the polycistronic *iscRSUA* transcript, encoding the components required for the synthesis of the iron-sulphur clusters (Massé *et al*, 2005a). RyhB binds to the *iscRSUA* mRNA at the RBS of *iscS*, resulting in the cleavage of the 3' part while the 5' part remains un-cleaved and stable (Figure 2D). The presence of an iron-responsive element-like strong secondary structure between *iscR* and *iscS* contributes to the protection against degradation by RNase E. Such regulation prevents the production of Fe-S cluster synthesis machinery in the absence of available iron, while maintaining production of the transcriptional regulator that facilitates survival under conditions of iron-limitation. (Desnoyers *et al*, 2009).

1.2.4 Modulation of sRNA activity by a decoy target

The levels of sRNAs can be modulated by their targets using cellular strategies such as molecular decoy, anti-sRNA, sequence mimicry or by acting as RNA sponges (Göpel & Görke, 2014). For example, ChiX sRNA silences *chiP* mRNA in the absence of chitosugars. However, in the presence of these sugars, ChiX repression of ChiP is alleviated by the production of anti-ChiX that acts as a decoy to degrade ChiX (Figueroa-Bossi *et al*, 2009). Likewise, AgvB RNA antagonizes GcvB by mimicking its sequence (Tree *et al*, 2014). Interestingly, GcvB is also controlled by an RNA sponge SroC that is derived from one of GcvB's own target mRNAs in *Salmonella*. (Miyakoshi *et al*, 2015).

RyhB is constantly made at significant basal levels even when the sRNA is not required. Its levels are modulated by the sponging activity derived from a tRNA precursor (Figure 2E). External transcribed spacers (ETS) are excised elements that are often rapidly degraded upon maturation of ribosomes or tRNAs. One such ETS derived from the 3′end of the *glyW–cysT–leuZ* tRNA precursor sequesters excess RyhB through base-pairing interactions in the absence of stress. However, upon stress induction, the production of RyhB is increased such that the ability of the 3′ETS^{leuZ} RNA to prevent accumulation is saturated (Lalaouna *et al*, 2015). Accordingly, ETS enable to set a threshold for sRNAs to be made, thus allowing for a control of transcriptional noise.

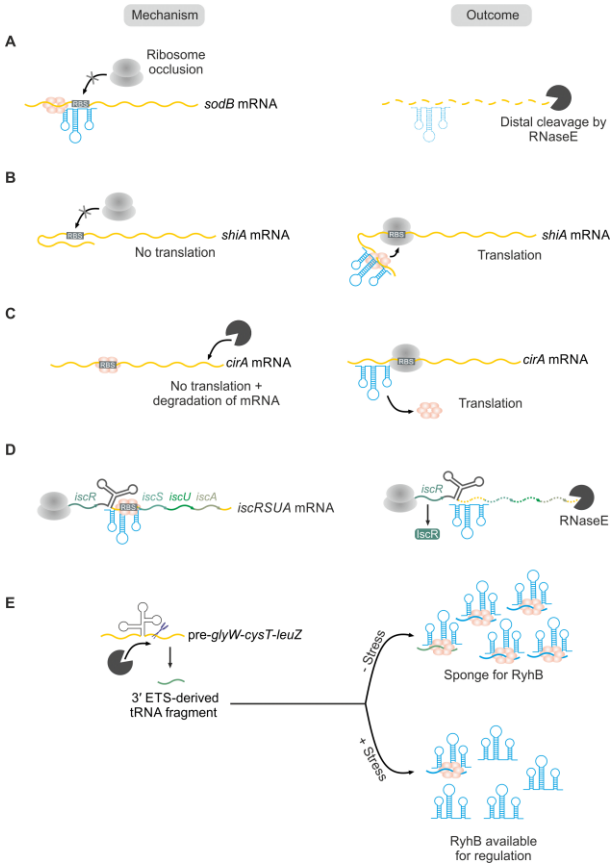

Figure 2. General regulatory principles of sRNAs in the context of RyhB. A) sRNA-mediated repression of target mRNAs is chaperoned by RBPs like Hfq, by binding to and occluding the RBS. Concomitantly ribonucleases like RNAse E gets recruited to mediated coupled degradation of both the sRNA and mRNA. RyhB represses expression of *sodB* by inhibiting translation initiation and inducing mRNA degradation**. B-C)** sRNAs not only repress their target mRNAs but can also activate them. Through an anti-antisense mechanism (B), binding of RyhB to the 5′UTR of *shiA* mRNA unwinds and the inhibitory loop to allow unmasking of the RBS, consequently allowing translation initiation. Alternatively, by the displacement of Hfq from the RBS by RyhB (C), the translation of *cirA* is de-repressed **D)** sRNAs can also discoordinately regulate an operon. Here, RyhB binds specifically to the RBS of the *iscS* gene of the *iscRSUA* operon. resulting in the cleavage of the 3′ end while the 5′ end remains un-cleaved and stable due to the presence of an iron-responsive element-like strong secondary structure, which contributes to the protection against degradation by RNase E. **E)** The activity of the sRNA can by modulated by a decoy target**.** Regulation of RyhB activity is modulated by the 3′ETS of *glyW-cysE-leuZ* tRNA. Based on the absence or presence of stress, the 3′ element base-pairs with RyhB to maintain its basal levels.

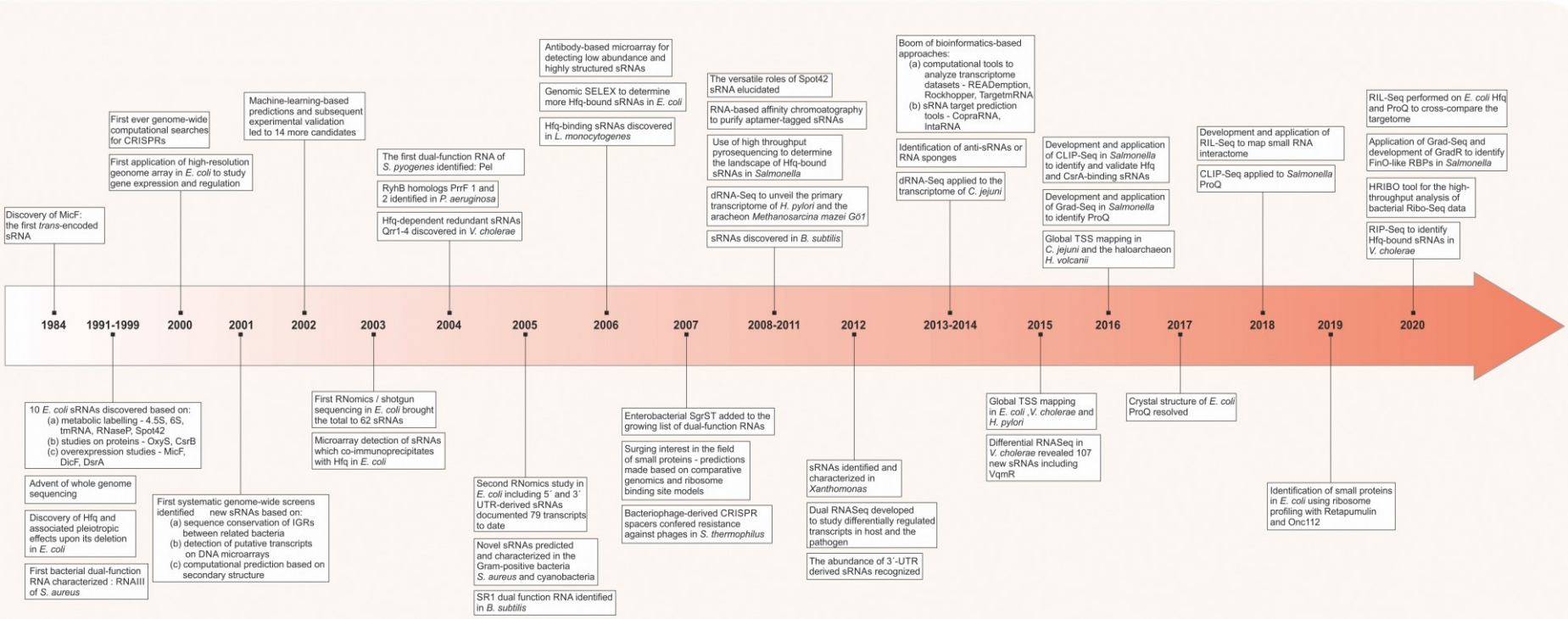

Figure 3. A timeline of the discovery of bacterial sRNAs and the techniques employed

1.3 Global discovery of sRNAs: what, when and how?

The first bacterial antisense RNA MicF was fortuitously discovered in *E. coli* as a regulator of OmpF in the mid-80s (Mizuno *et al*, 1984). Subsequent findings were made inadvertently while analyzing transcriptional regulation of adjacent protein-coding genes. However, the new millennium marked the turning point for methodical genome-wide searches for the identification of new sRNAs in *E. coli*, a model organism that still remains very relevant in the 2020s (Wassarman *et al*, 2001; Eddy, 2002). Following this, several hundreds of sRNAs have been identified and characterized in other bacterial species, emanating from novel technologies as well as modifications of long-standing standard screens and have been covered by several comprehensive reviews. The timeline of their discovery as well as the techniques that have been employed over the decades are outlined in Figure 3. In the context of this thesis, dual-function RNA regulators and small proteins are discussed in detail in the following sections.

1.4 Bacterial dual-function RNA regulators

The unprecedented importance of sRNAs in providing an additional layer of post-transcriptional regulation of bacterial genes has been comprehensively studied over the decades. However, the long-standing assumption has been that sRNAs are ''non-coding'' and lack an open reading frame (ORF). Conversely, a few exceptions have emerged that contain short ORFs that are also translated. This growing class of sRNAs encoding small proteins have come to be referred to as ''dual-function sRNAs''. To date, there are only five such validated dual regulators and in all the cases, the regulatory function of the RNA component was detected first, and only further characterization of the sRNA unveiled the translation of a small protein. Remarkably, there are several other potential dual regulators that have been reported, but the functions of their respective small proteins remain uncharacterized (Sonnleitner *et al*, 2011; Roberts & Scott, 2007; Engel *et al*, 2020).

RNAIII of *S. aureus* was the first dual-function RNA regulator identified (Novick *et al*, 1993; Morfeldt *et al*, 1995). It is 541 nt long with 14 stem loops (SL), and also contains an ORF that encodes the 26 amino acids (aa) cytotoxic peptide δ-heamolysin (*hld*) (Benito *et al*, 2000; Verdon *et al*, 2009). RNAIII is encoded in the quorum sensing *agr* (accessory gene regulator) loci comprising P2 and P3 operons encoding RNAII and RNAIII, respectively. The former transcript produces four proteins AgrA-D. The transmembrane protein AgrB processes pro-autoinducer peptide (AIP) encoded by AgrD (Figure 4). Consequently, RNAIII is induced at high cell density upon secretion of AIP, which is sensed by the two-component system comprising AgrC and AgrA. The RNA element of the dual regulator base-pairs with and regulates 12 mRNAs, all of which are important for the virulence and pathogenicity of the human pathogen *S. aureus*. RNAIII has multiple base-pairing sites, the 3´ SLs all repress gene expression whereas the SLs in the 5´end all activate genes. Interestingly, the former set of genes encode virulence factors that are crucial for early infection, while the latter genes encode secreted factors paramount for late infection (Gupta *et al*, 2015). On the other hand, the small protein Hld targets host cell membrane and causes lysis. Together, RNAIII plays a central role in the pathogenesis of the organism.

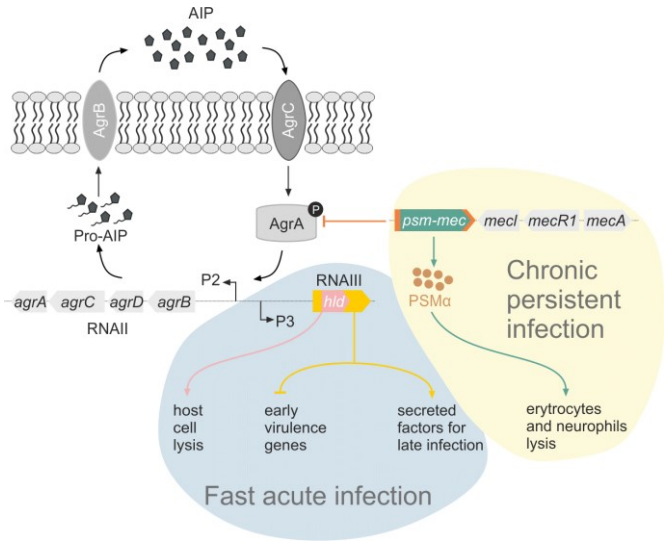

Figure 4. RNAIII and Psm-Mec dual-function RNA regulators of *S. aureus*. RNAIII is encoded on the *agr* loci and is activated by phosphorylated *agrA* upon secretion of autoinducer peptide (AIP). It contains an ORF that encodes *hld* responsible for host cell lysis, while the sRNA itself triggers fast acute infection by secretion of late virulence factors (blue). Psm-mec dual regulator controls *agrA* and represses its translation. Additionally, the PSM-mec ORF encodes PSMα protein that plays a role in immune evasion and elicits chronic persistent infection in the host (yellow).

The staphylococcal cassette chromosome is reponsible for MRSA (methiciliin resistant *S. aureus*) infections (Queck *et al*, 2009). Interestingly, this genetic element harbors the second dual-function regulator Psm-mec, and also contributes to the pathogenecity and virulence of the organism (Kaito *et al*, 2011). The sRNA base-pairs with *agrA* and represses its translation, concomitantly affecting RNAIII levels. The 22 aa PSM-mec ORF encoded within the sRNA makes up most of the transcript and produces cytolytic toxin PSMα. This small protein affects the host's immune system by lysing erythrocytes and neutrophils (Kaito *et al*, 2013). While RNAIII controls the fast acute infection, Psm-mec triggers a persistent chronic infection in the host (Figure 4).

Pel of *S. pyogenes* is very similar to RNAIII of *S. aureus* in that both transcripts are of comparable length and is expressed from an operon (the first gene of the *sagA-I* operon). Additionally, Pel also encodes the 53 aa haemolytic peptide streptolysin S, which targets the host cell membrane causing cell lysis (Nizet *et al*, 2000). Analogous to RNAIII, the regulatory element of Pel is a positive regulators of streptococcal virulence factors. However, the mechanism of this base-pairing remains to be investigated (Mangold *et al*, 2004).

SR1 of *B. subtilis* was initially found through a bioinformatics screen while searching for sRNAs in the IGRs of the bacteria (Licht *et al*, 2005). The 205 nt transcript also harbors a 39 aa ORF that encodes the SR1P small peptide. SR1 was found to be maximally expressed under gluconeogenic conditions and repressed under glycolytic conditions, mediated by CcpA and CcpN proteins that are responsible for sugar-mediated carbon catabolite repression in *B. subtilis*. The regulatory element base pairs with *ahrC* mRNA that encodes the transcriptional activators of genes for arginine catabolic operons, *rocABC* and *rocDEF* (Heidrich *et al*, 2006, 2007). On the other hand, the peptide SR1P interacts with GapA – the glyceraldehyde-3P dehydrogenase, which is active under glycolytic conditions. As the dual regulator is expressed under conditions when GapA is not needed, the peptide promotes the interaction of the protein with RNAse J1 consequently resulting in RNA degradation, thus hinting at a possible link between nutritional state of the cell and RNA degradosome machinery (Figure 5).

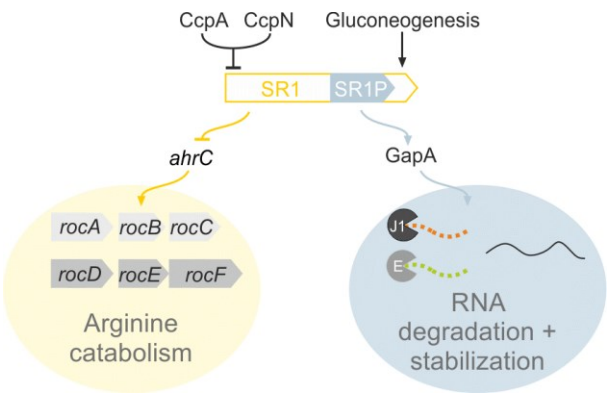

Figure 5. The dual-function RNA SR1 of *B. subtilis*. SR1 is expressed under conditions of gluconeogenesis and is repressed by the carbon catabolite repressors CcpA and CcpN. The sRNA also harbors an ORF that encodes the small protein SR1P. While the riboregulator controls arginine catabolism by inhibiting *ahrC* mRNA (yellow), the peptide acts on GapA to stabilize it as well as recruit RNAses J1 and E to degrade their target RNAs (blue).

Perhaps the most extensively studied dual-function RNA regulator is SgrS of Gram- negative *E. coli*. It is a 227 nt sRNA and also contains a 43 aa ORF *sgrT* that is translated (Vanderpool & Gottesman, 2004; Wadler & Vanderpool, 2007). Sugars such as glucose and mannose are imported into the cell via PTS-dependent transporters and released into the cytosol to enter the glycolysis pathway after their phosphorylation. However, when these phosphor sugars cannot be metabolized fast enough, sugar uptake and its subsequent metabolism become uncoupled, leading to sugar-phosphate stress. Consequently, this stress is sensed by the transcription factor SgrR, leading to the expression of SgrS. To relieve this stress, the sRNA represses the translation of mRNAs encoding sugar transporters – PtsG and ManXYZ, and other mRNAs involved in various metabolic pathways – *pur*, *adiY*, *folE* and *asd* (Rice & Vanderpool, 2011). While the action of SgrS on *ptsG* is via direct base pairing with consequent occlusion of the RBS and subsequent recruitment of RNaseE, on the other hand, the sRNA uses two distinct base-pairing sites on the *manXYZ* polycistronic mRNA. The first site lies within the *manX* ORF, whereas the second site lies in the UTR upstream of *manY*. Interestingly, both base-pairing sites use non-canonical modes of Hfq-based repression. The former site entails SgrS-dependent recruitment of Hfq to a binding site that overlaps the *manX* SD sequence, thus making the sRNA a chaperone-like partner and Hfq the direct repressor of translation (Azam & Vanderpool, 2018). In contrast, the latter base-pairing site is dependent on Hfq for the stabilization of SgrS as well as to act as a matchmaker to enable duplex formation. Additionally, an AU-rich enhancer element found upstream of *manY* mRNA serves as a binding site for the r-protein S1, thereby making the SD more accessible. Notably, SgrS repression of *manY* is through interference with S1 binding to the enhancer sequence, and as a result, reducing its translation efficiency (Azam & Vanderpool, 2020; Fröhlich & Papenfort, 2020). The small protein SgrT acts specifically to inhibit the transport activity of PtsG, by binding to the $EIIC^{Glc}$ domain, thereby relieving inducer exclusion (Figure 6). Together, both SgrS and SgrT block further accumulation of glucose 6-P levels and promote the utilization of alternate carbon sources (Lloyd *et al*, 2017; Raina & Storz, 2017). Orthologs of SgrS are also found in other bacteria such as *Klebsiella pneumoniae*, *Yersinia pestis* and *Salmonella*. In *Salmonella*, SgrS activates the synthesis of the enzyme YigL that is required for the detoxification upon sugar-phosphate stress, by facilitating the discoordinate expression of *pldB-yigL* operon (Papenfort *et al*, 2013).

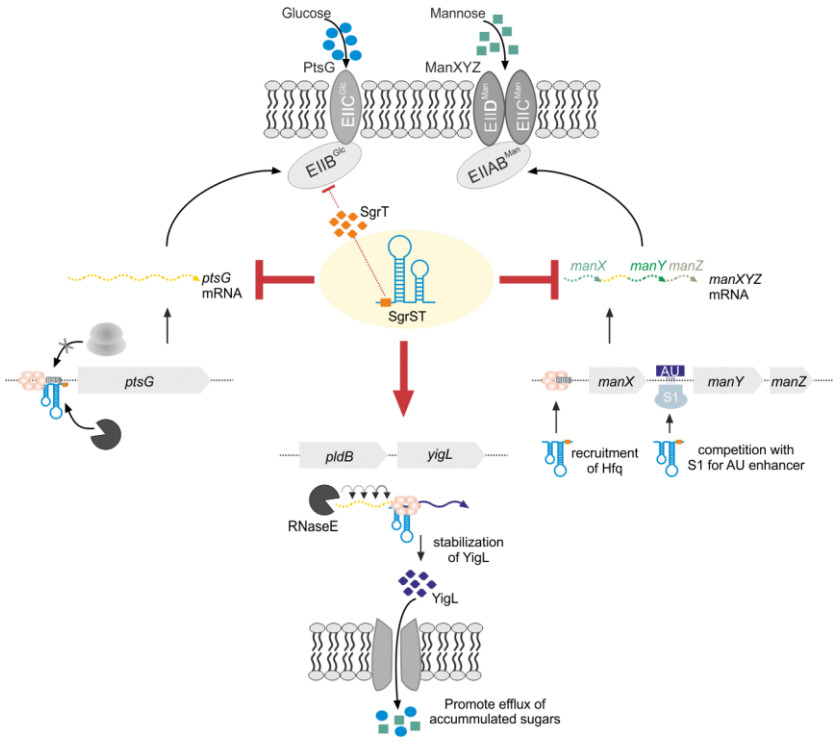

Figure 6. The dual-function RNA SgrS of Gram-negative enterobacteria. SgrS is produced under conditions of sugar-phosphate stress and represses sugar uptake by binding to the RBS of *ptsG* mRNA and recruiting RNases to cleave the transcript, thus repressing PtsG production. SgrS also acts on the *manXYZ* mRNA by binding to the SD of *manX* and recruiting Hfq or by competing with S1 r-protein for the AU-rich enhancer element upstream of *manY*, in both cases repressing the expression of ManXYZ. The sRNA also has an ORF that encodes SgrT (orange), which binds to the EIIB component of PtsG to repress its activity. The activation arm of SgrS discoordinately regulates the *pldB-yigL* operon to promote efflux of accumulated sugars.

1.5 Bacterial small proteins: the road to its discovery

How small is too small? Unlike a peptide that is derived via processing of a longer polypeptide / protein precursor, bacterial small proteins are defined as those that are below 50 amino acids in size and directly result from the translation of a small ORF (Storz *et al*, 2014; Garai & Blanc-Potard, 2020). The conventional one gene, one protein concept has been challenged by the growing list of proteins that emanate from internal translation initiation sites (iTIS), thus directing the ribosomes to out-of-frame (OOF) start codons (Meydan *et al*, 2018). Most genome annotations are based on size cut-offs to distinguish random annotations of ORFs from *bona fide* protein-coding ORFs, thus often overlooking nested ORFs (Orr *et al*, 2020). Additionally, the absence of RBS in some bacterial genomes and the presence of leaderless of mRNAs often makes it difficult to discern small ORFs (Dandekar *et al*, 2000; Goyal *et al*, 2017). While a handful of small proteins were discovered inadvertently, to fully capture the diversity of the translatome, systemic genome-wide approaches have been developed over the last few years and are reviewed below.

Standard proteomics methods often rely on gel electrophoresis or other chromatography-based methods, which are biased towards proteins larger than about 30 kDa (kilo Dalton), thereby precluding detection of very small proteins (Zhang *et al*, 2013b). The concept of proteogenomics was initially used to describe studies in which proteomic data are used for improved genome annotation and characterization of the protein-coding potential (Jaffe *et al*, 2004). The term has since been broadened to include the reannotation of genomes using mass-spectrometry (MS)-based proteomics.

While the approach has proven to be immensely useful for eukaryotic and plant genomes, MS has been particularly challenging for identifying bacterial small proteins (Nesvizhskii, 2014). Proteolytic cleavage of some small proteins often results in peptides of a length that is not detectable by MS. Secondly, the low abundance as well the hydrophobic nature of small proteins often leads to loss of tryptic peptides from being detected (Müller *et al*, 2010). To circumvent these limitations, quantitative MS analysis has been employed to examine the changes in peptide abundance between samples generated under various growth conditions. Although, the results of this approach have been limited, it has been successfully implemented to identify few small proteins expressed under cold shock conditions (D'Lima *et al*, 2017).

The shortcomings of MS can be overcome by combining its power with bioinformatic predictions to harness the potential of small ORFs discovery. Using comparative genomics as well biochemical identification of small proteins, (Hemm *et al*, 2008) systematically identified experimentally validated 20 new, 6 previously unconfirmed as well as 16 small proteins encoding toxins in *E. coli* K-12. Additionally, they also employed homology-based searches using *tblastx* and *blastn* to identify 58 potential ORFs as well as an RBS model that accounts for the SD sequence upstream of an ORF as well as the distance of it from the start codon to identify 13 more small ORFs. The group also went on to validate 51 more small proteins and later 36 more in two subsequent studies (Hemm *et al*, 2010; VanOrsdel *et al*, 2018).

A mammoth study involving an analysis across all the sequenced bacterial genomes at the time resulted in 1,153 candidate ORFs with the majority encoding proteins under 100 amino acids. Not surprisingly, they also show that a lot of these predicted ORFs often are overlooked in gene prediction tools like Glimmer, GeneMark and EasyGene. (Warren *et al*, 2010). The RNAcode tool was developed aimed at predicting protein coding regions from a set of homologous nucleotide sequences without relying on species-specific signatures or machine learning techniques. Alignments were created for *E. coli* K-12 with 53 other completely sequenced enterobacterial genomes and identified 35 potential candidate small proteins, 7 of which they could be validated by an improved MS approach that was optimized for small proteins (Washietl *et al*, 2011).

Subsequently, a hybrid ensemble learning algorithm called AdaBoost.M1 was developed by considering multiple features of genes such as codon bias, codon frequency, hydrophobicity, overall GC content as well as the GC content of the first codon, length distribution and entropy distance profile to predict small proteins in *E. coli, K. pneumoniae, Y. pestis* and *Enterobacter* 638 with high confidence. They report that their algorithm was substantially more accurate in identifying already known small proteins in *E. coli* (Goli & Nair, 2012).

Resources like AGMIAL (Bryson *et al*, 2006), AGeS (Kumar *et al*, 2011) and MicroScope (previously MaGe, (Vallenet *et al*, 2013)) and SearchDOGS (for yeast genomes (ÓhÉigeartaigh *et al*, 2011)) were all developed aimed at the manual curation of prokaryotic genomes. SearchDOGS helped in the systematic discovery of 594 previously unannotated genes across 11 yeast genomes. This tool was subsequently implemented for bacteria to identify genes that prove most tricky for automated annotation programs. The software relies on extensive gene synteny across species combined with sequence similarity searches to map coordinates of coding regions that may have been overlooked otherwise owing to their small size. Using this approach, it was reported that 58% of small proteins were incorrectly annotated in the 8 *E. coli* genomes they tested. Conserved, but previously overlooked ORFs were also found in *Shigella, Xanthomonas* and *Yersinia* (ÓhÉigeartaigh *et al*, 2014).

A flexible machine learning algorithm called DiSCO-Bac was developed based on conserved sequence features as well as comparative genomics to mine experimental data for small proteins that are likely to be functional (Friedman *et al*, 2017). As a result, it was reported that sRNAs containing ORFs in *B. subtilis* are enriched for high expression in biofilm growth, while in *S. pneumoniae* they are involved in virulence, and ORFs with transmembrane domains often contribute to type I toxin/anti-toxin systems.

In recent years, more prediction tools have emerged for several other bacteria such as RanSEPs to study the proteome of *Mycoplasma pneumoniae*, that predicted 109 small proteins (Miravet-Verde *et al*, 2019). The efficiency of this forest-based iterative tool was also tested on 12 other bacterial genomes, including 570 previously reported small proteins. Next, a computational machine learning model called sPepFinder was applied to predict novel small ORFs in *E. coli* not just in the IGR, but also 5'- and 3'UTR-derived with a remarkable 92.8% accuracy (Li & Chao, 2020). They also extended their study to other *Enterobacteriaceae* family to predict an incredible 0.1 million small ORFs in hundreds of bacterial strains. However, the tool is limited to only ATG and GTG start codons, ignoring non-canonical start codons as well as leaderless mRNAs. Another prediction tool called OCCAM uses a machine learning technique combined with the use of bait subject sequences to improve elimination of false BLAST alignments to identify bacterial small ORFs (R. Cerqueira & Vasconcelos, 2020).

Hitherto, the most effective approach to identify small ORFs is by employing ribosome profiling or Ribo-Seq, a deep sequencing technique aimed at monitoring the precise position of translating ribosomes (Ingolia *et al*, 2009). By revealing the precise locations of ribosomes on each mRNA, Ribo-Seq aims to identify potential protein-coding regions. However, the possibility of iTIS within a single transcript makes it challenging to define all ORFs, especially in complex transcriptomes. Additionally, Ribo-Seq only provides a static snapshot of ribosome positions but does not report on the dynamics of translational elongation or distinguish stalled ribosomes from those involved in active elongation. The technique was initially developed for mouse embryonic and mammalian cells by using the antibiotics harringtonine and lactimidomycin to trap newly initiated 80S ribosomes at start codons and identify initiation sites (Ingolia *et al*, 2011; Lee *et al*, 2012). However, both these antibiotics are not suitable for bacteria and led to the use of Tetracycline (Tet) instead (Nakahigashi *et al*, 2016). This broad spectrum antibiotic binds reversibly to the translating 30S ribosome subunit and prevents tRNA binding on the A site. TetRP (tetracaycline-inhibited ribosome profiling) revealed over 70% start site reannotations in the most recent version of the *E. coli* K-12 genome at the time, all of which had to be otherwise curated from several individual studies. Interestingly, the analysis was also performed using Chloramphenicol (Cm), which blocks translation elongation by targeting the peptidyl transferase center on the large ribosomal subunit. However, both these antibiotics have their limitations – Tet traps ribosomes imperfectly at the start codons and Cm produces broad high-density peaks from the initiation codon to ~50 nucleotides downstream of the coding region, thus blurring the real start site.

As an alternative to Tet and Cm, (Meydan *et al*, 2019; Weaver *et al*, 2019) used Onc112 and retapumulin (Ret) for identifying small proteins in *E. coli* using Ribo-Seq. Onc112 is a proline-rich antimicrobial peptide that blocks and destabilizes the initiation complex, whereas Ret exclusively stalls ribosomes at the start codons. Over 100 iTIS is were found, suggesting that OOF is more widespread than estimated. Additionally, 36 predicted small proteins were also validated on immunoblots. Ribosome profiling has also been used to identify the protein sequence motifs that are

susceptible and resistant to azithromycin (AZ) in *S. aureus* (Davis *et al*, 2014) as well as in *Salmonella* to detect 130 small proteins (Baek *et al*, 2017) and 31 out of 230 sRNAs that were ribosome-occupied in *S. coelicolor* (Jeong *et al*, 2016). Interestingly, even the well-studied λ phage with a very small genome of shows translation of 50 non-annotated ORFs (Liu *et al*, 2013). However, Ribo-Seq as a stand-alone technique has its limitations. For example, the abundance of stalled ribosomes does not directly reflect the abundance of the small protein. Also, the footprint intensity is not a quantitative measure of the true initiation rate.

Nevertheless, the true number of reliable small ORF predictions can be enhanced by combinatorial approaches. For instance, ribosome profiling and LC-MS was used to map translation start sites in *C. crescentus* with near complete coverage (Schrader *et al*, 2014). Ribo-Seq and machine learning algorithms were used to predict 465 *bona fide* and previously unannotated genes with small ORFs in pathogenic EHEC (Hücker *et al*, 2017). An integrative approach linking Ribo-Seq dataset with computational predictions made using sPepFinder was used to identify virulence-associated small proteins in *Salmonella* (Venturini *et al*, 2020). More recently, a new algorithm called smORFer was developed by integrating genomic information as well as structural features with existing Ribo-seq datasets to accurately predict bacterial ORFs (Bartholomäus *et al*, 2021).

1.6 Functions of bacterial small proteins

While several putative small ORFs have been predicted and their translation has been validated over the years, there are fewer small proteins whose functions have been examined. They have broadly been classified into two groups – small proteins that encode toxins and bacteriocins, and small proteins involved in the regulation of cellular processes (Figure 7). In addition to the small proteins encoded by the five characterized dual-function regulators described in section 1.4, other stand-alone small proteins have been characterized, a subset of which are described below.

The cell division machinery of the divisome is composed of ten core proteins that mediate cell constriction. Small proteins, among other regulators, regulate these core proteins to allow spatio-temporal control of cytokinesis. In *C. crescentus*, the primary SOS-induced division inhibitor is a 29 aa inner membrane protein called SidA (SOS-induced inhibitor of cell division A) that inhibits division by interacting with the late-arriving division protein FtsW (Modell *et al*, 2011). SidA does not directly disrupt the assembly or stability of the central cytokinetic ring protein FtsZ, nor does it affect the recruitment of other components of the cell division machinery. Instead, by inhibiting FtsW, the final constriction of the cytokinetic ring is prevented (Modell *et al*, 2014). In contrast, SulA of *E. coli* directly inhibits polymerization of FtsZ (Mukherjee *et al*, 1998). Similarly, the 40 aa small protein MciZ (mother cell inhibitor of FtsZ) of *B. subtilis* also inhibits FtsZ polymerization (Handler *et al*, 2008). MciZ regulated by the sigma factor RpoE (σ^E), and a recent study showed that MciZ can affect *B. subtilis* sporulation: Excessive amounts of MciZ is produced intracellularly or added exogenously can not only decrease spore formation efficiency but also inhibit spore germination (Araújo-Bazán *et al*, 2019).

Bacteria have several transporters to import nutrients as well as efflux pumps to secrete small molecules (Maloney, 1994; Delmar *et al*, 2014). The 29 aa KdpF protein was one of the first small proteins described to affect a transporter (Gassel *et al*, 1999). Limiting potassium concentrations in the medium leads to the two-component KdpD/KdpE-mediated signal transduction, resulting in the expression of *kdpABC* operon. The small protein KdpF is encoded in the promoter region of the

operon and interacts with the Kdp-ATPase potassium ion transporter complex and stabilizes it (Hamann *et al*, 2008). MntS is a 42 aa small protein and is induced by low manganese (Waters *et al*, 2011). The small protein helps to enlarge the manganese pool by facilitating manganese binding to a variety of enzymes. The manganese exporter MntP and MntS reciprocally inhibit each other to maintain manganese homeostasis (Martin *et al*, 2015). The two-component PhoPQ system of *E. coli* induces the small protein MgtS when magnesium is limiting (Wang *et al*, 2017). MgtS binds to and stabilizes the major PitA phosphate/magnesium symporter, thereby leading to increased intracellular levels of the ion (Yin *et al*, 2019).

Small proteins can be found in type 1 toxin/antitoxin (TA) systems (Brielle *et al*, 2016). These TA systems are encoded by two genes, one that encodes a toxic protein, and one that prevents its expression or its activity, and is comparatively less stable. The *tisAB* locus of *E. coli* encodes an SOS-induced toxic function that is repressed in the presence of the *istR* locus. The *tisAB* operon encodes two putative peptides, TisA (37 aa) and TisB (29 aa), of which only TisB is conserved in Enterobacteria and confers toxicity (Vogel *et al*, 2004). *S. aureus* encodes two TA systems comprising the toxin proteins PepA1 / Pep A2 and the RNA antitoxins $SprA1_{AS}$ and $SprA2_{AS}$, respectively (Sayed *et al*, 2012; Germain-Amiot *et al*, 2019). Under conditions of oxidative stress, the levels of the antitoxin is reduced, which allows the production of the corresponding toxin proteins.

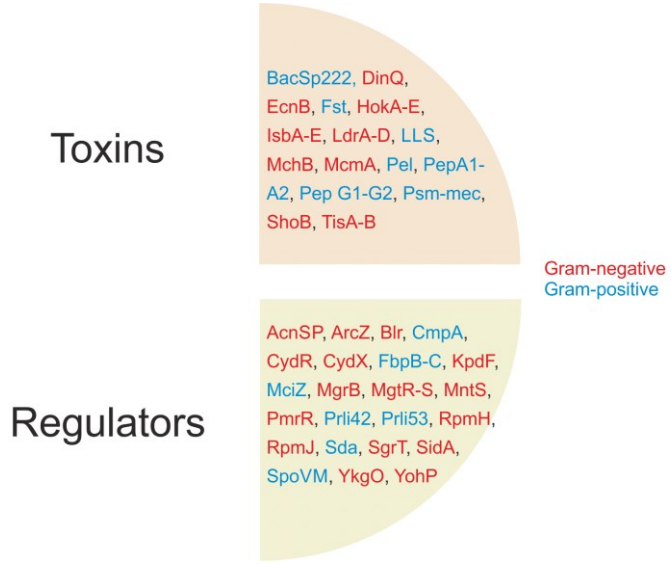

Figure 7. Bacterial small proteins with characterized functions. Gram-negative (green) and Gram-positive (blue) small proteins with known functions are depicted according to their regulatory functions: that encode toxins (pink) or regulators of cellular processes (yellow). The mechanism of action of a few of these small proteins are described in detail in the text.

1.7 *Vibrio cholerae*: the study organism

Vibrio cholerae is a Gram-negative γ-proteobacteria, that is the causative agent of the acute diarrheal disease, cholera (Lekshmi *et al*, 2018). Although there are over 200 different serotypes based on the surface somatic O antigen, pandemic cholera is exclusively associated with serotypes O1 and O139, while the other serotypes have no impact on public health. The O1 serotype exists as two biotypes – classical and El Tor, with three further sub-types based on the type of antigens – Ogawa (strains expressing A and B antigens, and a small amount of C antigen), Inaba (strains expressing A and C antigens only) or Hikojima (strains expressing all three antigens, although found rarely). Until

1961, there were six cholera pandemics, all caused by the classical biotype, whereas the seventh (and current) pandemic is caused by El Tor. The O139 or the Bengal serotype only emerged in 1991 in parts of India and Bangladesh (Morris, 2011; Hu *et al*, 2016). The most recent outbreak was in Yemen in late 2016 with 1.2 million cases reported as of 2018 (Camacho *et al*, 2018). With the first documented instance of cholera in 1817, this disease has marked over 200 years and is far from being eradicated, therefore making *V. cholerae* a very relevant organism to study.

The organism maintains a bi-phasic lifestyle: in the environment as well as within the host, often switching between motile and sessile modes. Primarily found in rivers, estuaries and coastal water, *V. cholerae* are often found in association with aquatic zooplanktons, microphytes and chitin-rich exoskeletons of crustaceans (Islam *et al*, 2020). Upon colonization of these chitinous surfaces, the bacteria are naturally competent to take up free DNA via horizontal gene transfer (HGT) (Meibom *et al*, 2005; Matthey & Blokesch, 2016). Regulation of chitin-induced natural competence has been linked to (a) quorum sensing and (b) catabolite repression (Scrudato & Blokesch, 2012).

1.7.1 Quorum sensing: a numbers game

Quorum sensing (QS) is bacterial communication involving sensing the number of chemical signaling molecules called autoinducers (AIs) to regulate gene expression. The pathogenesis of *V. cholerae* is attributed to its virulence cascade responsible for colonization and subsequent disease mediated by cholera toxin (CT) and toxin co-regulated pilus (TCP), both of which are regulated by QS. Additionally, the sessile lifestyle of the organism is credited to the formation of biofilms, which is also controlled by QS (Federle & Bassler, 2003; Conner *et al*, 2016).

The *V. cholerae* QS system secretes and responds to three AIs via three parallel circuits. The first AI is cholera autoinducer 1 (CA-1), synthesized by the enzyme CqsA. Homologs of this AI are only conserved among other *Vibrio* species, thus allowing inter-genus communication (Miller *et al*, 2002). The second AI is AI-2, synthesized by LuxS and is conserved among many Gram-negative and Gram-positive bacteria, thus serving as an interspecies communication molecule (Xavier & Bassler, 2005). Both these AIs are detected by the two-component systems CqsS and LuxPQ, respectively. In the absence of these AIs or when their numbers are relatively low *i.e.*, at low cell density (LCD), both the receptors function as kinases, and transfer phosphate to the shared response regulator LuxO. Phosphorylated LuxO activates four redundant Hfq-dependent sRNAs Qrr1-4 (Lenz *et al*, 2004), that represses *hapR* and activates *aphA* mRNAs. The former regulates formation of biofilms, while the latter controls the suite of genes responsible for the virulence cascade, both of which are hallmarks of LCD state. However, when the levels of AIs are high, *i.e.*, at high cell density (HCD), the signaling molecules bind to their cognate receptors, resulting in their conformational change. Consequently, they no longer act as kinases, resulting in the dephosphorylation of LuxO and subsequent repression of Qrr1-4. As a result, *hapR* is expressed while *aphA* is repressed, thus ensuing in the dispersal of bacteria back into the environment (Neiditch *et al*, 2005; Rutherford & Bassler, 2012).

The third limb of the QS pathway involves the recently discovered third AI, DPO, and is synthesized by the enzyme threonine dehydrogenase (Tdh). At HCD, the transcription factor VqmA in conjunction with DPO activates the expression of the Hfq-dependent sRNA VqmR. The sRNA post-transcriptionally represses genes for biofilm formation as well as the virulence cascade by acting on AphA (Papenfort *et al*, 2015, 2017; Herzog *et al*, 2019).

1.7.2 Quorum sensing-mediated biofilm formation

Biofilms are comprised of sessile bacterial communities characterized by cells that are attached to a surface, an interface, or to each other. They are often embedded in a self-produced matrix and remain anchored to abiotic substrata or to biotic surfaces such as human intestinal lumen. Such attached biofilms are pertinent for persistence of infection (Mukherjee & Bassler, 2019). The steps leading up to biofilm formation was first investigated by (Watnick & Kolter, 1999), involving the bacteria to approach the surface, attach and then become restrained on it. This and subsequent studies all involved examining the process at the population level, until, (Yan *et al*, 2016) employed single-cell live imaging to track a biofilm as it develops from one single founder cell to a mature biofilm of 10,000 cells. Biofilm formation augments the environmental fitness of the organism by providing protection against environmental stresses, predation by grazing protozoa as well as bacteriophages (Matz *et al*, 2005). Additionally, biofilm-grown cells have been reported to be able to better colonize mice intestines than planktonic-grown free swimming cells, regardless of the nature of the surface the biofilms were formed on: plastic, glass or chitin (Gallego-Hernandez *et al*, 2020).

Formation of mature biofilms requires specific extracellular matrix components, among which the exopolysaccharide, Vps is most crucial. In addition to Vps, the matrix protein called rugosity and biofilm structure modulator A and C (RbmA, RbmC) attach mother-daughter cells together at their poles while biofilm-associated protein 1 (Bap1) adheres cells to the contact surface (Duperthuy *et al*, 2013; Fong *et al*, 2006; Giglio *et al*, 2013). Moreover, enhanced production of the virulence factor regulators *toxT*, *toxR*, *tcpP*, *tcpH* and *ctxA* accounts for the hyperinfectivity of biofilms (Tamayo *et al*, 2010). The overall transcript abundance of these regulators was higher in biofilm-grown cells even before they have infected the host, suggesting that they might already be primed for host infection. It was recently shown that owing to enhanced TCP production, biofilms are likely to be hotbeds for phage transduction (Gallego-Hernandez *et al*, 2020).

Furthermore, cyclic-di-GMP (c-di-GMP) and HapR reciprocally regulate *V. cholerae* biofilm formation. While HapR directly binds to represses the biofilm transcriptional activator, *vpsT,* it also controls the transcription of 14 genes encoding a group of proteins that synthesize and degrade c-di-GMP (Beyhan *et al*, 2006; Waters *et al*, 2008). Thus, high levels of c-di-GMP enhance biofilm formation, repress virulence factor expression and motility, while low levels of c-di-GMP have the opposite consequence. The net effect entails sensing and assimilation of extracellular QS signals as well as intracellular chemical signals by c-di-GMP to regulate biofilm formation (Tischler & Camilli, 2004). Recently, the role of ferric uptake regulator (Fur) was also linked to the regulation of biofilm formation by inhibiting the intracellular levels of c-di-GMP (Gao *et al*, 2020).

The cell surface structure type IV mannose-sensitive hemagglutinin pili (MSHA) is critical for the initial attachment to the surface to allow biofilm formation. These pili are dynamic macromolecular appendages and can extend and retract, thus facilitating the attachment as well as surface-associated twitching motility that helps in dissemination. The genes encoding MshA are distributed over two operons: *mshI-F* and *mshB-Q.* The former cluster encodes proteins for assembly and secretion while the latter cluster encodes the structural components for the pilus (Thelin & Taylor, 1996). MshA pili in conjunction with the polar flagellum contribute to the 'orbiting' and 'roaming' trajectories of *V. cholerae* biofilms, allowing to synergistically scan the surface prior to attachment. Orbiting entails tight circular tracks over the same region as a result of strong MshA-surface interactions whereas roaming entails meandering over a large distance owing to weak interactions (Utada *et al*, 2014).

Interestingly, MshE, the ATPase that polymerizes the pilus has a novel binding motif for c-di-GMP, thus promoting MSHA production and facilitating the transition from motile to surface-attached biofilms (Jones *et al*, 2015; Floyd *et al*, 2020). Unlike rod-shaped bacteria, the comma-shaped curvature of *V. cholerae* only allows a small surface area that can potentially attach to a surface, thus underscoring the contribution of MSHA-dependent anchoring of the cells to the surface.

1.7.3 Quorum sensing-mediated virulence cascade

The major virulence factors responsible for *V. cholerae*'s pathogenesis are (a) toxin co-regulated pilus (TCP), a type IV pilus that mediates adherence and is required for intestinal colonization, (b) cholera toxin (CT), a bipartite AB_5 family ribosyl transferase that is responsible for the profuse rice-watery diarrhea and associated fluid loss and (c) accessory colonization factor (*acf*) genes that is essential for colonization as well as efficient biogenesis of TCP (Kaper *et al*, 1995; Brown & Taylor, 1995; Thelin & Taylor, 1996). The genes required to produce CT and TCP are encoded on two distinct pathogenicity islands, both having phage ancestry. The CT genes are located on a filamentous phage CTXφ (Waldor & Mekalanos, 1996) whereas TCP operon is encoded on the *Vibrio* pathogenicity island (VPI), from the genome of the filamentous phage VPIφ (Karaolis *et al*, 1999). Despite their distinct origins, both the virulence determinants are coordinately regulated by the same regulatory gene, *toxR* and is hence referred to as the ToxR regulon (Bina *et al*, 2003; Childers & Klose, 2007).

AphA is a PadR/MarR family transcription factor which serves as the master regulator that triggers the entire virulence cascade upon host colonization (Rutherford *et al*, 2011). On the other hand, AphB is a LysR-type transcription factor that functions in cooperation with AphA to activate the expression of *tcpPH* on the VPI (Kovacikova & Skorupski, 1999). TcpP and ToxR are membrane-bound sensory proteins that act in concert with accessory transmembrane / periplasmic proteins TcpH and ToxS, respectively to activate ToxT. The *toxT* gene is encoded within the TCP biogenesis operon downstream of *tcpA*, thus autoregulating its own expression as well as the other genes in the operon (Brown & Taylor, 1995). The activation of CT and TCP is directly controlled by ToxT, belonging to the AraC family. ToxT activates transcription by biding to a 13 bp DNA sequence called the 'toxbox', located upstream in the promoter of ToxT-activated genes (Withey & DiRita, 2006). In contrast, the activation of genes by ToxT is counteracted by the histone-like nucleoid-associated protein (H-NS), that also competes for the toxbox sequence. At the *ctxAB* promoter that produces CT, H-NS strongly represses transcription. However, at the *tcpA* promoter, the effect of H-NS is less pronounced, and ToxT is primarily a direct activator of transcription (Kazi *et al*, 2016). H-NS further reduces virulence gene expression by binding to the *toxT* promoter region and subsequently repressing its transcription (Nye *et al*, 2000).

Adding to the complexity of virulence regulation, ToxR also regulates the outer membrane proteins OmpT and OmpU, independent of TcpP. This regulation is reciprocal: in the absence of *toxR*, OmpT is maximally expressed whereas upon overexpression of *toxR*, OmpU is activated (Crawford *et al*, 1998; Li *et al*, 2000). This way, ToxR also contributes to the porin regulon, by differentially influencing the levels of porin proteins in response to environmental signals. This is of importance especially in the context of maintaining membrane integrity in response to changing osmotic conditions during intestinal colonization (Provenzano & Klose, 2000). Upon *V. cholerae* infection, profuse diarrhea is ensued by the activation of *ctxAB,* which is coordinately regulated by ToxT and TCP, to stimulate CT production. Upon binding to the receptors of GM1 gangliosides of the host through subunit B,

the entire AB_5 toxin complex is endocytosed. Then, the A subunit is cleaved off, which subsequently moves retrograde into the endoplasmic reticulum and is refolded and released into the cytosol (Spangler, 1992). Here, it induces adenosine diphosphate (ADP) ribosylation of the small G protein alpha-subunit, with consequent activation of the enzyme adenylate cyclase. This results in increased levels of cyclic AMP (cAMP), leading to the efflux of water and electrolytes into the lumen of the host, thus establishing the diarrheal disease and associated fluid loss (Sahyoun & Cuatrecasas, 1975; Field, 1978).

1.7.4 sRNAs associated with *V. cholerae* virulence

With the extensive post-transcriptional command that sRNAs offer for the robust regulation of a surfeit of cellular processes, it is not surprising that they also contribute to control of virulence and pathogenesis. One of the first examples identified in *V. cholerae* are the four redundant sRNAs Qrr1–4, that are transcriptionally activated at LCD by phosphorylated LuxO in concert with σ^{54} (Lenz *et al*, 2004). The sRNAs directly regulate transcription of *aphA* as well as via HapR (Kovacikova & Skorupski, 2002; Rutherford *et al*, 2011). Additionally, strains lacking Qrr sRNAs, do not produce TcpP or CT, and are severely impaired for colonization in an infant mouse model of *V. cholerae* pathogenesis (Zhu *et al*, 2002). Interestingly, termination of Qrr synthesis is induced by AIs of the quorum sensing system. This coupled with the concomitant increase in the levels of HapR positively regulate genes required for HGT and natural competence. Thus, QS-induced natural competence in *V. cholerae* can also promote acquisition of the CTXφ prophage carrying *ctxAB* (Udden *et al*, 2008).

A genetic screen to identify additional regulators contributing to *V. cholerae* QS revealed another set of redundant sRNAs CsrBCD, that control *qrr* expression and stability of *hapR* by altering the activity of LuxO (Lenz *et al*, 2005). These sRNAs are activated by VarS/VarA two component system and act as sponges to sequester CsrA (section 1.1.2). While there is no evidence of altered transcription of *luxO*, CsrA appears to have a positive effect on LuxO activity. VarA has also been directly implicated in modulating the levels of CT and *tcpA*, and subsequent intestinal colonization, independent of HapR (Wong *et al*, 1998). Intriguingly, the regulation of virulence by VarA is independent of ToxR and overexpression of *toxT* and *tcpPH* can suppress a *varA* mutant phenotype.

A transposon mutagenesis screen identified the sigma factor (RpoE or σ^E)-driven sRNA controlling formation of outer membrane vesicles (OMVs) called VrrA (Vaitkevicius *et al*, 2006; Song *et al*, 2008). The sRNA promotes the release of OMVs by the downregulation of the OmpA porin, independent of Hfq. Production of OMVs is critical to bacterial survival, nutrient acquisition, biofilm formation and pathogenesis (Kulp & Kuehn, 2010). Strains lacking *vrrA* displayed an increased colonization of an infant mouse, whereas an *ompA* deletion strain shows defective colonization, which could partly be attributed to the concomitant increase in TCP production. Recently, another σ^E-dependent sRNA in *V. cholerae* was discovered called MicV, that was shown to act redundantly with VrrA. Both sRNAs share a conserved seed-pairing domain allowing them to regulate multiple target mRNAs, therefore acting as global regulators of envelope stress response (Peschek *et al*, 2019).

A genome-wide deep sequencing analysis for the detection of ToxT-regulated sRNA transcripts one unknown and 17 potential new non-coding regulators (Bradley *et al*, 2011). Two of these sRNAs were found to be encoded in the IGR of the VPI and were upregulated upon ToxT expression, of which one was independently characterized prior to this study and was called TarA (Richard *et al*, 2010). Accordingly, the second sRNA was named TarB. TarA is an Hfq-dependent sRNA with a high

degree of similarity to *tcpA* toxbox and are expressed synchronously during infection. Analogous to *E. coli* SgrS, TarA negatively regulates *ptsG,* however in contrast, it is not a dual-function regulator. The original TarA study implicated a decrease in fitness in the absence of the sRNA, however the second study documented no virulence defect. One explanation could be the difference in impact of the sRNA on different serotypes- while the first study was performed on the classical serotype, the second one was studied on El Tor. On the other hand, the role of TarB in modulating pathogenesis is better understood. Like VrrA, TarB also negatively regulates virulence by interfering with *tcpF* expression independent of Hfq. However, it is counterintuitive that a positive regulator of virulence (ToxT) drives the expression of a factor (TarB) that negatively regulates the TCP operon (*tcpF*). It was proposed that TarB is only expressed under conditions where a fitness advantage is necessitated to counteract the effects of anti-TcpF antibodies, thereby allowing a tight repression of TcpF by the sRNA.

A high throughput RNA sequencing approach identified VqmR as a *trans*-acting Hfq-dependent sRNA encoded in proximity to the *vqmA* gene (Papenfort *et al*, 2015). The sRNA regulates *vpsT*, the *rtx* cluster as well as a homolog of siderophore receptor. Subsequently, the autoinducer DPO was identified and the third parallel QS pathway of *V. cholerae* was characterized (Papenfort *et al*, 2017). Successively, another RNA-Seq analysis revealed *aphA* inhibition by VqmR. Additionally, DPO significantly inhibited the promoter activities of *tcpP, toxT, tcpA*, and *ctxA* (Herzog *et al*, 2019). Interestingly, while all the sRNAs described in this section act directly or indirectly to regulate the production of TCP, only VqmR directly represses CT production (Figure 8).

1.7.5 Carbon catabolite repression (CCR)

With an alternating *V. cholerae* life cycle between a planktonic state and a biofilm state as well as between the environment and the host, the bacteria are presented with ample opportunities to imbibe free DNA through HGT. As noted in the beginning of this section, natural competence is not only driven by quorum sensing, but also via carbon catabolite repression (CCR). CCR refers to the repression of gene expression by preferred carbon sources. When the levels of such a carbon source is low, the organism adapts to using slowly metabolizable carbon sources, concomitantly activating adenylate cyclase (CyaA) (Brückner & Titgemeyer, 2002). CyaA then controls the expression of CRP-regulated genes by the activation of cAMP. CRP directly inhibits *tcpP* expression and activates the transcription of *ompT,* while cells lacking *crp* produce lower levels of HapR. (Skorupski & Taylor, 1997; Li *et al*, 2000).

A recent ChIP-Seq analysis to map CRP (cAMP receptor protein) binding sites in *V. cholerae* revealed substantial overlap between the ToxR regulon and control of other virulence factors not regulated by the ToxR system, further confirming the role of CRP for the specific induction of gene expression during intestinal colonization (Manneh-Roussel *et al*, 2018). The lack of cAMP and / or CRP also interferes with natural transformation at three stages when grown on a chitin-rich surface: chitin surface colonization, chitin degradation/metabolism and competence gene expression (Meibom *et al*, 2005). Competence in *V. cholerae* is induced by the central competence regulator TfoX, that is activated by the sRNA TfoR. The expression of competence genes *comEA* and *comEC* are critical for efficient DNA uptake and subsequent transformation, and are directly under the control of TfoX through its influence on the transcription factor QstR (Seitz & Blokesch, 2013; Lo Scrudato & Blokesch, 2013). cAMP-CRP positively regulates *tfoX* expression through direct biding on its promoter, thus establishing a link between CCR and competence induction (Wu *et al*, 2015).

To better understand the influence of environmental niche on the genetic composition of *V. cholerae*, (Baharoglu *et al*, 2012) studied the influence of SOS response on integron integrase. The chromosomal superintegron encodes an array of genes whose rearrangement is influenced by the activity of integrase, IntlA. Growth on preferred and non-preferred sugars as well as transcriptional regulation by CRP on *intlA* expression elucidated the impact of CCR on the cassette rearrangements within the integron, and its implicated role in pathogenicity.

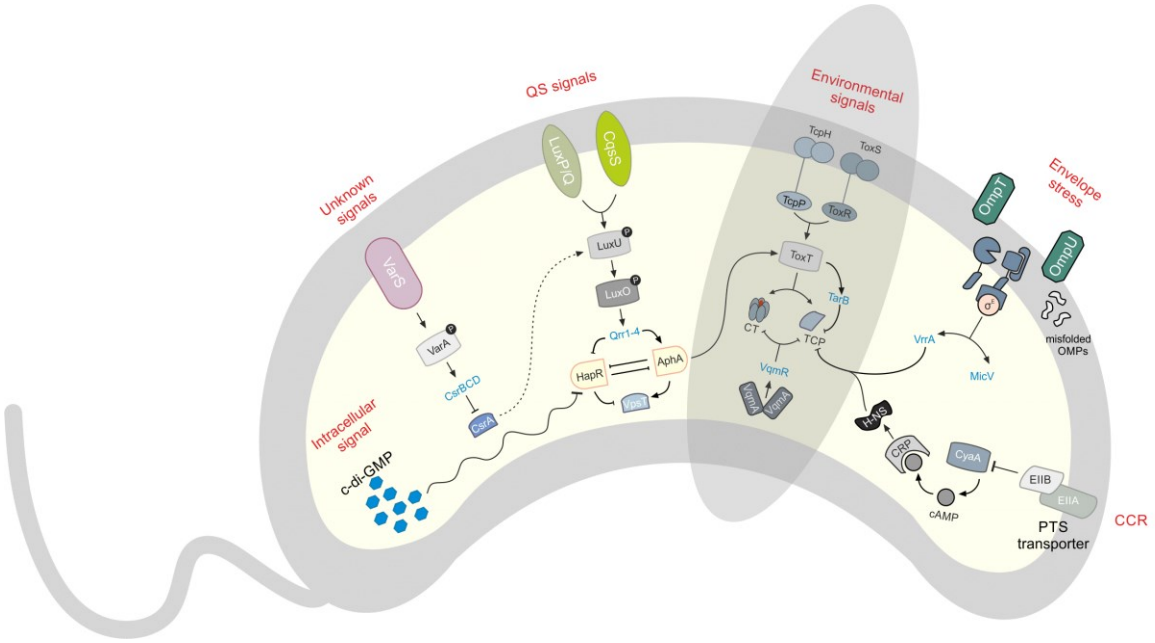

Figure 8. Control of virulence in *V. cholerae*. The complex network of different regulatory pathways (refer text for details) directly or indirectly converges on the virulence cascade (grey ellipse). Based on the nature of the signal sensed (in red), the corresponding membrane receptor respond to and transduce the signal into the cytosol. Additionally, post-transcriptional regulation by sRNAs (in blue) also contribute to this multifaceted mode of virulence regulation.

1.8 Aims of the thesis

Vibrio cholerae is the etiological agent of the human diarrheal disease cholera. Two virulence factors that are produced by *V. cholerae* and are essential for disease are toxin-coregulated pilus (TCP) and cholera toxin (CT). Expression of the genes encoding TCP and CT in *V. cholerae* is controlled by a complex regulatory cascade that is influenced by both specific regulators, such as ToxR/S, TcpP/H, and ToxT, and global regulators, such as cAMP-CRP (Ramamurthy *et al*, 2020). Although the mechanisms by which these factors influence virulence production has been widely studied, the known roles of sRNAs in modulating virulence in *V. cholerae* is still limited. A handful of characterized sRNAs that affect TCP production are known, however to date only one known sRNA that influences CT levels has been studied (Herzog *et al*, 2019). Virulence factors need to be tightly regulated to support the bacterium's bi-phasic lifestyle. Additionally, adaptation to the aquatic environment and / or host entails sensing and responding to nutrient availability to tailor the gene expression patterns accordingly. Therefore, the first study aims at identifying and characterizing additional post-transcriptional regulators of CT production by employing a forward genetic screen to monitor CT repression.

The localization of RNA to subcellular compartments provides a mechanism for regulating gene expression with exquisite temporal and spatial control. In bacteria, RNAs can be targeted to specific cellular locations affecting stability and translation efficiency of these molecules. *V. cholerae* quorum sensing (QS) involves regulation of gene expression in response to fluctuations in the bacterial population density (Zhu *et al*, 2002). A preliminary screen to identify additional regulators affecting QS transition identified a putative RNA-binding protein called Vc0159. The second study is aimed at the characterization of this protein and understanding its contribution to RNA localization in *V. cholerae* globally and how it affects QS specifically.

Chapter 2

The first dual-function RNA regulator of *V. cholerae*

Parts of the results presented in this chapter were performed in collaboration with the following people:

- Dr. Mona Hoyos (former member of the lab)
- Dr. James R. J. Haycocks and Dr. David C. Grainger (University of Birmingham, UK)
- Dr. Liam Cassidy and Prof. Dr. Andreas Tholey (University of Kiel, Germany)
- Dr. Beatrice Engelmann, Dr. Ulrike Rolle-Kampczyk and Prof. Dr. Martin von Bergen (UFZ Leipzig, Germany)

Data presented in this chapter have been adapted from Venkat *et al*., 2021 Under the Creative Commons Attribution 4.0 Unported Licence (CC BY)

2.1 Post-transcriptional regulators that repress CT production

While the regulatory hierarchy controlling the expression of *ctxAB* is almost identical between both the pandemic serotypes of *V. cholerae*, they have distinguishing properties and have disparate factors that contribute to efficient CT production under laboratory conditions. The classical strains are relatively more permissive, producing CT under different conditions, but induction of CT synthesis in the El Tor strains is more arduous. Several defined media and conditions have been tested such as syncase medium (Finkelstein *et al*, 1966), TCY medium (Evans & Richardson, 1968), CAYE medium (Kusama & Craig, 1970), and yeast extract peptone water (Iwanaga & Kuyyakanond, 1987). However, the CT production by El Tor strains in these media was still rather poor. To mimic the intestinal niche *in vitro,* (Iwanaga *et al*, 1986) developed the so-called AKI condition to stimulate CT production. In addition to the components found in the media used prior to AKI, sodium bicarbonate was supplemented to mimic the intestinal composition. This proved to be very effective to produce up to several micrograms of the toxin per ml. Under AKI conditions, cells are grown in a biphasic state: involving a static growth under oxygen limiting condition, followed by vigorous shaking with aeration upon entering exponential phase of growth. Growth under microaerophilic conditions reflects the actual physiological conditions that *V. cholerae* encounters during infection in the oxygen-limiting small intestine. Studies have shown the temporal nature of virulence gene expression such that there is distinct expression of the *tcpA* operon early and *ctxAB* operon later to ensure that only bacteria that are adherent to the intestinal epithelial cells secrete the toxin (Lee *et al*, 1999; Merrell *et al*, 2002; Schild *et al*, 2007). Although the AKI medium cannot mimic the full complexity of these spatiotemporal interactions, (Kanjilal *et al*, 2010) have shown that the overall gene expression under AKI conditions is similar to the that of *in vivo* studies.

As outlined in section 1.6.4, to date all the known sRNAs of *V. cholerae* implicated in the regulation of virulence do not directly act on CT. So far, there is only one known sRNA that represses *ctxA*, which is VqmR (Herzog *et al*, 2019). To determine additional sRNAs that repress CT, a forward genetic screen was employed to score for the phenotype. To exclude HapR-mediated repression of CT (Kovacikova & Skorupski, 2002), the screen was performed in cells deficient of *hapR*. To this end, an overexpression plasmid library comprising 28 sRNAs (including VqmR) as well a control plasmid were grown under AKI conditions. These sRNAs are part of the laboratory collection and were identified from a previous RNA-Seq study (Papenfort *et al*, 2015). Secreted protein fractions were examined on Western blots for CT production using α-CT antibody. The normalized band intensities from three independent biological replicates were quantified and plotted as fold repression with respect to the control plasmid (Figure 9). As expected, VqmR strongly reduced CT production (~8-fold, in grey), whereas 26 sRNAs only moderately repressed CT (under 3-fold, in blue). Remarkably, the Vcr082 sRNA led to a considerable reduction of CTX (~5-fold, in pink). Therefore, Vcr082 was selected for further characterization.

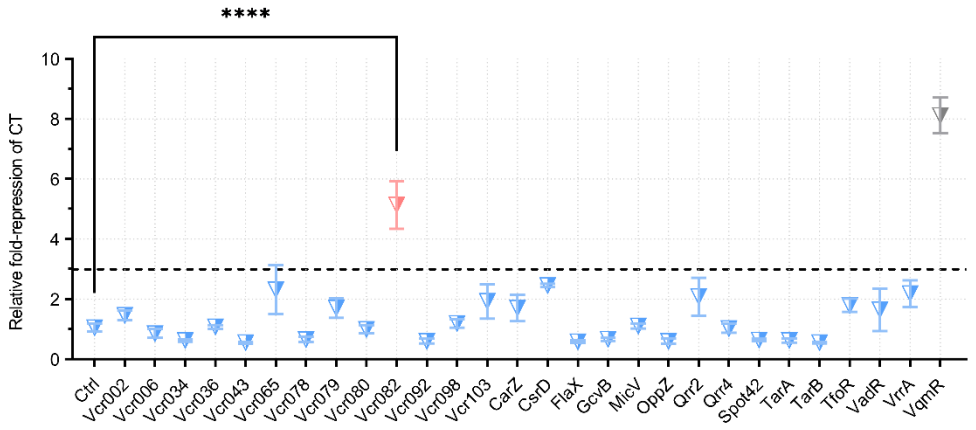

Figure 9. Forward genetic screen to score for repression of CT production. *V. cholerae* Δ*hapR* cells expressing the indicated sRNAs (*x-axis*) plotted as relative fold-repression (*y-axis*) of secreted CT protein fractions in comparison with a control strain (Ctrl) grown under AKI conditions. Indicated in grey is the only known repressor of CT to date (VqmR) and indicated in pink is the yet uncharacterized sRNA, Vcr082. Data are presented as mean ± SD, n = 3. Statistical significance was determined using one-way ANOVA and post hoc Dunnett's test. **** corresponds to p-value ≤ 0.0001.

2.2 Vcr082 is a dual-function RNA regulator

The *vcr082* gene is encoded antisense on the main chromosome of *V. cholerae* between the genes coding for a permease (*vc2278*) and a dipeptidase (*vc2279*), and gene synteny analysis reveals conservation among numerous *Vibrios* (Figure 10A). Structure prediction for the sRNA identified multiple stem loops and an alignment of *vcr082* sequence showed several highly conserved elements including a Rho-independent terminator, that is typical for many sRNAs (Figure 10B, C) (Régnier & Hajnsdorf, 2013; Chen *et al*, 2019). The transcription start site (TSS) of Vcr082 has already been determined, as is its dependency on Hfq (Papenfort *et al*, 2015; Huber *et al*, 2020). In contrast to most Hfq-binding sRNAs which are about 50-250 nt long (Vogel & Luisi, 2011), the length of *vcr082* is 306 nt. This unusually long transcript prompted further assessment to search for additional conserved feature(s), that might reflect a specific function(s). Interestingly, an ORF is predicted within the sRNA, initiating translation at position 104 of the *vcr082* transcript and terminating at position 193, with a well-defined Shine-Dalgarno sequence upstream (Figure 10C). Remarkably, this makes Vcr082 a potential dual-function RNA candidate to characterize further and was therefore re-named VcdRP for *Vibrio cholerae* dual RNA and protein, eponymic to their respective roles.

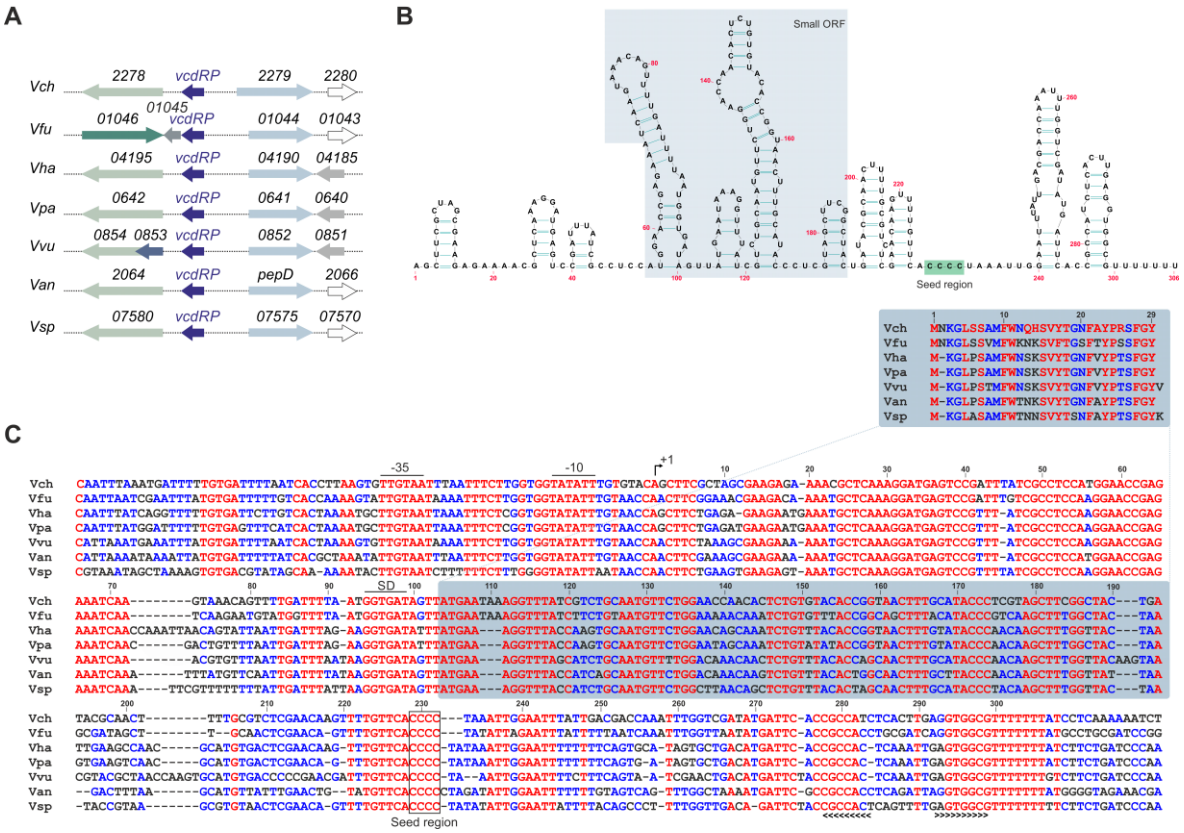

Figure 10. VcdRP is a conserved dual-function regulator. A) Gene synteny analysis of *vcdRP* and its flanking genetic loci among different *Vibrio* species. Homologous genes are depicted using the same colors. **B)** The predicted structure of VcdRP comprising 11 stem loop structures. The small ORF and seed regions are indicated in blue and green, respectively. **C)** Lower panel: Alignment of *vcdRP* sequences, including the promoter regions from various *Vibrio* species: *V. cholerae* (Vch), *V. furnissi* (Vfu), *V. harveyi* (Vha), *V. parahaemolyticus* (Vpa), *V. vulnificus* (Vvu), *V. anguillarum* (Van) and *V. splendidus* (Vsp). The -35 box, -10 box, transcriptional start site (TSS; arrow), the start and stop positions of VcdP (blue box), the highly conserved base-pairing seed region (black box) and Rho-independent terminator (brackets) are indicated. Upper panel: sequence conservation of the 29 aa small protein, VcdP, in different *Vibrio* species.

To evaluate the expression pattern of VcdRP, a *V. cholerae* wild-type strain was grown in rich LB medium as well as defined minimal M9 medium supplemented with either glucose or glycerol as the carbon source. Total RNA samples were prepared and examined on Northern blots. The full-length *vcdRP* transcript accumulates at 306 nt and is further processed into several shorter isoforms (Figure 11A). The expression of VcdRP is highest at LCD in rich medium (lanes 1-2). However, in the presence of glucose, expression of VcdRP remained high regardless of cell density (lanes 5-8). In contrast, upon switching the carbon source to glycerol, the expression was constantly low over cell density (lanes 9-12). To confirm the translation of the small ORF VcdP, a sequential peptide affinity (SPA) tag was chromosomally integrated upstream of its stop codon (Zeghouf *et al*, 2004). The SPA tag comprises the 3×FLAG epitope and a calmodulin binding protein, thereby adding 8 kDa to the small protein. Using a relatively large tag thus circumvents the size limitations for detection of small proteins on Western blots. Accordingly, protein samples of SPA-tagged VcdP were collected under the same conditions as above and examined on a Western blot. Indeed, VcdP is translated, and its abundance is in line with the observations made in the Northern blot analysis above, however the abundance is slightly shifted over cell density in LB medium, by the prominent band observed even at OD_{600} of 1.0. (Figure 11B). Taken together, this novel dual-function regulator is most abundant at LCD and its expression is dependent on the availability of suitable carbohydrates.

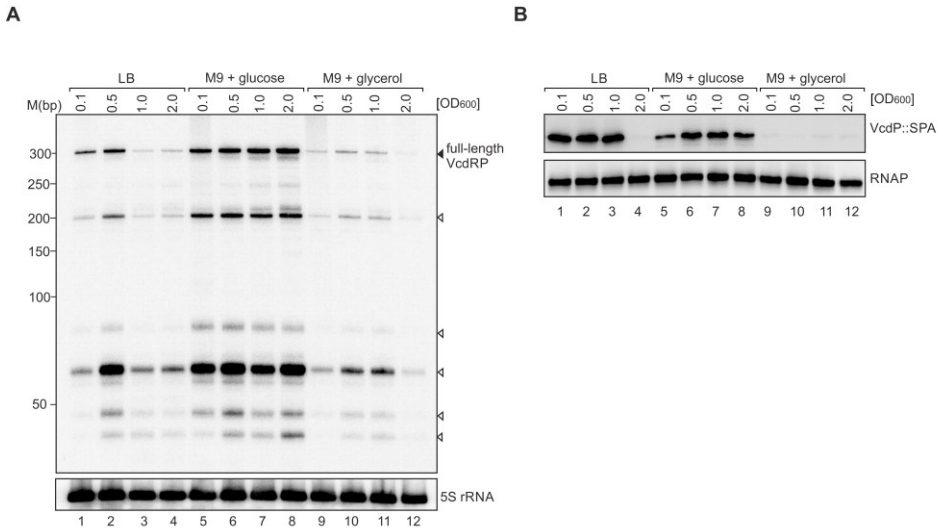

Figure 11. VcdRP is expressed at LCD. A) Northern blot analysis of *V. cholerae* wild-type strain examined for the expression of VcdRP monitored over bacterial growth in the indicated media. The solid triangle refers to the band corresponding to the full length primary *vcdRP* transcript, whereas the open triangles correspond to the different processed shorter isoforms. Probing with 5S RNA confirmed equal loading. **B)** Western blot analysis of the production of chromosomally-encoded VcdP::SPA, corresponding to the same growth conditions as the Northern blot in (A). RNAP served as a loading control.

2.3 Transcriptional control of VcdRP

To assess the impact of carbohydrate utilization on VcdRP expression, the promoter of the dual regulator was screened for possible regulatory motifs. Indeed, the promoter of *vcdRP* harbors a conserved two-box binding site for CRP (Figure 12A). The predicted motif for CRP biding in *E. coli* comprises a palindromic sequence separated by a non-conserved 6 nt linker sequence: 5′-TGTGANNNNNNTCACA-3′ (Shimada *et al*, 2011; Tsai *et al*, 2018). The motif upstream of *vcdRP* of *V. cholerae* has a near-perfect CRP box sequence. This bioinformatic prediction was also bolstered by the findings of (Manneh-Roussel *et al*, 2018), who used ChIP-Seq to map the distribution of CRP sites on the genome of *V. cholerae* (Figure 12B). CRP is a global transcriptional factor that contributes to carbon metabolism and its importance has already been outlined in the preceding section.

An electrophoretic mobility shift assay (EMSA) was performed using the *vcdRP* promoter sequence. Migration of the [^{32}P] end-labelled *vcdRP* promoter fragment in the absence and presence of increasing concentrations of purified CRP protein (0, 0.35, 0.7, 1.4, 2.1, 2.8, 3.5, and 4.2 µM) was determined by native polyacrylamide gel electrophoresis and autoradiography (Figure 12C). Indeed, CRP directly binds to the promoter of *vcdRP*. Further, to test how transcription of *vcdRP* is regulated by CRP, *in vitro* transcription was performed. To this end, a DNA template from the intergenic region (IGR) of *vc2278 / vc2279* (including the promoter of *vcdRP*) was generated by PCR. The template was pre-incubated with increasing amounts of purified CRP along with (α-^{32}P)-UTP. The *in vitro* transcription was started with RNAP and transcripts were analyzed on a denaturing polyacrylamide gel (Figure 12D). Remarkably, a decrease in *vcdRP* transcription upon elevated CRP levels was observed, when normalized to control RNA-I transcript levels (lanes 2-5).

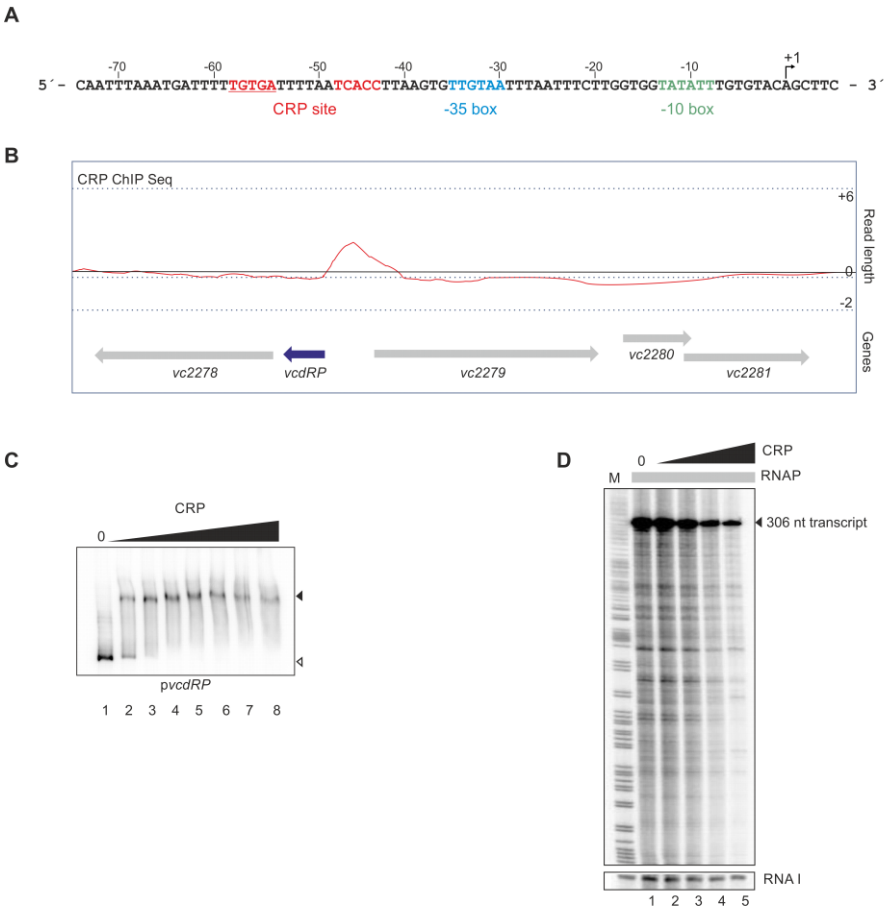

Figure 12. CRP controls the transcription of *vcdRP*. A) The promoter of *vcdRP* harbors a consensus motif for the two box CRP binding site. The binding site is shown in red, and the promoter -10 and -35 elements are shown in green and blue, respectively. The different promoter positions are numbered relative to the transcription start site (+1). The underlined sequence corresponds to the deletion introduced in the reporter fusion used in Figure 13C. **B)** The ChIP-Seq analysis by (Manneh-Roussel *et al*, 2018) determined a CRP peak upstream of the *vcdRP* gene. C) An electrophoretic mobility shift assay performed on the promoter of *vcdRP* in the absence or presence of increasing concentrations of purified CRP protein (black triangle above the gel). The solid triangle indicates CRP-bound *vcdRP*, whereas the open triangle indicates free unbound *vcdRP*. **D)** Autoradiograph of *in vitro* transcription assay performed using RNAP and *vcdRP* promoter with increasing amounts of CRP protein (black triangle above the gel). The gel is calibrated with a Maxim-Gilbert DNA sequencing reaction (M). The RNA-I transcript served as loading control for normalizing.

Further, to evaluate the extent of transcriptional control by CRP, the expression of VcdRP was monitored in cells lacking *crp* and *cyaA*. To this end, total RNA was collected at LCD and HCD in rich medium and compared with expression in a wild-type strain using Northern blotting. In contrast to wild-type cells, VcdRP was de-repressed at HCD in both *crp*- and *cyaA*-deficient cells (Figure 13A, lane 3 vs. lanes 6 and 10). Adenylate cyclase encoded by the *cyaA* gene is known to catalyze the synthesis of cAMP, which together with the receptor protein, CRP, exerts its regulatory effects to modulate carbon utilization (Skorupski & Taylor, 1997). Therefore, the effects of externally supplementing cAMP to wild-type as well as mutants of *crp* and *cyaA* was also tested. The wild-type cells showed no striking differences in VcdRP expression without or with cAMP at both the optical densities (lanes 1-2 vs. 3-4). However, for *cyaA*-deficient cells, the elevated levels of VcdRP levels at HCD in the absence of cAMP could be suppressed by the addition of cAMP (lanes 6 vs. 8). Interestingly, addition of cAMP to the *crp* mutant also had no effect on the expression of VcdRP at HCD (lanes 4 and 12). Notably, a similar expression pattern was observed when probed for the sRNA Spot42, whose promoter has previously been known to be repressed by CRP (Rice & Dahlberg, 1982; Polayes *et al*, 1988). In summary, the promoter of *vcdRP* harbors a binding motif for CRP and is transcriptionally repressed by it.

Additionally, to corroborate the findings above, a plasmid-based transcriptional reporter of the promoter of *vcdRP* fused to *mkate2* was generated. The background-corrected fluorescence was measured in a wild-type strain as well as strains lacking *crp* and *cyaA*. In line with the Northern blot analysis, the absence of *crp* and *cyaA* resulted in elevated levels of mKate2 (Figure 13B). However, this effect was abrogated upon introducing a deletion (*vcdRP**) in the CRP box of the plasmid-based transcriptional reporter (Figure 12A, 13C). In conclusion, VcdRP is repressed by CRP.

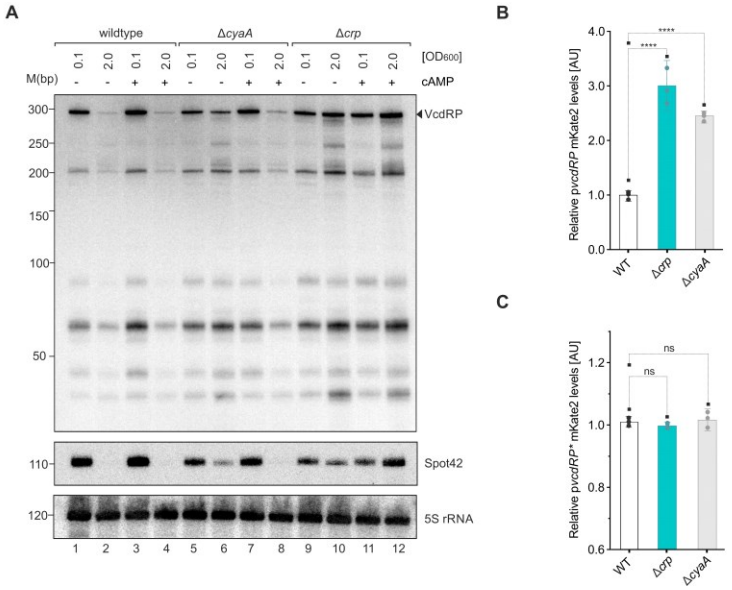

Figure 13. VcdRP is repressed by CRP. A) Northern blot analysis of the expression of VcdRP and Spot42 sRNA monitored at low cell density (OD_{600} 0.1) and high cell density (OD_{600} 2.0), in the absence (minus) or presence (plus) of cAMP supplemented externally. The solid triangle indicates the band corresponding to the full-length primary transcript. Probing with 5S RNA confirmed equal loading. **B, C)** Relative fluorescence intensities of *V. cholerae* wild-type or mutants lacking *crp* or *cyaA* harboring a transcriptional reporter plasmid of *vcdRP* or *vcdRP** fused to *mKate2*. Cells were grown in LB to OD_{600} of 1.0 and fluorophore production was measured. The fluorescence of WT was set to 1 after correcting for autofluorescence. Data are presented as mean ± SD, n = 3. Statistical significance was determined using one-way ANOVA and post hoc Dunnett's multiple comparisons test. The p-value is summarized as follows - **** for $p \leq 0.0001$ and ns for $p > 0.05$.

2.4 Disentangling the dual functions of VcdRP

To study the contributions of the base-pairing RNA element and the small protein, different variants of VcdRP were generated. First, a stop codon was introduced in the third position of the ORF encoding the small protein, VcdP (Figure 14B, D). As a result, this variant served as a readout of the base-pairing riboegulator only (hereafter, VcdR). For sole expression of VcdP, the nucleotide sequence of the ORF was scrambled, however maintaining the same amino acid sequences (Figure 14C, D). Additionally, the sequences flanking the *vcdP* ORF were altered to include an artificial 5´UTR and terminator. This was done to avoid potential regulatory effects originating from the native contiguous VcdRP sequence (Figure 14A).

The native VcdRP as well as the modified VcdR / VcdP variants along with an empty control plasmid were expressed from an arabinose inducible promoter. Total RNA samples were collected prior to as well as after induction with L-arabinose. The expression of each variant was monitored by Northern blotting (Figure 15A). As expected, in the absence of induction with arabinose, no expression of VcdRP was observed (lane 1 vs. 3 vs. 5). Introduction of the stop codon at the third position of the ORF did not affect the expression of the full-length *vcdRP* transcript. (lane 2). Since the nucleotide sequence of VcdP variant differs from the native VcdRP sequence, a different probe specific to the scrambled sequence was used to monitor its expression (lane 6). To further investigate

if the scrambled variant of VcdP is translated *in vivo*, a translational reporter fusion of this variant fused to *sfGFP* was generated. Strikingly, strong GFP production in the presence of the VcdP ORF was observed (~60-fold when compared to an empty control). In addition, upon mutation of the third codon of the ORF (Figure 14B and D), this effect could be abrogated (Figure 15B).

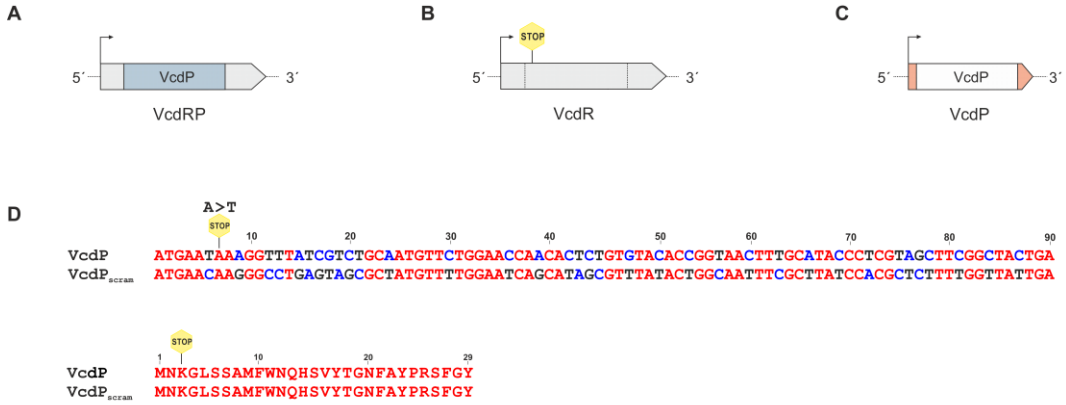

Figure 14. The different VcdRP variants generated for the study. A) A schematic representation of the native VcdRP locus comprising the TSS (indicated by arrow) and the small protein VcdP (in blue). **B)** A mutation from A to T was introduced in the third position of the VcdP ORF, thereby introducing a stop codon in the native VcdP locus (upper panel of D). **C)** The sequences flanking the *vcdP* ORF were altered to include an artificial 5´UTR and terminator (orange). **D) T**he nucleotide sequence of the ORF was scrambled such that the nucleotide sequence is altered with respect to the native sequence (upper panel), however retaining the same amino acid sequence as the native version (lower panel).

To test which element of VcdRP contributed to the repression of CT, over-expression plasmids harboring either an empty vector control or the variants described above were tested in cells lacking *hapR* (analogous to the screen described in section 2.1). Interestingly, both VcdRP and VcdR reduced CT levels, but VcdP had no effect on the production of the protein (Figure 15C), suggesting that only the RNA element in VcdRP is responsible for CT inhibition.

A parallel study in the lab revealed potential RNase E-dependent processing of VcdRP (Hoyos *et al*, 2020). To confirm this, *V. cholerae rne*^{TS} strain that exhibits full RNase E activity at permissive temperatures (30°C) but is rendered inactive under non-permissive temperatures (44°C) was used alongside a wild-type control. Both strains were grown to HCD of OD_{600} of 2.0. Cultures were divided in half and continuously grown at either 30°C or 44°C for 30 minutes. Cleavage patterns of VcdRP was monitored by Northern blotting (Figure 15D). Transfer of the *rne*^{TS} strain to non-permissive temperatures led to a change in mature full-length VcdRP levels as well as its processing intermediates, thereby confirming RNase E-mediated processing of VcdRP.

To further assess the impact of the stop codon mutation on the stability of the full-length VcdRP transcript, samples were treated with rifampicin to measure transcript stability. To this end, over-expression plasmids harboring either VcdRP or VcdR were tested in a *V. cholerae* strain deficient of *vcdRP*. Cells were grown to log-phase and transcription was arrested by the addition of rifampicin. Samples for RNA extraction were collected prior to as well as 2, 4-, 8-, 16- and 32-minutes after addition of the antibiotic. The stability of the transcripts was monitored by Northern blotting (Figure 15E). Interestingly, introduction of the stop codon at the third position (*i.e.*, the VcdR variant) displayed a mildly reduced stability of the full-length transcript when compared to VcdRP (lane 11 vs. 5). However, the shorter processed isoforms were of comparable stability. This suggests that the lack of translation in the VcdR strain does not interfere with the RNase E-mediated processing of VcdRP.

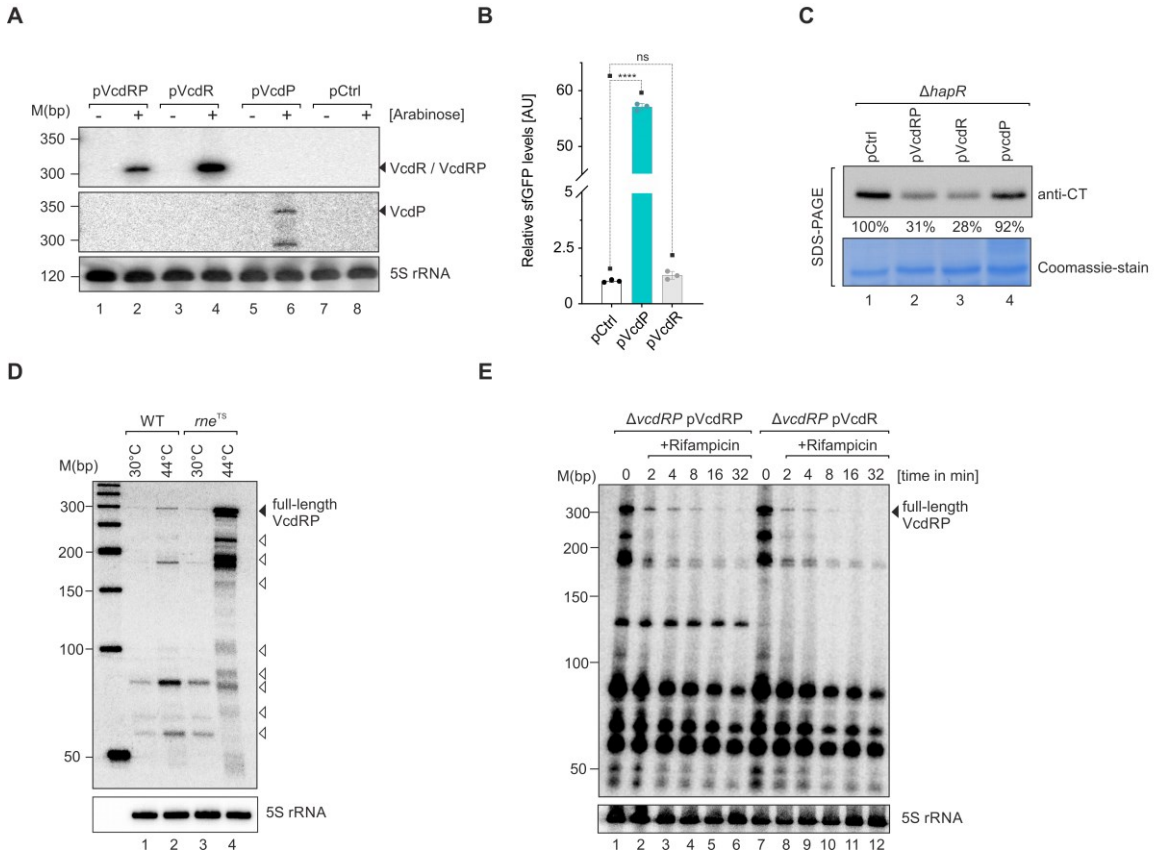

Figure 15. Expression and stability of the different VcdRP variants. A) The different VcdR/P variants were expressed from an arabinose-inducible promoter and RNA samples were analyzed by Northern blotting, before and after induction. Probing for 5S rRNA served as loading control. **B)** Translational fusions of *vcdP* and *vcdR* fused to *sfGFP* were measured at OD_{600} of 1.0. The sfGFP levels were corrected for autofluorescence and fold changes were compared relative to an empty control plasmid set to 1. **C)** Western blot analysis (top) of CTX levels detected in secreted protein fractions of *V. cholerae* Δ*hapR* cells grown under AKI conditions and carrying the indicated plasmids. Coomassie-stained SDS gel (bottom) confirmed equal loading of the protein fractions. The normalized band intensities are indicated for each expression plasmid relative to pCtrl set to 100%. **D)** RNase E-mediated processing of VcdRP**.** *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 were allowed to continue growing at either 30°C or 44°C for 30 min. The cleavage patterns of VcdRP were monitored on a Northern blot. The solid triangle refers to the band corresponding to the full length primary *vcdRP* transcript, whereas the open triangles correspond to the different processed isoforms. Probing with 5S rRNA confirmed equal loading. Northern blot analysis of rifampicin treated RNA samples at the indicated time points. Over-expression strains of VcdRP and VcdR were compared in a *vcdRP* mutant background. Probing for 5S rRNA served as loading control. **D)** Data information: For (B), data are presented as mean ± SD, n = 3. Statistical significance was determined using one-way ANOVA and post hoc Dunnett's multiple comparisons test. The p-value is summarized as follows - **** for $p \leq 0.0001$ and ns for $p > 0.05$.

2.5 Global transcriptome analyses of VcdR/P functions

To study the regulation conferred by the dual-function regulator VcdRP as well as by the individual base-pairing element and the small protein alone, an RNA-Seq analysis was performed on the variants outlined in Figure 14A-C. To this end, *V. cholerae* Δ*vcdRP* strains carrying pBAD-*vcdRP*, pBAD-*vcdR*, pBAD-*vcdP*, or an empty vector control (pCtrl) were cultivated to early exponential phase (OD_{600} of 0.1) in LB medium. Cells were treated with L-arabinose and RNA samples were harvested 15 minutes after induction. Strand-specific cDNA libraries were generated and subjected to Illumina sequencing. Genes with an absolute fold change ≥2.0 and an FDR-adjusted p-value ≤0.05 were considered as differentially expressed.

The study revealed in the differential regulation of 103 genes (Figure 16A). Of these, 84 were upregulated (↑) while 19 were repressed (↓). In contrast, VcdR led to a change in 8 genes (2 ↑ / 6 ↓), whereas VcdP modulated the expression of 49 genes (41 ↑ / 8 ↓). Interestingly, most genes regulated

by either VcdP or VcdR were also differentially expressed in response to VcdRP (Table 1). However, regulation of 64 genes was specific only to VcdRP, suggesting that simultaneous expression of VcdR and VcdP could have a cooperative effect on the cellular function. Enrichment of GO (gene ontology) terms was performed using DAVID tool v6.8 (Huang *et al*, 2009a, 2009b). Interestingly, a high proportion of DEG were involved in central metabolic processes, particularly carbon metabolism (Figure 16B). To further corroborate these findings, a KEGG mapper (v4.3) analysis was performed to infer cellular pathways (Kanehisa & Sato, 2020). The pathways depicted are in line with the CRP-mediated transcriptional control observed in the preceding section.

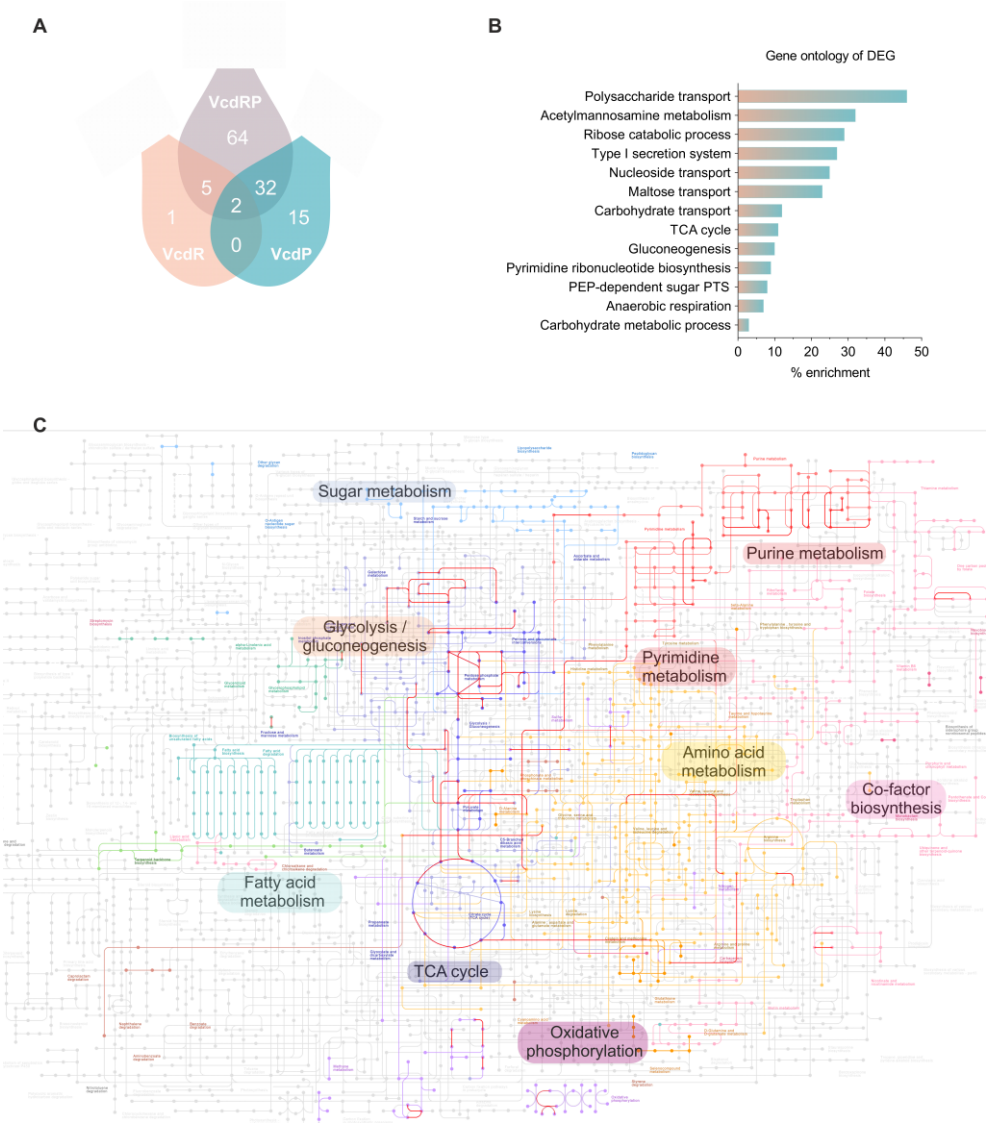

Figure 16. Global transcriptome changes in response to VcdR/P expression. A) A Venn diagram of the differentially expressed genes (DEG) among VcdRP (brown), VcdR (orange) and VcdP (blue). RNA samples were harvested from *V. cholerae* Δ*vcdRP* strains carrying pBAD-based plasmids depicted in (Fig. 14A), including an empty vector control. Genes with a total count cutoff >10 in all samples, with an absolute fold-change ≥ 2.0 and an FDR adjusted *p*-value ≤ 0.05 were considered as differentially expressed. **B)** Gene enrichment analysis of the DEG shown in (A) using gene ontology analysis. Data information: the differentially expressed genes in (A) are shown in Table 1. **C)** KEGG Mapper analysis of the DEG depicting the metabolic pathways that are regulated, in line with the enrichment analysis shown in (B).

2.6 Target spectrum of VcdR

The genes regulated by the riboregulator VcdR revealed 8 potential targets, 6 which were downregulated (*ptsG*, *pstH*, *ptsI*, *nagE*, *treB*, and *vc0177*) and 2 which were upregulated (*lamB* and *vc1779*) (Table 1). Notably, *ptsG*, *nagE*, and *treB* all encode phosphoenolpyruvate phosphotransferase system (PTS) transporters. Additionally, *ptsH* and *ptsI* encode phospho-carrier proteins, which transfer phosphate to the PTS transporters during glycolysis (Houot *et al*, 2010a; Hayes *et al*, 2017).

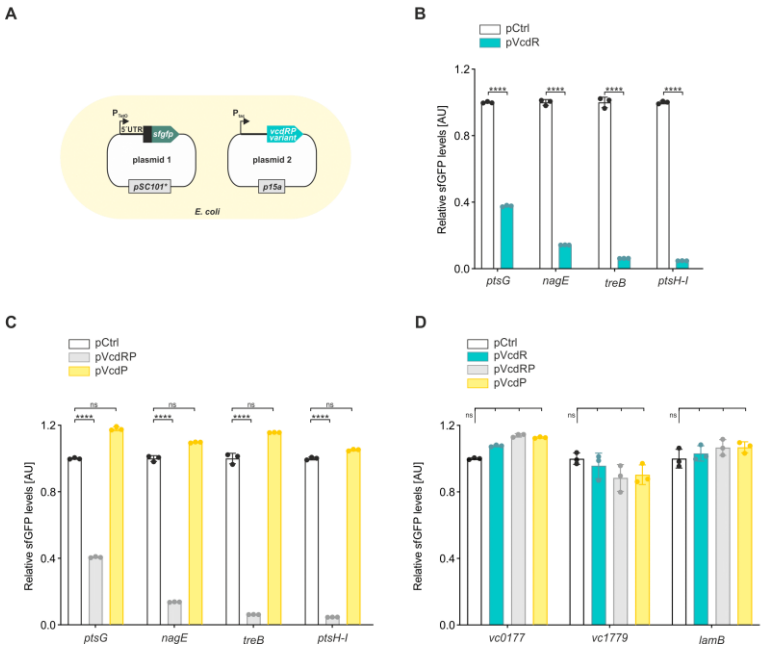

Figure 17. VcdR regulates genes involved in PTS transport. A) A two-plasmid system in *E. coli* to score for post-transcriptional regulation of the target by the riboregulatory component of the dual RNA regulator VcdRP. Plasmid '1' harbors the 5'UTR and the first 20 amino acids of the target genes fused to *sfGFP*, driven by P_{TetO} promoter. On the other hand, plasmid '2' harbors *vcdRP* variants under the control of P_{Tac} promoter. **B)** Plate reader measurements of *ptsG*-, *nagE*-, *treB*- or *ptsHI*- fused to *sfGFP*, combined with either an empty control plasmid (pCtrl) or pVcdR. **C)** Plate reader measurements of the same set of targets fused to *sfGFP*, combined with either an empty control plasmid (pCtrl) or pVcdRP or pVcdP. **D)** Plate reader measurements of *vc0177*-, vc1779-, or *lamb*- fused to *sfGFP*, combined with either an empty control plasmid (pCtrl) or pVcdR, pVcdRP or pVcdP. Data information: fluorophore measurements in (B-D) were monitored from cells that were grown in LB to OD_{600} of 1.0. The levels of sfGFP were calculated as relative fold changes with respect to pCtrl set to 1. The p-value is summarized as follows - **** for $p \leq 0.0001$ and ns for $p > 0.05$.

A previously developed *sfgfp*-based translational fusion system was used to validate the post-transcriptional control of the predicted sRNA targets (Corcoran *et al*, 2012). In this system, the 5′ UTR and the sequence corresponding to the first 20 amino acids of the target genes were fused to *sfgfp* under the control of the P_{TetO} promoter. These plasmids were transferred into a heterologous *E. coli* host along with a second plasmid transcribing variants of *vcdRP* (shown in Figure 14A-C) or an empty control from a P_{Tac} promoter (Figure 17A). Previously determined TSS of *ptsH* and *ptsI* showed that they form an operon (Papenfort *et al*, 2015). Therefore, they were tested in tandem such that the 5′UTR of *ptsH*, its entire gene, and the first 20 amino acids of *ptsI* were fused to *sfGFP*. All the targets tested were repressed significantly by VcdR and VcdRP, but not VcdP (Figure 17B-C). This therefore suggests that these targets are not regulated by VcdP and are specific to the riboregulatory element of VcdRP. However, this assay could only validate all the targets involved in the glycolytic pathway mentioned above, but showed no direct regulation of *vc0177*, *lamB* and *vc1779* (Figure 17D).

To identify the seed region involved in the above base-pairing interaction, the *vcdRP* gene was truncated from its 5′ end, maintaining the terminal 256 bp, 156 bp, 87 bp, and 71 bp sequence,

respectively (Figure 18A). These variants were used to test the repression of sfGFP-based PtsG, TreB, NagE and PtsHI reporters in *E. coli*, as described above. While the 256, 156, and 87 bp variants all repressed GFP production, truncation of *vcdRP* to 71 bp abrogated repression (Figure 18C-F). Therefore, that base-pairing with *ptsG, nagE, treB* and *ptsHI* requires a VcdR sequence element located in the 3′ end of the transcript. Of note, the 156 bp, 87 bp, and 71 bp VcdRP variants correspond to the endogenous VcdRP isoforms associated with RNase E-mediated cleavage (Figure 15D) indicating that the 156 bp and 87 bp processed transcript variants could act post-transcriptionally to control target gene expression, whereas cleavage at position 71 renders the sRNA inactive (Figure 18B).

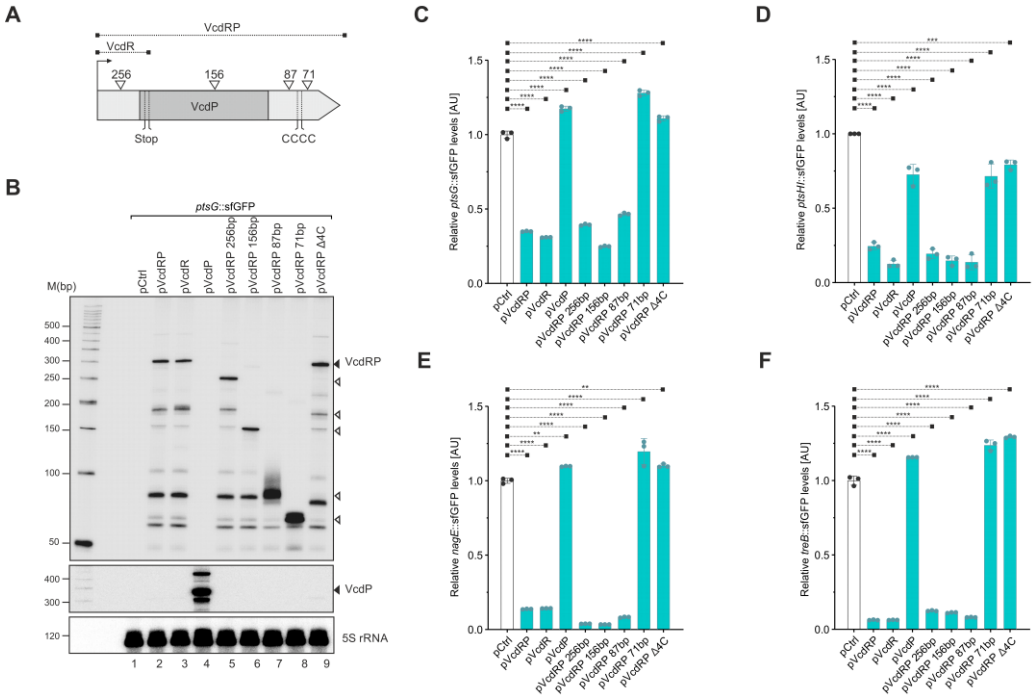

Figure 18. VcdR regulates through a conserved element in the 3´region. A) A schematic representation of the different truncation plasmids of VcdRP tested in (B-F). These include pVcdRP, pVcdR, pVcdP and pVcdRP Δ4C or plasmids truncated from the 5′ end maintaining the terminal 256 bp, 156 bp, 87 bp, and 71 bp of vcdRP, respectively. **B)** The expression of the different VcdRP variants were tested in the sfGFP-reporter of PtsG in *E. coli* by Northern blotting. The solid triangle refers to the band corresponding to the full-length primary VcdRP transcript, whereas the open triangles indicate the sizes of the respective truncations. Probing with 5S rRNA confirmed equal loading. **C-F)** Relative fluorescence intensities of *E. coli* translational reporters of *ptsG* (C), *ptsHI* (C), *nagE* (D) and *treB* (E) fused to sfGFP harboring either an empty control vector (pCtrl) or VcdRP expression plasmids described in (A). Cells were grown in LB to OD600 of 1.0 and fluorophore production was measured. The fluorescence of pCtrl (for each reporter fusion) was set to 1. Data information: data in (C-F) are presented as mean ± SD, *n* = 3. Statistical significance was determined using one-way or two-way ANOVA and post hoc Dunnett's multiple comparisons test. The p-value is summarized as follows - ns for $p > 0.05$, * for $p \leq 0.05$, ** for $p \leq 0.01$, *** for $p \leq 0.001$ and **** for $p \leq 0.0001$.

2.7 Molecular basis for the target mRNA recognition by VcdR

To pinpoint the exact binding site of the base-pairing interaction of VcdR with all its targets, the 3′ end of transcript between the terminal 87 bp and 71 bp stretch was scanned for a conserved stretch of sequence. Indeed, the distal end of *vcdR* harbors four consecutive cytosine residues, conserved among all species tested (Figure 10C). Therefore, this region was deleted (indicated as Δ4C) in *vcdRP* and the repression of each of the target translational fusions was analyzed. In all cases, this mutation abolished regulation of the reporters (Figure 19A). This indicates that VcdRP employs a single conserved base-pairing site to regulate multiple target mRNAs. To bolster these observations, RNA duplexes between VcdR and each of the validated targets from the previous section were predicted using the RNA Hybrid algorithm (Rehmsmeier *et al*, 2004). As expected, the

translational initiation regions of all four tested genes harbor a conserved stretch of four guanine residues (Figures 19B-E). Subsequent mutational analysis validated the base-pairing predictions. Mutation of two of the four cytosine residue to guanine in VcdRP (indicated as M2) abrogated repression of all four targets and could be compensated by the complementary mutation (indicated as M2*) in the targets (Figures 19F-I). Thus, VcdRP employs a conserved stretch of four cytosines at the 3′ end of the transcript to base-pair with all its targets.

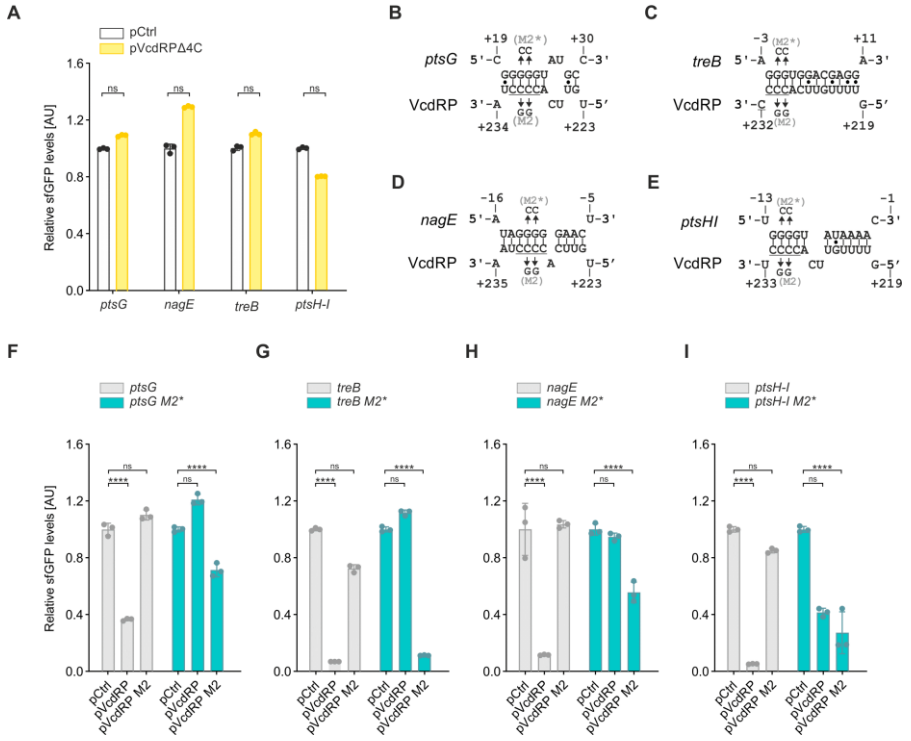

Figure 19. VcdR base pairs via four consecutive cytosines at the 3′end. A) Plate reader measurements of the target 5′ UTRs fused to *sfGFP*, combined with either an empty control plasmid (pCtrl) or pVcdRP Δ4C. **B-E)** Prediction of RNA duplex formation between selected mRNAs and VcdRP. Numbers indicate the distances from the TSS for VcdRP and the start codons of the target mRNA sequences, respectively. Arrows indicate the mutations tested in (F-I). **F-I)** Validation of the predicted mRNA-sRNA duplexes using compensatory base-pair mutations. Data information: fluorophore measurements in (A, F-I) were monitored from cells that were grown in LB to OD_{600} of 1.0. The levels of sfGFP were calculated as relative fold changes with respect to pCtrl set to 1. The p-value is summarized as follows - **** for $p \leq 0.0001$ and ns for $p > 0.05$.

2.8 VcdR modulates sugar homeostasis

Although the glycolytic pathway entails production of energy by the breakdown of glucose, excess accumulation of phosphosugars such as glucose 6-phosphate can be detrimental. High levels of these sugars can result in cell toxicity, impaired growth as well as DNA damage (Irani & Maitra, 1977; Lee & Cerami, 1987; Kadner *et al*, 1992). This damage is in part due to the formation of toxic by-products such as methylglyoxal (Hopper & Cooper, 1971). Similarly, many non-metabolizable sugars can also cause phosphosugar stress. For example, the α-methyl-glucoside (α-MG) is a metabolically inert glucose analog which is also efficiently imported by PtsG in *E. coli* and *Salmonella* (Chou *et al*, 1994; Song *et al*, 2004). Therefore, the levels of these phosphosugar intermediates must be under tight control.

In most bacteria, glucose is transported into cell mainly by a PTS system, comprising two general sugar transport proteins enzyme I (EI) and histidine protein (HPr). Additionally, there are two glucose-specific proteins glucose-specific enzyme IIA ($EIIA^{Glc}$) and $EIICB^{Glc}$ (Deutscher *et al*, 2006). The *V. cholerae* genome encodes 25 PTS components, including two EI homologs, three HPr

homologs, and nine EIIA homologs (Houot *et al*, 2010a). VcdRP represses the production of the PTS-specific transporters PtsG, TreB and NagE that import glucose, trehalose, and N-acetyl glucosamine, respectively (Figure 20A). The expression of the genes encoding these transporters is extensively regulated transcriptionally as well as post-transcriptionally. Upon depletion of glucose, adenylate cyclase is produced, which activates cAMP. CRP acts in concert with cAMP to regulate a plethora of other regulatory processes (Siebold *et al*, 2001). Additionally, both *ptsG* and *ptsHI* are also transcriptionally activated by both cAMP-CRP and glucose (Kimata *et al*, 1997). Since the levels of CRP and cAMP are reduced by glucose, the expression of the glucose PTS should be regulated not only by CRP–cAMP but also by other factors that mediate the effect of glucose. Post-transcriptional control of PTS is exerted by sRNAs like the enterobacterial SgrS in *E. coli* and *Salmonella* (section 1.4).

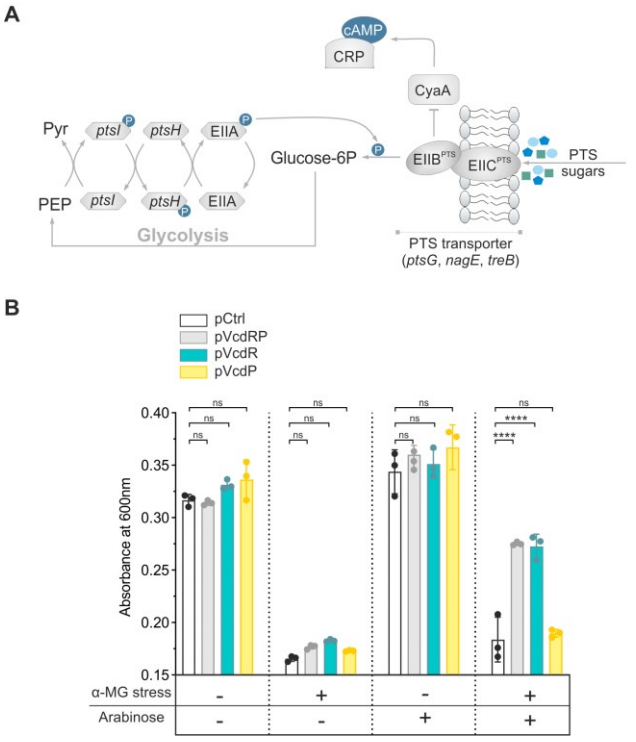

Figure 20. VcdR protects cells from α-MG stress. A) Typical glycolytic pathway in bacteria. PTS sugars such as glucose, trehalose and N-acetyl glucosamine are imported into the cell via their respective transporters PtsG, NagE and TreB. These sugar-specific two component systems convert the imported sugars into phosphsugars. The phosphate themselves are generated from a sugar-unspecific phospho-relay system involving Hpr (encoded by *ptsH* and *ptsI*) and EIIA components. The overall transcriptional control of PTS-dependent sugar import is under the control of cAMP-CRP through CyaA. **B)** The base-pairing element of the dual regulator confers protection against the sugar analog α-MG. *V. cholerae* Δ*vcdRP* strain harboring either an empty vector control (pCtrl) or inducible *vcdRP* expression plasmids (pVcdRP, pVcdR, pVcdP on *x*-axis) were grown to early log phase at which 0.1% α-MG and / or 0.2% arabinose was added. Absorbance at 600nm (*y*-axis) was measured after 5h of growth without (-) or with (+) αMG and / or arabinose (indicated in the lower panel). Data is presented as mean ± SD, *n* = 3. Statistical significance was determined using one-way or two-way ANOVA and post hoc Dunnett's multiple comparisons test. The p-value is summarized as follows - **** for $p \leq 0.0001$ and ns for $p > 0.05$.

Both glucose and its sugar analog α-methyl glucoside (α-MG) are taken up and phosphorylated by the PTS system (Deutscher *et al*, 2006). If the metabolism of glucose-6-phosphate (G6P) is blocked or if cells accumulate non-metabolizable α-MG6P, *sgrS* is induced by the transcription factor SgrR. The sRNA regulates numerous mRNA targets via base pairing interactions that result in alterations in mRNA translation and stability. Regulation of target mRNAs allows cells to reduce import of additional sugars and increase sugar efflux (Wadler & Vanderpool, 2007; Rice & Vanderpool, 2011). To understand the role of VcdRP analogous to SgrS, *V. cholerae* strains were grown in the presence of α-MG. Accordingly, *V. cholerae* Δ*vcdRP* strain harboring either an empty vector control (pCtrl) or inducible *vcdRP* expression plasmids (pVcdRP, pVcdR, pVcdP) were grown to early log phase at which 0.1% α-MG and / or 0.2% arabinose was added. Absorbance at 600nm was measured after 5h

of growth without (-) or with (+) α-MG and / or arabinose. Indeed, the sugar analog strongly reduced the growth of cells, and this effect could be offset when VcdRP or VcdR were over-expressed. In contrast, VcdP did not rescue this phenotype (Figure 20B). This effect could be attributed to the VcdR-mediated repression of PtsG, that suppress further import and consequently promote cell replication.

2.9 Regulation by VcdP

In contrast to VcdR, VcdP regulates far more targets: 49 genes (41 ↑ / 8 ↓). Of the downregulated targets, two of them have been annotated as hypothetical proteins (*vca0743* and *vca0052*, Table 1). The remaining 6 genes are involved in: (a) providing a link between glycolytic and the pentose phosphate pathway (*vca0623*, *vca0624* and *vc2350*), (b) nucleoside catabolism (*vca0053*, *vc1953*), and transport activities (*vc2761*). However, only *vca0053* was specifically and most strongly downregulated by VcdP and not VcdR (~7-fold). Accordingly, further validation was performed on this gene alone. The gene *vca0053* catalyzes the first step in either purine base salvage or nucleoside catabolism (Bzowska *et al*, 2000).

To correlate the expression pattern observed in the transcriptome data, a chromosomal 3xFLAG tag was introduced at the C-terminal end of *vca0053*. Total protein and RNA samples were collected from *V. cholerae vca0053*::3xFLAG strains harboring either an empty control plasmid (pCtrl) or over-expression VcdRP plasmids (pVcdRP, pVcdR, pVcdP and pVcdP-SPA). Protein samples were examined on a Western blot by probing with anti-FLAG antibody (Figure 21A). RNA samples were analyzed using qRT-PCR to examine the transcript levels of *vca0053* in response to each variant (Figure 21B). As expected, VcdP and VcdP-SPA repressed the levels of *vca0053* transcript and Vca0053 protein levels. This reduction in levels could also be observed for VcdRP, but not VcdR, therefore making *vca0053* a VcdP-specific target.

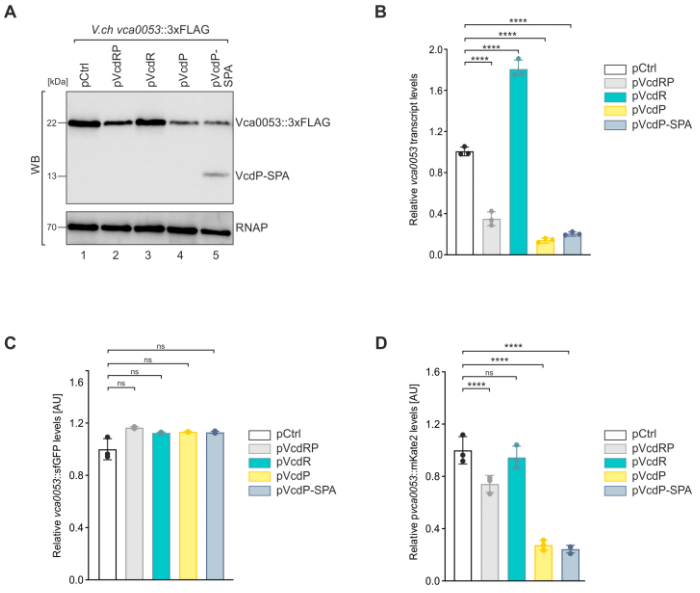

Figure 21. VcdP represses purine nucleoside phosphorylase production. A) Western blot analysis of chromosomally 3x-FLAG tagged Vca0053 in response to either an empty control (pCtrl) or overexpression VcdRP plasmids (pVcdRP, pVcdR, pVcdP, pVcdP-SPA). RNAP served as loading control. **B)** qRT-PCR analysis of the RNA samples harvested in parallel to (A). The *vca0053* transcript levels are plotted as relative fold-changes with respect to pCtrl set to 1. **C-D)** Fluorescence measurements of translational or transcriptional fusions of *vca0053* fused to *sfGFP* (C) or *mKate2* (D), combined with the same set of plasmids as in (A-B). Cells were grown in LB to OD_{600} of 1.0 and the levels of each fluorophore was calculated as relative fold change with respect to pCtrl set to 1. The p-value is summarized as follows - **** for $p \leq 0.0001$ and ns for $p > 0.05$.

To decipher the nature of control, both translational and transcriptional reporter fusions of *vca0053* were generated. To this end, the 5'UTR and the first 20 amino acids of the gene was fused to *sfGFP* to generate a translational reporter (similar to section 2.6). Alternatively, a transcriptional reporter was generated by fusing the promoter of *vca0053* with *mKate2*. Production of fluorescence in response to co-expression of an empty control plasmid or VcdRP over-expression plasmids (pVcdRP, pVcdR, pVcdP and pVcdP-SPA) was measured in LB medium. No significant regulation of the translational reporter in response to any of the variants was observed (Figure 21C). In contrast, the production of Vca0053::mKate2 was inhibited by VcdP and VcdRP. As expected, VcdR failed to repress this reporter in line with the transcriptome analysis. Notably, adding a C-terminal SPA tag to VcdP did not interfere with the readout of either of the assays, suggesting that the relatively large tag by itself did not affect the regulatory role of the small protein VcdP (Figure 21D).

2.10 LC-MS analysis of VcdP-SPA

To gain further insights into the potential interaction partners of the small protein VcdP, a co-immunoprecipitation (co-IP) experiment followed by mass spectrometry (LC-MS) analysis was performed. To achieve this, *V. cholerae* strains harboring either an empty vector control or an over-expression strain of VcdP-SPA were cultivated in LB medium to mid-log phase in biological triplicates. Cells equivalent to 50 OD_{600} units were collected and lysed using bead ruptor. The cleared lysates were subsequently subjected to IP in the presence of monoclonal anti-FLAG antibody and protein G Sepharose. The resulting protein samples were examined on by Western blotting using anti-FLAG antibody (Figure 22A). Samples for LC-MS were processed using the single-pot solid-phase-enhanced sample preparation (SP3) protocol (Hughes *et al*, 2019) and has been described in detail in chapter 6. The raw MS data were searched with semi-trypsin specificity with a maximum of 4 missed cleavages. The results were filtered using the following criteria: protein level FDR <5% with a minimum of two high confidence peptides (FDR <1%) with at least one is unique peptide (Table 2).

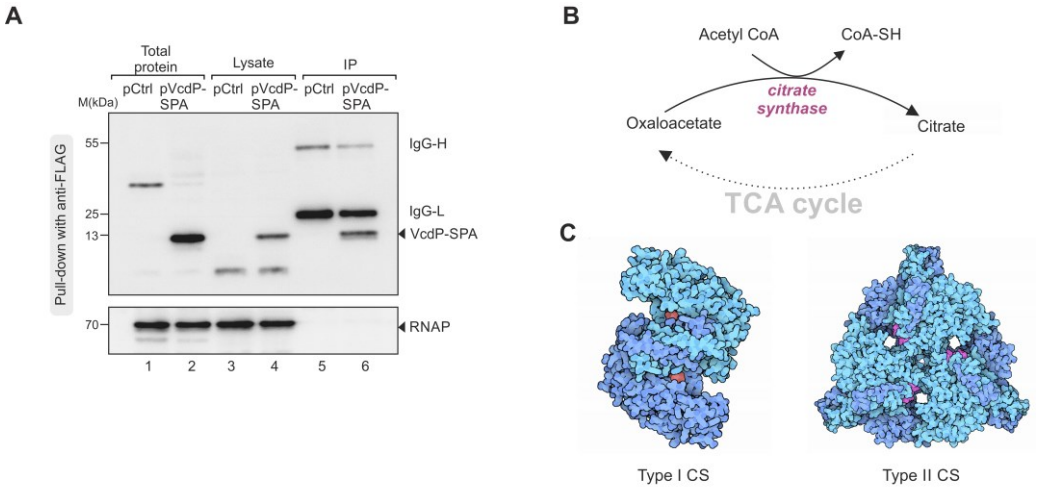

Figure 22. LC-MS identifies citrate synthase as an interacting partner of VcdP. A) Co-IP experiment performed on *V. cholerae* strains harboring either an empty vector control or an over-expression strain of VcdP-SPA. Protein samples corresponding to the total input and cell lysates before and after subjecting to immunoprecipitation with anti-FLAG antibody were loaded on an SDS-PAGE gel. Western blotting with anti-FLAG antibody confirmed successful pull-down of VcdP. RNAP served as loading control. The solid triangles indicate the corresponding protein sizes. **B)** Citrate synthase was the protein that was identified in the LC-MS analysis among all the SPA-tagged replicates but absent in the control strains. CS catalyzes the first step of the TCA cycle and converts oxaloacetate and acetyl-coA to form citrate, along with a thiol group that gets released. **C)** CS exist in two forms: type I (left) exist as dimers whereas type II (right) exist as hexamers.

Interestingly, among the biological replicates that were analyzed, only one protein was found in all three replicates of VcdP-SPA and was therefore considered for further analysis. This protein was citrate synthase (CS, encoded by *gltA*), which provides the gateway into the TCA cycle (LaNoue *et al*, 1972). CS is the enzyme catalyzing the condensation of acetyl-CoA and oxaloacetate to form citrate (Figure 22B). The enzyme exists in two forms: Type I CS exist as dimers and found are in eukaryotes, archaea and Gram-positive bacteria, whereas Type II CS are hexamers (dimer of trimers) and are limited to Gram-negative bacteria (Figure 22C; Tong & Duckworth, 1975; Remington, 1992; Russell *et al*, 1994).

2.11 VcdP interacts with citrate synthase

To validate the interaction of VcdP with CS, another co-IP experiment was setup. *V. cholerae* strains carrying a chromosomal HA tag at the C-terminus of *gltA* harboring either an empty vector control or an over-expression strain of VcdP-SPA were cultivated in LB medium to mid-log phase in biological triplicates. Cellular lysates were subjected to immunoprecipitation with either anti-FLAG or reciprocally with anti-HA antibodies. VcdP co-precipitated when GltA-HA was used as bait (Figure 23A). Reciprocally, GltA co-precipitated when VcdP::SPA was used as bait (Figure 23B). To evade tag-specific false-positive results, a chromosomal 6xHis tag was introduced in the C-terminus of *gltA* and the co-IP /reciprocal IP experiments were carried out using anti-FLAG and anti-His antibodies, respectively. Consistent with the previous IP, VcdP-SPA could also be recovered using GltA::6xHis as bait and vice-versa (Figures 23 C and D).

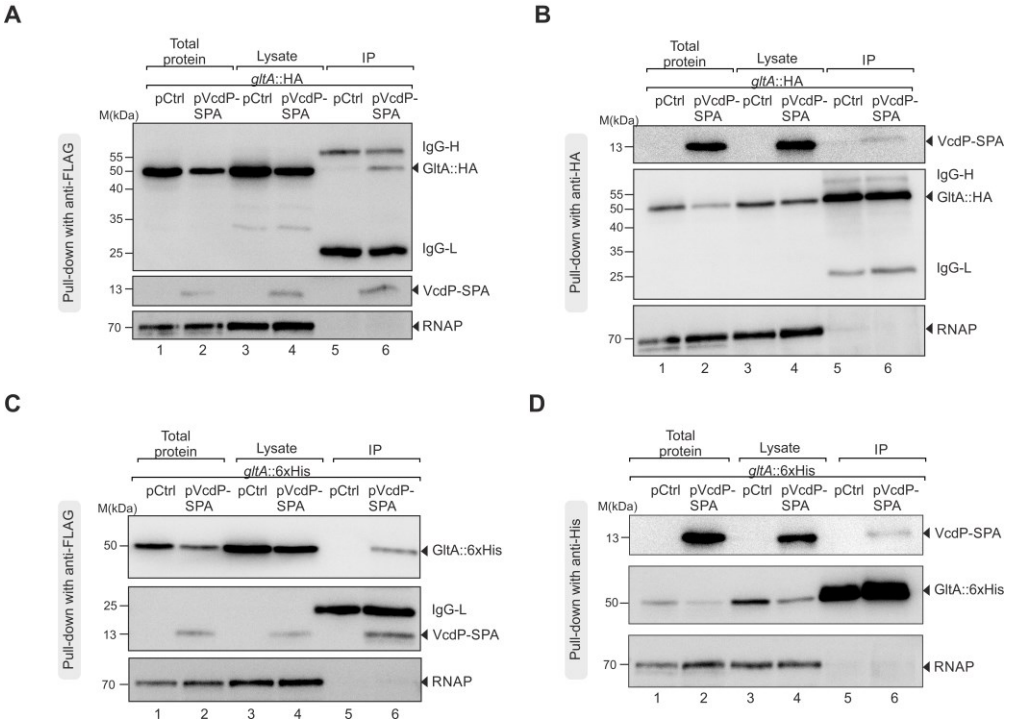

Figure 23. Citrate synthase co-precipitates with VcdP. A-B) IP of *V. cholerae gltA*::HA strains harboring an empty vector control or an over-expression strain of VcdP-SPA. Protein samples corresponding to the total input and cell lysates before and after subjecting to immunoprecipitation with anti-FLAG (A) and anti-HA (B) antibodies were loaded on an SDS-PAGE gel. Western blotting with the corresponding antibodies confirmed successful pull-down of GltA and VcdP, respectively. RNAP served as loading control. The solid triangles indicate the corresponding protein sizes. **C-D)** IP and reciprocal IP performed in line with (A, B), however using *V. cholerae gltA*::6xHis strains harboring the corresponding plasmids, subjected to pull-down using anti-FLAG (C) and anti-His (D) antibodies.

2.12 VcdP enhances the activity of citrate synthase

To test whether interaction of VcdP affects the activity of the enzyme, a colorimetric assay was employed. The by-product CoA-SH formed during the condensation step (Figure 22B) reacts with Ellman's reagent, DTNB (5,5'-dithiobis-(2-nitrobenzoic acid)) to form TNB (Figure 24A). This reaction exhibits a measurable absorbance at 412nm, the intensity of which is proportional to the citrate synthase activity (Srere, 1969).

To examine how VcdP modulates CS activity, cell extracts were obtained from *V. cholerae* wild-type and Δ*vcdRP* cells carrying either an empty control vector (pCtrl), the VcdP over-expression plasmid, or a modified version of this plasmid in which amino acids 15-18 of VcdP were all mutated to alanine (Figure 24B, indicated as VcdP*). These cultures were grown to stationary phase and lysed using bead ruptor. After determining their protein concentrations, same amount of each sample was subjected to a coupled enzymatic reaction catalyzed by oxaloacetate and acetyl CoA as substrates. DTNB was added to the reaction mix to serve as a readout of the colorimetric product (Figure 24A).

Interestingly, over-expression of VcdP led to an increase in CS activity in both the wild-type as well as Δ*vcdRP* strains (Figure 24C). In contrast, VcdP* did not activate CS in both backgrounds. Comparison of the control plasmids revealed a mild reduction in enzyme activity in the *vcdRP* mutant than the wild-type, suggesting that chromosomal VcdP also facilitates GltA activity. The effect of VcdP was also tested in *V. cholerae* cells lacking *gltA*. As expected, the enzyme activity was drastically reduced, however the basal levels that are observed despite the deletion of the *gltA* gene could be attributed to the residual CS activity emanating from a paralog encoding for methyl-CS) (*vc1337*; Heidelberg *et al*, 2000).

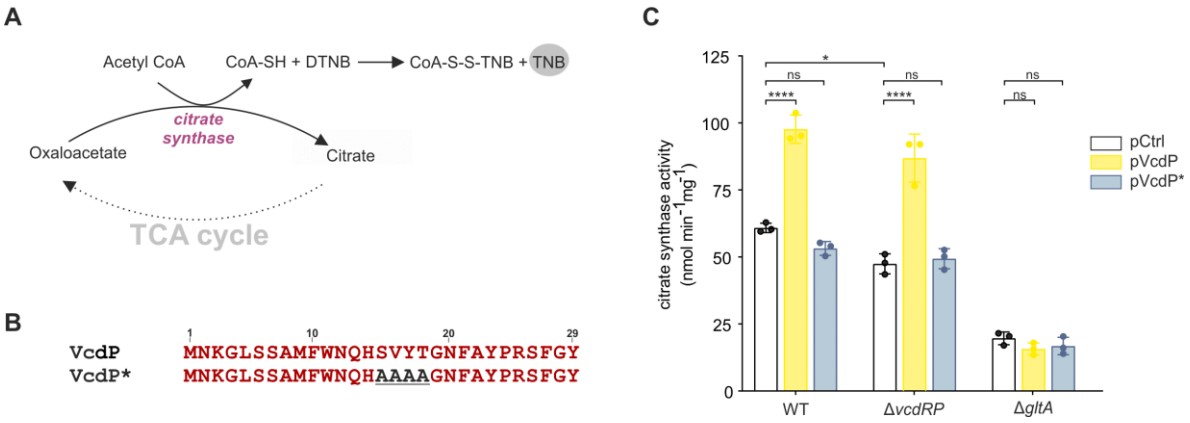

Figure 24. VcdP accelerates the activity of CS enzyme. A) The principle of CS activity measurement is based on the thiol group that is formed during the conversion of oxaloacetate and acetyl coA to citrate by CS. This thiol group reacts with Ellman's reagent (DTNB) to form yellow TNB. This absorbance measured at 412nm is proportional to the amount of CS present. **B)** A mutated version of VcdP was generated by replacing the amino acids from positions 15-18 of the small protein with alanine as underlined in black. **C)** CS enzyme activity measurements performed on cellular lysates. An empty vector control (pCtrl) or *vcdP* expression plasmids (pVcdP and pVcdP*) were conjugated into *V. cholerae* wild-type, Δ*vcdRP* and Δ*gltA* backgrounds. Cells were grown in LB medium to stationary phase and subsequently lysed. These cellular extracts served as input for measuring the activity of citrate synthase enzyme colorimetrically at an absorbance of 412nm.

2.13 VcdP specifically regulates type II CS

As indicated in Figure 22C, the structural basis for type I and type II CS are different: type I CS exist as dimers, whereas type II CS are hexamers (dimer of trimers) (Tong & Duckworth, 1975). To study how VcdP affects the two types of CS, the *V. cholerae* VcdP sequence was cloned into compatible expression plasmids for *Vibrio natriegens* and *Bacillus subtilis* (Brockmeier *et al*, 2006; Schleicher *et al*, 2018). The CS activity of cell lysates was measured in these organisms in comparison to an empty vector control. While over-expression of VcdP substantially upregulated CS activity in *V. natriegens* (type II), it did not have any effect on *B. subtilis* (type I; Figure 25A). These results suggest that VcdP-mediated activation of CS might be limited to the hexameric type II of Gram-negative bacteria.

In addition to the structural differences between type I and type II CS, only type II CS are specifically inhibited by NADH (nicotinamide adenine dinucleotide with hydrogen). To test if VcdP affects CS activity by counteracting inhibition by NADH, a previously studied mutant of *gltA* was used in which a phenylalanine residue at position 383 of the protein was changed to alanine (F383A). This GltA variant is capable of binding NADH, however, its enzymatic activity is not inhibited by the cofactor (Nguyen *et al*, 2001; Maurus *et al*, 2003). This mutation was introduced in the *V. cholerae* chromosome and cell extracts were obtained from chromosomal F383A mutation in *gltA,* harboring either a control vector or VcdP expression plasmids (pVcdP and pVcdP*). In contrast to wild-type, the F383A variant was not activated by VcdP (Figure 25B). This indicates that that VcdP might suppress the inhibitory effect of NADH on CS, and consequently exert its regulatory role specifically on type II hexamers.

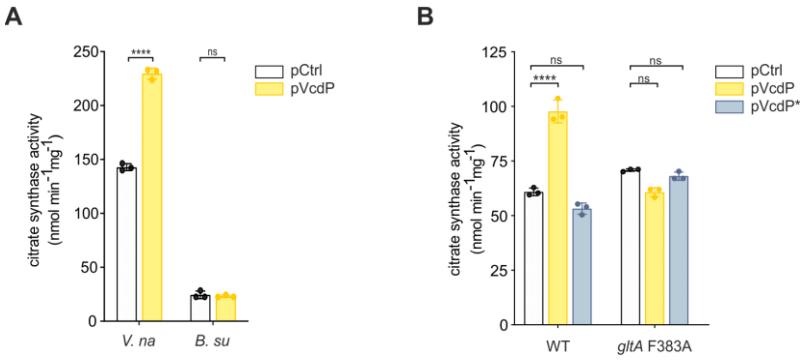

Figure 25. VcdP binds to type II CS. A-B) CS enzyme activity measurements performed on cellular lysates. An empty vector control (pCtrl) or VcdP expression plasmids (pVcdP and pVcdP*) were conjugated into, *V. natriegens* and *B. subtilis* wild-type (A) or *V. cholerae* wild-type and *gltA* F383A backgrounds (B). Data information: data in (A-B) are presented as mean ± SD, n = 3. Statistical significance was determined using one-way ANOVA and post hoc Dunnett's multiple comparisons test. The p-value is summarized as follows - ns for $p > 0.05$ and **** for $p \leq 0.0001$.

To examine the impact of VcdP binding to CS *in vitro*, native *V. cholerae* GltA and F383A proteins were purified as previously described for *V. cholerae* VqmR (Papenfort *et al*, 2017). Briefly, N-terminal 6xHis tagged proteins were expressed in *E. coli* inducible plasmids. Pellets from cells were induced with IPTG, treated with protease inhibitor, and lysed via sonication. The cleared lysates were then applied to Ni-NTA resin and incubated at 4°C. After several wash steps, the washed mixture was loaded on a polypropylene column where on-column cleavage was induced using elution buffer. Protein purification was verified by SDS-PAGE analysis by the prominent band observed at the expected sizes (Figure 26A). A pre-determined amount of each protein was treated with increasing concentrations (1x, 2x, 5x, 10x, 15x and 20x) of commercially synthesized peptide variants (VcdP and VcdP*). In line with the previous measurements, addition of VcdP activated CS activity in a

concentration-dependent manner (Figure 26B). Interestingly, VcdP also activated F383A, however to a weaker extent. In contrast, VcdP* was unable to activate either of the proteins even when added at a twenty-fold excess. Thus, VcdP modulates CS activity both *in vivo* as well as *in vitro* and this might also affect the allosteric inhibition by NADH.

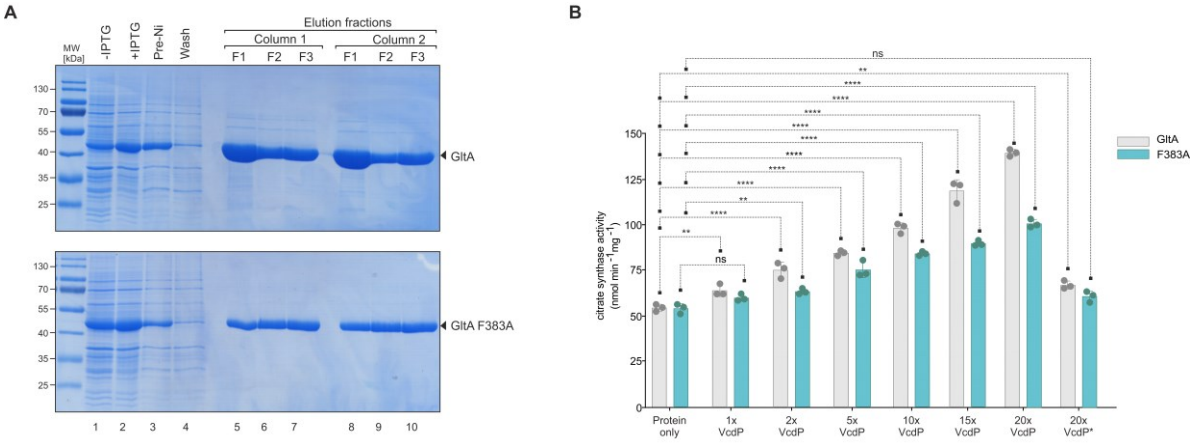

Figure 26. VcdP accelerates the *in vitro* activity of hexameric CS. A) *V. cholerae* GltA and F383A proteins were purified in *E. coli* and verified on an SDS-PAGE gel. The solid triangle corresponds to the expected size of the proteins. **B)** CS enzyme activity measurements performed *in vitro*. The purified proteins from (A) were treated without or with increasing concentrations of synthesized VcdP or VcdP* peptides (1x, 2x, 5x, 10x, 15x and 20x with respect to the protein monomer). The CS activity because of this interaction was measured. Data are presented as mean ± SD, n = 3. Statistical significance was determined using one-way ANOVA and post hoc Dunnett's multiple comparisons test. The p-value is summarized as follows - ns for $p > 0.05$, ** for $p \leq 0.01$ and **** for $p \leq 0.0001$.

2.14 Metabolome analysis in response to VcdR/P expression

By now it is evident that VcdRP modulates sugar uptake via VcdR and the TCA cycle by altering CS activity. This control of central metabolism is not limited to energy production but is also involved in the generation of various metabolic intermediates including reduced purine nucleotides through oxidative phosphorylation (Table 1). To better understand how VcdR/P over-expression affects the glycolytic and TCA cycle pathways, a targeted metabolome analysis was performed. The typical workflow involves three main steps: (a) sample preparation, (b) formal LC/MS detection of metabolites, and (c) data analysis to normalize and quantify these metabolites (Wang & Huang, 2021). The strains tested were *V. cholerae* wild-type, Δ*vcdRP and* Δ*gltA* each harboring an empty control plasmid (pCtrl) and Δ*vcdRP* with pVcdRP expression plasmids (pVcdRP and pVcdP) as well as Δ*gltA* harboring pVcdP. Cells were grown in LB medium to exponential and stationary phase. LC/MS samples were prepared and detected in accordance with a previously established protocol (Buescher *et al*, 2010).

2.14.1 Impact on glucose metabolism

While glycolysis is a ubiquitous pathway in bacteria and serves as the primary route for carbohydrate breakdown, not all of them use the canonical Embden–Meyerhoff–Parnass (EMP) glycolytic pathway. Glucose metabolism in prokaryotes is particularly diverse and relies on alternate pathways like Entner–Doudoroff (ED) and phosphoketolase (PK) pathways to also generate phosphoenolpyruvate (PEP), pyruvate and acetyl-CoA (Peekhaus & Conway, 1998; Sauer & Eikmanns, 2005). Although *V. cholerae* can utilize both EMP and ED pathways, the former is considered to be more efficient (Patra *et al*, 2012). Thus, the metabolites of the EMP pathway were analyzed in detail.

Intriguingly, in contrast to wild-type cells harboring a control plasmid, the lack of *vcdRP* resulted in elevated pyruvate and PEP levels under exponential growth conditions. This effect could in turn be offset by the over-expression of VcdRP, resulting in reduced levels of both the metabolites (Figure 27, white vs. green vs. grey bars). Additionally, the glucose-mediated effect on VcdRP observed earlier (Figure 10A), could be correlated with the increased amounts of glucose-6-phosphate upon over-expression of VcdRP under exponential growth (grey bars). In contrast, the levels of these metabolites show no striking differences during stationary phase growth. This could be attributed to the transcriptional control exerted by cAMP-CRP on VcdRP (section 2.3). Remarkably, over-expression of VcdP (yellow bar) had little impact on any of the glycolytic metabolites. This thus reinforces the hypothesis that VcdR acts in regulating sugar uptake and glucose metabolism, while VcdP acts further downstream on the TCA cycle.

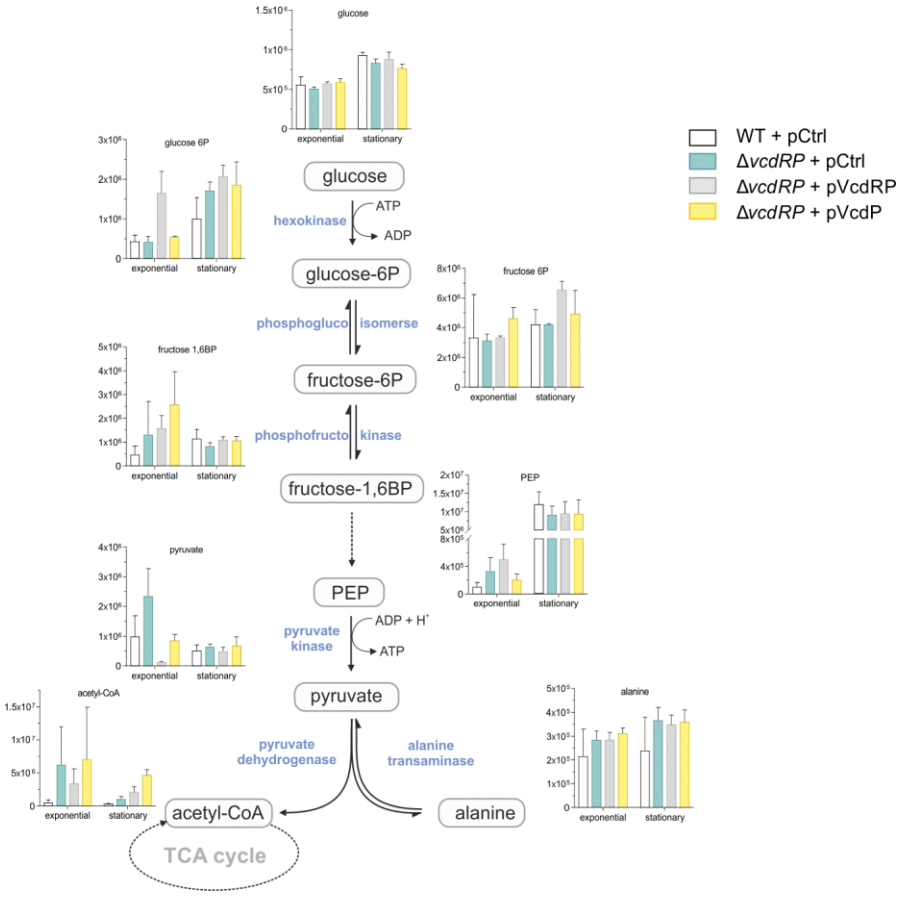

Figure 27. Abundance of glycolytic metabolites. *V. cholerae* wild-type and Δ*vcdRP*, each harboring an empty vector control (pCtrl) or Δ*vcdRP* with VcdRP expression plasmids (pVcdRP and pVcdP) were grown in exponential and stationary phase of growth. LC-MS analyses on these samples are plotted as normalized metabolite abundance. The y-axis represents the peak area of each metabolite. The main operative EMP pathway including the fates of certain amino acids are indicated. Data are presented as mean ± SEM, n = 3. Abbreviations: glucose-6P (glucose 6-phosphate), fructose-6P (fructose 6-phosphate), fructose-1,6BP (fructose 1,6-bisphosphate), PEP (phosphoenolpyruvate).

2.14.2 Impact on TCA cycle

The TCA cycle (a.k.a. Krebs cycle or citric acid cycle) was identified nearly 80 years ago and is now recognized to be the final pathway in aerobic organisms for oxidation of carbohydrates, fatty acids, and amino acids (Krebs, 1937; Korla & Mitra, 2014). The TCA cycle not only serves to catalyze the complete oxidation of acetyl-CoA to CO_2, but also provides carbon precursors and NADPH for biosynthetic processes (Martin, 2020). The TCA cycle involves at least 10 enzymes and

several amino acids feeding into as well as flushing out excess metabolites from the cycle. Additionally, the glyoxylate shunt entails the conversion of 2-carbon compounds to be converted to a 4-carbon molecule, succinate, which in turn can be converted to sugar or to amino acids or can replenish the TCA cycle (Ensign, 2006).

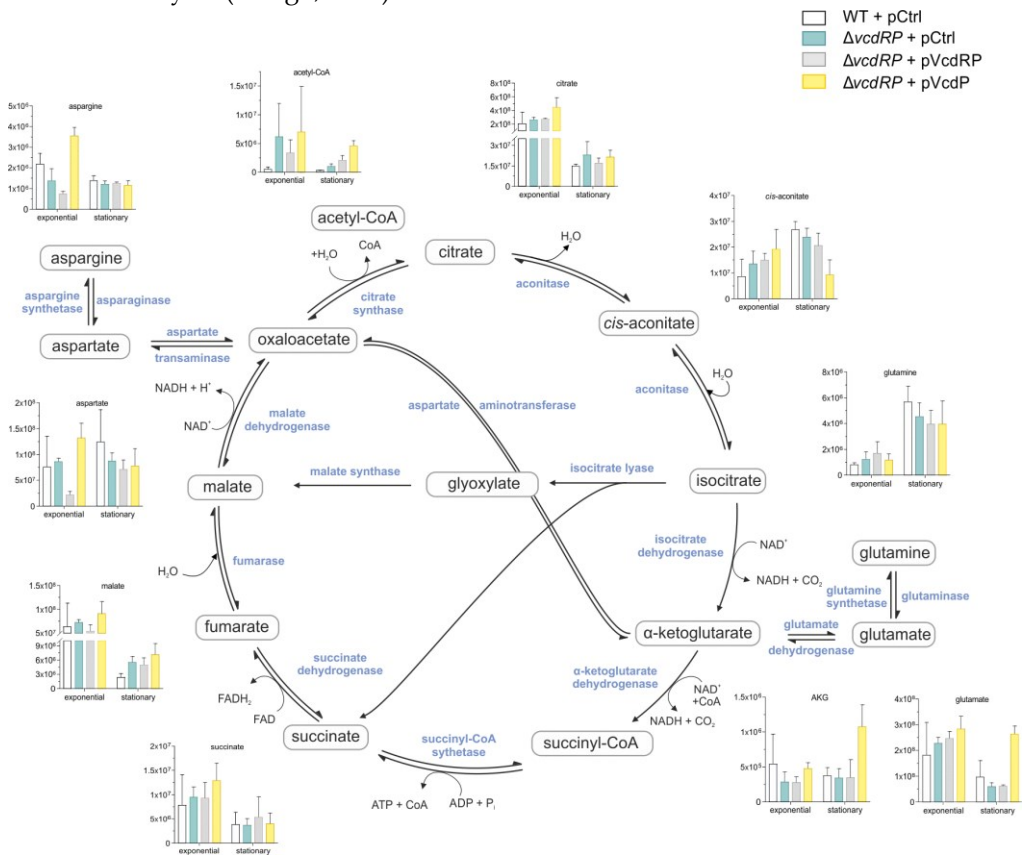

Figure 28. Abundance of TCA cycle metabolites. *V. cholerae* wild-type and Δ*vcdRP*, each harboring an empty vector control (pCtrl) or Δ*vcdRP* with VcdRP expression plasmids (pVcdRP and pVcdP) were grown in exponential and stationary phase of growth. LC-MS analyses on these samples are plotted as normalized metabolite abundance. The y-axis represents the peak area of each metabolite. The main operative pathway including the fates of certain amino acids are indicated. Data are presented as mean ± SEM, n = 3.

Analysis of metabolites belonging to the TCA cycle revealed elevated citrate levels upon VcdP over-expression, in line with the results from the preceding section (Figure 28). Interestingly, all the metabolites with the exception of α-ketoglutarate show higher levels when compared to the wild-type under exponential growth. α-ketoglutarate can feed into the cycle as well as be quenched out of the cycle by interconversion to glutamate by the enzyme glutamate dehydrogenase (Yang *et al*, 2014). Another key observation is the inverse trend between aspartate and glutamate upon over-expressing the small protein under both exponential and stationary phases. Notably, aspartate and α-ketoglutarate reversibly convert to glutamate and oxaloacetate by the enzyme aspartate aminotransferase (Salerno *et al*, 1982). A previous study in *E. coli* has underscored the importance of the fluctuating α-ketoglutarate pool in response to carbon availability (Huergo & Dixon, 2015). Not surprisingly, the α-ketoglutarate node coordinates multiple carbon dedicated pathways including fatty acid production and carbon uptake. Notably, the metabolic profiles of *V. cholerae* Δ*gltA* cells carrying the pVcdP plasmid or a control plasmid did not change levels of the levels of α-ketoglutarate and glutamate (Figures 29, 30). The unchanged levels of all metabolites in these strains also suggest that VcdP needs to exert its effect on citrate synthase to modulate carbon flux.

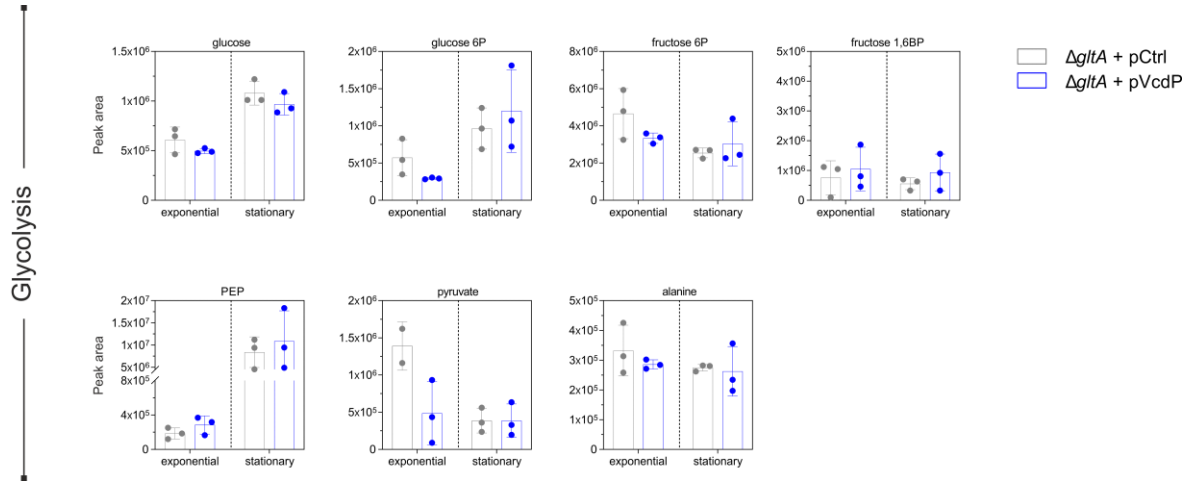

Figure 29. Abundance of glycolytic metabolites in the absence of CS. *V. cholerae* Δ*gltA* cells harboring an empty vector control (pCtrl) or an over-expression plasmid of the small protein (pVcdP) grown in exponential and stationary phase of growth. LC-MS analyses on these samples are plotted as normalized metabolite abundance. The y-axis represents the peak area of each metabolite. Data are presented as mean ± SEM, n = 3. Abbreviations: glucose-6P (glucose 6-phosphate), fructose-6P (fructose 6-phosphate), fructose-1,6BP (fructose 1,6-bisphosphate), PEP (phosphoenolpyruvate).

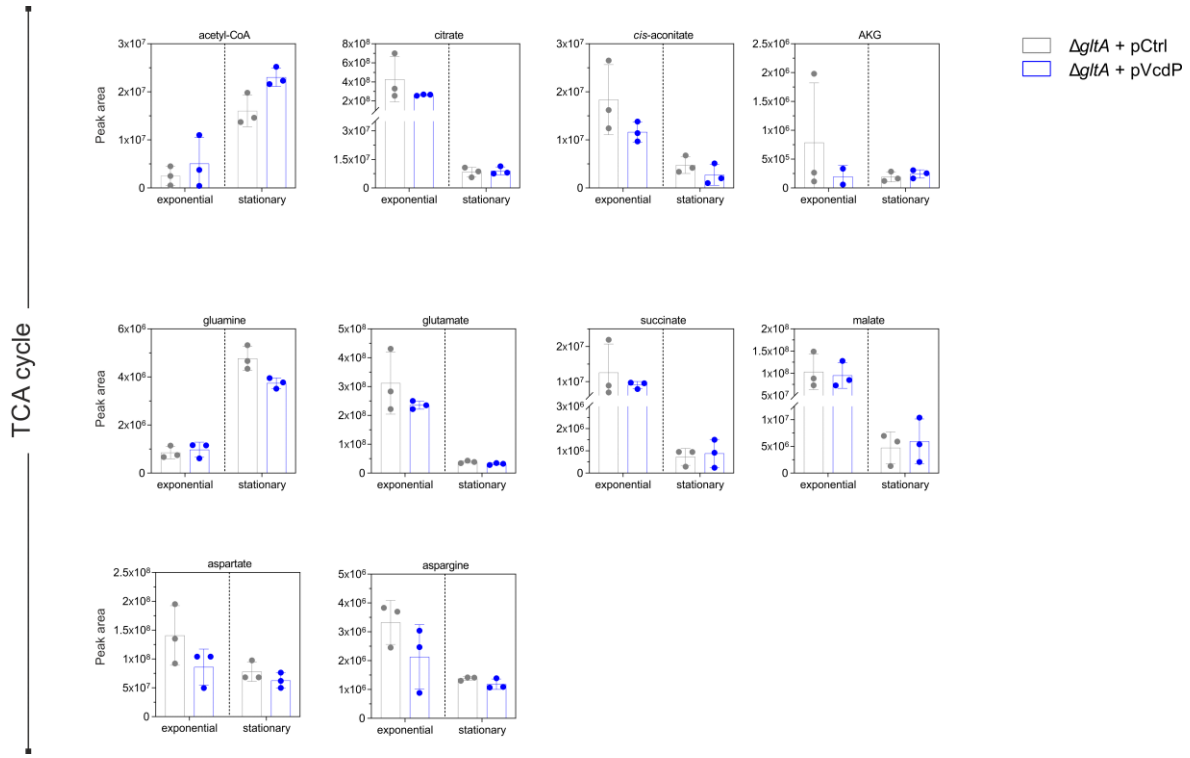

Figure 30. Abundance of TCA cycle metabolites in the absence of CS. *V. cholerae* Δ*gltA* cells harboring an empty vector control (pCtrl) or an over-expression plasmid of the small protein (pVcdP) grown in exponential and stationary phase of growth. LC-MS analyses on these samples are plotted as normalized metabolite abundance. The y-axis represents the peak area of each metabolite. Data are presented as mean ± SEM, n = 3.

2.15 Concluding summary

To make the most of their environment, bacterial control of nutrient uptake and utilization is precisely regulated. In addition, pathogenic microorganisms frequently couple the production of their virulence factors with nutrient availability and the overall metabolic status of the cell. The pathogen *V. cholerae* colonizes and infects the upper intestines by producing two key virulence determinants – TCP and CT toxin. While all the known small regulatory RNAs (sRNA) of *V. cholerae* act directly or indirectly to regulate the production of TCP, the sRNA VqmR is the only known direct repressor of CT production to date. Therefore, a forward genetic screen was employed to score for CT repression. This screen identified another promising candidate called Vcr082. Although sRNAs are assumed to be generally non-coding, there is a growing list of regulators in bacteria that also harbor an open reading frame (ORF), thereby making them dual-function regulators. Interestingly, Vcr082 also encodes a 29 aa ORF and hence was re-named VcdRP, for *V. cholerae* dual RNA regulator and protein, eponymous to their roles.

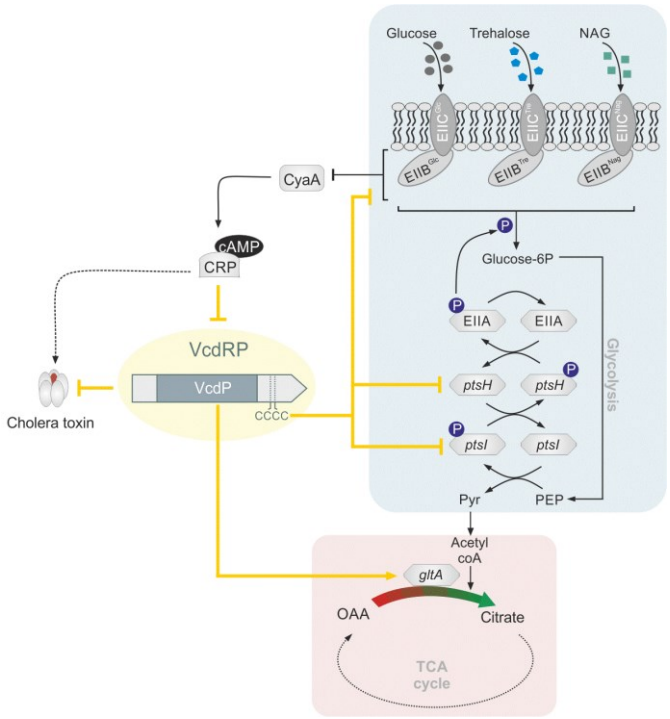

Figure 31. Model depicting the dual function of VcdRP in balancing carbon metabolism. VcdRP is a dual-function regulator that represses CT production in *V. cholerae*. Additionally, CRP and cAMP act on the regulator at the transcriptional level and suppress the production of VcdRP. The sRNA element (VcdR) acts via a conserved stretch of four consecutive cytosines to base pair with the promoters of PTS-specific and non-specific sugars, resulting in the inhibition of carbon uptake. The peptide VcdP binds to and increases the activity of citrate synthase (*gltA*) enzyme, thereby modulating the metabolite concentration within the cycle. Thus, VcdR and VcdP together synchronize the carbon uptake (blue) and its subsequent utilization (pink), thereby balancing carbon metabolism.

VcdRP is transcriptionally repressed by the global transcription factor of carbon utilization, cAMP and CRP. The riboregulatory component (VcdR) is conserved at the 3' end of the dual regulator. By employing a conserved stretch of four cytosines, VcdR represses mRNAs that encode for transporters that import PTS sugars: PtsG, NagE and TreB, that imports glucose, N-acetyl glucosamine, and trehalose, respectively. Additionally, VcdR also downregulates the phosphor-carrier proteins PtsH and PtsI that are involved in the phospho-relay during glycolysis. LC-MS analysis of the tagged small protein VcdP identified citrate synthase as its interacting partner. VcdP exerts its regulatory role by interacting with and accelerating the activity of citrate synthase enzyme, opening the gateway into the TCA cycle. This way, both VcdR and VcdP act to block sugar uptake and modulate the flux through the TCA cycle, thereby striking a balance to maintain overall carbon metabolism (Figure 31).

Chapter 3

A novel RNA-binding protein in *V. cholerae*

Parts of the results presented in this chapter were performed in collaboration with the following people:

- Annika Sprenger (former member of the lab)
- Prof. Dr. Kai Papenfort (Friedrich-Schiller-University Jena, Germany)
- Dr. Patrick Pausch[#] and Prof. Dr. Gert Bange (SYNMIKRO, University of Marburg, Germany)
- Dr. Charlotte Michaux[&], Dr. Lars Barquist and Prof. Dr. Jörg Vogel (University of Würzburg, Germany)

[#]Present affiliation: Vilnius University, Lithuania

[&]Present affiliation: Imperial College, London, United Kingdom

3.1 A forward genetic screen to identify factors affecting QS transition

The diverse environments that *V. cholerae* inhabits necessitates that the organism rapidly perceives changes in its external environment and appropriately tailors its gene expression patterns. To achieve this, the bacteria employ QS to communicate and coordinate a suitable response. While this mechanism of census taking has been well-documented early on in several marine bacteria, more recent studies have identified additional QS systems in *V. cholerae* (Bassler, 2002; Winans & Bassler, 2002; Liu *et al*, 2006; Papenfort *et al*, 2015).

To identify further regulators of the *V. cholerae* QS pathway, a forward genetic screen in a *V. cholerae* mutant library (Cameron *et al*, 2008) was employed to score for an altered QS phenotypic transition. This large sequence-defined transposon insertion library comprises interruptions in 3,096 out of 3,885 annotated ORFs, with 92.8% insertion mutants being unique. To correlate the impact of these mutants on QS at a phenotypic level, a luciferase reporter plasmid encoding the *lux* operon from *V. harveyi* was moved into this mutant library (Figure 32, left panel). QS of *V. harveyi* is analogous to *V. cholerae*, however the former species additionally produce bioluminescence in response to AI signals (Henke & Bassler, 2004). LuxR of *V. harveyi* drives light production and HapR of *V. cholerae* can

replace LuxR to culminate in the the same regulatory outcome as LuxR (Ball *et al*, 2017). Hence, with this assay, production of light served as a readout for altered QS phenotype.

The wild-type strain exhibits a light curve (Figure 32B, black line) typical for QS-dependent bioluminescence (Nackerdien *et al*, 2008). The light emission is high at the onset of the experiment as the samples were inoculated from HCD pre-cultures. Within two hours, light production decreases by approximately three-fold and thereafter returns to maximum levels two hours later. As expected, a mutant of *hfq* resulted in a constantly 'bright' phenotype, regardless of cell-density (Lenz *et al*., 2004, yellow line). In the absence of *hfq*, the sRNAs Qrr1-4 are unstable and result in constant expression of HapR (section 1.7.1). Notably, a mutant of *hapR* leads to a perpetually 'dim' phenotype, regardless of cell-density (grey line). In addition to these known players in regulating QS, another gene was identified that also resulted in a phenotype similar to *hapR*-deficient cells. This previously uncharacterized gene is *vc0159*, and its increased bioluminescence levels at LCD links its contribution to a defective QS-pathway.

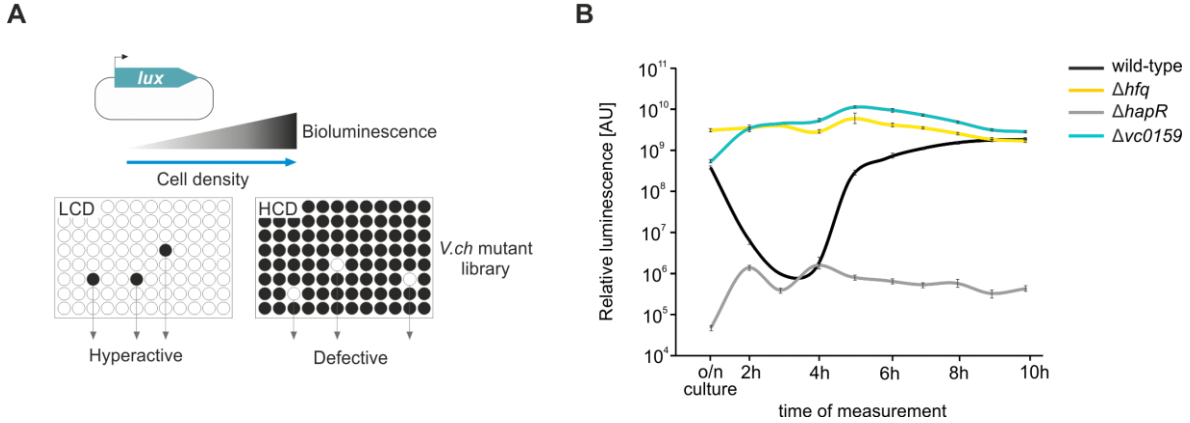

Figure 32. A forward genetic screen for QS transition identifies a new player. A) The experimental setup employed to score for altered QS phenotype. Production of light in response to a mutant from the transposon library upon expression of a luciferase plasmid encoding the *V. harveyi lux* operon was measured. The amount of bioluminescence is proportional to the cell density state. **B)** *V. cholerae* bioluminesence assays were performed as previously described (Miller et al, 2002). Briefly, the indicated strains carrying the *V. harveyi lux* operon were inoculated from overnight cultures and grown at 30°C with aeration. The OD_{600} of each culture was measured, and the cultures were diluted to equal cell density. Diluted cultures were grown, and light production was measured at the indicated time-points using a scintillation counter. Data are presented as mean ± SD, n = 3.

3.2 Vc0159 is an RNA-binding protein that localizes to the membrane

Closer inspection of the domain architecture of Vc0159 revealed that the 153 aa long protein carries two transmembrane domains at the N-terminus and a conserved RRM-like (RNA recognition motif-like) RNA-binding domain close to the C-terminus (Figure 33A). These RNA-binding motifs are highly prevalent across all kingdoms of life and are most abundant in eukaryotes (Bateman *et al*, 2002). Among prokaryotes, RRM domains are widespread in cyanobacteria (Maruyama *et al*, 1999). The RRM folds into an αβ sandwich structure with a symmetric $\beta_1\alpha_1\beta_2$ topology (Maris *et al*, 2005). Interestingly, the Vc0159 protein is also conserved among other marine bacteria such as *Photobacterium*, *Salnivibrio*, *Shewanella*, *Pseudoalteromonas* and *Aeromonas* species. In all cases, two patches of hydrophobic amino acid residues near the N-terminus of the protein have been maintained, predicting that Vc0159 could be a membrane-bound protein.

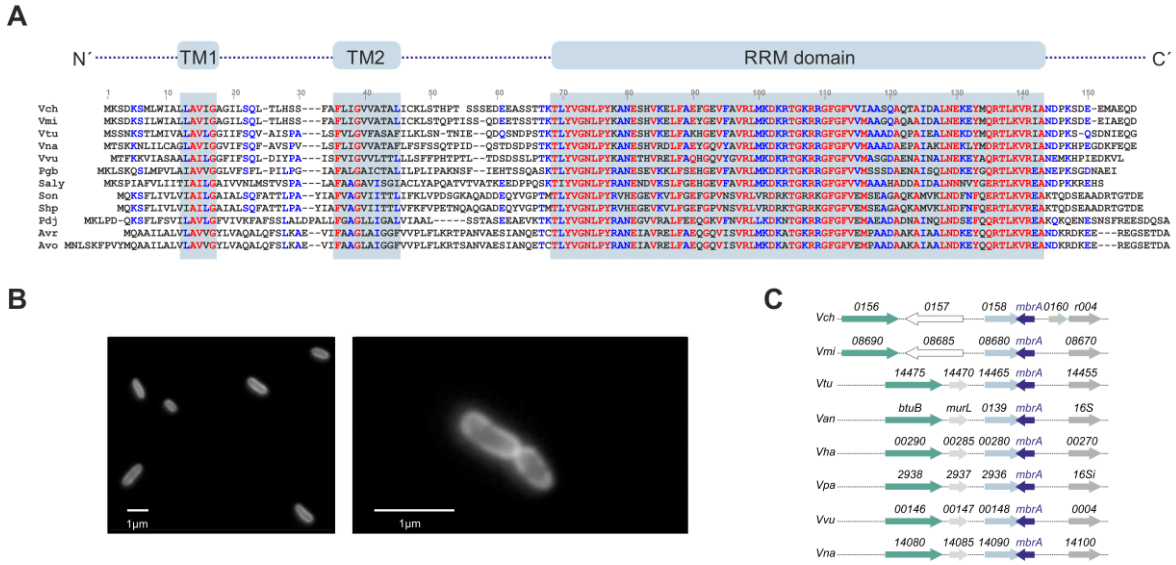

Figure 33. Domain architecture and gene synteny analysis of Vc0159. A) A protein BLAST analysis of Vc0159 revealed its conservation not just among Vibrio species, but also other marine bacteria. Notably, all the bacteria carry conserved patches for two transmembrane (TM1 and 2) domains at the N-terminus and an RNA recognition motif (RRM) at the C-terminus. **B)** Fluorescent microscopy analysis of plasmid-based sfGFP-tagged protein in *V. cholerae* showed localization at the membrane. Therefore, Vc0159 was renamed MbrA, for membrane-bound RNA-binding protein A. C) Gene synteny analysis of *mbrA* and its flanking genetic loci among different *Vibrio* species. Homologous genes are depicted using the same colors. Data information: For (A) and (C), the abbreviations used are as follows – Vch (*Vibrio cholerae*), Vmi (*Vibrio mimicus*), Vtu (*Vibrio tubiashii*), Vna (*Vibrio natriegens*), Vvu (*Vibrio vulnificus*), PgB (*Photobacterium gaetbulicola*), Saly (*Salinivibrio sp. YCSC6*), Son (*Shewanella oneidensis*), Shp (*Shewanella putrefaciens*), Pdj (*Pseudoalteromonas donghaensis*), Avr (*Aeromonas veronii* B565) and Avo (*Aeromonas veronii* TH0426).

To test the localization of the protein, a plasmid-based fluorescent reporter was cloned such that the full-length *vc0159* gene was fused to sfGFP at its C′ terminus and expressed in *V. cholerae*. Subsequent microscopy analysis confirmed membrane localization of the RNA-binding protein (Figure 33B). Hence, the gene was renamed *mbrA*, for membrane-bound RNA-binding protein A. The *mbrA* gene is encoded antisense on the main chromosome of *V. cholerae*, and this gene partially overlaps with the gene coding for glutamate racemase (*vc0158*). In close proximity to *mbrA* is also a gene encoding a subunit of 16S ribosomal RNA (Figure 33C).

3.3 Sub-cellular fractionation of MbrA

The physical demarcation of transcriptional and translational machineries in eukaryotes has contributed to a better overall understanding of spatio-temporal control of gene regulation (Palacios & Johnston, 2001). In contrast, due to a lack of typical membrane-bound organelles and a nuclear compartmentalization, prokaryotic cells were long assumed to lack complex subcellular localization of macromolecules. Moreover, spatial localization has not been considered to play a significant role in expression and post-transcriptional regulation of bacterial mRNAs (Keiler, 2011; Irastortza-Olaziregi & Amster-Choder, 2021). Bacterial RNAs typically localize diffused in the cytoplasm, in specific patches across the cell, at the poles or at the membrane (Figure 34A). Distinct localization patterns within the cell minimizes the energy in translocating each protein molecule individually to the subcellular region where the protein functions. Additionally, co-localization of mRNAs encoding different protein constituents of a single complex facilitates efficient complex formation. Also, targeting of mRNAs to specific sites may protect them from being exposed to ribonucleases and help in maintaining their proper levels in the cell (Keiler, 2011; Irastortza-Olaziregi & Amster-Choder, 2021).

To corroborate the preliminary findings of the membrane-bound nature of MbrA, a sub-fractionation experiment was carried out. Cellular sub-fractions were prepared based on a previously established protocol (Thein *et al*, 2010). Briefly, the inner and outer membranes were prepared by cell lysis and separated by selective detergent treatment followed by differential centrifugation. To control for the outer membrane fraction, a C-terminal 3xFLAG epitope was introduced in the chromosomal locus of *ompT* (Chakrabarti *et al*, 1996). Either an empty control plasmid (pCtrl) or a C-terminal 3xFLAG epitope-tagged MbrA plasmid (p*mbrA*::3xFLAG) was introduced in the above strain. The different fractions were examined on an SDS-blot probed with anti-FLAG antibody. RNAP served as a control for the cytoplasmic fraction. As expected, OmpT was enriched in the outer membrane fraction (Figure 34B, lanes 1 and 5), whereas MbrA was enriched in the inner membrane fraction (lane 7). As expected, RNAP localized exclusively in the cytoplasm (lanes 4 and 8)

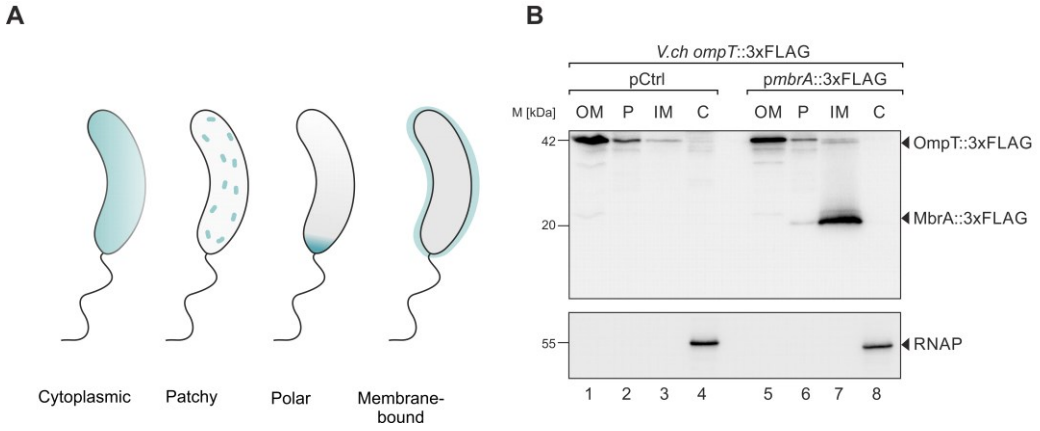

Figure 34. Sub-fractionation of MbrA confirms localization in the inner membrane. A) Typical RNA localization patterns observed in bacteria: cytoplasmic, in specific patches within the cell, at the poles or bound to the membrane. **B)** SDS-PAGE analysis of different protein fractions collected from the outer membrane (OM), periplasm (P), inner membrane (IM) and cytoplasm (C). These fractions were extracted from *V. cholerae* strains with a 3xFLAG tag at the C-terminus of *ompT* harboring either an empty control plasmid (pCtrl) or over expression of tagged-MbrA (p*mbrA*::3xFLAG). The solid triangle corresponds to the expected sizes of the indicated proteins.

3.4 The two transmembrane domains are critical for localization of MbrA

Transmembrane (TM) domains are typically composed of hydrophobic non-polar residues (Wayne Albers, 2012). While the exact sequence of both the TM domains of MbrA are not identical among the marine bacteria analyzed (Figure 33A), they all harbor conserved patches of non-polar residues. Eukaryotic TM domains have emerged as major determinants of intracellular localization and transport processes (Cosson *et al*, 2013). Moreover, the length of the domains also serve as signatures for their respective intracellular locations/organelles. (Singh & Mittal, 2016) have reported an evolutionary pressure in modulating length of TM domains of membrane proteins with increasing complexity of communication between sub-cellular compartments. To study the localization pattern of MbrA in the absence of its TM domains, over-expression plasmid-based fluorescent reporters were cloned. To this end, either of the TM domains of MbrA were deleted and the mutant protein was fused to sfGFP at its C-terminus and expressed in *V. cholerae*. Subsequent microscopy analysis showed that cells lacking these TM domains were de-localized from the membrane (Figure 35A). Moreover, the lack of membrane-localization was more severe with the second TM domain deleted.

Next, to examine the protein levels of MbrA, a 3xFLAG epitope was introduced in the C-terminal locus of *mbrA*. A second tagged variant with both the TM domains deleted was also generated. Both

these strains were grown in LB medium and protein samples were harvested over growth and examined on Western blots with anti-FLAG antibody. The levels of MbrA and the mutated version of the protein were low at LCD and increased in a cell-density dependent manner (Figure 35B). In conclusion, the TM domains are crucial for correct membrane localization, however, is dispensable for its translation.

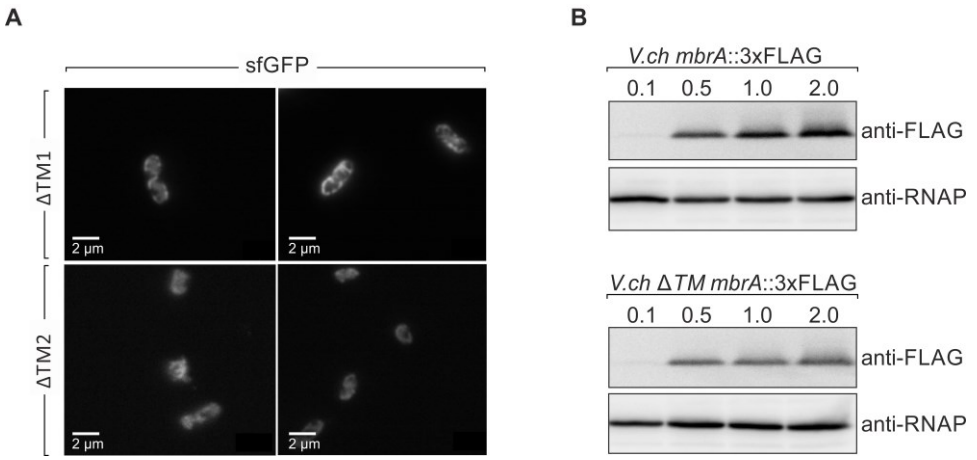

Figure 35. Impact of the deletion of the two TM domains of MbrA. A) Fluorescent microscopy analysis of plasmid-based sfGFP-tagged versions of MbrA lacking either of the TM domains in *V. cholerae* shows de-localization from the membrane **B)** Western blot analysis of chromosomally 3xFLAG-tagged versions of the full-length protein as well as a variant lacking both the TM domains. RNAP served as loading control.

Further, to examine the levels of protein production in the absence of *mbrA*, a chromosomal 3xFLAG epitope was introduced in the C-terminus of *hapR* and *luxO*, in *V. cholerae* wild-type and *mbrA*::Tn5 mutant strains. Western blot analysis was performed on these strains at different stages of growth in LB medium. In contrast to wild-type levels, there was a stark increase in HapR protein levels in cells lacking *mbrA* (Figure 36C, lanes 1-4 vs. 5-8). However, the production of LuxO had little-to-no impact upon deletion of *mbrA* when compared to its wild-type counterpart (Figure 36B, lanes 9-12 vs. 13-16). In summary, MbrA seems to modulate QS-behavior by interfering with HapR production.

3.5 CRP modulates MbrA production

In addition to the CRP binding site found in the promoter of *vcdRP* (section 2.3), the ChIP-seq data also indicated a CRP binding peak in the promoter of *mbrA* (Figure 37A; Manneh-Roussel *et al*, 2018). To confirm this, the promoter of *mbrA* was screened for a suitable CRP binding motif conserved among *Vibrio* species. Indeed, *mbrA* harbors a conserved two-box binding site for CRP (Figure 37B). The predicted motif for CRP biding in *E. coli* comprises a palindromic sequence separated by a non-conserved 6 nt linker sequence: 5′-TGTGANNNNNNTCACA-3′ (Shimada *et al*, 2011; Tsai *et al*, 2018). The motif in the promoter of *mbrA* has a similar CRP box sequence among all the tested *Vibrio* species.

Next, to examine the production of MbrA in response to deletion of either *crp* or *cyaA*, chromosomally tagged protein samples were analyzed on Western blots. The samples were collected over growth and probed using anti-FLAG antibody. In contrast to wild-type cells, MbrA production was abolished in both the mutant strains (Figure 37C, lanes 1-4 vs. 5-8 and 9-12). In conclusion, MbrA production is activated by CRP.

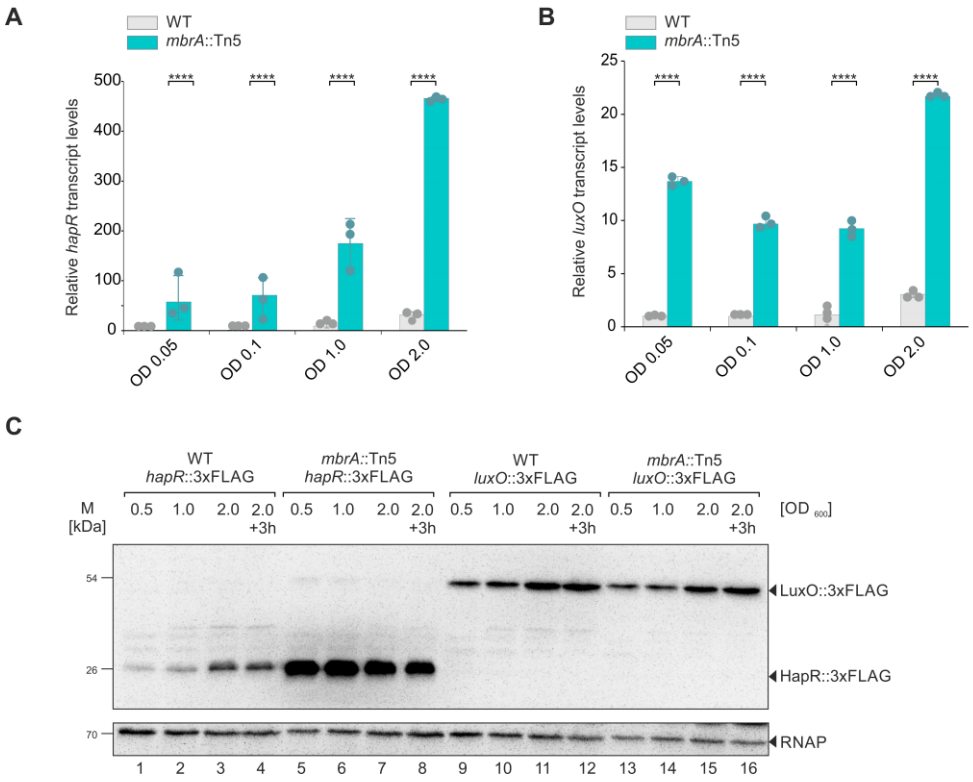

Figure 36. MbrA interferes with QS by regulating the *hapR* levels. A-B) The transcript levels of *hapR* and *luxO* mRNAs were measured by quantitative real-time PCR. Total RNA of wild-type and *mbrA*::Tn5 mutant cells was collected from various stages of growth and relative expression changes were plotted. Data are presented as mean ± SD, n = 3. The p-value is summarized as follows - **** for $p \leq 0.0001$. **C)** Western blot analysis of cells with chromosomal 3xFLAG epitope at the C-terminus of *hapR* and *luxO* in *V. cholerae* wild-type and *mbrA*::Tn5. The solid arrow represents the expected sizes of the indicated protein. RNAP served as loading control.

The Western blot analysis revealed that MbrA production is activated by CRP. To understand how CRP regulates *mbrA* transcription, *V. cholerae* wild-type cells expressing a plasmid-borne transcriptional reporter of the promoter of *mbrA* fused to *sfGFP* was tested in different media: LB or M9 medium supplemented with either glucose or glycerol. The availability of nutrients in the growth medium is known to affect the cAMP pool: the presence of glucose results in low adenylate cyclase and consequently low cAMP levels, whereas the in the presence of glycerol, the cAMP pool is shifted to favor CRP regulation (Notley-McRobb *et al*, 1997; Green *et al*, 2014).

In contrast to LB medium, the sfGFP levels were significantly elevated in the presence of glycerol and decreased in the presence of glucose (Figure 37D). Additionally, a second set of fluorescent reporters were generated such that the CRP-binding motif in the promoter of *mbrA* were either mutated or deleted as shown in Figure 38E. These plasmids were tested along with the native *mbrA* promoter fused to *sfGFP* in *V. cholerae* wild-type cells grown in M9 medium supplemented with glycerol. As expected, mutations in the CRP box decreased the MbrA-sfGFP levels (Figure 37F). Taken together, CRP activates *mbrA* transcription and deletion or mutation of the CRP-binding motif in the promoter of *mbrA* abrogates this effect.

3.6 Ligands binding to MbrA (CLIP-seq analysis)

Although RIP-seq (RNA immunoprecipitation followed by deep sequencing) offers a quick overview of major RNA regulons and has been employed to study the global profile of the three major RBPs of bacteria – Hfq, CsrA and ProQ (Chao *et al*, 2012; Huber *et al*, 2020; Timmermans & Van Melderen,

2010; Smirnov *et al*, 2017), the approach has its limitations – it is semi-quantitative and does not provide positional information of where the RBP binds (Saliba *et al*, 2017). To circumvent this, another method called cross-linking followed by immunoprecipitation and sequencing (CLIP-seq) was employed to identify ligands bound to MbrA. This method has previously been employed to *Salmonella* Hfq, CsrA and ProQ (Holmqvist *et al*, 2016, 2018). Unlike RIP-seq, this approach involves a UV cross-linking step that ensures the covalent cross-link of the ligand to the RBP prior to co-IP. This crosslinking enables trimming by ribonucleases to yield protein-protected RNA fragments, thereby pinpointing binding regions at single nucleotide resolution.

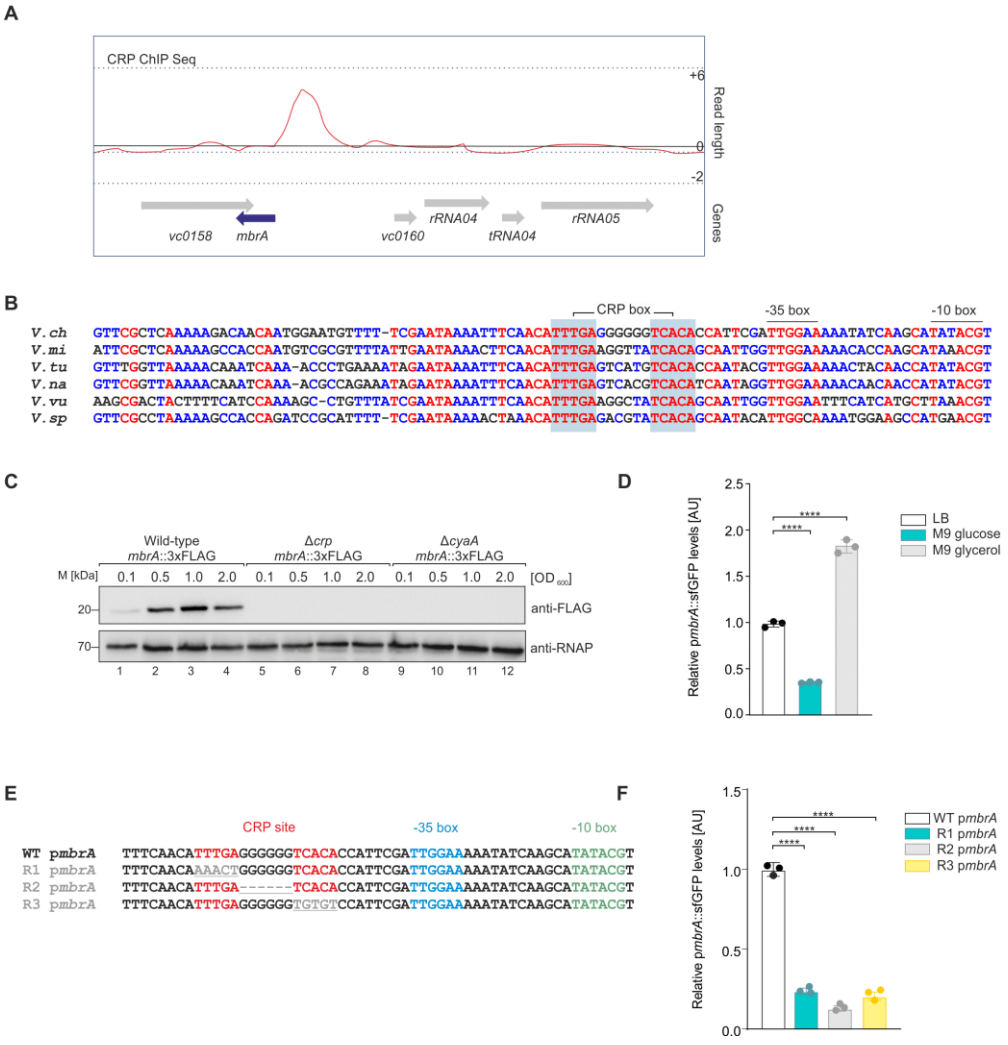

Figure 37. CRP regulates the production of MbrA A) The ChIP-Seq analysis by (Manneh-Roussel *et al*, 2018) determined a CRP peak upstream of the *mbrA* gene. **B)** The promoter of *mbrA* harbors a consensus motif for the two box CRP binding site. The binding site is shown in blue box, and the -10 and -35 elements are indicated by a black line. The abbreviations used are as follows – Vch (*Vibrio cholerae*), Vmi (*Vibrio mimicus*), Vtu (*Vibrio tubiashii*), Vna (*Vibrio natriegens*), Vvu (*Vibrio vulnificus*), and Vsp (*Vibrio splendidus*). **C)** Western blot analysis of wild-type as well as cells lacking *crp* and *cyaA*, all harboring a chromosomal a 3xFLAG epitope fused to the C-terminus of *mbrA*. Samples were tested at the indicated OD_{600} using anti-FLAG antibody. RNAP served as loading control. **D)** Relative fluorescence measurements of a transcriptional reporter of the promoter of *mbrA* fused to *sfGFP* expressed in *V. cholerae* wild-type cells. The samples were grown in the indicated media and expressed as relative fold changes with respect to measurements from LB medium set to 1. **E)** The promoter of *mbrA* harbors a consensus motif for the two box CRP binding site. The binding site is shown in red, and the promoter -10 and -35 elements are shown in green and blue, respectively. The underlined sequences shown in grey correspond to the mutations (R1 and R3) or deletion (R2) introduced in the reporter fusions tested in F. **F)** Relative fluorescence measurements of the different transcriptional reporters of the promoter of *mbrA* shown in (B) tested in *V. ch* wild-type cells. The samples were grown in M9 medium supplemented with glycerol and expressed as relative fold changes with respect to measurements from the native promoter sequence set to 1. Data information: data shown in (D-F) are presented as mean ± SD, n = 3. The p-value is summarized as follows - **** for $p \leq 0.0001$.

To comprehensively analyze the targets of MbrA *in vivo*, *V. cholerae* wild-type and chromosomally 3xFLAG-tagged *mbrA* and *hfq* strains were cultivated to early stationary phase of OD_{600} of 1.0. Hfq-FLAG served as a positive control for the setup because the global profile of Hfq-binding RNAs in *V. cholerae* was already known (Huber *et al*, 2020). Accordingly, one half of each culture was irradiated with UV light (indicated as +) while the other half was left untreated (indicated as -). Strong radioactive signals after labelling of co-IP and RBP-associated RNA were dependent on UV treatment, indicating that non-specifically bound transcripts were successfully depleted (Figure 38A). The signals in the cross-linked Hfq sample were far more enriched in comparison to MbrA hinting that the latter is not likely to be a global RBP and binds only to few ligands. Importantly, the protein recovery of the tagged MbrA and Hfq was unaffected by UV treatment (Figure 38B).

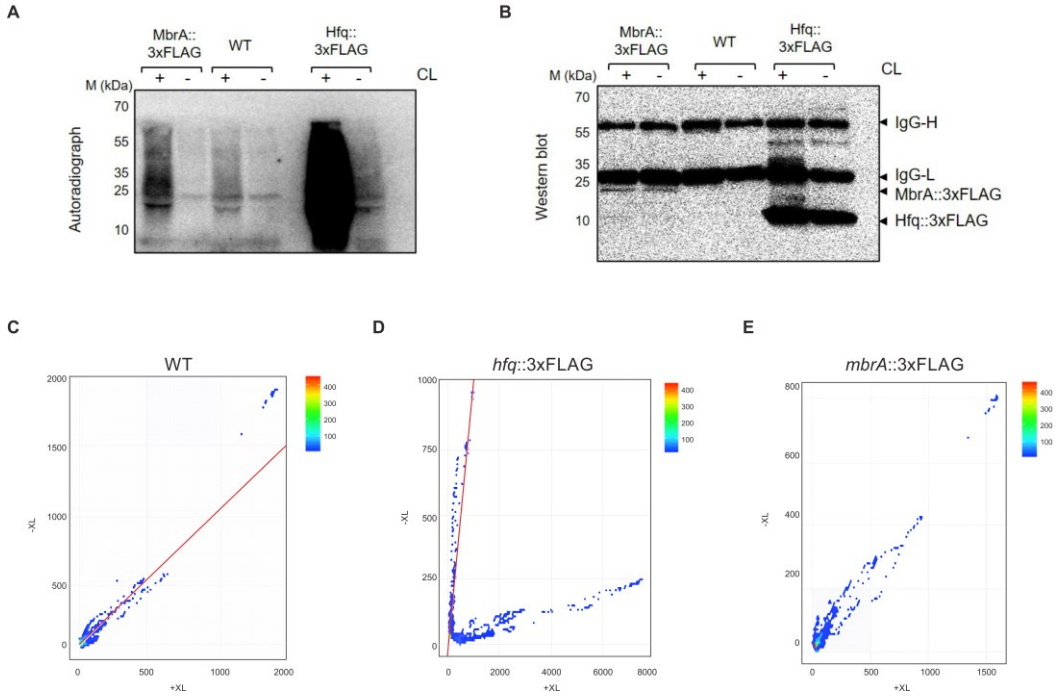

Figure 38. CLIP-seq analyses of MbrA. A-B) Detection of cross-linked (CL), immunoprecipitated, and radioactively labelled RNA-protein complexes after separation on denatureing SDS-PAGE and after transfer to nitrocellulose membranes. Radioactive signals were detected by phosphorimaging (A) and the protein samples were detected using anti-FLAG antibody on a Western blot **(B)** to confirm successful pull-down. + and – refer to the presence or absence of cross-linking with UV. **C-E)** Scatter plots showing read counts in non-crosslinked (-XL) versus crosslinked (+XL) libraries for wild-type (C), *hfq*::3xFLAG (D) and *mbrA*::3xFLAG (E).

Next, purified RNA from replicates was converted to cDNA and subjected to Illumina sequencing, including the non-crosslinked samples to account for background correction. Using a peak-calling algorithm specifically designed for CLIP-seq analyses (Holmqvist *et al*, 2016), specific RBP-binding sites can be determined from the transcriptome. Each peak was tested for significant enrichment in the cross-linked sample versus non-crosslinked samples using DESeq2 (Love *et al*, 2014) and visualized using scatter plots. As expected, the wild-type strain revealed no specific enrichment between the two tested conditions (Figure 38C). Interestingly, tagged Hfq showed two different scatter patterns between (-/+) crosslinked conditions, confirming that the Hfq pull-down was successful (Figure 38D). However, the density plot for MbrA looks like wild-type and does not recapitulate the enrichment pattern observed for Hfq (Figure 38E). There may be multiple reasons why this approach did not yield in any evidence for RNA-binding for MbrA: a) MbrA only binds to very few ligands, and thus needs to be sequenced deeper, b) the incubation time of the antibody as well as the duration of treatment with the nucleases needs to be optimized c) the nature of MbrA in

that it is a membrane protein may lead to poor enrichment because of solubility problems, hence contributing to deteriorating yield (in comparison to Hfq) at various steps of the CLIP-seq procedure.

3.7 Generation of unmarked *mbrA* mutant

The proximity of *mbrA* gene to the highly abundant 16S rRNA as well as its partial overlap with an essential gene that encodes glutamate racemase (*vc0158*; Figure 33C) contributed to the difficulties associated with generating a clean mutant in the *V. cholerae* genome. Standard RK2/RP4-based conjugal transfer using homology-mediated recombination (Skorupski & Taylor, 1996) did not work in the case of *mbrA*. Therefore, an alternate approach called TransFLP method based on the chitin-mediated induction of natural competence in *V. cholerae* (section 1.6; (Blokesch, 2012)) was employed.

Natural competence allows the organism to take up free DNA (such as PCR-generated fragments with homologous flanking regions at the site of deletion or mutation, including a selection marker). Once taken up, the DNA recombines in the chromosome and can be easily selected on suitable antibiotics to screen for the deletion or mutation. To delete *mbrA*, a PCR-fragment was generated from chromosomal homologous flanking regions of 1000 bp on either side of *mbrA* (without affecting the translation of the overlapping *vc0158* gene) along with a kanamycin cassette flanked by FLP sites. *V. cholerae* wild-type was induced for competence by growing on chitin flakes in artificial seawater and selected on LB-agar plates with kanamycin. This proved successful to knock out *mbrA* from the *V. cholerae* genome and the deletion was confirmed by qRT-PCR as well as by Northern blotting to measure *mbrA* transcript levels and expression, respectively (Figure 39A, B).

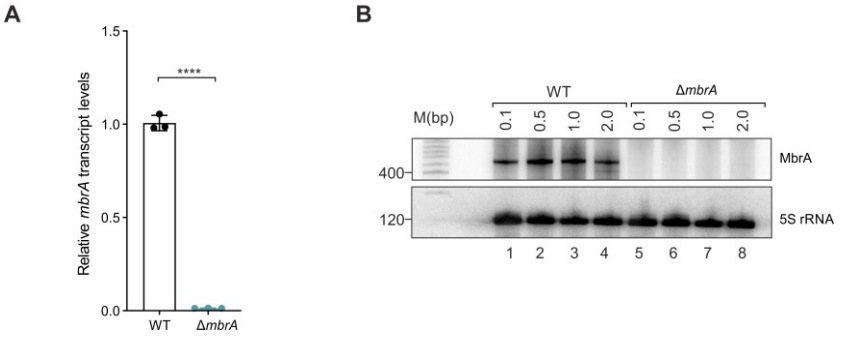

Figure 39. *mbrA* mutant was generated using TransFLP method. A) relative transcript levels of *mbrA* measured by qRT-PCR, when comparted to *V. cholerae* wild-type levels. RNA samples were collected from early stationary phase samples (OD_{600} of 1.0). *recA* served as the housekeeping gene. **B)** Northern blot analysis of *V. cholerae* wild-type and Δ*mbrA* strains grown in LB medium and monitored over growth. Probing with 5S rRNA served as loading control. Data information: data shown in (A) are presented as mean ± SD, n = 3. The p-value is summarized as follows - **** for $p \leq 0.0001$.

3.8 Transcriptome analyses of unmarked *mbrA* deletion

Since CLIP-seq did not give any indication about the role of MbrA in the cells, a genome-wide transcriptomics approach was used to compare differentially expressed genes among *V. cholerae* wild-type and the newly generated Δ*mbrA* cells harboring an empty control plasmid (pCtrl) or Δ*mbrA* complemented with a plasmid-based version of *mbrA*, driven from its native promoter (pMbrA). RNA samples were collected at early stationary phase (OD_{600} of 1.0) grown in LB medium and the expression of *mbrA* was monitored on a Northern blot (Figure 40A). Subsequently, the samples were depleted of rRNA, cDNA libraries were prepared and sequenced using Illumina sequencing.

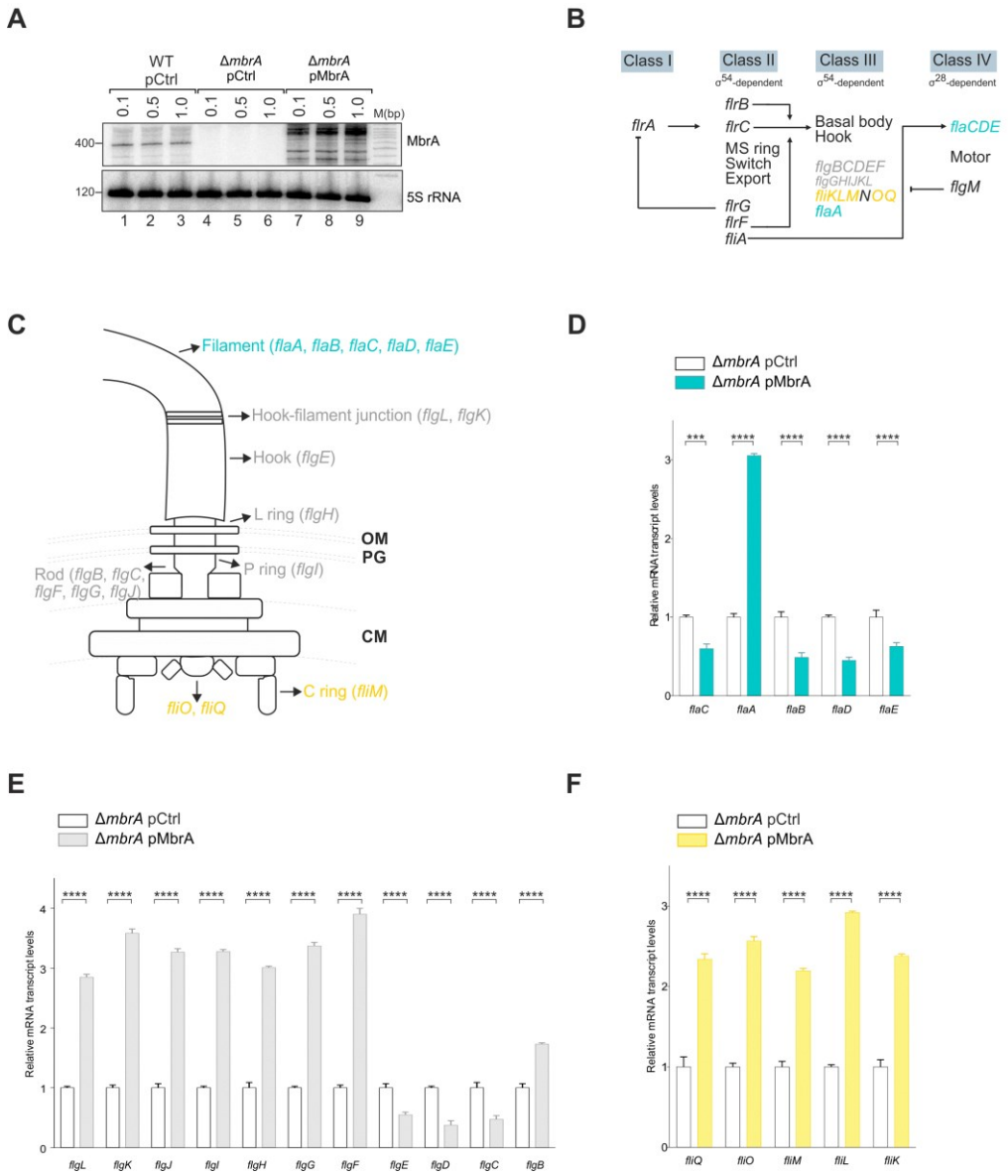

Figure 40. Transcriptome analysis of *mbrA* deletion and overexpression. A) Northern blot analysis to monitor the expression of *mbrA* over growth in *V. cholerae* wild-type and Δ*mbrA* harboring an empty vector control (pCtrl) or Δ*mbrA* harboring an overexpression plasmid of *mbrA* (pMbrA). Probing with 5S rRNA served as loading control. **B)** The flagellar transcriptional hierarchy of *V. cholerae* adapted from (Syed *et al*, 2009). The genes are color coded as *fla* genes (blue), *flg* genes (grey) and *fli* genes (yellow) corresponding to their functions as indicated in (C). **C)** A schematic of a typical polar flagellum and the various genes involved in itsy synthesis and assembly. **D-F)** Relative transcript levels of the indicated class of transcripts (color coded as in B and C) measured by qRT-PCR by comparing V. cholerae Δ*mbrA* harboring an empty vector control (pCtrl) or an overexpression plasmid of *mbrA* (pMbrA). *recA* served as the housekeeping gene for these measurements. Data information: data shown in (D-E) are presented as mean ± SD, n = 3. The p-value is summarized as follows - *** for $p \leq 0.001$ and **** for $p \leq 0.0001$.

Genes with an absolute fold change ≥2.0 and an FDR-adjusted p-value ≤0.05 were considered as differentially expressed. Interestingly, the only gene that was differentially regulated between *V. cholerae* wild-type and Δ*mbrA* cells harboring an empty control plasmid (pCtrl) was *mbrA* itself. However, upon overexpression of *mbrA*, 165 genes were differentially expressed in comparison to both wild-type as well as the mutant variants (Table 3). Upon careful inspection of these DEGs, over a third of the genes were involved in the synthesis and assembly of *V. cholerae* flagellar apparatus (Boin *et al*, 2004; Echazarreta & Klose, 2019). The flagellar transcriptional hierarchy is organized over four tiers (Syed *et al*, 2009). A single class I gene *flrA* activates the alternate sigma factor σ^{54}-dependent class II genes, that encodes the two-component system FlrBC. Phosphorylated FlrC activates class III genes and finally class IV genes are activated in a σ^{28}-dependent manner (Figure 40B, C). To validate

the RNA-seq findings, qRT-PCR analyses was carried to measure the transcript levels of these DEGs tested under the same conditions as the transcriptome experiment (Figure 40D-F). All the transcripts displayed comparable fold-changes to that from the transcriptome analysis. In summary, deletion of *mbrA* seems to have no impact on the cells, however, over-expression of *mbrA* interferes with the synthesis and assembly of the polar flagellum.

3.9 The 3xFLAG epitope interferes with MbrA function

The transcriptome experiment of MbrA overexpression revealed its role in regulation of flagellar genes. To test what impact MbrA has on motility, *V. cholerae* wild-type and Δ*mbrA* cells harboring an empty control (pCtrl) or Δ*mbrA* harboring a plasmid over-expressing *mbrA* (pMbrA) were spotted on a soft LB-agar plate and allowed to incubate at 30°C for 16h. As expected, deletion of *mbrA* had little-to-no impact on chemotaxis, and its overexpression led to increased motility (spot #2 increased by~2.3-fold, Figure 41A). To test whether addition of a tag had an impact on motility, *V. cholerae* wild-type and Δ*mbrA* cells harboring a plasmid-based 3xFLAG-tagged *mbrA* were also spotted. In contrast to what was expected, this did not lead to a significant increase in motility (spot #4 and #5 vs. #3, Figure 41A). A subsequent Northern blot analysis comparing Qrr4 sRNA levels at OD_{600} of 0.1 and OD_{600} of 1.0 in *V. cholerae* wild-type and chromosomally 3xFLAG-tagged MbrA revealed altered Qrr4 production. In wild-type cells, the expression of this sRNA is dependent on QS signals and is highest at LCD and decreases at HCD. In contrast, the levels of Qrr4 remained the same at both tested conditions, further confirming the loss of function of MbrA in the presence of the 3xFLAG tag. This may well explain why the RIP-seq and CLIP-seq analysis did not yield reliable results. The presence of the 3xFLAG tag possibly interferes with how MbrA integrates in the cell membrane.

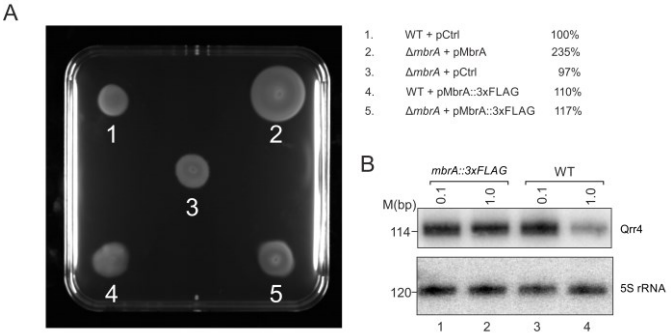

Figure 41. 3xFLAG tag interferes with MbrA function. A) Motilty assay performed on soft LB-agar plate. Spots from the indicated strains (right panel) were incubated at 30°C for 16h and diameters were quantified relative to WT pCtrl set to 100%. **B)** Northern blot analysis of *V.cholerae* that was chromosomally tagged with 3xFLAG epitope at the *mbrA* locus and wild-type strains tested for Qrr4 expression. Probing with 5S rRNA served as loading control.

3.10 CLIP-seq using Spot®-Tagged MbrA

To circumvent 3xFLAG-related problems, another tag was used to tag the chromosomal locus of *mbrA* at its C-terminus. The Spot®-Tag is a short peptide tag (12 amino acids long, 1.4kDa) and is inert, highly stable and robust (Virant *et al*, 2018). In addition to including the tag only at the C-terminus locus, the tag was also inserted at multiple positions within the protein. These positions were determined based on the crystal structure obtained for MbrA (detailed in section 3.11). Although the Spot® nanobody is specifically designed for IP and purification of proteins, detection on a Western blot using chemiluminescence has been challenging mainly because of multiple unspecific bands that also get detected. Therefore, the different plasmid-based Spot®-tagged variants of MbrA were instead analyzed on a soft LB-agar plate for altered motility. These tagged plasmids

were expressed in *V. cholerae* cells lacking *mbrA* and spotted for 8h at 30°C. Interestingly, four variants were significantly upregulated: Spot® tags that were added at positions S23, Q86, K99 and D141 of MbrA (Figure 42A).

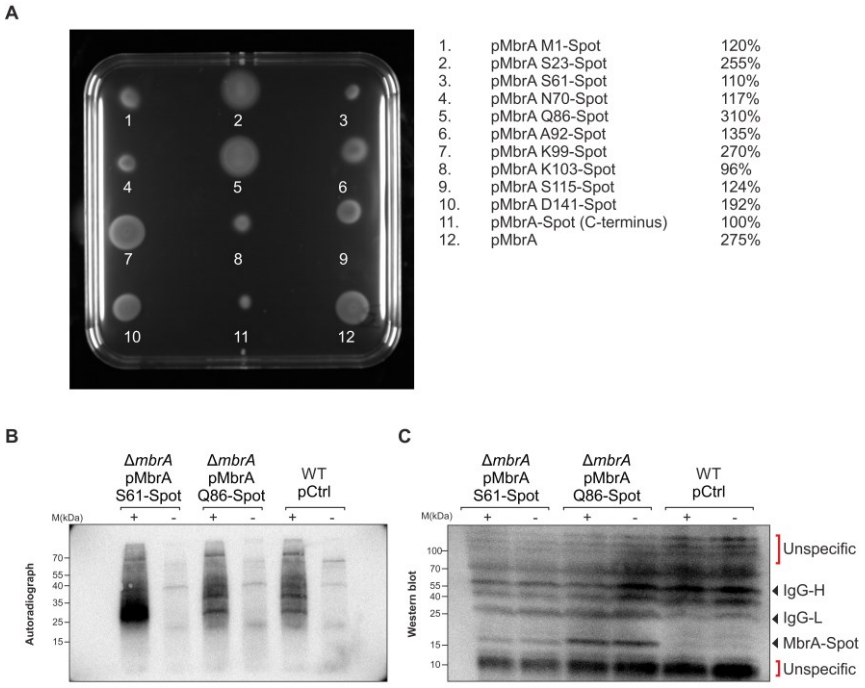

Figure 42. Spot-tagged MbrA varaints that were used for another CLIP-seq analysis. A) Motilty assay performed on soft LB-agar plate. Spots from the indicated plasmid-based Spot-tagged MbrA strains (right panel) were expressed in *V. cholerae* cells lacking *mbrA* and were incubated at 30°C for 8h and diameters were quantified relative to the C-terminal Spot-tagged variant set set to 100%. M1, S23, S61, N70, Q86, A92, K99, K103, S115 and D141 (relative to +1 of the ORF) refer to the amino acid position at which the tag was introduced. **B-C)** Two variants from this motility screen (S61 and Q86) were selected for another round of CLIP-seq analysis. Detection of cross-linked, immunoprecipitated, and radioactively labelled RNA-protein complexes after separation on denatureing SDS-PAGE and after transfer to nitrocellulose membranes. Radioactive signals were detected by phosphorimaging (B) and the protein samples were detected using anti-Spot nanobody on a Western blot **(C)** to confirm successful pull-down. + and – refer to the presence or absence of cross-linking with UV.

Two of these variants, namely S61 (unchanged motility) and Q86 (increased motility) were chosen for another round of CLIP-seq analysis. *V. cholerae* wild-type strain harboring an empty control plasmid (pCtrl) and the two plasmid-based tagged variants (S61 and Q86) were cultivated to early stationary phase of OD_{600} of 1.0. One half of each culture was irradiated with UV light (indicated as +) while the other half was left untreated (indicated as -). Strong radioactive signals after labelling of co-IP and RBP-associated RNA were dependent on UV treatment, indicating that non-specifically bound transcripts were successfully depleted (Figure 42B). As expected, the protein recovery of the tagged MbrA variants were unaffected by UV treatment (Figure 42C). However, unlike the previous 3xFLAG-based CLIP-seq setup (section 3.6), there were a lot more unspecific bands that were detected in the Western blot using anti-Spot® nanobody (in red, Figure 42C). Subsequent deep sequencing analysis revealed no differential regulation among any of the strains as well as between non-crosslinked and cross-linked samples. In conclusion, the tag probably interferes with the folding of MbrA, and hence consequently affects its regulatory function.

3.11 Crystal structure of MbrA

While affinity tags like polyhistidine (polyHis), maltose binding protein (MBP), Strepdavidin (Strep) and glutathione S-transferase (GST) are used commonly to facilitate protein

purification (Lichty *et al*, 2005), in order to evade tag-related problems to purify MbrA, the intein-based IMPACT® (intein mediated purification with an affinity chitin-binding tag) system was used (Banki & Wood, 2005; Sharma *et al*, 2006). It is a novel protein purification strategy that utilizes the inducible self-cleavage activity of protein splicing elements (termed inteins) to separate the target protein from the affinity tag. Each intein tag contains a chitin binding domain (CBD) for the affinity purification of the fusion protein on chitin resin. Induction of on-column cleavage, using thiol reagents such as dithiothreitol (DTT), releases the target protein from the intein tag (Figure 45A, schematic above the gels). *E. coli* ER 2566 was used the recombinant host and the expression vector was derived from the pTYB1 plasmid. The full-length MbrA protein was cloned into this plasmid, induced with isopropyl β-d-1-thiogalactopyranoside (IPTG), purified on a chitin column, and subjected to on-column cleavage. While the induction of the protein worked, no protein was eluted in the elution fractions (Figure 45A, upper panel). One caveat here could be the lack of solubilization of the membrane fraction. Therefore, another variant of the MbrA protein lacking both the transmembrane domains was cloned into the expression vector and tested again. This resulted in successful pull-down of the purified protein (Figure 43A, lower panel).

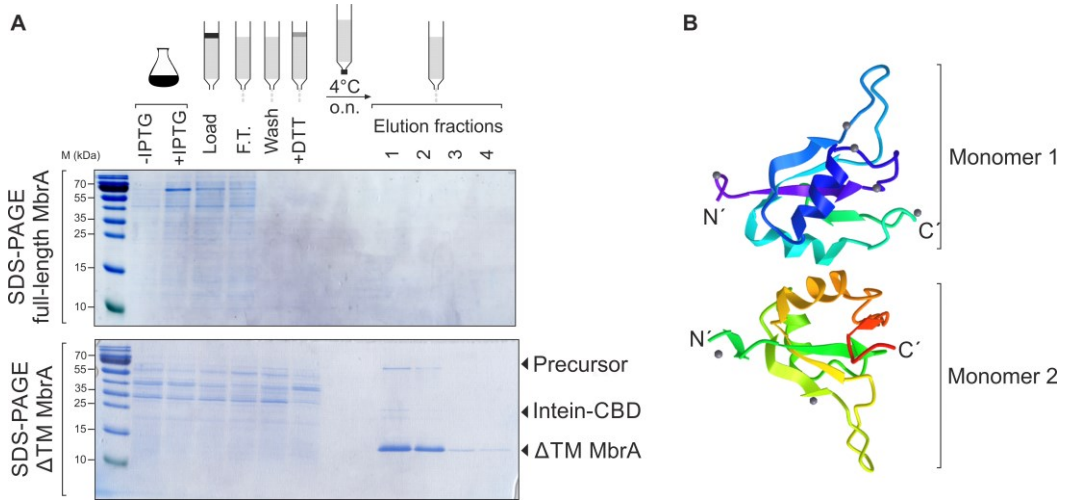

Figure 43. Purification and determination of the crystal structure of MbrA. A) Upper schematic: the various steps involved in the intiein-based IMPACT® purification system. Cells are induced with IPTG and crude cell extracts are clarified prior to loading onto a chitin column. The flow through (F.T.) samples determines the binding efficiency of the chitin resin. After several rounds of washing, cleavage is introduced by the addition of DTT. This allowed to incubate overnight at 4°C and finally the elution fractions are collected. The samples collected from the various stages of purification were analyzed on SDS-PAGE: from the purification of the full length MbrA (upper gel) and a version of MbrA devoid of its TM domains (lower gel) were stained using Coomasssie Brilliant blue. The full-length variant did not yield any protein in the elution fraction. However, the indicated arrows correspond to the expected sizes of ΔTM MbrA protein. **B)** Crystal structure of ΔTM MbrA protein determined using single drop vapor diffusion method. The structure was solved by molecular replacement using the human RBP HuR as a search model. MbrA exists as a dimer and in the absence of the TM domains, does not bind to any ligands. The color scheme used is VIBGYOR (N-terminus to C-terminus).

The version of the protein devoid of the TM domains was further used to determine its crystal structure. The crystals were generated by sitting drop vapor diffusion method (Kowalinski *et al*, 2007) and appeared after 7 months. The structure was solved by molecular replacement using PDB 4ED5 was a search model. The search model used is the ubiquitously expressed human RNA-binding protein (HuR), that contains three RRM domains. The two N-terminal tandem RRM domains can selectively bind AU-rich elements, while the third RRM domain contributes to interactions with the poly-A tail of target mRNA and other ligands (Wang *et al*, 2013). Interestingly, MbrA lacking its TM domains exists as a dimer in its unbound state (Figure 43B). Each monomer comprises two alpha-helices and four beta-sheets. However, when complexed with a mRNA, no structure could be obtained. This suggests that MbrA likely behaves as an RBP only in the presence of its TM domains.

3.12 Concluding summary

In this study, a novel putative RNA-binding protein of *V. cholerae* was discovered while screening for additional factors affecting QS transition. This gene was *vc0159* and is encoded antisense on the main chromosome of the *V. cholerae* genome and partially overlaps with the gene encoding glutamate racemase. An alignment of this RBP among different marine bacteria revealed conserved patches of two TM domains at the N-terminus and an RRM-like RNA-binding domain at the C-terminus. Microscopy and sub-fractionation analysis of the protein confirmed that it localizes to the inner membrane. Thus, it was renamed MbrA, short for membrane-bound RNA-binding protein A.

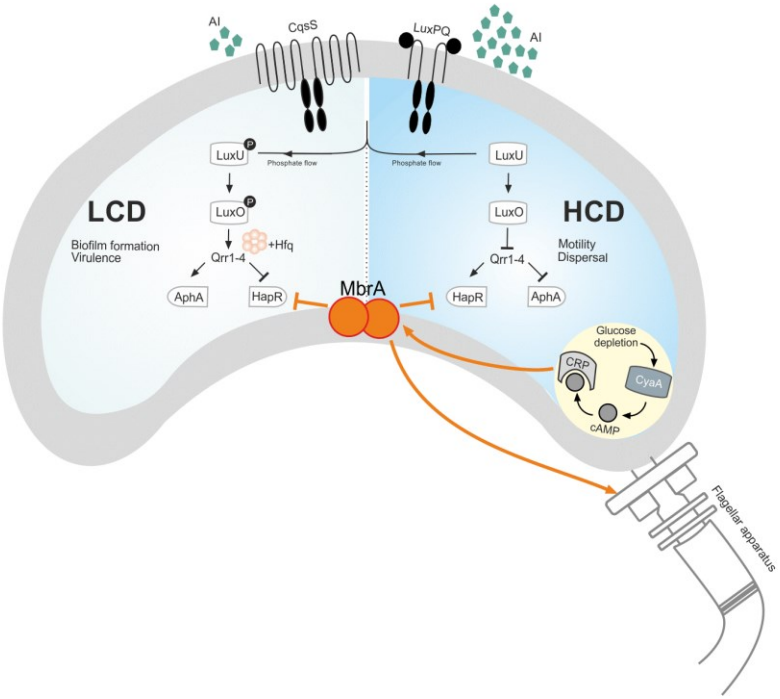

Figure 44. A model summarizing the role of MbrA in *V. cholerae*. MbrA was discovered when screening for additional factors influencing QS transition. At low cell density (LCD, in pale grey), low levels of autoinducer (AI) molecules are secreted. The two-component systems comprising CqsS and LuxPQ function as kinases to transfer phosphor-groups to LuxU and LuxO. This results in the activation of the sRNAs Qrr1-4. Along with Hfq, these sRNAs activate AphA and repress HapR. Biofilm formation and virulence gene production are hallmarks of LCD phenotype. In contrast, at high cell density (HCD, in pale blue), CqsS and LuxPQ act as phosphatases, resulting in reduced levels of Qrr1-4. This causes the dispersal and activates the motility genes in the organism. MbrA exists as a homo-dimer (in orange). It localizes at the inner membrane (in orange) and interferes with HapR production. The promoter of *mbrA* is activated by cAMP-CRP, and the effect is most pronounced in glucose depleted medium (yellow inset). Over-expression of MbrA positively regulates the synthesis and assembly the polar flagellum.

In the absence of *mbrA*, *hapR* transcript levels are significantly upregulated, suggesting that MbrA might play a role in the QS transition by interfering with HapR production. Analogous to *vcdRP*, the promoter of *mbrA* harbors a consensus motif for CRP-like promoters. However, in contrast to *vcdRP*, *mbrA* is activated by the global regulator, CRP. In the absence of CRP or adenylate cyclase (that activates cAMP), the transcript levels of *mbrA* are reduced. In line with this, the levels of *mbrA* are upregulated in a glucose depleted medium. CLIP-seq method was employed to determine the ligands that bind to MbrA. However, this failed to offer any specifically enriched genes in the cross-linked samples. Subsequently, an unmarked mutant of *mbrA* was generated using the Trans-FLP method. Transcriptome analysis on *V. cholerae* wild-type and the mutant of *mbrA* harboring an empty control plasmid were compared with Δ*mbrA* complemented with a plasmid-based version of *mbrA*, driven from its native promoter. Interestingly, the only gene differentially regulated between the

wild-type and the mutant was *mbrA* itself. In contrast, upon over-expression of *mbrA*, the genes involved in the synthesis and assembly of the polar flagellar apparatus were differentially regulated. A motility screen was then performed on soft LB-agar plates and overexpression of *mbrA* contributes to increased motility. But the presence of a 3xFLAG tag in the same plasmid abrogates this effect. Therefore, a different short, inert tag called Spot® was used for another CLIP-seq analysis. Again, there were no significantly enriched peaks, suggesting that presence of a tag probably interferes with the folding of MbrA, and hence consequently affects its regulatory function. The crystal structure of MbrA lacking both its TM domains was resolved. In its unbound form, the protein exists as a homo-dimer (shown in orange, Figure 44) comprising two alpha-helices and four beta-sheets. Conversely, when complexed with a mRNA, no structure could be obtained. This suggests that the presence of TM domains is crucial for its RNA-binding ability. However, the RNA-binding ability of MbrA could not be conclusively determined.

Chapter 4

Discussion

Bacteria thrive in diverse dynamic niches, often coupled with alternating feast and fast cycles depending on nutrient availability. Accordingly, they have adapted to alter their metabolic capabilities rapidly and precisely based on their surrounding niche. Such adjustments entail complex regulatory networks that regulate expression of genes involved in nutrient uptake and metabolism. Several protein regulators of metabolism have been characterized in-depth, such as CadC of *E. coli* that senses low pH and converts lysine to cadaverine via decarboxylation reaction, which ultimately restores the intracellular pH (Küper & Jung, 2005), or PhoR-PhoB of *E. coli* that controls the phosphate regulon (Wanner & Wilmes-Riesenberg, 1992) or NepR of Alphaproteobacteria that regulates general stress response genes (Fiebig *et al*, 2015) or GmaR of *L. monocytogenes* that regulates motility at low temperatures (Kamp & Higgins, 2011), to name a few. However, over the last two decades, posttranscriptional mechanisms involving sRNAs have emerged as an additional layer of control in these networks. Extensive cross-talk of sRNAs with transcriptional regulators ensures a fine-tuned and coordinated metabolic output (Bobrovskyy *et al*, 2015). The coupled degradation of many sRNA-mRNA pairs permits sRNAs to achieve responses that are different than those of transcription factors (Shimoni *et al*, 2007). In addition, the ability to regulate a target at two levels, with both a transcription regulator and base pairing sRNA, coordinates regulatory processes especially when cells must quickly and robustly respond to a sudden change in environmental conditions (Beisel & Storz, 2010). Simple mathematical models to probe how regulatory performance is affected by sRNAs and transcription factors have revealed that network motifs of regulatory circuits in which sRNAs replace transcription factors as nodes confer different, often advantageous, kinetic properties (Mehta *et al*, 2008; Mitarai *et al*, 2009). sRNAs are better than transcription factors at mediating regulation because a large pool of sRNAs shortens the effective mRNA lifetime and buffers against target mRNA fluctuations (Mehta *et al*, 2008).

4.1 sRNAs at the crossroads of metabolism and virulence control

The crosstalk between metabolism and virulence is often mediated by sRNAs that are transcribed in response to metabolism- or infection-related cues. For example, the Hfq-dependent *Salmonella* SgrS not only represses *ptsG* to regulate glucose uptake, but also represses the SopD protein, which is an acquired virulence factor (Papenfort *et al*, 2012). Likewise, Spot 42 of *V. parahaemolyticus* not only modulates central metabolism by regulating genes of the TCA cycle, but also represses the expression of *vp1682* mRNA, which functions as a chaperone for one of the type 3 secretion systems (T3SS), that triggers the virulence cascade of the bacteria (Tanabe *et al*, 2015).

In vivo transcriptome analysis of *Yersinia pseudotuberculosis* revealed four sRNAs (RyhB1, RyhB2, SgrS, GlmZ) that were significantly upregulated during infection of lymphatic tissues (Nuss *et al*, 2017). The *V. cholerae* sRNA TarA is activated by the virulence transcription factor ToxT, and in-turn controls the levels of *ptsG* and thus, regulating uptake of glucose (Richard *et al*, 2010). In addition to the canonical regulation of the cell envelope by GlmY and GlmZ, these sRNAs of Enterohemorrhagic *E. coli* (EHEC) also regulate the locus of enterocyte effacement (LEE) operons that are key for the organism's pathogenesis and virulence repertoire (Gruber & Sperandio, 2015). IsrE sRNA of *Salmonella* is encoded on an island and is dispensable for murine virulence (Hébrard *et al*, 2012), however acts redundantly with RyhB to downregulate *sodB* mRNA (Vogel, 2009). InvR is another example of horizontally acquired sRNA that mediates repression of the core genome-encoded *ompD* mRNA in *Salmonella* (Pfeiffer *et al*, 2007). The binding of S-adenosylmethionine to its cognate riboswitch leads to the accumulation of the sRNAs SreA and SreB, which repress the expression of the virulence regulator PrfA in *L. monocytogenes* (Loh *et al*, 2009). In line with this observation, VcdRP of *V. cholerae* is yet another regulator that not only modulates carbon flux through the glycolytic and TCA cycle, but also represses the virulence determinant of the pathogen, CT. (Waldor & Mekalanos, 1996); Figure 9 and 15D).

The underscored importance of sRNAs at the crossroads of virulence and metabolism poses an interesting question about the evolutionary hierarchy of what came first: if the sRNA may have been initially involved exclusively in one process – virulence or metabolism, and only then subsequently acquired the other novel function (Updegrove *et al*, 2015)? An added constraint in the context of evolution is the functional redundancy of sRNAs. For instance, the QS regulating sRNAs of *V. cholerae*, Qrr1-4 all need to be deleted to abolish repression of *hapR* (Lenz *et al*, 2005). In contrast, the Qrr sRNAs of *V. harveyi* act additively (Tu & Bassler, 2007). The redundancy is not just limited to 'sibling' sRNAs all regulating the same target, but also multiple unrelated sRNAs controlling a common target. For example, the outer membrane porin OmpD of *S. enterica* is repressed by four different sRNAs – MicC, SdsR, RyhB and InvR (Fröhlich *et al*, 2012). The need for nuanced modulation of key factors in response to wide regulatory cues may well be the driving force for the evolution of sRNAs.

The spatio-temporal control of *V. cholerae* pathogenesis genes is another prime example of the need for precise gene regulation in response to environmental cues. In the planktonic state, the virulence cascade remains repressed, while the genes for motility and chemotaxis are transcribed. However, upon entry into a human host, the bacteria encounter different signals, hence inversing their transcriptome profile (Klose, 2001; Childers & Klose, 2007). The direct transcriptional activator of the virulence determinants TCP and CT is ToxT (Yu & DiRita, 2002). This activation is dependent on the presence of binding sequences called toxboxes upstream of the -10 and -35 promoter elements. The

sRNAs TarA and TarB were initially discovered when screening for these toxboxes (Bradley *et al*, 2011). The regulation of *ptsG* by TarA may be linked to the accumulation of glycogen granules in the stool samples of cholera patients. By reducing *ptsG* levels, TarA re-directs the glucose away from the TCA cycle, and allows glycogenesis (Schild *et al*, 2007; Kamp *et al*, 2013). The regulation of *tcpF* by TarB on the other hand appears to have a positive effect on colonization (Bardill & Hammer, 2012). Expression of TarB is highest under micro-aerobic conditions, therefore could repress *tcpF* expression prior to penetration of the mucosal barrier of the small intestine. VqmR is another sRNA that controls virulence by directly base-pairing with *aphA* (Herzog *et al*, 2019). Interestingly, the screen for toxboxes by (Bradley *et al*, 2011) also identified VcdRP as one of the 16 other potential sRNAs transcribed from intergenic regions with *cis* ToxT binding sites.

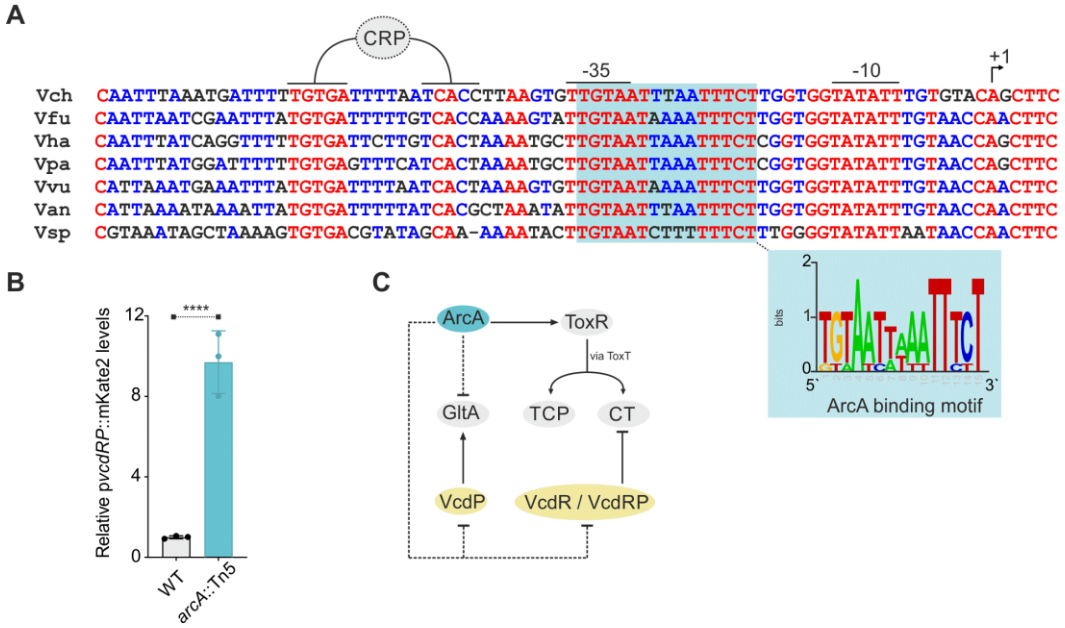

Figure 45. *vcdRP* harbors a consensus motif for ArcA in its promoter. A) The two-box CRP binding motif (grey), -35 box, -10 box, the ArcA bindng motif (blue box) and transcriptional start site (TSS; arrow) are indicated. Lower inset (blue): the consensus motif of ArcA motif. Abbreviations used: *V. cholerae* (Vch), *V. furnissi* (Vfu), *V. harveyi* (Vha), *V. parahaemolyticus* (Vpa), *V. vulnificus* (Vvu), *V. anguillarum* (Van) and *V. splendidus* (Vsp). B) Relative fluorescence measurements of mKate2-based transcriptional reporter fusion of *vcdRP* in *V. cholerae* strain harboring a transposon insertion in *arcA* and its corresponding wild-type control. Data are presented as mean ± SD, *n* = 3. Statistical significance was determined using one-way ANOVA and post hoc Tukey's multiple comparisons test. The p-values are summarized as follows: **** for p ≤ 0.0001. **C)** Proposed model for the ArcA transcriptional control on VcdRP. ArcA activates the ToxR regulon and in-turn positively regulates TCP and CT via ToxT. VcdRP, on the other hand repressed CT production and VcdP activates GltA activity. *E. coli* ArcA is known to repress *gltA*. Therefore, ArcA may serve as a transcriptional repressor of *vcdRP*. The solid lines depict validated regulation, whereas the dashed line corresponds to the speculated regulation that needs to be validated.

4.2 Transcriptional control of VcdRP

Anaerobiosis has been linked to the production of virulence factors in several bacterial pathogens. For instance, *Salmonella typhi* mutants defective in anaerobic respiration are compromised in their ability to replicate within epithelial cells (Contreras *et al*, 1997). Global regulators like fumarate and nitrate reductase (FNR) and the anoxic redox control (ArcAB) two component systems coordinate the aerobic / anaerobic interface (Jordan *et al*, 1997; Loui *et al*, 2009). Notably, ArcA also has a positive regulatory effect on the ToxR regulon and mutants of *arcA* exhibit impaired CT and TCP production, with a significant decrease in *toxT* expression (Sengupta *et al*, 2003). Interestingly, closer inspection of the promoter of *vcdRP* also revealed a conserved ArcA binding motif (Figure 45A). To validate this potential binding, a plasmid-borne transcriptional reporter of the *vcdRP* promoter fused to the *mKate2* fluorescent protein gene was used to test the levels of mKate2 in a strain with a transposon insertion in *arcA* and its corresponding wild-type control. The production

of mKate2 was significantly upregulated (~10-fold) in the absence of *arcA*. This suggests that ArcA acts as a repressor of VcdRP. In *E. coli*, ArcA is known to repress citrate synthase (*gltA*) expression (Park *et al*, 1994; Perrenoud & Sauer, 2005). It is therefore tempting to speculate that the activation of GltA by VcdP and repression of CT by VcdR/VcdRP is driven by the transcriptional repression of *vcdRP* by ArcA (Figure 45B).

The transcriptional control of VcdRP by cAMP-CRP provides an interesting new link between virulence and central carbon metabolism in *V. cholerae*. cAMP-CRP forms a bi-stable switch along with ToxT to control production of TCP and CT during infection (Nielsen *et al*, 2010). The global regulator directly controls a plethora of other additional genes that are key for virulence of *V. cholerae* in multiple modes of infection (Skorupski & Taylor, 1997; Notley-McRobb *et al*, 1997; Manneh-Roussel *et al*, 2018). The repression of CT by VcdRP may be attributed to the repression of *ptsH* and *ptsI* by its riboregulatory element (Table 1 and Figure 17). It has previously been shown that cells lacking PtsHI express reduced TcpA levels and consequently had diminished CT production as well as were compromised in their ability to colonize mouse infant intestine (Wang *et al*, 2015). A subsequent study showed that mutations in the PTS-transporters PtsG, NagE and TreB showed no significant changes in virulence gene expression, suggesting that utilization of PTS-dependent carbohydrates is dispensable for infection (Hayes *et al*, 2017). Moreover, biofilm-associated cells with a functional PTS exhibited reduced growth and this effect disappeared when the *ptsI* gene was disrupted (Houot & Watnick, 2008). The exact molecular mechanism of PtsHI-mediated repression of CT still remains unclear, however, (Wang *et al*, 2015) speculated that mutations in these phospho-carrier proteins may increase cAMP levels, that would then activate CRP. Consistent with this hypothesis, addition of external cAMP to cells lacking *cyaA* indeed resulted in significantly elevated *ctxA* and *ctxB* transcript levels (in comparison with *V. cholerae* wild-type) when tested under AKI conditions (Figure 46)

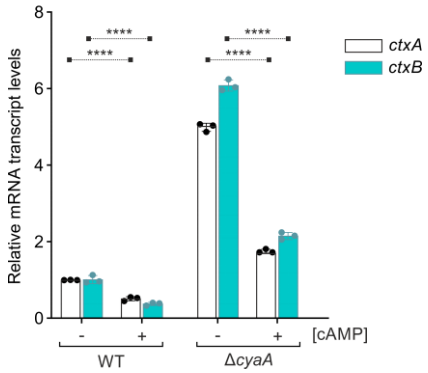

Figure 46. Increased cAMP leads to reduction of *ctxAB* transcript levels. *V. cholerae* wild-type and Δ*cyaA* strains were grown without (-) or with (+) cAMP (f.c. 5mM) in AKI medium to stimulate production of CT. Growth under AKI conditions involve biphasic cultures. In the first phase, the cultures were grown in a still tube for 4h at 37°C. Subsequently, in the second phase, the cultures were poured into a flask to continue growing with shaking. RNA samples equivalent to OD600 of 2.0 was harvested after 16h of continuous shaking followed by qRT-analyses of the *ctxA* and *ctxB* transcripts. The fold changes of transcript levels were calculated relative to wild-type (-) cAMP sample set to 1. *recA* served as the reference house-keeping gene for these measurements. Data are presented as mean ± SD, $n = 3$. Statistical significance was determined using one-way ANOVA and post hoc Tukey's multiple comparisons test. The p-values are summarized as follows: **** for $p \leq 0.0001$.

4.3 PTS-mediated carbohydrate transport

The PTS catalyzes the uptake and concomitant phosphorylation of a variety of carbon sources. Upon uptake, the phosphoryl groups are derived from PEP and transferred sequentially along the general cytoplasmic proteins (including PtsH and PtsI) and membrane-bound enzyme

IICB^{PTS} to glucose (Figure 31). The presence of PTS sugars causes dephosphorylation of the enzyme IIA^{PTS}, resulting in PTS-mediated inducer exclusion and subsequent decrease in adenylate cyclase levels (Brückner & Titgemeyer, 2002). A long-standing hypothesis was that only carbon sources that are transported by the PTS positively regulated the dephosphorylated state of IIA^{PTS} and support for this model came from the observation that αMG, a non-metabolizable glucose analogue, caused dephosphorylation of enzyme IIA^{PTS}, whereas a non-PTS carbohydrate such as galactose caused no dephosphorylation (Mitchell *et al*, 1982; Vadeboncoeur & Pelletier, 1997). However, (Hogema *et al*, 1998) showed that the non-PTS carbon source glucose 6-phosphate also induced strong dephosphorylation. Interestingly, the driving force behind this dephosphorylation was linked to a significant decrease in the PEP to pyruvate ratio. Consistent with this, the metabolome data upon overexpression of VcdRP indicated increased glucose-6-phosphate levels whereas its deletion dramatically increased pyruvate levels (Figure 29). Also, the RNA-seq dataset indicated upregulated levels of the genes encoding PEP carboxykinase (*vc2738*, ~2-fold) and PEP synthase (*vc0987*, 4.5-fold) upon overexpression of VcdRP (Table 1). PEP carboxykinase catalyzes the reversible interconversion of oxaloacetate and PEP, whereas PEP synthase mediates the conversion of pyruvate to PEP (Koendjbiharie *et al*, 2021). In the presence of PEP, the PTS transporters of *Enterobacteriaceae* are mainly found in the phosphorylated state (Deutscher *et al*, 2006; Yeh *et al*, 2009). Moreover, phosphorlylated IIA^{PTS} binds to and increases the activity of the enzyme glycerol kinase (Rohwer *et al*, 1998; Siebold *et al*, 2001). Accordingly, over-expression of VcdRP resulted in 4-fold increase of glycerol kinase (*glpK*, Table 1). Therefore, VcdRP seems to balance sugar uptake through glycolysis by modulating the PEP to pyruvate ratio.

To cope with nutrient availability, bacteria often have to sense and adapt to environmental cues and accordingly alter their physiological traits to exploit their ecological niche (Scheuerl *et al*, 2020). The metabolism of carbon and nitrogen compounds are fundamental to all forms of life and recently, the metabolite α-ketoglutarate (AKG) has emerged as a key regulator at the hub of nitrogen assimilation as well as central carbon metabolism in bacteria (Huergo & Dixon, 2015). AKG is a key metabolite of the TCA cycle and also serves as a carbon skeleton for reactions within nitrogen metabolism, including assimilation of ammonia (Commichau *et al*, 2006). Several studies in bacteria have demonstrated the fluctuating levels of AKG in response to changing carbon sources. Both HPLC and LC-MS studies have shown that the levels of AKG dropped significantly in carbon-starved cells of *E. coli* (Brauer *et al*, 2006; Yan *et al*, 2011). Quantitative monitoring of AKG levels in real-time under different growth conditions using a FRET-based biosensor also confirmed accumulation of AKG upon addition of glucose *in vivo* (Zhang *et al*, 2013a). *In vivo* accumulation of AKG has been shown to also reduce glucose uptake by directly inhibiting *ptsI* (Doucette *et al*, 2011). Interestingly, molecular docking simulations to derive the structural complex of *ptsI*-AKG also revealed the competition with PEP for the same binding site on the enzyme (Venditti *et al*, 2013). Modulation of the PEP to AKG ratio therefore serves as an elegant link between central metabolism and nitrogen metabolism

Pulse expression of VcdRP regulated the expression of significantly more genes (103) than the sum of genes affected by either VcdR (49) or VcdP (8; Figure 16A and Table 1). In addition, the functional inter-dependency of VcdR with VcdP is also reflected in the conservation of both the riboregulatory element as well as the small peptide among all tested species (Figure 10). Therefore, the physiological role of VcdP might only become fully evident in combination with VcdR and *vice-versa*. This is especially underscored within the TCA cycle, where although VcdP alone activates the CS enzyme (Figure 24C and 26B), the overexpression of the dual regulator VcdRP also increases the levels of AKG dehydrogenase (*sucA*) and glutamate dehydrogenase (*vc1492*) by ~2-3 fold (Table 1).

sucA converts AKG to succinyl coA, whereas glutamate dehydrogenase reversibly interconverts AKG to glutamate (Yang *et al*, 2014). While increased flux through a pathway requires increased flux through each individual enzyme, the glycolytic and TCA reactions always operate close to equilibrium (Miller & Smith-Magowan, 1990). Therefore, a combination of small changes in allosteric control (likely mediated by VcdP on NADH binding with CS; discussed below) and substrate/product occupancy of the active site (competition of PEP and AKG on *ptsI*) may collectively produce a substantial flux change to ultimately re-route central metabolism. The repression of *ptsI* itself by VcdR further bolsters the synergistic control of the metabolic flux by functionally inter-dependent elements of a dual-function regulator in *V. cholerae*.

4.4 The regulatory role of PTS^{Ntr}

Proteobacteria are known to harbor PTS transporters not only specific for carbohydrate transport, but also for nitrogen metabolism (PTS^{Ntr}). The *EIIA^{Ntr}* and *EI^{Ntr}* are paralogous to *ptsI* and *ptsH*, respectively, and the phosphorelay cascade in PTS^{Ntr} is analogous to that in the canonical PTS (Cases *et al*, 2007). However, a key difference between the two is the absence of an identified final acceptor of the phsohoryl group from EIIA^{Ntr} (Pflüger-Grau & Görke, 2010). The genomic contexts of *EIIA^{Ntr}* and *EI^{Ntr}* encoded in close proximity to *rpoN*, which encodes the alternative RNA polymerase sigma factor σ^{54}, is conserved among proteobacteria, except epsilon proteobacteria (Deutscher *et al*, 2006). *V. cholerae* encodes two homologs of EIIA^{Ntr} and EI^{Ntr} (Houot *et al*, 2010b) and interestingly, the transcriptome data indicates upregulation of *EIIA^{Ntr}* as well as *rpoN* upon overexpression of VcdRP (by ~2.5 fold, Table 1). It has been shown previously for *E. coli* that AKG and glutamine reciprocally regulate the phosphorylation of *EI^{Ntr}*. Glutamine inhibits phosphorylation whereas AKG stimulates this phosphorylation (Lee *et al*, 2013). Both the homologs of *EIIA^{Ntr}* also repress biofilm accumulation in *V. cholerae* (Houot *et al*, 2010b).

In the aquatic environment, *V. cholerae* degrades the chitinous exoskeletons of the associated zoopanktons to N-acetylglucosamine (NAG), a sugar transported exclusively by the PTS (Meibom *et al*, 2005). VPS, the extracellular polysaccharide of the biofilm matrix is activated when PTS substrates are abundant and is repressed when PTS substrates become scarce (Houot & Watnick, 2008). The authors proposed that the augmentation of the bacteria by means of biofilm formation is likely to be regulated by PTS-dependent pathways. In line with this hypothesis, it is worth speculating the role of VcdRP in indirectly mediating biofilm formation. It is possible that based on the nutritive potential of the environmental surface, the riboregulatory element VcdR inhibits the production of the NagE transporter, coupled with the concomitant modulation of the flux through the TCA cycle by VcdP as well as positively acting on *EIIA^{Ntr}* to repress biofilm formation (Figure 47)

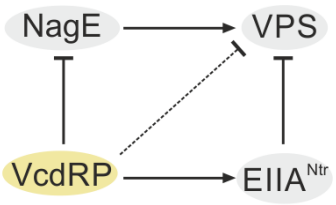

Figure 47. VcdRP indirectly influences biofilm formation. Depending on the nutrient availability, PTS substrates positively regulate biofilm formation. VcdR inhibits NagE, a PTS transporter. Additionally, EIIA^{Ntr} represses biofilm formation by downregulating VPS synthesis. Transcriptome analysis of VcdRP indicated the upregulation of EIIA$^{Ntr.}$ Therefore, by influencing the levels of EIIA^{Ntr}, it is possible that VcdRP indirectly regulates biofilm formation.

4.5 Citrate synthase-associated fitness and its allosteric inhibition

Citrate synthase (CS) enzyme serves as the pace-making enzyme in the first step of the TCA cycle (Weitzman & Jones, 1975). As an almost ubiquitous enzyme, CS is found in most, though not all, microorganisms, including all eukaryotes and most prokaryotes. As gatekeepers of the TCA cycle, CSs play a significant role in controlling the energy flux and metabolic rate of the cell by linking nutrient assimilation with conversion of subsequent intermediates and with the generation of precursors for various biosynthetic pathways (Liao *et al*, 2014). As a node affecting many pathways, CS appears to be tightly regulated in microorganisms.

The deletion of the citrate synthase gene (*gltA*) in *Klebsiella pneumoniae* has been shown to dramatically reduce metabolic flexibility as well as its overall fitness. Additionally, *gltA* is needed specifically for spleen and gut colonization of mice, but was dispensable for replication in the bloodstream (Vornhagen *et al*, 2019). The organ-specific fitness factor was attributed to the differences in nutrient composition. Likewise, deletion of the *gltA* gene in *V. cholerae* has previously been linked to the altered fitness during different stages of its bi-phasic lifestyle (Kamp *et al*, 2013). Virulence was attenuated specifically in infant rabbits, but not in infant mice. The authors link this host-specific diminished fitness with differences in carbon and energy sources that are available in both hosts. This hypothesis is further bolstered by differentially expressed *tcpPH* levels between infant mice and rabbits (Mandlik *et al*, 2011).

Interestingly, the structural basis for CS regulation also differs based on the organism. In eukaryotes, archaea, and Gram-positive bacteria, CS exists as a homodimer, with two active sites. This kind of CS, which does not show regulatory properties, is classified as type I (Figure 23C). Its crystal structure has been resolved, with several studies of such enzymes from vertebrates (Remington *et al*, 1982; Wiegand & Remington, 1986) and archaea (Russell *et al*, 1994, 1997). These studies have enabled detailed analyses of active sites, substrate binding, catalysis through physical measurements as well as by site-directed mutagenesis studies, not limited to eukaryotes and archaea (Handford *et al*, 1988; Anderson & Duckworth, 1988; Pereira *et al*, 1994).

In contrast, the CS of Gram-negative bacteria are allosterically inhibited by NADH (Duckworth & Tong, 1976; Pereira *et al*, 1994). Such type II CS were first discovered by Weitzman in 1966 in *E. coli*. However, the complexity of its structure made it particularly challenging to decipher until the groundbreaking study by (Maurus *et al*, 2003). The three-dimensional structure of *E. coli* CS was deduced from a mutated variant, in which the active site phenylalanine, Phe383, was replaced with alanine (F383A). A previous study showed that the kinetic properties of this variant were strongly shifted toward the low-affinity "tense" T state (Pereira *et al*, 1994). Since NADH is an allosteric inhibitor, it binds selectively to the T state. The mutated F383A variant therefore remained bound to NADH more tightly, thereby allowing the co-crystallization of F383A in complex with NADH. Interestingly, the enzyme activity of this variant is not affected by NADH (Nguyen *et al*, 2001).

Comparison of CS from representative organisms of both types revealed a high degree of sequence conservation in the active sites (Figure 48, indicated in red, adapted from (Nguyen *et al*, 2001)). Type II CS have been shown to exhibit substrate saturation for acetyl CoA (Donald *et al*, 1991). Interestingly, this has been linked to specific amino acid residues that are only found in type I, but are missing in type II sequences : arginine residue at position 46 as well as the VVPGY conserved patch (Figure 48, indicated in black, adapted from (Ner *et al*, 1983)).

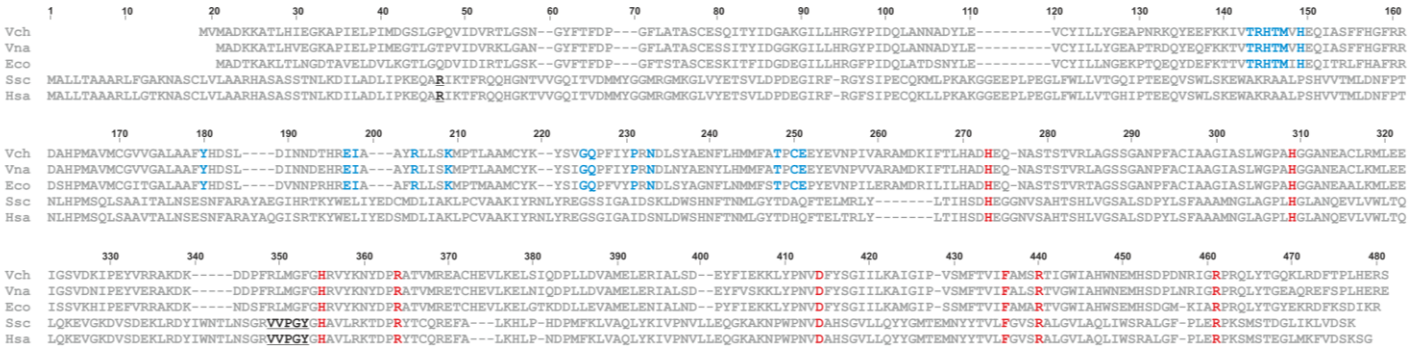

Figure 48. Conservation among type I and type II citrate synthase from representative species. Type I CS sequences are from Ssc (*Sus scrofa* (pig)) and Hsa (*Homo sapiens*), whereas type II sequences are from Vch (*V. cholerae*), Vna (*V. natriegens*) and Eco (*E. coli*). Indicated in blue are the 22 NADH-interacting residues found in the hexameric type II CS structure deduced by (Maurus *et al*, 2003). The residues indicated in red are the active site residues conserved among both types. The underlined sequences in black are exclusive to type I and correspond to the main chain atoms which bind to the adenine ring of acetyl CoA (Ner *et al*, 1983).

To specifically address how the increase in activity of GltA by VcdP correlated with the type of CS, the enzyme activity of cell lysates was measured in *B. subtilis* and a different species of *Vibrio*, *V. natriegens*. Remarkably, VcdP over-expression readily upregulated the GltA activity in *V. natiengens*, but not in *B. subtilis* (Figure 25A). This suggests the specificity of VcdP-mediated activation is only limited to type II hexameric CS of Gram-negative bacteria. The activation also likely affects the interaction of GltA with NADH and this hypothesis is further bolstered by the unchanged CS activity in the lysates of *V. cholerae* F383A upon over-expression of VcdP (Figure 27B).

4.6 Evolution of citrate utilization

The fine-tuning of CS activity levels has been associated with the overall competitive fitness of *E. coli*. The Lenski-long term evolution experiment (LTEE) that began in 1988 is an ongoing study that initially started with twelve identical populations of *E. coli* aimed at tracking the course of genetic mutations that have been acquired over tens of thousands of generations (Lenski *et al*, 1991; Blount *et al*, 2018). Natural selection works systematically to adapt populations to their surrounding environments. Mutations in genes are therefore acquired to thrive in niches that had previously been unoccupied (Lenski, 2017).

Although citrate is an atypical carbon source, *E. coli* can ferment citrate under anoxic conditions. However, the only known documented study of citrate consumption under aerobic conditions was reported for *E. coli* K-12 (Hall, 1982). The study hypothesized that a single complex mutation, or a combination of multiple mutations activated genes that jointly expressed a citrate transporter, although these genes were not identified. Therefore, despite having the potential to evolve a mutation that would allow the bacteria in the LTEE to grow on citrate aerobically, none of the twelve populations acquired mutation in the *gltA* gene for 15 years (Blount *et al*, 2008).

The first mutation (*gltA1*) alleviated repression of CS by NADH, thereby increasing its enzymatic activity, however, this mutation critically affected the overall fitness. A subsequent analysis to 'replay' this event, combined with whole-genome sequencing suggested that the bacteria had to acquire two other prior mutations, that were initially not hypermutable, and therefore went amiss previously. These preceding mutations refined the *gltA1* mutation to allow the complete aerobic utilization of citrate as the sole carbon source (Blount *et al*, 2012).

Acquiring multiple mutations in the same gene has been very rare among LTEE lineages, especially among those like *gltA* that had retained a low ancestral mutation rate (Wielgoss *et al*, 2011). Therefore, it was surprising when the *gltA* gene acquired a subsequent mutation (*gltA2*). Unlike, *gltA1*, this newly acquired mutation was beneficial for growth on citrate. Its improved fitness has been attributed to the concomitant decrease in the activity of CS. Upon growth on citrate, the CS reaction is detrimental since it consumes acetyl-CoA and diverts oxaloacetate that is otherwise needed for gluconeogenesis back into the TCA cycle. Therefore, the *gltA2* mutation emerged to reverse the change in enzyme activity caused by *gltA1* (Quandt *et al*, 2014, 2015). It is therefore possible that the VcdP provides a transient increase in CS activity, and VcdR may compensate for the associated negative fitness. Indeed, the *gltA1*-associated increase in CS activity has been shown to be detrimental especially when grown on glucose and acetate. Therefore, VcdR seems to regulate the incoming carbohydrates by inhibiting the PTS transporters as well as PtsH and PtsI.

4.7 Post-transcriptional control of bacterial metabolism

The pivotal role of sRNAs as an additional layer of regulation in the uptake and utilization of specific carbohydrates has been extensively investigated in various bacterial species (Gorke & Vogel, 2008; Durica-Mitic *et al*, 2018). Some of these include the sRNA Spot 42 of *Enterobacteriaceae* and *Vibrionaceae*, that represses genes for utilization of secondary carbon sources (Møller *et al*, 2002; Beisel & Storz, 2011; Bækkedal & Haugen, 2015). Like VcdR, Spot 42 is also repressed by cAMP-CRP and therefore only active in the presence of glucose generating low cAMP levels (Figure 13A). Upon growth on glucose, *E. coli* Spot 42 participates in a multi-output feed forward loop to decrease the expression of genes involved in central and secondary metabolism. These include *gltA* of the TCA cycle, *maeA* which contributes to malate catabolism, *fucI* involved in fucose catabolism, and *sthA*, which aids in the oxidation of NADPH, to name a few (Beisel & Storz, 2011). The sRNA GcvB of *E. coli* and *Salmonella* regulate amino acid and oligopeptide metabolism by base-pairing with *dppA* and *oppA* operons (Pulvermacher *et al*, 2008; Sharma *et al*, 2011). sRNAs also coordinate carbohydrate metabolism with oxygen and iron availability. For example, in the absence of oxygen, sRNAs such as FnrS in *E. coli* and RoxS in *B. subtilis*, metabolism gets redirected from oxidative phosphorylation to anaerobic respiration or fermentation (Durand & Storz, 2010; Durand *et al*, 2015). Upon iron starvation, the sRNA RyhB represses TCA cycle enzymes to save iron for essential processes (Massé *et al*, 2005b). In *E. coli*, two homologous sRNAs GlmY and GlmZ regulate the key enzyme GlmS to achieve homeostasis of glucosamine-6-phosphate, an essential precursor for cell envelope synthesis (Göpel *et al*, 2011). The sRNA SdhX is processed from TCA cycle genes and adjusts carbon flux by discoordinately regulating acetate kinase levels in *E. coli*, thereby linking TCA cycle with acetate metabolism (De Mets *et al*, 2019).

The Hfq-dependent sRNA SgrS regulates sugar-phosphate stress by impeding the production of PtsG and ManXYZ transporters (Rice & Vanderpool, 2011). However, inhibiting new synthesis of glucose transporters by sRNAs like SgrS and VcdR alone may not be sufficient to relieve stress under conditions where preexisting glucose transporters remain competent for further glucose uptake. Therefore, it is possible that sRNAs have subsequently also acquired additionally coding functions to bridge the gap and serve as dual-function regulators. Accordingly, the synergistic contribution of their cognate small proteins SgrT and VcdP redundantly or concomitantly reduce the accumulated sugars. Dual-function regulators like VcdRP therefore provides a nuanced and robust response to balance overall carbon metabolism.

4.8 Dual-function RNAs (two are better than one)

Although only a handful of bacterial dual-function RNAs have been characterized, their overarching importance in modulating carbon metabolism and /or driving virulence gene expression is becoming increasingly clear. Despite this commonality, there exist a plethora of differences that make them fascinating candidates for further studies. All, but the recently discovered dual-function regulator SR7 of *B. subtilis*, act in *trans* on their target mRNAs (Ul Haq *et al*, 2021). The riboregulatory component of SR7 (formerly called S1136) acts in *cis* to downregulate the *rpsD* gene via an antisense mechanism (Mars *et al*, 2015). Interestingly, the spatial separation of the ORF from the base-pairing element may determine if they function in concert or in two mutually exclusive pathways. SR1 of *B. subtilis* employs the same nucleotide sequence for both its riboregulation and coding function, and therefore can only act under distinct environmental conditions: SR1 during glycolysis, whereas SR1P during gluconeogenesis (Gimpel *et al*, 2012). In contrast, both SgrST and VcdRP harbor regulatory domains that are physically separate, and their riboregulatory elements protect cells from the toxic effects of αMG (Richards *et al*, 2013 and Figure 21B). Similarly, RNAIII also encodes its riboregulatory element and its small protein Hld in distinct loci (Morfeldt *et al*, 1995). However, while SgrS and SgrT act independently in the same physiological pathway in response to glucose-phosphate stress response (Wadler & Vanderpool, 2007; Rice & Vanderpool, 2011), VcdR and VcdP on the other hand independently modulate sugar uptake and the TCA cycle, respectively (Figures 29, 30), while the RNAIII riboregulator mediates virulence production, its small protein encodes δ-hemolysin that is essential for cell lysis. Interestingly, the translation of *hld* is delayed by one hour after RNAIII transcription (Balaban & Novick, 1995) and this delay was attributed to intramolecular interactions of the 3′ end of RNAIII with the RBS of *hld* resulting in the RBS of being occluded in a secondary structure. Deletion of the 3′ end of RNAIII eliminated the delay between RNAIII production and the appearance of δ-hemolysin, corroborating the hypothesis that there is a translation-inhibitory structure between the 3′ end of RNAIII and the *hld* RBS.

Unlike Gram-negative sRNAs, Gram-positive bacteria do not typically need Hfq to base-pair (Bohn *et al*, 2007), although both RNAIII and SR1 bind to Hfq (Heidrich *et al*, 2006). The role of Hfq is particularly interesting in the context of Gram-negative SgrST and VcdRP dual regulators, because although both the sRNAs are Hfq-dependent, yet how the translation initiation / elongation and how the structures of their peptide counterparts are re-modelled in response to Hfq remain to be assessed (Wadler & Vanderpool, 2007; Huber *et al*, 2020). The base-pairing of PTS-dependent *ptsG*, *nagE* and *treB* as well as PTS-independent *ptsH* and *ptsI* with VcdR occur in close proximity to each of their RBS (Figure 20B-E), and therefore there may be competition for the binding of either ribosomes or the sRNA with Hfq. The atypical molecular mechanism of SgrS binding at the *manXYZ* mRNA was recently elucidated (Azam & Vanderpool, 2020). This non-canonical regulation entails the recruitment of Hfq by SgrS at the binding site that overlaps with the *manX* SD sequence, making the RBP the actual repressor and SgrS the chaperone. In addition, the evolution of dual-function RNA regulators appears to be driven in both directions. For example, SgrT homologs are limited to only enteric bacteria, even though SgrS is conserved among other phyla such as *Yersinia*, *Erwinia* and *Klebsiella*, suggesting that the mRNA element probably pre-evolved before the ORF (Horler & Vanderpool, 2009). In contrast, the δ-haemolysin of RNAIII is more broadly conserved than its riboregulatory component (Verdon *et al*, 2009).

4.9 MbrA interferes with Fur regulon

The role of MbrA in the cell has been difficult to elucidate. Although it is annotated as an RNA-binding protein and is conserved among many marine bacteria, the RNA-binding ability of MbrA could not be conclusively determined in *V. cholerae*. However, over-expression of *mbrA* in the cell seems to regulate about 165 transcripts (Table 3). Closer inspection of the differentially expressed genes upon overexpression of *mbrA* indicates several genes involved in the Ferric uptake regulator (Fur) regulon.

The transcription factor Fur is the major iron-responsive regulatory factor in Gram-negative bacteria (Hassan & Troxell, 2013). Based on the availability of iron, Fur forms complexes with iron and binds to sites in the promoters called Fur boxes (Escolar *et al*, 1999). Although most of the regulation by iron and Fur is negative, instances of positive gene regulation have also been documented. Often this is mediated by the sRNA RyhB (Massé & Gottesman, 2002). The Hfq-dependent RyhB modulates the expression of several genes that control motility, chemotaxis, and biofilm formation in the pathogen (Mey *et al*, 2005). Interestingly, RyhB was downregulated by ~2.2-fold upon over-expression of *mbrA*.

In *V. cholerae*, both Fur and iron positively regulate *ompT* expression, independent of RyhB (Craig *et al*, 2011). *V. cholerae* has approximately 10 major outer membrane protein (OMPs) that function as porins (Kelley & Parker, 1981). The differential regulation of these porins are subject to nutrient availability, and are controlled both at the transcriptional as well as post-transcriptional level (Provenzano & Klose, 2000). The post-transcriptional control of *omp* expression is mediated by the sRNA VrrA in a σ^E-dependent manner (Song *et al*, 2008). The transcriptome data also shows a ~2-fold upregulation of the genes *vc1742* and *vc1743*. VrrA is encoded in the intergenic region (IGR) of *vc1741* and *vc1743*; and *vc1742* is a very small 138 bp predicted ORF that has no clear SD sequence and 13 of the 46 codons overlap with the *vrrA* locus. MicV was recently shown to act redundantly with VrrA to regulate multiple target genes in *V. cholerae* (Peschek *et al*, 2019) and was also found to be upregulated by 2.5-fold upon overexpression of MbrA. Interestingly, mutants of the sRNA VrrA have previously been shown to increase infant mouse colonization (Song *et al*, 2008) whereas the *fur* mutant of *V. cholerae* exhibits significant defect in colonization of infant mice (Mey *et al*, 2005). It is therefore possible that MbrA counteracts the Fur-mediated regulation to decrease colonization.

The reciprocal regulation of MbrA and Fur is also evident in the antagonistic regulation of the gene *vca0734* encoding a hypothetical protein, which has previously been linked to be negatively regulated by Fur, independent of iron availability (Mey *et al*, 2005), and is upregulated upon overexpression of MbrA. Additionally, the same study also identified a Furbox upstream of the CDS of *vc0205*, encoding a hypothetical protein and is significantly upregulated by ~16-fold by MbrA. The gene *vc0519* encoding a hypothetical protein was downregulated by ~2-fold in the transcriptome

4.10 MbrA regulates outer membrane vesicle-associated hemolysins

Outer membrane vesicles (OMVs) are naturally released from the outer membrane of Gram-negative bacteria (Beveridge, 1999). OMV formation is regulated by the phospholipid transporter VacJ/Yrb in Gram-negative bacteria. Upon iron depletion, a Fur-dependent repression of the transporter is triggered (Roier *et al*, 2016). Accordingly, mutants of *vacJ* and *yrbE* in *H. pylori*, *E. coli*, *Salmonella* and *V. cholerae* resulted in increased production of OMVs. The *V. cholerae* specific gene that encodes for *vacJ* is *vc2048*, is downregulated by ~2-fold upon overexpression of MbrA.

OMVs are known to contain membrane proteins, lipopolysaccharides, peptidoglycan, phospholipids, metabolites as well as signaling molecules (Jan, 2017). Among the prominent roles in diverse physiological and pathological functions, OMVs have been recognized for their role in acquisition of nutrients, stress responses, adhesion and virulence factors to evade host defense system (Zingl *et al*, 2020). In addition, OMVs also contribute to the delivery of toxins and hemolysins to the host cells: cytotoxic necrotizing factor type 1 (CNF1) is a Uropathogeneic *E. coli* (UPEC) virulence factor that tightly associates with OMVs and is delivered via the bound-vesicles to the environment or infected tissues (Kouokam *et al*, 2006). Similarly, α-hemolysin is an important virulence factor expressed in EHEC and is found inside OMVs that are released by the bacteria. The toxin is found in its cytotoxically active form and is capable of lysing red blood cells (Balsalobre *et al*, 2006). The cytolysin protein A (ClyA) of *E. coli* K-12 is a pore-forming toxin that is accumulated and delivered via OMVs and confers its cytotoxicity upon contact with mammalian cells (Wai *et al*, 2003).

This vesicle-mediated transport mechanism that is responsible for the activation and delivery of pathogenic effector proteins may also be extended to *V. cholerae* and is possibly mediated by MbrA. Upon over-expression of MbrA, the genes *vca0218* and *vca0219* were upregulated between ~2.3 to 2.8-fold. These genes encode cholera haemolysin HlyA (Heidelberg *et al*, 2000). It has been recently shown that the expression of HlyA is regulated under the collective actions of HapR and Fur in *V. cholerae*. Repression of *hlyA* by HapR at the transcriptional level was achieved through direct binding of HapR to the its promoter, while at the posttranscriptional level was mediated via the hemagglutinin protease HapA (Tsou & Zhu, 2010). The Fur box-like sequence found upstream of *hlyA* is responsible for the Fur-mediated repression of the organism's hemolytic activity (Gao *et al*, 2018). Interestingly, both HlyA and MbrA were highly induced when *V. cholerae* was cultured in rabbit ileal loops (Xu *et al*, 2003). Furthermore, *vc1798* encoding the transcriptional factor Eha similar to *Edwardsiella tarda* and *S. enterica,* which binds to and activates the cytolytic component of hemolysin, is also upregulated upon overexpression of MbrA. Therefore, MbrA may also contribute to the regulation of hemolysin production and / or delivery of OMVs.

4.11 Regulation of biofilm formation by MbrA

Historically, *V. cholerae* classical strains could be differentiated from El Tor by their sensitivity to polymyxin B (Gangarosa *et al*, 1967). Pathogenic bacteria in the host are often presented with a variety of stresses imposed upon them by the host immune system. Production of antimicrobial peptides is one such defense mechanism (Mahlapuu *et al*, 2016). Bacteria have evolved several strategies to counteract the effect of these peptides such as by secreting proteases that would degrade the peptides or by encoding efflux systems that actively pump antimicrobial peptides back to the environment (Groisman *et al*, 1997). The lipopolysaccharide of Gram-negative bacteria is composed of the O-antigen polysaccharide, the core polysaccharide and lipid A that is anchored in the bacterial outer membrane and is often recognized by the host (Wang & Quinn, 2010). The altered sensitivity of polymyxin B between the two serotypes could therefore be due to a modified lipid A structure (Matson *et al*, 2010).

Three amino acid lipid modification (Alm) proteins were found to be essential for this unique lipid A composition: AlmG (Vc1577), AlmF (Vc1578) and AlmE (Vc1579). Collectively, they contribute to the glycine modification of lipid A that is essential for polymyxin B resistance (Hankins *et al*, 2012). AlmG encodes a lipid A hydroxyacyltransferase whereas AlmF and AlmE catalyze the glycine and

diglycine modification of the lipid (Hankins *et al*, 2011). AlmE generates glycyl-AMP and pyrophosphate from glycine and ATP. The AMP is subsequently released during the ligation of glycine with AlmF. The glycyl-AlmF then serves as the aminoacyl donor to AlmG, that esterifies glycine to the lipid A (Henderson *et al*, 2014). It was proposed that the overall decrease in cell surface charge from the glycine metabolism impacted the antimicrobial resistance. Moreover, deletion of the operon has been shown to inhibit biofilm formation, however did not exhibit any defects in colonization (Bilecen *et al*, 2015). Upon exposure to polymyxin B, the transcript levels of the *alm* operon was significantly elevated (Matson *et al*, 2017). Interestingly, AlmG and AlmF are both upregulated by ~2.7-fold upon overexpression of MbrA. Therefore, MbrA seems to be indirectly involved in the repression of biofilm formation via regulating the genes of the *alm* operon.

4.12 Regulation of Cpx regulon by MbrA

The bacterial Cpx two-component signal transduction pathway mediates adaptation to envelope protein misfolding. *E. coli* Cpx regulon comprises at least 50 genes and over 34 operons (Price & Raivio, 2009). A comprehensive microarray study of the Cpx regulon of *V. cholerae* under virulence-inducing conditions showed significantly increased expression of the genes *vca0732* (~3.4-fold) and *vca0733* (~4.7-fold) upon overexpression of CpxR (Acosta *et al*, 2015). Interestingly, the MbrA transcriptome data also indicates a significant upregulation of *vca0733* (5-fold) as well as directly or indirectly regulates other genes in its vicinity. The gene upstream of *vca0733*, namely *vca0732* encodes *sipA* (Saul-McBeth & Matson, 2019). The study demonstrated that SipA binds to antimicrobial peptides in the bacterial periplasm. Although mutants of *sipA* showed no obvious defect in colonization of El Tor strains, they are required for resistance of antimicrobial peptides in classical biotypes. Much like their homologs YgiW of *E. coli* and YdeI of *S. enterica* (Pilonieta *et al*, 2009), SipA also associates with outer membrane porins (specifically, OmpA). SipA chaperones antimicrobial peptides to porins, for efflux out of the cells, thereby relieving stress. Accordingly, *sipA* transcript levels were also significantly upregulated upon polymyxin B exposure (Matson *et al*, 2017). Of note, a fifth of the genes involved in iron acquisition, usage and storage were also significantly upregulated in the same study. Interestingly, the gene downstream of *vca0733*, *i.e.*, *vca0734* is repressed by Fur (Mey *et al*, 2005), and is upregulated by 2.2-fold by MbrA. Although the roles of Fur and Cpx appear to have unique functions in OMP biogenesis, modulating antimicrobial peptide resistance, biofilm formation and periplasmic misfolding, how MbrA cumulatively connects these two regulatory pathways is an important direction for future research.

4.13 Regulatory interplay between MbrA and CosR

Bacteria need to adapt to changing osmolarity by either importing or producing compatible solutes to counteract osmotic pressure. It is energetically favorable to import rather than produce these solutes and accordingly, bacteria encode multiple osmoregulated transporters (Kempf & Bremer, 1998). CosR (short for compatible solute regulator) is a global regulator of osmotic stress response in *Gammaproteobacteria* (Gregory *et al*, 2020). The bi-phasic nature of *V. cholerae* has made the bacterium more halo-tolerant (Singleton *et al*, 1982). Although the bacteria are capable of growing and surviving in fresh water, saline-rich brackish water has been the primary epicenter of cholera epidemics in the past (Louis *et al*, 2003; Huq *et al*, 2005). In the host, *V. cholerae* must adapt to a range of salinities and changing osmolartities as it passes through the digestive system and is dispersed back into the environment. Moreover, production of virulence factors as well as biofilm formation are also modulated by osmolarity (Tamplin & Colwell, 1986; Shikuma & Yildiz, 2009).

CosR activates biofilm formation and represses motility (Shikuma *et al*, 2013). Comparing the transcriptome profiles of a *ΔcosR* strain to that of a wild-type *V. cholerae* strain in this study revealed 38 additional DEGs. Of note, *ompT* and *cpxR* were inversely regulated upon deletion of *cosR*. Of the remaining DEGs, *vc1657*, *vc2151* and *vca0467*, all encoding hypothetical proteins were upregulated by nearly 2-fold in the absence of *cosR*. Interestingly, these genes are also significantly upregulated when MbrA is overexpressed: *vc1657*, *vc2151* and *vca0467* by 2.5-, 3.2- and 13-fold, respectively. Overexpression of MbrA positively regulates flagellar synthesis and motility (Figure 40). Cumulatively, CosR seems to behave antagonistic to MbrA in that while the former activates biofilm formation and repress motility, MbrA seems to reverse the effect directly or indirectly.

4.14 Regulation of VPI-2 by MbrA

Acquisition of virulence genes encoded on mobile genetic elements has been crucial for the emergence of pathogenic cholera-producing serotypes of *V. cholerae* (Morris, 2011). The virulence determinants of the disease, CT and TCP are both derived from phage origins – CTXϕ and VPIϕ, respectively (Karaolis *et al*, 1999; Mukhopadhyay *et al*, 2001). A pathogenicity island has been defined as a large unstable chromosomal region that encodes several virulence genes, is present only in pathogenic isolates, has a G+C content that differs from the rest of the genome, is associated with a tRNA gene, has insertion and/or repeat sequences near the site of integration and contains a bacteriophage-like integrase (Hacker *et al*, 1997). The bacterial gene encoding neuraminidase (*nanH*) has been thought to be inherited from phages via horizontal gene transfer (Roggentin *et al*, 1993). A subsequent study of *nanH* of *V. cholerae* and its flanking region revealed another pathogenicity island named *Vibrio* pathogenicity island-2 (VPI-2; Jermyn & Boyd, 2002).

Table 4. ORFs within the *V. cholerae* VPI-2 pathogenicity island that are upregulated in the MbrA transcriptome data. Homology is based on BLASTP analysis.

ORF	Transcriptome fold-change	Length (aa)	Homologous protein	Organism	Amino acid identity (%)	*E*-value
Vc1791	2.07	346	Mu-like GP42	Mu phage	35	4.00E-48
Vc1792	2.56	119	Mu-like GP41	Mu phage	33	1.00E-05
Vc1793	2.26	125	Transposase	*C. elegans*	32	3.40
Vc1795	2.16	106	Mor protein	Mu phage	28	0.26
Vc1797	2.78	153	No significant match	-	-	-
Vc1798	2.17	383	Eha protein	*S. typhi*	43	6.00E-37
Vc1799	2.54	585	Integrase	*S.* Typhimurium	21	0.02
Vc1800	2.66	323	Plasmid replication protein C	*C. butyricum*	27	4.00
Vc1801	3.99	120	No significant match	-	-	-
Vc1802	4.96	78	Transcriptional regulator	*S. typhi*	37	3.00

The VPI-2 is a 57.3 kb region that possesses all the characteristics of a pathogenicity island as defined by (Hacker *et al*, 1997) and harbors 52 ORFs of which 10 are upregulated by 2 to 5-fold upon overexpression oy MbrA (Table 4). Among these, 3 ORFs showed similarity to Mu bacteriophage genes. The amino acid sequence of the ORFs Vc1791 and Vc1792 share similarity with the GP42 and GP41 protein of Mu phage, whereas Vc1797 closely resembles the Mor protein of the bacteriophage. Two of the genes *vc1797* and *vc1801* showed no significant matches and the remaining five genes have been linked to phage origins and have known functions from other organisms, including *vc1793* and *vc1799* that show similarity to genes encoding transposase and integrase proteins with known roles in *C. elegans* and *S.* Typhimurium, respectively. These genes could potentially be involved in

the mobilization and integration of this region (Jermyn & Boyd, 2002). Although no direct link has been established between the VPI-2 cluster and virulence, the *nan-nag* cluster of genes has been implicated in the catabolism of sialic acid of the gut and intestine, thus providing a competitive advantage to *V. cholerae* to utilize these aminosugars as carbon sources (Almagro-Moreno & Boyd, 2009). It would therefore by interesting to analyze the fitness landscape of the organism in the presence and absence of MbrA for *in vivo* survival.

4.15 Regulation of motility by MbrA

Although the transcriptional data provides some evidence for MbrA's role in repressing biofilms as well as reducing colonization through its effect on the Alm operon and acting on OMV biogenesis, the regulation of motility seems to be the most discernable phenotype.

During early stages of infection, a large inoculum of *V. cholerae* is necessary for successful infection, given that the bacteria encounter a variety of hostile compounds such as gastric acids and bile salts during transmission (Svenningsen *et al*, 2008). Therefore, the bacterial numbers are greatly reduced by the time they reach the small intestine. Furthermore, a drop in HapR levels is crucial for the expression of virulence factors (Rutherford *et al*, 2011). *hapR* expression is also modulated by the regulatory network that governs flagellar assembly (Liu *et al*, 2008). Specifically, FliA, which is the alternative σ-factor (σ^{28}) that activates late-class flagellin genes in *V. cholerae* (Figure 40B), represses *hapR* expression. In addition, the study also shed light on the role of removal of flagellar rod proteins, which results in high levels of active σ^{28} because of increased secretion of the anti-σ^{28} protein FlgM, inhibits *hapR* transcription. Although motility is essential when the bacteria must transit through the digestive tract until they reach the small intestine, upon mucosal penetration, the flagella break (Echazarreta & Klose, 2019). Concomitantly, the anti-sigma factor FlgM is secreted during this stage and leads to the activation of FliA. FliA in turn represses *hapR*, thus leading to maximal expression of CT and TCP (Liu *et al*, 2008). Late stages of infection often involve the detachment of the bacteria from the site of colonization to either exit the host or find another site to initiate a new infectious cycle. Therefore, motility conferred on the bacteria by its single polar flagellum is necessary for this process (LaRocque *et al*, 2005). MbrA could therefore play a role in the integration of quorum sensing signals and flagellar regulatory networks by repressing *hapR* levels at LCD (Figure 36) and promoting motility and chemotaxis at HCD (Figure 40).

4.16 Outlook

The present work expands the otherwise limited list of bacterial dual-function regulators. Pathogenic bacteria like *V. cholerae*, at least transiently, derive their carbon and energy parasitically or destructively from a host organism. This is accomplished in part by the synthesis of a wide assortment of virulence determinants that can kill host cells and catabolizing macromolecules. Therefore, it is not surprising that the regulation of many virulence determinants is controlled by nutrient availability. VcdRP is the first identified dual-function regulator in *V. cholerae*. The riboregulatory element VcdR uses a canonical base-pairing mechanism typical for sRNAs to repress the mRNAs that encode PTS-dependent sugars as well as the two phosphor-carrier proteins involved in glycolysis. Although the functions of only a few bacterial small proteins have been described, most of them are inhibitory (Storz *et al*, 2014). Therefore, it is unusual that the small protein VcdP interacts with citrate synthase and increases its enzymatic activity. Together, they strike a balance between sugar uptake and its utilization. Additionally, the riboregulatory element is also responsible for the

repression of the virulence determinant, CT of *V. cholerae*. Although the present work elegantly demonstrates the roles of VcdR and VcdP, the exact molecular mechanism by which each of these components repress sugar uptake and increase citrate synthase activity remain to be addressed. VcdRP has previously shown to be Hfq-dependent (Huber *et al*, 2020). It would be interesting to know how binding of Hfq to VcdRP modulates the structure of VcdP. Furthermore, it would be interesting to note how the interaction of citrate synthase with VcdP interferes with its allosteric inhibition by NADH. Hydrogen deuterium exchange mass spectrometry (HDX-MS) studies of the small protein and citrate synthase would shed light on the exact binding sites of the peptide-protein interaction.

Although the RNA-binding nature of the putative RNA-binding protein MbrA could not be conclusively determined, it is interesting to note how the regulatory patterns of over 160 genes change upon overexpression of MbrA. Pathogenic bacteria like *V. cholerae* that coexist in competitive environments with other species, must develop different survival strategies to compete for space, nutrients, and ecological niches. MbrA positively regulates flagellar synthesis and motility and through its repression of HapR at low cell density serves as a bridge between quorum sensing and control of motility. The hypothesis that MbrA counteracts Fur-mediated regulation is an interesting new direction that needs further validation. A transcriptome analysis comparing the impact of bipyridyl treatment (to mimic iron-limiting conditions) upon overexpression of MbrA would provide insights into the regulatory interplay between Fur and MbrA. Although the generation of a clean knock-out of *mbrA* from the *V. cholerae* has been challenging, it would be interesting to test a CRISPR/Cas-based targeted editing of the genome to knock out *mbrA*.

Chapter 5

Appendix tables

Table 1. List of differentially expressed genes identified from the global transcriptome analysis of VcdRP

[#]Description is based on the annotation at KEGG (https://www.genome.jp/kegg)

[§]Genes with a total count cutoff >10 in all samples, with an absolute fold-change ≥ 2.0 and a FDR adjusted p-value ≤ 0.05 were considered to be differentially regulated

Gene	Description[#]	Fold change[§]		
		pVcdRP vs. pCtrl	pVcdR vs. pCtrl	pVcdP vs. pCtrl
astD	Succinylglutamic semialdehyde dehydrogenase	2.30	1.10	1.70
carB	Carbamoyl-phosphate synthase large subunit	1.40	1.00	2.00
cpdB	2',3'-cyclic-nucleotide 2'-phosphodiesterase / 3'-nucleotidase	2.20	1.30	1.40
galM	Aldose 1-epimerase	2.10	1.40	1.40
glpK	Glycerol kinase	4.00	1.50	2.50
glpQ	Glycerophosphoryl diester phosphodiesterase	3.30	1.70	1.40
glpT	Glycerol-3-phosphate transporter	3.60	1.50	1.20
glpX	Fructose-1,6-bisphosphatase II	1.30	1.10	2.10
gltA	Citrate synthase	2.10	1.40	1.40
lamB	Maltoprotein	10.00	2.00	4.40
lldD	L-lactate dehydrogenase	2.10	1.20	1.50
malE	Maleylacetate reductase	7.40	1.80	3.40
malF	Maltose/maltodextrin transport system permease protein	4.00	1.60	2.70
malG	Maltose/maltodextrin transport system permease protein	4.00	1.60	2.70
malQ	4-alpha-glucanotransferase	4.50	1.50	3.10
malS	Alpha-amylase	2.60	1.20	2.20

Table continues on the next page

Table continued from the previous page

Gene	Description[#]	Fold change[§]		
		pVcdRP vs. pCtrl	pVcdR vs. pCtrl	pVcdP vs. pCtrl
mglC	Methyl-galactoside transport system permease protein	4.20	1.50	2.10
nagE	PTS system N-acetylglucosamine-specific transporter subunit IIABC	-3.10	-3.30	1.10
potE	Putrescine transporter	3.40	1.30	1.60
ptsH	Phosphocarrier protein HPr	-4.00	-3.50	1.30
ptsI	Phosphoenolpyruvate-protein phosphotransferase	-2.50	-3.40	1.40
purN	Phosphoribosylglycinamide formyltransferase 1	1.20	1.10	2.10
purT	Phosphoribosylglycinamide formyltransferase 2	1.50	1.00	2.60
pyrB	Aspartate carbamoyltransferase	-1.00	-1.10	2.30
rbsC	Ribose transport system permease protein	6.80	1.40	2.80
sdhA	Succinate dehydrogenase, flavoprotein subunit	2.10	1.30	1.60
sdhC	Succinate dehydrogenase / fumarate reductase	2.10	1.30	1.20
sucA	2-oxoglutarate dehydrogenase EI component	2.10	1.30	1.50
sucC	Succinyl CoA synthetase, beta subunit	3.30	1.50	1.90
tnaC leader peptide	Tryptophanase leader peptide	3.00	1.70	1.50
treB	Trehalose PTS system EIIBC	-1.50	-2.40	1.60
ushA	5'-nucleotidase / UDP-sugar diphosphatase	2.00	1.40	1.40
vc0177	Hypothetical protein	-2.00	-2.10	1.30
vc0216	Methyl-accepting chemotaxis protein	2.30	1.30	1.30
vc0282	Methyl-accepting chemotaxis protein	2.10	1.40	1.30
vc0338	Solute carrier family 13 (sodium-dependent dicarboxylate transporter)	2.00	1.20	1.30
vc0384	Sulfite reductase (NADPH) flavoprotein alpha-component	1.70	1.10	2.00
vc0432	Malate dehydrogenase	2.30	1.40	1.80
vc0706	Sigma-54 modulation protein	2.30	1.50	1.10
vc0931	Hypothetical protein	-1.10	1.40	2.00
vc1325	Galactoside ABC transporter periplasmic D-galactose/D-glucose-binding protein	6.10	1.80	3.10
vc1327	Galactose/methyl galaxtoside transporter ATP-binding protein	4.30	1.50	2.40
vc1446	ATP-binding cassette, subfamily B, bacterial RtxE	-2.40	-1.40	-1.70
vc1447	Membrane fusion protein, RTX toxin transport system	-2.70	-1.30	-1.50
vc1448	ATP-binding cassette, subfamily B, bacterial RtxB	-3.30	-1.50	-2.00
vc1449	Hypothetical protein	-2.20	-1.80	-1.40
vc1450	RTX toxin activating protein	-2.00	-1.60	-1.30
vc1492	Glutamate dehydrogenase	2.70	1.40	1.70
vc1539a	Hypothetical protein	2.20	1.30	1.40
vc1595	Galactokinase	2.20	1.40	1.60
vc1596	Galactose-1-phosphate uridylyltransferase	2.10	1.30	1.60
vc1658	Serine transporter	-2.00	-1.20	-1.50
vc1696	DNA-binding protein	2.10	1.30	1.20
vc1741	TetR family transcriptional regulator	2.00	1.50	1.30
vc1774	N-acetylneuraminic acid mutarotase	2.10	1.10	1.40
vc1776	N-acetylneuraminate lyase	4.10	1.20	1.60
vc1777	Hypothetical protein	4.70	1.20	1.90

Table continues on the next page

Table continued from the previous page

Gene	Description[#]	Fold change[§]		
		pVcdRP vs. pCtrl	pVcdR vs. pCtrl	pVcdP vs. pCtrl
vc1778	Hypothetical protein	6.90	1.40	2.20
vc1779	C4-dicarboxylate-binding protein	9.90	2.10	3.70
vc1781	N-acetylmannosamine-6-phosphate 2-epimerase	6.80	1.40	3.10
vc1782	N-acetylmannosamine kinase	2.70	1.20	2.00
vc1783	N-acetylglucosamine-6-phosphate deacetylase	2.20	1.00	1.50
vc1784	Neuraminidase	2.70	1.20	1.90
vc1822	PTS system fructose-specific transporter subunit IIABC	2.50	1.50	1.40
vc1823	PTS system fructose-specific transporter subunit IIB	2.30	1.40	1.40
vc1824	PTS system nitrogen regulatory subunit IIA	2.50	1.50	2.00
vc1898	Methyl-accepting chemotaxis protein	2.80	1.40	1.50
vc1905	Alanine dehydrogenase	2.40	1.40	2.10
vc1953	Concentrative nucleoside transporter, CNT family	-4.80	-1.50	-4.50
vc1998	Methionine sulfoxide reductase B	-2.30	-2.00	-1.20
vc2013	PTS system glucose-specific transporter subunits IIBC	-2.70	-4.90	1.50
vc2084	Succinyl-coa synthetase subunit alpha	3.30	1.60	2.00
vc2086	Dihydrolipoamide succinyltransferase	2.20	1.30	1.70
vc2277	Xanthine-guanine phosphoribosyltransferase	-2.20	-2.00	-1.10
vc2305	Outer membrane protein OmpK	2.80	1.30	1.60
vc2338	Pseudogene	2.10	1.30	1.60
vc2350	2-deoxyribose-5-phosphate aldolase	-1.30	1.20	-2.20
vc2352	Concentrative nucleoside transporter, CNT family	2.10	1.50	1.30
vc2416	2′,3′-cyclic-nucleotide 2′-phosphodiesterase	2.10	1.30	1.40
vc2511	Aspartate carbamoyltransferase	-1.00	-1.10	2.00
vc2544	Fructose-1,6-bisphosphatase	2.20	1.30	2.00
vc2600	Hypothetical protein	-2.00	-1.30	-1.40
vc2656	Fumarate reductase flavoprotein subunit	2.50	1.40	2.20
vc2657	Fumarate reductase iron-sulfur subunit	2.40	1.30	2.30
vc2658	Fumarate reductase subunit C	2.20	1.30	2.10
vc2659	Fumarate reductase subunit D	2.60	1.30	2.10
vc2667	Hypothetical protein	1.10	1.50	2.40
vc2738	Phosphoenolpyruvate carboxykinase	2.20	1.10	1.80
vc2761	Inner membrane transport protein YdhC	-3.60	1.20	-5.70
vca0013	Maltodextrin phosphorylase	3.50	1.40	2.50
vca0015	Pseudogene	2.70	1.50	2.10
vca0025	NadC family protein	3.60	1.60	1.90
vca0037	Periplasmic copper chaperone A	2.00	1.30	1.20
vca0052	Hypothetical protein	-1.40	1.30	-4.10
vca0053	Purine nucleoside phosphorylase	-2.00	1.40	-7.00
vca0087	Hypothetical protein	-2.10	-1.70	-1.70
vca0127	D-ribose pyranase	9.40	1.60	3.80
vca0128	D-ribose transporter ATP-binding protein	8.90	1.40	3.20
vca0130	D-ribose transporter subunit RbsB	8.90	1.50	2.60
vca0131	Ribokinase	3.90	1.00	1.90

Table continues on the next page

Table continued from the previous page

Gene	Description[#]	Fold change[§]		
		pVcdRP vs. pCtrl	pVcdR vs. pCtrl	pVcdP vs. pCtrl
vca0179	Concentrative nucleoside transporter, CNT family	-2.30	-1.30	-1.60
vca0205	Anaerobic C4-dicarboxylate transporter	2.50	1.30	1.50
vca0276	Pseudogene	2.60	1.50	2.40
vca0277	Glycine cleavage system protein H	2.10	1.20	1.50
vca0280	Pseudogene	2.30	1.40	1.50
vca0556	Hypothetical protein	-2.20	-1.20	-1.40
vca0610	Isoprenoid biosynthesis protein	2.40	1.70	1.50
vca0623	Transaldolase B	-1.50	1.30	-2.30
vca0624	Transketolase	-1.40	1.20	-2.40
vca0743	Hypothetical protein	-1.10	1.20	-2.00
vca0745	Pseudogene	4.50	1.50	2.30
vca0843	Glyceraldehyde-3-phosphate dehydrogenase	3.10	1.30	1.90
vca0867	Outer membrane protein W	3.10	1.60	1.30
vca0946	Maltose/maltodextrin transporter ATP-binding protein	2.40	1.20	2.30
vca0985	Oxidoreductase/iron-sulfur cluster-binding protein	2.10	1.30	1.50
vca0987	Phosphoenolpyruvate synthase	4.50	1.30	3.70
vca1027	Maltose operon periplasmic protein	2.60	1.00	1.50
vca1063	Ornithine decarboxylase	5.50	1.70	2.70
vca1069	Methyl-accepting chemotaxis protein	2.50	1.30	1.40

Table 2. Proteins identified via LC-MS as potential interaction partners of VcdP::SPA.

Only proteins identified co-precipitating with VcdP::SPA in at least two out of three biological replicates, and that were absent from control samples, are listed. A minimum of two high confidence peptides with at least one unique peptide, and a protein level FDR <5%, was required for a protein to be classed as identified.
#Description is based on UniProt annotations (www.uniprot.org/)
§Identified in all three biological replicates (all other proteins were identified in two out of three biological replicates).
*The VcdP::SPA protein was given the arbitrary accession number A00055_SPA, to allow for database searches.
(PSM: peptide spectral matches, MW: molecular weight)

Accession	Description#	Unique Peptides	PSM	MW [kDa]
Q9KQA8§	Citrate synthase	15	55	48.6
A00055_SPA*	VcdP::SPA	11	54	11.3
Q9KTY5	Inositol-1-monophosphatase	6	24	29.1
Q9KV04	Peptidyl-prolyl cis-trans isomerase	7	22	28.1
Q9KR21	PTS system, fructose-specific II ABC component	11	20	65.6
Q9KLJ9	Glycerol kinase	8	19	55.6
P0C6C3	Flagellin A	8	18	40.4
Q9KV30	DNA-directed RNA polymerase subunit beta	15	16	149.4
O34242	Chaperone protein DnaJ	7	15	40.8
P0C6C5	Flagellin C	3	15	39.9
Q9KRZ1	Uncharacterized protein	10	15	183.4
Q9KSW2	ATP-dependent Clp protease, ATP-binding subunit ClpA	12	14	84
Q9KLA3	Glyceraldehyde 3-phosphate dehydrogenase	10	14	55.6
Q9KUT5	Immunogenic protein	7	14	35.2
Q9KQL3	Long-chain-fatty-acid-CoA ligase	8	14	62.8
Q9KQB6	Succinate-CoA ligase [ADP-forming] subunit alpha	10	14	29.9
Q9KUR3	A/G-specific adenine glycosylase	4	12	40.1
Q9KU42	Carbon starvation protein A, putative	4	12	53.1
Q9KKM6	Uncharacterized protein	8	12	67.1
Q9KPH4	Protein translocase subunit SecA	6	11	102.4
Q9KNH4	ATP synthase gamma chain	7	10	31.8
Q9KU11	DNA-binding response regulator PhoB	4	10	26.2
Q9KT50	Nucleoid-associated protein VC_1055	3	10	12
Q9KT11	Oligopeptide ABC transporter, ATP-binding protein	6	9	36.1
Q9KQB4	Dihydrolipoyllysine-residue succinyltransferase component of 2-oxoglutarate dehydrogenase complex	8	8	44.1
Q9KQT0	Peptidylprolyl isomerase	4	8	68.4
Q9KTB7	Adenylate kinase	4	7	23.3
Q9KUT3	Malate dehydrogenase	5	7	32
Q9KR88	Paraquat-inducible protein B	3	7	60.9
Q9KTW0	Uncharacterized protein	3	7	21.1
Q9KUZ2	30S ribosomal protein S6	3	6	14.2
Q9KUR8	Co-chaperone protein DjlA	5	6	31.9
Q9KQU1	Cys regulon transcriptional activator	3	6	36.2
Q9KN38	D-ribose pyranase	2	6	15.3
Q9KSF2	Fumarate hydratase class I	4	6	54.7
Q9KLH1	Methyl-accepting chemotaxis protein	1	6	52.1

Table continues on the next page

Table continued from the previous page

Accession	Description[#]	Unique Peptides	PSM	MW [kDa]
Q9KLH1	Methyl-accepting chemotaxis protein	1	6	52.1
Q9KVH6	Peptide ABC transporter, ATP-binding protein	6	6	63.5
Q9KL10	Transcriptional regulator, DeoR family	4	6	28
Q9KTX1	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase (flavodoxin)	3	5	40.6
Q9KSD1	Galactose/methyl galactoside import ATP-binding protein MglA	5	5	56.4
Q9KNG6	ParA family protein	5	5	28
Q9KSW1	Putative transport protein VC_1145	3	5	60.5
Q9KPK5	Threonine synthase	4	5	46.2
Q9KNG4	tRNA uridine 5-carboxymethylaminomethyl modification enzyme MnmG	5	5	70.1
Q9KL28	Uncharacterized protein	5	5	42.4
Q9KNC8	Uncharacterized protein	4	5	16.4
Q9KPT0	Uncharacterized protein	4	5	39.8
Q9KUY1	Uncharacterized protein	3	5	16.7
P23247	Aspartate-semialdehyde dehydrogenase 2	2	4	37.4
Q9KUJ8	Beta-ketoadipate enol-lactone hydrolase, putative	3	4	30
Q9KR62	Putative N-acetylmannosamine-6-phosphate 2-epimerase	2	4	25.5
Q9KSF0	Uncharacterized protein	2	4	51.9

Table 3. List of differentially expressed genes upon overexpression of MbrA

[#]Description is based on the annotation at KEGG (https://www.genome.jp/kegg)

[§]Genes with a total count cutoff >10 in all samples, with an absolute fold-change ≥ 2.0 and a FDR adjusted p-value ≤ 0.05 were considered to be differentially regulated

Gene / regulatory element	Description[§]	Fold-change[#]
CsrC	sRNA	-2.17
FlaX	sRNA	2.55
RyhB	sRNA	-2.19
Vcr015	sRNA	-2.56
Vcr019	sRNA	-2.79
Vcr039	sRNA	13.24
Vcr058 (CoaR)	sRNA	-2.35
Vcr069	sRNA	2.00
Vcr078	sRNA	-2.83
Vcr080	sRNA	-2.03
Vcr089 (MicV)	sRNA	2.52
Vcr102	sRNA	-2.84
vc0159	MbrA	11.52
vc0380	Phage shock protein G	8.00
vc0497	Prophage regulatory protein	-2.21
vc0572	Protein with putative formate dehydrogenase activity	-2.09
vc1008	sodium-type flagellar protein MotY	3.31
vc1064	Lipoprotein-like protein	2.13
vc1252	Competence damage protein CinA	2.23
vc1309	Alanine N-acetyltransferase	2.00
vc1576	N4-acetylcytidine amidohydrolase	2.83
vc1650	Collagenase	2.00
vc1693	Cytochrome c-type protein TorC	2.09
vc1798	Eha protein	2.17
vc2048	VacJ lipoprotein	-2.28
vc2122 (fliQ)	Flagellar biosynthesis protein	2.03
vc2124 (fliO)	Flagellar protein	2.04
vc2126 (fliM)	Flagellar motor switch protein	2.05
vc2127 (fliL)	Flagellar basal body protein	2.29
vc2128 (fliK)	Flagellar hook-length control protein	3.09
vc2142 (flaA)	Flagellin	-2.05
vc2143	Flagellin	-2.23
vc2144	Flagellin	-2.49
vc2187	Flagellin	-2.02
vc2188	Flagellin	2.68
vc2190 (flgL)	Flagellar hook-associated protein	2.89
vc2191 (flgK)	Flagellar hook-associated protein	3.44
vc2192 (flgJ)	Flagellar rod assembly protein/muramidase	3.46
vc2193 (flgI)	Flagellar basal body P-ring biosynthesis protein	2.90
vc2194 (flgH)	Flagellar basal body L-ring protein	3.89
vc2195 (flgG)	Flagellar basal body rod protein	3.88
vc2196 (flgF)	Flagellar basal body rod protein	3.69
vc2197 (flgE)	Flagellar hook protein	-2.13
vc2198 (flgD)	Flagellar basal body rod modification protein	-2.02
vc2200 (flgB)	Flagellar basal-body rod protein	2.18
vc2206 (flgP)	Outer membrane protein	2.49
vc2207 (flgO)	Flagellar H-ring protein	2.47
vc2208 (flgT)	Flagellar H-ring protein	2.22

Table continues on the next page

Table continued from the previous page

Gene / regulatory element	Description§	Fold-change#
vc2211	Ferric vibriobactin receptor	-2.07
vc2641	Argininosuccinate lyase	4.05
vc2750	GGDEF family protein	2.25
vcA0027	Chtinase	3.44
vcA0178	FrnE protein	2.18
vcA0218	Thermolabile hemolysin	2.39
vcA0219	Hemolysin	2.74
vcA0260	Pseudogene	-2.28
vcA0282	IS5 transposase	-3.93
vcA0289	50S ribosomal protein L35	-2.02
vcA0602	Iron(III) transport system ATP-binding protein	2.10
vcA0984	L-lactate dehydrogenase	2.00
vc0121	Hypothetical protein	3.67
vc0205	Hypothetical protein	16.15
vc0226	Hypothetical protein	-2.25
vc0279	Hypothetical protein	-2.20
vc0427	Hypothetical protein	2.24
vc0507	Hypothetical protein	-3.13
vc0511	Hypothetical protein	3.15
vc0519	Uncharacterized protein	-2.04
vc0555	Hypothetical protein	2.00
vc0598	Hypothetical protein	3.52
vc0599	Hypothetical protein	2.68
vc0686	Hypothetical protein	2.01
vc0712	Hypothetical protein	-3.99
vc0733	Hypothetical protein	2.02
vc0738	Hypothetical protein	3.64
vc0740	Uncharacterized protein	5.66
vc0782	Hypothetical protein	-2.30
vc0816	Hypothetical protein	3.85
vc0874	Hypothetical protein	2.11
vc0915	Hypothetical protein	2.96
vc1032	Hypothetical protein	2.99
vc1035	Hypothetical protein	2.46
vc1090	Hypothetical protein	3.71
vc1262	Hypothetical protein	2.64
vc1317	Uncharacterized protein	2.12
vc1381	Hypothetical protein	-2.00
vc1383	Hypothetical protein	-2.36
vc1384	Hypothetical protein	2.43
vc1385	Hypothetical protein	2.17
vc1404	Hypothetical protein	6.11
vc1514	Hypothetical protein	-2.08
vc1530	Hypothetical protein	2.16
vc1569	Uncharacterized protein	2.63
vc1575	Hypothetical protein	2.10
vc1577	Hypothetical protein	2.69
vc1578	Hypothetical protein	2.70
vc1582	Uncharacterized protein	2.07
vc1648	Hypothetical protein	-2.14
vc1654	Hypothetical protein	2.70
vc1657	Hypothetical protein	2.52

Table continues on the next page

Table continued from the previous page

Gene / regulatory element	Description§	Fold-change#
vc1691	Hypothetical protein	4.09
vc1699	Hypothetical protein	2.14
vc1728	Hypothetical protein	3.28
vc1733	Hypothetical protein	4.43
vc1742	Hypothetical protein	2.07
vc1743	Hypothetical protein	2.17
vc1744	Hypothetical protein	2.22
vc1747	Hypothetical protein	2.97
vc1791	Hypothetical protein	2.07
vc1792	Hypothetical protein	2.56
vc1793	Hypothetical protein	2.26
vc1795	Hypothetical protein	2.16
vc1797	Hypothetical protein	2.78
vc1799	Hypothetical protein	2.54
vc1800	Hypothetical protein	2.66
vc1801	Hypothetical protein	3.99
vc1802	Hypothetical protein	4.96
vc1818	Hypothetical protein	-2.80
vc1891	Hypothetical protein	-2.29
vc1946	Hypothetical protein	-2.84
vc1954	Hypothetical protein	-3.82
vc1982	Hypothetical protein	3.01
vc2079	Hypothetical protein	-3.44
vc2114	Hypothetical protein	2.52
vc2151	Hypothetical protein	3.29
vc2158	Hypothetical protein	2.37
vc2189	Uncharacterized protein	3.17
vc2218	Hypothetical protein	-2.69
vc2243	Hypothetical protein	4.08
vc2263	Hypothetical protein	2.72
vc2306	Hypothetical protein	3.25
vc2313	Hypothetical protein	2.57
vc2315	Hypothetical protein	2.66
vc2601	Hypothetical protein	-2.11
vc2639	Hypothetical protein	3.45
vcA0028	Hypothetical protein	-2.19
vcA0062	Hypothetical protein	2.24
vcA0138	Hypothetical protein	5.48
vcA0149	Hypothetical protein	2.21
vcA0195	Hypothetical protein	2.00
vcA0315	Hypothetical protein	-13.38
vcA0321	Hypothetical protein	-7.14
vcA0376	Hypothetical protein	-2.11
vcA0404	Hypothetical protein	-2.10
vcA0448	Hypothetical protein	-4.25
vcA0456	Hypothetical protein	4.74
vcA0466	Hypothetical protein	3.60
vcA0467	Hypothetical protein	13.07
vcA0515	Hypothetical protein	2.11
vcA0553	Hypothetical protein	-9.87
vcA0613	Hypothetical protein	2.13

Table continues on the next page

Table continued from the previous page

Gene / regulatory element	Description[§]	Fold-change[#]
vcA0651	Hypothetical protein	2.70
vcA0733	Uncharacterized protein	5.24
vcA0734	Hypothetical protein	2.20
vcA0750	Hypothetical protein	2.49
vcA0787	Hypothetical protein	3.36
vcA0810	Hypothetical protein	-3.64
vcA0816	Hypothetical protein	2.29
vcA0868	Pseudogene	2.03
vcA0934	Hypothetical protein	-2.96
vcA0951	Hypothetical protein	2.11
vcA1007	Hypothetical protein	2.23
vcA1009	Uncharacterized protein	-2.63
vcA1016	Hypothetical protein	2.21
vcA1064	Hypothetical protein	3.30

Chapter 6

Materials and methods

6.1 General equipment, consumables, commercial kits, and chemicals

Table 6.1: General equipment and instruments.

Equipment and instruments	Manufacturer	Model
Biosafety cabinet	Telstar	Bio II
Cell lysis instrument	Omni international	Bead ruptor 4
Centrifuge	VWR International	Microstar 17R
Centrifuge	VWR International	Megastar 3.0R
Chemi imaging system	Vilber	Fusion FX
Electroporator	BioRad	MicroPulser
Electrophoresis system	Peqlab	PerfectBlue™ system
Gel documentation system	Biometra	BioDoc Analyze
Horizontal blotting equipment	Peqlab	PerfectBlue™ system
HPLC system	ThermoScientific	Dionex U3000 nano
Hybridization oven	UVP	HB-1000 hybridizer
Imaging screen eraser	Amersham	Multipurpose standard
Incubator	VWR International	IncuLINE
Magnetic stirrer	Heidolph Instruments	MR3001
Mass spectrometer	Sciex	QTRAP 6500+
Microcentrifuge	Eppendorf	5424, 5415R
Nanodrop	ThermoScientific	Nano2000C
pH meter	VWR International	OrionStar
Phosphoimager	Amersham Biosciences	Typhoon 9400
Phosphor screen	Cytiva	BAS storage

Pipette boy	Eppendorf	Easypet3
Pipettes	Gilson and Eppendorf	Researchplus
Plate reader	Tecan	Spark10M
Power supply unit	VWR International and Biometra	VWR250 and P25
Radiation counter	Graetz Strahlungsmeßtechnik GmbH	CoMo 170
Rocking platform	VWR International	Model 100
RT-qPCR system	Biorad	CFX96
Shaker incubator	ThermoScientific	MaxQ 6000
Sonication system	MRC laboratory instruments	Ultrasonic frequency 25kHz
Spectrophotometer	GE Healthcare	Ultrospec 2100™ prep
Thermal cycler	Peqlab	PeqSTAR
Thermoblock	ThermoScientific	TS100
Ultracentrifuge	Sorvell	WX series
Ultracentrifuge rotor	Sorvell	WX series
UV-crosslinker	VWR International	230V
Vacuum pump	IBS Integra Biosciences	VacuSafe
Vertical blotting equipment	Peqlab	PerfectBlue™ system
Vortexer	Scientific industries	Vortex genie 2
Weighing balance	Sartorius	CPA4202S

Table 6.2: Consumables.

Consumables	Manufacturer	Type
96-well plates	Greiner	Flat transparent, black
Blotting paper	VWR International	Grade 703
Centrifuge tubes	Eppendorf, Sarstedt	Safe-lock 1.5ml, 2 ml
Coverslips	Hatrenstein	#1, 150 µm
Electroporation cuvettes	Cell projects	2mm gap width
Falcon tubes	Corning	Falcon® 15ml, 50 ml
Glass beads	Carl-Roth	0.1mm
Gloves	Sempermed	Sempercare
Inoculation loops	VWR International	10µl
Microscope glass slides	Hartenstein	1mm thickness
Micro cuvettes	Sarstedt	Semi micro, 1.5ml
Microspin clean-up columns	GE Healthcare	G-25, 5-50
Multi-channel pipette	Eppendorf	10µl, 100 µl
Multi dispenser	Eppendorf	Repeater Plus
Multi dispenser combitips	Eppendorf	5ml, 10ml
Nylon blotting membrane	GE Healthcare	Hybond™ XL
Nylon blotting membrane	Sigma	Nylon66, 0.45µm
PCR tubes	Corning thermowell	0.2ml
Petri dishes	Sarstedt	92 x 16 mm transparent,

Phase-lock tubes	QuantaBio	Phase lock gel ™
Pipette tips	Starlabs international	TipOne®
PVDF blotting membrane	Carl Roth	Roti PVDF 0.45µm
Serological pipettes	Greiner	5-, 10-, 25-, 50-ml plastic
Sterile filters	Sarstedt	0.22 µm
Syringe	BD Plastipak	5ml, 10ml, 50ml

Table 6.3: Commercially available kits

Kit	Manufacturer
DNA clean-up and concentration	Zymo research
HiYield® mini plasmid DNA extraction kit	Süd-laborbedarf
Luna® universal one-step RT-qPCR	NEB
MEGAscript T7 kit	Ambion
NEBNext® Small RNA library prep set for Illumina®	NEB
NEBNext® Ultra II Directional library prep kit for Illumina®	NEB
Pierce™ BCA protein assay kit	ThermoFisher Scientific
QIAquick PCR purification kit	Qiagen
RiboZero rRNA depletion kit for Gram-negative bacteria	ThermoFisher Scientific
SV40 total RNA isolation kit	Promega

Table 6.4: Antibiotics and chemicals.

Chemical	Manufacturer	Chemical	Manufacturer
1kb plus DNA ladder	NEB	Ni-NTA resin	ThermoFisher Scientific
50bp DNA ladder	NEB	Oxaloacetate	Sigma-Aldrich
Acetyl CoA	Sigma-Aldrich	Paraformladehyde	Carl-Roth
Agar-agar	Roth	PBS	Sigma-Aldrich
Ampicilin sodium salt	Carl-Roth	P:C:I (25:24:1)	Carl-Roth
Beta-mercaptoethanol	Sigma-Aldrich	p-coumaric acid	Sigma-Aldrich
Boric acid	Carl-Roth	PEG	Sigma-Aldrich
Bromophenol blue	Carl-Roth	Peptone	Sigma-Aldrich
BSA	Sigma-Aldrich	Polymyxin-B	Carl-Roth
Calcium chloride	Carl-Roth	Potassium acetate	Carl-Roth
Casaminoacids	VWR International	Potassium chloride	Carl-Roth
Chitin flakes	Sigma-Aldrich	Pre-stained protein ladder	ThermoFisher Scientific
Chitin resin	NEB	Proteinase K	Carl-Roth
Chloramphenicol	Carl-Roth	RedSafe	iNtRON
Chloroform	Carl-Roth	Rifampicin	Carl-Roth

D(+)-glucose	Sigma-Aldrich	Roti-aqua phenol	Carl-Roth
Di-sodium hydrogen phosphate	Carl-Roth	Roti-hybri-quick	Carl-Roth
DMSO	Carl-Roth	Rotiphorese gel 30 (37.1:1)	Carl-Roth
dNTPS	ThermoFisher Scientific	Rotiphorese gel 40 (19:1)	Carl-Roth
EDTA	Carl-Roth	SDS pellets	Sigma-Aldrich
Ethanol	Carl-Roth	Sepharose G	Carl-Roth
EXTRAzol	Blirt	Silver nitrate	Carl-Roth
Glycerol	Carl-Roth	Spermidine	ThermoFisher Scientific
GlycoBlue	Invitrogen	Sodium acetate	Sigma-Aldrich
HEPES	Carl-Roth	Sodium bicarbonate	Sigma-Aldrich
Hydrochloric acid	Carl-Roth	Streptomycin sulfate	Sigma-Aldrich
Hydrogen peroxide	Carl-Roth	Sodium dihydrogen phosphate	Sigma-Aldrich
Kanamycin sulfate	Carl-Roth	Trichloroacetic acid	Carl-Roth
L(+) arabinose	Sigma-Aldrich	Tris	Sigma-Aldrich
LB (Lennox)	Sigma-Aldrich	Triton X-100	Carl-Roth
LB-agar	Sigma-Aldrich	Tween 20	Carl-Roth
Luminol	Sigma-Aldrich	Urea	Sigma-Aldrich
M9 minimal salts	Sigma-Aldrich	Yeast extract	Sigma-Aldrich
Magnesium sulfate	Sigma-Aldrich	αMG	Sigma-Aldrich
Methanol	Carl-Roth	γ^{32} ATP	Hartmann Analytic
Milk powder (blotting grade)	Sigma-Aldrich	γ^{32} UTP	Hartmann Analytic

6.2 Enzymes, antibodies, buffers and solutions

Table 6.5: Enzymes.

Enzyme	Manufacturer
Benzonase nuclease	ThermoFisher Scientific
Calf intestinal phosphate (CIP)	NEB
GoTaq polymerase	Promega
Lysozyme	Sigma-Aldrich
Q5 Hi-Fidelity polymerase	NEB
Restriction enzymes	Restriction enzymes
RNase inhibitor	ThermoFisher Scientific
Shrimp alkaline phosphatase (SAP)	Shrimp alkaline phosphatase (SAP)
SuperaseIN	ThermoFisher Scientific
T4 ligase	NEB
T4 polynucleotide kinase (PNK)	NEB
Turbo DNase	ThermoFisher Scientific

Table 6.6: Antibodies.

Antibodies	Working dilution	Manufacturer + ID
Anti-6xHis (rabbit)	1:5000	AbCam #200537
Anti-cholera toxin (rabbit)	1:2000	AbCam #123129
Anti-Flag (mouse)	1:1000	Sigma-Aldrich #F1804
Anti-HA (mouse)	1:5000	AbCam #18181
Anti-Llama (goat)	1:1000	ChromoTek #etb-250
Anti-mouse, HRP-conjugated	1:10,000	ThermoFisher Scientific #34130
Anti-Rabbit, HRP-conjugated	1:10,000	ThermoFisher Scientific #16104
Anti-RNAP (rabbit)	1:10,000	BioLegend #WP003
Anti-Spot (llama)	1:1000	AbCam #112786

Table 6.7: Buffers and solutions.

Buffer / solution	Composition
10x PBS	2g KCl, 2.4g KH_2PO_4, 80g NaCl, 14.4g Na_2HPO_4, pH ad. to 7.4 with HCl; H_2O ad. 1l
10x SDS running buffer	30.275 Tris, 144g glycine, 10g SDS; H_2O ad. 1l
10x TB	30g Tris, 144g glycine; H_2O ad. 1l
10x TBE	108g Tris, 5g boric acid, 20 mM EDTA pH 8.0, H_2O ad. 1l
10x TBS	24.11g tris, 72.6g NaCl, pH ad. to 7.4 with HCl; H_2O ad. 1l
1x TBS-T	1xTBS, 0.1% (w/v) Tween-20
1x transfer buffer (SDS-PAGE)	100ml 10xTB, 200ml methanol; H_2O ad. 1l
20x SSC buffer	3M NaCl, 0.3M sodium citrate, pH ad. to 7.0 with HCl
2x RNA loading buffer	0.025% (w/v) each bromophenol blue and xylene cyanol, 18µM EDTA (pH 8.0), 0.1% (w/v) SDS, 95% formamide
30:1 precipitation mix	3 vol. ethanol, 0.1 vol 3M sodium acetate, pH 5.2 or 6.5
50x TAE	242g Tris, 51.7ml acetic acid, 10mM EDTA pH 8.0; H_2O ad. 1l
5x M9 salts	85.7g Na_2HPO_4, 15g KH_2PO_4, 2.5g NaCl, 5g NH_4Cl, H_2O ad. 1l
5x protein loading buffer	15g SDS, 46.9ml 1M Tris-HCl (pH 6.8), 75ml glycerol, 11.56g DTT, 0.075g bromophenol blue; H_2O ad. 150ml
Agarose gel solution	1%, 1.5% or 2% agarose in 1xTAE
Chemiluminescence detection solution	1.8ml chemiluminescence solution A, 200µl chemiluminescence solution B, 3µl H_2O_2 (from a 30% w/v stock)
Chemiluminescence solution A	0.1M Tris-HCl (pH 8.6), 0.025% (w/v) luminol
Chemiluminescence solution B	0.1% (w/v) para- coumaric acid in DMSO
CIP buffer	100 mM NaCl, 50 mM Tris–HCl pH 7.4, 10 mM $MgCl_2$
High salt buffer	50 mM NaH_2PO_4, 1 M NaCl, 0.05% Tween, pH 8.0
Lower buffer (SDS-PAGE)	1.5 M Tris-HCl, pH 8.8, 0.4 % (w/v) SDS
NP-T buffer	50 mM NaH_2PO_4, 300 mM NaCl, 0.05% Tween, pH 8.0
PAA solution (Northern blotting)	100 ml 10xTBE, 420 g urea (7M), 100 or 150 ml Rotiphorese gel 40 (19:1) for 4 or 6% gels; H_2O ad 1 l
PK buffer	50 mM Tris pH 7.4, 75 mM NaCl, 6 mM EDTA, 1% SDS, 10 U of

	SUPERaseIN, 1 mg/ml proteinase K
PNK buffer	50 mM Tris–HCl pH 7.4, 10 mM $MgCl_2$, 0.1 mM spermidine
Stop-mix	95% ethanol, 5% acidic phenol
TE buffer	100mM Tris (pH 8.0), 10mM EDTA
Upper buffer (SDS-PAGE)	0.5 M Tris-HCl, pH 6.8, 0.4 % (w/v) SDS

6.3 Media supplements

Table 6.8: Antibiotics and media supplements.

Antibiotic / supplement	Solvent	Stock	Working dilution
Ampicillin	H_2O	100mg/ml	1:1000
Chloramphenicol	Ethanol (absolute)	20mg/ml	1:1000
Kanamycin	H_2O	50mg/ml	1:1000
Polymyxin-B	H_2O	6.25mg/ml	1:1000
Rifampicin	DMSO	50mg/ml	1:1000
Streptomycin	H_2O	5000mg/ml	1:1000
Cas-amino acids	H_2O	10% (w/v)	1:50
L-arabinose	H_2O	20% (w/v)	1:1000
D(+)glucose	H_2O	20% (w/v)	1:50
Glycerol	H_2O	20% (w/v)	1:50

6.4 Software

Table 6.9: Software.

Software (version)	Purpose	Developer / Reference
CLC Main Workbench v20.0.3	*In silico* design for molecular cloning	Qiagen
BIO-1D	Western blot quantification	Vilber Lourmat
CLC Genomics Workbench v12.0.3	RNA-seq analysis tool	Qiagen
CorelDRAW X6	Vector Graphics Editor	Corel Corporation
IGV v2.6.3	Genome browser	(Robinson *et al*, 2011)
ImageJ v1.8.0	Image processing	(Rueden *et al*, 2017)
MS Office 365	Data processing and spreadsheet compiler	Microsoft Corporation
Prism v9.1.0	Statistical analysis and data reconfiguration	GraphPad Software, Inc.
PyMOL v2.0	Protein structure visualization	Schrödinger, LLC.
VARNA 3-93	RNA secondary structure visualization	(Darty *et al*, 2009)
Zotero	Reference manager	Corporation for Digital Scholarship

6.5 Bacterial strains, plasmids and oligonucleotides

Table 6.10: Bacterial strains

Strain	Relevant markers/ genotype	Reference/ source
V. cholerae		
KPS-0014	C6706 wild-type	(Thelin & Taylor, 1996)
KPS-0053	C6706 Δ*hapR*	(Svenningsen et al, 2009)
KPVC-10269	C6706 *hapR*:: *hapR* -3xFlag	This study
KPVC-10270	C6706 *luxO*:: *luxO* -3xFlag	This study
KPVC-10983	C6706 *mbrA*:: *mbrA* -3xFlag	This study
KPVC-10141	C6706 *rne*-3071	(Hoyos et al, 2020)
KPVC-10609	C6706 Δ*vcdRP*	This study
KPVC-10985	C6706 Δ*crp*	This study
KPVC-11023	C6706 Δ*cyaA*	This study
KPVC-12323	C6706 Δ*mbrA::kan*	This study
KPVC-13609	C6706 *gltA*::*gltA*-HA	This study
KPVC-13611	C6706 *gltA*::*gltA*-6xHis	This study
KPVC-13686	C6706 Δ*gltA*	This study
KPVC-13724	C6706 *gltA F383A*	This study
KPVC-13820	C6706 *vcdP*:: *vcdP*-SPA	This study
V. natriegens		
KPVC-14140	ATCC 14048 wild-type	(Baumann et al, 1980)
B. subtilis		
KPVC-14141	168 wild-type	(Burkholder & Giles, 1947)
E. coli		
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)
BL21(DE3)	F- *ompT hsdS*$_B$ (r_B-, m_B-) *gal dcm* (DE3)	Novagen
ER2566	*huA2 lacZ::T7 gene1 [Ion] ompT gal sulA11 R(mcr-73::miniTn10-TetS)2 [dcm] R(zgb-210::Tn10--TetS) endA1 Δ(mcrC-mrr) 114::IS10*	New England Biolabs

Table 6.11: Plasmids

Plasmid trivial name	Plasmid stock name	Purpose	Origin, marker	Reference
pEVS143	pEVS143	Constitutive over-expression plasmid	p15A, Kan^R	(Dunn *et al*, 2006)
pKAS32	pKAS32	Suicide plasmid for allelic exchange	R6K, Amp^R	(Skorupski & Taylor, 1996)
pXG10-SF	pXG10-SF	Control plasmid	pSC101*, Cm^R	(Corcoran *et al*, 2012)
pCMW-1	pCMW-1	Control plasmid	p15A, Kan^R	(Waters & Bassler, 2006)
pEVS143-*spot42*	pAS001	*spot42* expression plasmid	p15A, Kan^R	This study
pEVS143-*tarB*	pAS002	*tarB* expression plasmid	p15A, Kan^R	This study

pEVS143-*tarA*	pAS003	*tarA* expression plasmid	p15A, Kan^R	This study
pEVS143-*vqmR*	pKP333	*vqmR* expression plasmid	p15A, Kan^R	(Papenfort *et al*, 2015)
pKAS32-*mbrA*::3xFlag	pASp2	*mbrA*::3xFlag allelic replacement	R6K, Amp^R	This study
pEVS143-*mbrA* ΔTM::sfGFP	pASp4	*mbrA* ΔTM::sfGFP expression plasmid	p15A, Kan^R	This study
pBAD-1C-*mbrA*::3xFlag	pASp7	*mbrA*::3xFlag expression plasmid	p15A, Kan^R	This study
pBAD-1C-*mbrA*::sfGFP	pASp10	*mbrA*::sfGFP expression plasmid	p15A, Kan^R	This study
pBAD-1C-*mbrA* ΔTM2::sfGFP	pASp11	*mbrA* ΔTM2::sfGFP expression plasmid	p15A, Kan^R	This study
pBR-FRT-*kan*-FRT	pASp14	Template for TransFLP procedure	pBR322, Kan^R	(Blokesch, 2012)
pKAS32-*hapR*::3xFlag	pASp17	*hapR*::3xFlag allelic replacement	R6K, Amp^R	This study
pKAS32-*luxO*::3xFlag	pASp18	*luxO*::3xFlag allelic replacement	R6K, Amp^R	This study
pBAD-1K-*mbrA* ΔTM:: 3xFlag	pKV41	*mbrA* ΔTM::3xFlag expression plasmid	p15A, Kan^R	This study
pBAD-1K-*mbrA* ΔTM1:: 3xFlag	pKV42	*mbrA* ΔTM1::3xFlag expression plasmid	p15A, Kan^R	This study
pBAD-1K-*mbrA* ΔTM2:: 3xFlag	pKV43	*mbrA* ΔTM2::3xFlag expression plasmid	p15A, Kan^R	This study
pTYB11-*mbrA*	pKV44	MbrA protein purification	pMB1, Amp^R	This study
pTYB11- ΔTM *mbrA*	pKV45	MbrA mutant protein purification	pMB1, Amp^R	This study
pCMW-1C-p*mbrA*::sfGFP	pKV52	Transcriptional reporter for *mbrA*	p15A, Cm^R	This study
pCMW-1C-ΔR1 p*mbrA* ::sfGFP	pKV56	Transcriptional reporter for *mbrA*	p15A, Cm^R	This study
pCMW-1C-ΔR2 p*mbrA* ::sfGFP	pKV57	Transcriptional reporter for *mbrA*	p15A, Cm^R	This study
pCMW-1C-ΔR3 p*mbrA* ::sfGFP	pKV58	Transcriptional reporter for *mbrA*	p15A, Cm^R	This study
pCMW-1C-ΔCRP box p*mbrA*::sfGFP	pKV59	Transcriptional reporter for *mbrA*	p15A, Cm^R	This study
pCMW-1C-sfGFP	pKV69	Template for transcriptional reporters	p15A, Cm^R	This study
pCMW-1C-p*mbrA*	pKV70	*mbrA* expression plasmid	p15A, Cm^R	This study
pCMW-1C-p*mbrA*::3xFlag	pKV79	Transcriptional reporter for *mbrA*	p15A, Cm^R	This study
pCMW-1C-p*mbrA* M1-Spot	pKV99	Spot®-tagged variant to test motility	p15A, Cm^R	This study
pCMW-1C-p*mbrA* S23-Spot	pKV100	Spot®-tagged variant to test motility	p15A, Cm^R	This study
pCMW-1C-p*mbrA* S61-Spot	pKV101	Spot®-tagged variant to test motility	p15A, Cm^R	This study
pCMW-1C-p*mbrA* N70-Spot	pKV102	Spot®-tagged variant to test motility	p15A, Cm^R	This study
pCMW-1C-p*mbrA* Q86-Spot	pKV103	Spot®-tagged variant to test motility	p15A, Cm^R	This study
pCMW-1C-p*mbrA* A92-Spot	pKV104	Spot®-tagged variant to test motility	p15A, Cm^R	This study
pCMW-1C-p*mbrA* K99-Spot	pKV105	Spot®-tagged variant to test motility	p15A, Cm^R	This study
pCMW-1C-p*mbrA* K103-	pKV106	Spot®-tagged variant to test	p15A, Cm^R	This study

Spot		motility		
pCMW-1C-p*mbrA* S115-Spot	pKV107	Spot®-tagged variant to test motility	p15A, Cm^R	This study
pCMW-1C-p*mbrA* D141-Spot	pKV108	Spot®-tagged variant to test motility	p15A, Cm^R	This study
pCMW-1C-p*mbrA*-Spot	pKV109	Spot®-tagged variant to test motility	p15A, Cm^R	This study
pEVS143-*vcdP*-SPA	pKV114	*vcdP-SPA* expression plasmid	p15A, Kan^R	This study
pKAS32-*gltA*::3xFlag	pKV150	*gltA*::3xFlag allelic replacement	R6K, Amp^R	This study
pKAS32-*gltA*::HA	pKV152	*gltA*::HA allelic replacement	R6K, Amp^R	This study
pKAS32-*gltA*::6xHis	pKV153	*gltA*::6xHis allelic replacement	R6K, Amp^R	This study
pKAS32-Δ*gltA*	pKV154	Suicide plasmid for *gltA* knockout	R6K, Amp^R	This study
pKAS32-*gltA F383A*	pKV155	Suicide plasmid for *gltA F383A* allelic replacement	R6K, Amp^R	This study
pET15b-*gltA F383A*	pKV156	Over-expression construct for GltA F383A purification	pBR322, Amp^R	This study
pKAS32-*vcdP*::SPA	pKV157	*vcdP*::SPA allelic replacement	R6K, Amp^R	This study
pEVS143-*vcdP**	pKV159	*vcdP** expression plasmid	p15A, Kan^R	This study
pCMW1C-ΔCRP box-pVcdRP ::mKate2	pKV164	Transcriptional reporter for *vcdRP*	p15A, Cm^R	This study
pBSmul2	pKV168	*B. subtilis* expression plasmid	pUB110, Amp^R	*(Brockmeier et al*, 2006)
pBSmul2-*vcdP*	pKV169	*B. subtilis* expression plasmid	pUB110, Amp^R	This study
pEVS143-*gltA*	pKV175	*gltA* expression plasmid	p15A, Kan^R	This study
pBAD-1K-*gltA*	pKV176	*gltA* expression plasmid	p15A, Kan^R	This study
pKAS32-rne3071	pMD003	Suicide plasmid for temperature sensitive *rne* allelle	R6K, Amp^R	(Hoyos *et al*, 2020)
pBAD-1K-ctrl	pMD004	Control plasmid	p15A, Kan^R	Lab collection
pKAS32-Δ*vcdRP*	pMD054	Suicide plasmid for *vcdRP* knockout	R6K, Amp^R	This study
pEVS143-*vcdR*	pMD055	*vcdR* expression plasmid	p15A, Kan^R	This study
pEVS143-*vcdRP 87nt*	pMD062	*vcdRP* expression plasmid truncated at 87th nucleotide (from the 3′ end)	p15A, Kan^R	This study
pEVS143-*vcdRP 71nt*	pMD063	*vcdRP* expression plasmid truncated at 71st nucleotide (from the 3′ end)	p15A, Kan^R	This study
pCMW-1C-pVcdRP::mKate2	pMD064	Transcriptional reporter for *vcdRP*	p15A, Cm^R	This study
pEVS143-*vcdP*::3xFlag	pMD065	*vcdP*::3xFlag expression plasmid	p15A, Kan^R	This study
pBAD-1K-*vcdRP*	pMD072	*vcdRP* expression plasmid	p15A, Kan^R	This study
pBAD-1K-*vcdR*	pMD077	*vcdR* expression plasmid	p15A, Kan^R	This study
pEVS143 (protein)	pMD080	Modified expression plasmid with artificial 5′UTR, MCS and T1 terminator	p15A, Kan^R	This study
pEVS143-*vcdRP* Δ4C	pMD083	*vcdRP* Δ4C expression plasmid	p15A, Kan^R	This study
pBAD-1K-*vcdP*	pMD087	*vcdP* expression plasmid	p15A, Kan^R	This study
pEVS143-*oppZ*	pMD090	*oppZ* expression plasmid	p15A, Kan^R	(Hoyos *et al*, 2020)
pEVS143-*tfoR*	pMD104	*tfoR* expression plasmid	p15A, Kan^R	This study
pEVS143-*vcr078*	pMD105	*vcr078* expression plasmid	p15A, Kan^R	This study

pEVS143-*vcdP*	pMD111	*vcdP* expression plasmid	p15A, Kan^R	This study
pXG10-*ptsG*	pMD161	Translational reporter for *ptsG*	pSC101*, Cm^R	This study
pXG10-*nagE*	pMD162	Translational reporter for *nagE*	pSC101*, Cm^R	This study
pXG10-*ptsHI*	pMD164	Translational reporter for *ptsHI*	pSC101*, Cm^R	This study
pEVS143-*vcdRP M2*	pMD389	*vcdRP M2* expression plasmid	p15A, Kan^R	This study
pXG10-*treB M2**	pMD401	Translational reporter for *treB M2**	pSC101*, Cm^R	This study
pXG10-*ptsG M2**	pMD402	Translational reporter for *ptsG M2**	pSC101*, Cm^R	This study
pXG10-*nagE M2**	pMD403	Translational reporter for *nagE M2**	pSC101*, Cm^R	This study
pXG10-*ptsHI M2**	pMD405	Translational reporter for *ptsHI M2**	pSC101*, Cm^R	This study
pET15b-*gltA*	pMD408	Over-expression construct for GltA purification	pBR322, Amp^R	This study
pEVS143-*vcr002*	pNP001	*vcr002* expression plasmid	p15A, Kan^R	This study
pEVS143-*micV*	pNP002	*micV* expression plasmid	p15A, Kan^R	(Peschek *et al*, 2019)
pEVS143-*vcr036*	pNP003	*vcr036* expression plasmid	p15A, Kan^R	This study
pEVS143-*vcr043*	pNP004	*vcr043* expression plasmid	p15A, Kan^R	This study
pEVS143-*vadR*	pNP005	*vadR* expression plasmid	p15A, Kan^R	(Peschek *et al*, 2020)
pXG10-*treB*	pNP0058	Translational reporter for *treB*	pSC101*, Cm^R	This study
pEVS143-*vcr079*	pNP006	*vcr079* expression plasmid	p15A, Kan^R	This study
pEVS143-*vcr065*	pNP007	*vcr065* expression plasmid	p15A, Kan^R	This study
pEVS143-*vcr034*	pNP008	*vcr034* expression plasmid	p15A, Kan^R	This study
pEVS143-*vcdRP*	pNP009	*vcdRP* expression plasmid	p15A, Kan^R	This study
pEVS143-*vcr092*	pNP010	*vcr092* expression plasmid	p15A, Kan^R	This study
pEVS143-*vcr098*	pNP011	*vcr098* expression plasmid	p15A, Kan^R	This study
pEVS143-*vcr103*	pNP012	*vcr103* expression plasmid	p15A, Kan^R	This study
pEVS143-*carZ*	pNP015	*carZ* expression plasmid	p15A, Kan^R	(Hoyos et al, 2020)
pEVS143-*vrrA*	pRH001	*vrrA* expression plasmid	p15A, Kan^R	(Peschek *et al*, 2019)
pEVS143-*qrr2*	pRH002	*qrr2* expression plasmid	p15A, Kan^R	This study
pEVS143-*qrr4*	pRH003	*qrr4* expression plasmid	p15A, Kan^R	This study
pEVS143-*gcvB*	pRH006	*gcvB* expression plasmid	p15A, Kan^R	This study
pEVS143-*csrD*	pRH009	*csrD* expression plasmid	p15A, Kan^R	This study
pEVS143-*flaX*	pRH010	*flaX* expression plasmid	p15A, Kan^R	This study
pKAS32-Δ*crp*	pRH023	Suicide plasmid for *crp* knockout	R6K, Amp^R	This study
pKAS32-Δ*cyaA*	pRH024	Suicide plasmid for *cyaA* knockout	R6K, Amp^R	This study
pEVS143-*vcr080*	pSG002	*vcr080* expression plasmid	p15A, Kan^R	This study
pEVS143-*vcdRP 256nt*	pSG006	*vcdRP* expression plasmid truncated at 256th nucleotide (from the 3′ end)	p15A, Kan^R	This study
pEVS143-*vcdRP 156nt*	pSG008	*vcdRP* expression plasmid truncated at 156th nucleotide (from the 3′ end)	p15A, Kan^R	This study

pEVS143-*vcr006*	pYH002	*vcr006* expression plasmid	p15A, Kan^R	This study
pCMW-1C-mKate2	pYH010	Template for transcriptional reporters	p15A, Cm^R	Lab collection

Table 6.12: Oligonucleotides.
Sequences are given in 5′ →3′ direction; 'P-' denotes a 5′ monophosphate

ID	Sequence	Description
KPO-0009	CTACGGCGTTTCACTTCTGAGTTC	*E.coli* 5S rRNA oligoprobe
KPO-0092	CCACACATTATACGAGCCGA	pEVS143-derivates
KPO-0196	GGAGAAACAGTAGAGAGTTGCG	pBAD1K-derivates
KPO-0243	TTCGTTTCACTTCTGAGTTCGG	*V.ch.* 5S rRNA oligoprobe
KPO-0267	TAATAGGCCTAGGATGCATATG	pKAS32-derivates
KPO-0268	CGTTAACAACCGGTACCTCTA	pKAS32-derivates
KPO-0456	P-CAGAGCATGAGTTGCATGAC	pKP333
KPO-0465	GTTTTTGGATCCAGCTTATCTTGCCTATTCGG	pKP333
KPO-0570	GGCTGAAAGCGATAATGATCTTG	*vc1449* qRT-PCR
KPO-0571	CGGCTTCCATTCTAGGATCTTC	*vc1449* qRT-PCR
KPO-0999	P-ACCACTGCTTTTTCTTAGAAGAC	pNP02
KPO-1000	GTTTTTTCTAGAGGATTAGAACCCGAATTAAACT	pNP02
KPO-1001	P-TCACAGAACCGCTGTGACCA	pNP10
KPO-1002	GTTTTTTCTAGATTGACTACTTCATTCGCCAC	pNP10
KPO-1003	P-GCAAACACATTGGTAAGATATTAG	pNP01
KPO-1004	GTTTTTTCTAGATATAACCTGTTCAGAATGTGCT	pNP01
KPO-1005	P-GTCATCTCGTTAGTCATTACGA	pNP04
KPO-1006	GTTTTTTCTAGACACTGACAAACCGGTGTTGG	pNP04
KPO-1007	P-GTAAGGTTAGTGAGAACATTTCT	pNP11
KPO-1008	GTTTTTTCTAGAAGTTTCAAATTTCGTGGACAGC	pNP11
KPO-1009	P-ACTTACTTGGATAAATATGCATTG	pNP08
KPO-1010	GTTTTTTCTAGAGTATTGTTTGTCTGTCATAAAGTT	pNP08
KPO-1011	P-ACCTGTCGCTAATTTCAGTATC	pNP07
KPO-1012	GTTTTTTCTAGAAGCCTAACCCTATCTTTCGT	pNP07
KPO-1013	P-TATTACAACAAGAGAGGCTCAA	pNP015
KPO-1014	GTTTTTTCTAGACAGACGCTACATCAAACTGAA	pNP015
KPO-1015	P-AATAGACAACCTTTTGTCCTATC	pNP05
KPO-1016	GTTTTTTCTAGAATAGAAAGCACTGAGTCAGGA	pNP05
KPO-1021	P-GTTTGAACCCCGGCGGCT	pNP06
KPO-1022	GTTTTTTCTAGAAAACCGACTCCTTGCAAGAA	pNP06
KPO-1023	GTTTTTTCTAGAGGATCCGGTGATTGATTGAG	pEVS143 derivates
KPO-1024	P-ACCCAAAGGGTAGAGCAAAC	pNP03
KPO-1025	GTTTTTTCTAGAGAAAACGAAGTAATCTTCACCTT	pNP03
KPO-1026	P-TGAATAATCAAAGACGAGGCTC	pNP12
KPO-1027	GTTTTTTCTAGAGAACAGCCAGTTAACTTGAGA	pNP12
KPO-1070	P-ACATGAGCGGTTACCTCAT	pAS03
KPO-1071	GTTTTTTCTAGATTATAGAGATAGGTTTGTGTGTG	pAS03
KPO-1072	P-GGTTTAGCACTCCCCCCTA	pAS02
KPO-1073	GTTTTTTCTAGAGTTTTTGTCTTTAGGAAAATAAAG	pAS02
KPO-1076	P-GCGTAGGGTACAGAGGTAA	pAS01
KPO-1077	GTTTTTTCTAGAAGTGCCAACGTGGAATAGC	pAS01
KPO-1082	P-GTGATTGACAGAGCTTTGAGA	pRH01
KPO-1083	GTTTTTTCTAGATCGCCAATGAACCGACTTG	pRH01
KPO-1084	P-GCAACGGCGGCCTGAACGG	pRH06
KPO-1085	GTTTTTTCTAGAAGCTCAGTATTTACTGGTTGGA	pRH06
KPO-1090	P-TGACCCTTCTAAGCCGAGG	pRH03
KPO-1091	GTTTTTTCTAGACCACGAAAGCCAAGATGCT	pRH03
KPO-1092	P-ACAAAGTATCACAAAAATCAGGG	pRH02
KPO-1093	GTTTTTTCTAGAAAAGCAGTGAAAATAGCGGG	pRH02
KPO-1219	P-AGCTTCGCTAGCGAAGAG	pNP09
KPO-1220	GTTTTTTCTAGAGAATGTTGCGATCAAGTTCG	pNP09
KPO-1226	TCGTATAATGTGTGGGTAAGGTTAGTGAGAACATTTCT	pRH05
KPO-1227	ACCGGATCCTCTAGAAGTTTCAAATTTCGTGGACAGC	pRH05
KPO-1282	GATCGCATTACGTAGACTACC	pMD054
KPO-1283	GTTCGAATAGGGAAGATTTTTTGAGACCACCAAGAAATTAAATTACAACAC	pMD054
KPO-1284	CTCAAAAAATCTTCCCTATTCGAAC	pMD054

KPO-1285	GTTAAACCAGTACGGCCACC	pMD054
KPO-1286	GTTTTTGGTACCGAATTCGACTAAGCGCAGATC	pMD054
KPO-1287	GTTTTTCCTAGGGTGGTGTTTGAAGTACCGAGC	pMD054
KPO-1375	TCGTATAATGTGTGGGTCAGCAGGAAGCGGACAC	pRH09
KPO-1376	ACCGGATCCTCTAGATACGGAAAGATGCCAAGAGA	pRH09
KPO-1383	TCGTATAATGTGTGGTTGCCAAACTCTGCAATCTCG	pRH10
KPO-1384	ACCGGATCCTCTAGAAGTAGGAAAGATAAAGGTGGGG	pRH10
KPO-1397	GATCCGGTGATTGATTGAGC	pBAD1K-derivates
KPO-1440	TAGAGGTACCGGTTGTTAACGCTTCGAGCTTGTATTCTGC	pMD003
KPO-1441	CATATGCATCCTAGGCCTATTACAGAGAAGGGCTCAAACG	pMD003
KPO-1442	GCTGAAAGACACGGTTTCtTCCCTCTCAAAGAAATCGC	pMD003
KPO-1443	GCGATTTCTTTGAGAGGGAaGAAACCGTGTCTTTCAGC	pMD003
KPO-1448	TCGGCTCGTATAATGTGTGGACAGAATGAGTAATCAACCAAAG	pYH02
KPO-1449	CTCAATCAATCACCGGATCCCAATAAAAAAGGACGCGATGC	pYH02
KPO-1484	GCTCAATCAATCACCGGATCAAGGCCCAGTCTTTCGAC	pMD004, pMD111
KPO-1485	CGCAACTCTCTACTGTTTCTCCTTTTTTCTAGATTAAATCAGAACGCAG	pMD004
KPO-1488	TTTTTTCTAGATTAAATCAGAACGCAG	pKV176
KPO-1492	GATAAAACGAAAGGCCCAGTCTTTCGACTGAGCCTTTCG	pMD080
KPO-1524	CAACGGGAATCCTGCTCTG	pMD083
KPO-1525	GCGGCCCTCTCACTTCC	pMD083
KPO-1702	ATGCATGTGCTCAGTATCTCTATC	pXG10SF-derivates
KPO-1703	GCTAGCGGATCCGCTGG	pXG10SF-derivates
KPO-1708	GAGATACTGAGCACATGCATAATTGATTTGGGACTGTTCCCAA	pNP58
KPO-1709	GAGCCAGCGGATCCGCTAGCCAATTCGATAAGACGCGTCAC	pNP58
KPO-1860	TCGGCTCGTATAATGTGTGGGCAAGTCAGTGGTGTTGG	pSG02
KPO-1861	CTCAATCAATCACCGGATCCGTACTGTCAATATCGACCAC	pSG02
KPO-1949	TCTAGAGGATCCGGTGATTG	pEVS143-derivates
KPO-2040	GTGAATCATATCGACCAAATTTG	VcdRP oligoprobe
KPO-2050	GTCAGCAGAATATGTGATACAGG	pKV159
KPO-2082	CTCAATCAATCACCGGATCCGAATGTTGCGATCAAGTTCG	pEVS143-derivates
KPO-2083	TCGGCTCGTATAATGTGTGGCTCCATGGAACCGAGAAATC	pSG06
KPO-2085	TCGGCTCGTATAATGTGTGG GTACACCGGTAACTTTGCATAC	pSG08
KPO-2090	GATTTTAATGGTGATAGTTATGAATTAAGGTTTATCGTCTGCAATGTTC	pMD055
KPO-2091	GAACATTGCAGACGATAAACCTTAATTCATAACTATCACCATTAAAATC	pMD055
KPO-2100	CGCAACTCTCTACTGTTTCTCCAGCTTCGCTAGCGAAGAG	pMD072
KPO-2101	GCTCAATCAATCACCGGATCGAATGTTGCGATCAAGTTCG	pMD072
KPO-2109	TCGGCTCGTATAATGTGTGGTTTTGTTCACCCCTAAATTGG	pMD062
KPO-2110	TCGGCTCGTATAATGTGTGGATTGGAATTTATTGACGACCAAATTTG	pMD063
KPO-2111	GTTTTTGCATGC GTAGGCAAATGCATCTTCATGC	pMD064
KPO-2112	GTTTTTGTCGACGAAGCTGTACACAAATATACCAC	pMD064
KPO-2229	CCTGTATCACATATTCTGCTGAC	pKV159
KPO-2243	AGAGGTACCGGTTGTTAACGGGAAGGTGTCTAAGTTAGCAC	pRH023
KPO-2244	TTAAATCTAGCGCATCATACCGTGTCAGTAGTCCAATTTGC	pRH023
KPO-2245	GTATGATGCGCTAGATTTAACGG	pRH023
KPO-2246	TATGCATCCTAGGCCTATTAGATACAACGCTGCTTCCTGC	pRH023
KPO-2249	AGAGGTACCGGTTGTTAACGCAGCATGGTTAAGAGGTGCG	pRH024
KPO-2250	GACGCATAATAAAAAAGAGCCTGATATGACAGATAAAAGCCGC	pRH024
KPO-2251	GCTCTTTTTTATTATGCGTCAGTG	pRH024
KPO-2252	TATGCATCCTAGGCCTATTACTCCGATTCCACGTTAAAGC	pRH024
KPO-2259	TCGGCTCGTATAATGTGTGGTACCCGTTTTTTGGGCTAACAGGAGGAATTAACC	pMD080
KPO-2260	GCCTCTAGATTATGAGACCAGGTCTCACATGGTTAATTCCTCCTGTTAGC	pMD080
KPO-2261	TGGTCTCATAATCTAGAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAG	pMD080
KPO-2322	GTCTCGAACAAGTTTTGTTCATAAATTGGAATTTATTGACGACC	pMD083
KPO-2323	TGAACAAAACTTGTTCGAGAC	pMD083
KPO-2329	ACTGGGCCTTTCGTTTTATCTCTAGAGGATCCGGTGATTGATTGAGC	pMD080
KPO-2330	CTAACAGGAGGAATTAACCATGAACAAGGGCCTGAGTAGCGCTATGTTTTGGAATCAGCA	pMD111
KPO-2356	CCAAAAGAGCGTGGATAAGC	VcdP oligoprobe
KPO-2378	GGTAACCCAGAAACTACCACTG	*recA* qRT-PCR
KPO-2379	CACCACTTCTTCGCCTTCTT	*recA* qRT-PCR
KPO-2410	ATAACCAAAAGAGCGTGGATAAG	pKV114
KPO-2471	CCCAGGTTGTTGTACAGAACTGAAATCTGACAAATCGATGTTATG	pKV44
KPO-2472	CGGATCCCCTTCCTGCAGTTAATCCTGTTCAGCCATTTCTTC	pKV44
KPO-2553	GCTCAATCAATCACCGGATCAGGCGATTGGTCGTGTTG	pMD090
KPO-2565	GCTCAATCAATCACCGGATCATTGAAGTGAGTGATGGTAATAG	pMD104
KPO-2568	TCGGCTCGTATAATGTGTGGATTCAGCCTGTGACGGG	pMD090
KPO-2570	TCGGCTCGTATAATGTGTGGGTTGAAAGGACATCCCTCC	pMD104
KPO-2571	TCGGCTCGTATAATGTGTGGATCAAAGATACAGACTTTGCC	pMD105
KPO-2572	GCTCAATCAATCACCGGATCCTTTTTAGGTTTCTCTGCCACA	pMD105

KPO-2715	GACCTGTTCCTGCTTTGTTTAC	*lamB* qRT-PCR
KPO-2716	CTGCTTTCGCTGTCGATTTC	*lamB* qRT-PCR
KPO-2747	AAGCAGTCAGGTGGTCTTATG	*ctxA* qRT-PCR
KPO-2748	ACAAATCCCGTCTGAGTTCC	*ctxA* qRT-PCR
KPO-2779	GAGATACTGAGCACATGCATCGGAAAATATAATGCAAAAAGTGG	pMD161
KPO-2780	GAGCCAGCGGATCCGCTAGCGATTAAGTTATTAGAATTGCTGGG	pMD161
KPO-2781	GAGATACTGAGCACATGCATGACTAAATTGGGCGACTAAAAAAAG	pMD162
KPO-2782	GAGCCAGCGGATCCGCTAGCCGTAGCGATAGGTAGCATC	pMD162
KPO-2792	GAGATACTGAGCACATGCATGTTACCAAGTTCAGGTGAACG	pMD164
KPO-2794	GAGCCAGCGGATCCGCTAGCTTGAAGAAGTAATGCTTTACCAATAG	pMD164
KPO-3010	ATATCTCCTTAACTAGGCCTGATTTCATATTGAATTACTACGCATC	pKV52
KPO-3011	AAAGACCCTTCATTGTCGACGGAGACAGAGTTTTATGACAAGC	pKV52
KPO-3236	TCTAGAGGCATCAAATAAAACGAAAGGC	pKV175
KPO-3302	AATGATACGGCGACCACCG	CLIP-seq library
KPO-3303	CAAGCAGAAGACGGCATACG	CLIP-seq library
KPO-3435	AATTTCAACAAAACTGGGGGGTCACACCATTCG	pKV56
KPO-3436	GGTGTGACCCCCCAGTTTTGTTGAAATTTTATTCGAAAAACATTCC	pKV56
KPO-3437	CATTTGAGGGTCACACCATTCGATTGGAAAAATATC	pKV57
KPO-3438	CGAATGGTGTGACCCTCAAATGTTGAAATTTTATTCGAAAAACATTC	pKV57
KPO-3439	GGGGGTGTGTCCATTCGATTGGAAAAATATCAAG	pKV58
KPO-3440	CAATCGAATGGACACACCCCCCTCAAATGTTGAAATTTTATT	pKV58
KPO-3726	CAGCCTAATCCAATAACGTGAAAC	Spot 42 oligoprobe
KPO-4168	GACTACAAAGACCATGACGG	pKV114
KPO-4202	ATGGCTGCCGCGCGGC	pMD408
KPO-4203	GGATCCGGCTGCTAACAAAG	pMD408
KPO-4320	GCTGCCTCCCGTAGGAGTC	TransFLP procedure
KPO-4321	ATCCCCCTCTACCGGGCAAT	TransFLP procedure
KPO-4322	AAGCAGCTCCAGCCTACGCTTAATTGAATTACTACGCATCAGAGAG	TransFLP procedure
KPO-4323	GCGTAGGCTGGAGCTGCTTCGAAGTTCCTATAC	TransFLP procedure
KPO-4324	CATATGAATATCCTCCTTAGTTCCTATTCCG	TransFLP procedure
KPO-4325	GAACTAAGGAGGATATTCATATGCGAATGATCCTAAATCGGATGAAG	TransFLP procedure
KPO-4326	GTCGCACTCGTGTCATCTCC	TransFLP procedure
KPO-4327	GTTCGTTTTTTGTTCTTCTTTTGTAAG	TransFLP procedure
KPO-4715	CCGTCATGGTCTTTGTAGTCGAGCTCACCCTGAAAATACAAATTC	pKV114
KPO-4716	CTTATCCACGCTCTTTTGGTTATTCCATGGAAAAGAGAAGATGG	pKV114
KPO-5017	TCATACCGCCTTTATCCATCTC	*flgL* qRT-PCR
KPO-5018	GAAATCTTCGTGCTGCTCTTC	*flgL* qRT-PCR
KPO-5019	TTACTTCAGCCTCAGCCTTATC	*flgK* qRT-PCR
KPO-5020	CAAGGTTTGCCGTGAACTTATC	*flgK* qRT-PCR
KPO-5021	TGTTCCAGTGGTGGAGAAAG	*flgJ* qRT-PCR
KPO-5022	CGCATTGGTATAACGTGGATTG	*flgJ* qRT-PCR
KPO-5023	CATAGTGGTTGGGCAGAATGTA	*flgI* qRT-PCR
KPO-5024	TTAGGCTGGCTGACGTTTAAG	*flgI* qRT-PCR
KPO-5025	CGGGCTCGCTGTTTAATTTG	*flgH* qRT-PCR
KPO-5026	GCTGCTTTGGTGCTTTCATT	*flgH* qRT-PCR
KPO-5027	GTTTGCCGAGTGGTTTGATG	*flgG* qRT-PCR
KPO-5028	CAAACTGTTGGAGGTCGTTTG	*flgG* qRT-PCR
KPO-5029	TTGCCTGTGCCGCTAAA	*flgF* qRT-PCR
KPO-5030	ATTCGATCCACCACTTCCATC	*flgF* qRT-PCR
KPO-5031	CAGTGGGATGCTACTCAGTTC	*flgE* qRT-PCR
KPO-5032	TCGATGTTCGACTGCTCTAATG	*flgE* qRT-PCR
KPO-5033	CTTGCTGGATGGAGAGAGTAAAG	*flgD* qRT-PCR
KPO-5034	GCCAGATTGAGTAGTACGTTACC	*flgD* qRT-PCR
KPO-5035	CGAGTCGGTTCGTCTCAATAC	*flgC* qRT-PCR
KPO-5036	ACTCAACTCTGCGCCAAATA	*flgC* qRT-PCR
KPO-5037	AGTCTTAGCCGTACCGATAGT	*flgB* qRT-PCR
KPO-5038	CATCCACCGTATTGCCATCA	*flgB* qRT-PCR
KPO-5041	CTGGTACTGATCATGGTGTGTG	*fliQ* qRT-PCR
KPO-5042	CGTCTGTTCGTTAATCGAGGTG	*fliQ* qRT-PCR
KPO-5043	CTCTGTTTCTCATCCCTGTCC	*fliO* qRT-PCR
KPO-5044	TAACCAAGCCAGTAGCAGTATC	*fliO* qRT-PCR
KPO-5045	GTGATTGTCGTCAGCTCCTT	*fliM* qRT-PCR
KPO-5046	GGATCGGTTCTACCATCGAATAC	*fliM* qRT-PCR
KPO-5047	GGATGCCAAAGATCGCTTAGT	*fliL* qRT-PCR
KPO-5048	CCACTAAAGGCGAGTGGTAAC	*fliL* qRT-PCR
KPO-5049	TCAAACCGTAGCGGTCAATC	*fliK* qRT-PCR
KPO-5050	TAAAGCTCGGCCACCTAATG	*fliK* qRT-PCR
KPO-5266	TTCTGCCGATCTGTTCTACAC	*vca0053* qRT-PCR

KPO-5266	TTCTGCCGATCTGTTCTACAC	*vca0053* qRT-PCR
KPO-5267	CTCGGCAATGTGATGATGTTTG	*vca0053* qRT-PCR
KPO-5267	AGCCGCTACACCGTAGATA	*vca0053* qRT-PCR
KPO-5268	CTCGGCAATGTGATGATGTTTG	*vc2761* qRT-PCR
KPO-5269	CCAATGTCCCGTGCTTGATA	*vc2761* qRT-PCR
KPO-5270	GCCTCTTTAGGTGTGGAATTGA	*vc1953* qRT-PCR
KPO-5271	CATCGACGTGTTCGGTTTCT	*vc1953* qRT-PCR
KPO-5520	GACTATAAAAAGGCGCACTGCGCTTATTCAGTTGAGTGTCAAAC	pKV70
KPO-5521	CTGCGTTCTGATTTAATCTAGACGCTGAACATAAAGCTTGAACAG	pKV70
KPO-5522	CTGATGCGTAGTAATTCAATATGAAATCTGACAAATCGATGTTATGG	pKV79
KPO-5523	GAACATAAAGCTTGAACAGTTAGGTTTTACTATTTATCGTCATCTTTGTAG	pKV79
KPO-5587	TGCGCGCAGTCTCTCACTGGAGCAGC AAATCTGACAAATCGATGTTATGG	pKV99
KPO-5588	GTGAGAGACTGCGCGCACGCGATCCGGCGAGCCCATATTGAATTACTACGCATCAGAG	pKV99
KPO-5589	TGCGCGCAGTCTCTCACTGGAGCAGCCAATTGACTCTACACTCTTCTTTC	pKV100
KPO-5590	GTGAGAGACTGCGCGCACGCGATCCGGCGAGCCCGATAAAATCCCGGCACCAAT	pKV100
KPO-5591	TGCGCGCAGTCTCTCACTGGAGCAGCAGCGAAGATGAAGAGGCGTC	pKV101
KPO-5592	GTGAGAGACTGCGCGCACGCGATCCGGCGAGCCTGATGAAAGAGTGGGGTGAGTAG	pKV101
KPO-5593	TGCGCGCAGTCTCTCACTGGAGCAGCCTGCCCTATAAAGCCAATGAATC	pKV102
KPO-5594	GTGAGAGACTGCGCGCACGCGATCCGGCGAGCCATTGCCCACATAGAGTGTTTTCG	pKV102
KPO-5595	TGCGCGCAGTCTCTCACTGGAGCAGCTTTGGCGAGGTCTTTGCGG	pKV103
KPO-5596	GTGAGAGACTGCGCGCACGCGATCCGGCGAGCCCTCGGCGAACAGCTCTTTAAC	pKV103
KPO-5597	TGCGCGCAGTCTCTCACTGGAGCAGCGTGCGTTTAATGAAAGACAAACG	pKV104
KPO-5598	GTGAGAGACTGCGCGCACGCGATCCGGCGAGCCCGCAAAGACCTCGCCAAAC	pKV104
KPO-5599	TGCGCGCAGTCTCTCACTGGAGCAGCCGAACAGGTAAACGGCGAG	pKV105
KPO-5600	GTGAGAGACTGCGCGCACGCGATCCGGCGAGCCTTTGTCTTTCATTAAACGCACCG	pKV105
KPO-5601	TGCGCGCAGTCTCTCACTGGAGCAGCCGGCGAGGGTTTGGATTTG	pKV106
KPO-5602	GTGAGAGACTGCGCGCACGCGATCCGGCGAGCCTTTACCTGTTCGTTTGTCTTTCATTAA	pKV106
KPO-5603	TGCGCGCAGTCTCTCACTGGAGCAGCCAAGCTCAAACAGCCATTGATG	pKV107
KPO-5604	GTGAGAGACTGCGCGCACGCGATCCGGCGAGCCGCTTGCGGCTATCACCAC	pKV107
KPO-5605	TGCGCGCAGTCTCTCACTGGAGCAGCCCTAAATCGGATGAAGAAATGG	pKV108
KPO-5606	GTGAGAGACTGCGCGCACGCGATCCGGCGAGCCATCATTCGCAATGCGAACTTTTAAC	pKV108
KPO-5607	TGCGCGCAGTCTCTCACTGGAGCAGCTAAAACCTAACTGTTCAAGCTTTATG	pKV109
KPO-5608	GTGAGAGACTGCGCGCACGCGATCCGGCGAGCCATCCTGTTCAGCCATTTCTTCATC	pKV109
KPO-6306	TAGAGGTACCGGTTGTTAACGATGGTTATGGCGGATAAGAAAGC	pKV150, pKV155
KPO-6307	CCGTCATGGTCTTTGTAGTCACTGCGTTCATGCAAGGGTG	pKV150
KPO-6308	CGACTACAAAGATGACGATAAATAGTCATTGATTCAAAAGCAAAAGAGCG	pKV150
KPO-6309	CATATGCATCCTAGGCCTATTAGGCAGTTGGTGGTTGAGCA	pKV150, pKV155
KPO-6476	TTTGTTCACGGCTAAATTGGAATTTATTGACGACC	pMD389
KPO-6477	CCAATTTAGCCGTGAACAAAACTTGTTCGAGAC	pMD389
KPO-6480	GCTACGGCCGTATGCCGTTAGTCCCAG	pMD402
KPO-6481	GCATACGGCCGTAGCCTAAATGTTAGCAC	pMD402
KPO-6484	CTAAGCGTAATCTGGAACATCGTATGGGTAACTGCGTTCATGCAAGGGTG	pKV152
KPO-6485	TGTTCCAGATTACGCTTAGTCATTGATTCAAAAGCAAAAGAGCG	pKV152
KPO-6486	CTAGTGATGGTGATGGTGATGACTGCGTTCATGCAAGGGTG	pKV153
KPO-6487	CACCATCACCATCACTAGTCATTGATTCAAAAGCAAAAGAGCG	pKV153
KPO-6524	TAGAGGTACCGGTTGTTAACGCATAGGCCAATCCAGCCATG	pKV154
KPO-6525	CTCCTTTGTTTATTATTTAATCCGTC	pKV154
KPO-6526	GACGGATTAAATAATAAACAAAGGAGTCATTGATTCAAAAGCAAAAGAGCG	pKV154
KPO-6527	CATATGCATCCTAGGCCTATTACCCTGATAGTCTGGCCACTG	pKV154
KPO-6558	AACTCATTGCCGTATAAAACATGTACGAGAAGC	pMD405
KPO-6559	GTTTTATACGGCAATGAGTTTATTTTTTGTAAGTTTG	pMD405
KPO-6560	AATTCAATACCGTGGACGAGGATATGAGC	pMD401
KPO-6561	CGTCCACGGTATTGAATTATTATCATTGCAGCAA	pMD401
KPO-6562	ATCCTATAGCCGGAACTTAAGGTGAATATTCTTG	pMD403
KPO-6563	TAAGTTCCGGCTATAGGATTTTTGTATAGTATTG	pMD403
KPO-6588	CCGCGCGGCAGCCATATGGTTATGGCGGATAAGAAAG	pMD408
KPO-6589	CTTTGTTAGCAGCCGGATCCTTAACTGCGTTCATGCAAGGG	pMD408
KPO-6669	GTCTATGGCGACGGTGATCTTTGCGATGTCTC	pKV155, pKV156
KPO-6670	GATCACCGTCGCCATAGACACAGGAATACCAATGG	pKV155, pKV156
KPO-6768	GAACCAACACGCGGCAGCGGCGGGTAACTTTGCATACCCTCG	pKV159
KPO-6769	GTTACCCGCCGCTGCCGCGTGTTGGTTCCAGAACATTGC	pKV159
KPO-6954	GACCAAGATCAATTTAAATGATTTTTTTTAATCACCTTAAGTGTTGTAATTTA	pKV164
KPO-6955	AAAATCATTTAAATTGATCTTGGTCAG	pKV164
KPO-7127	CTCGAGCACCACCACCAC	pKV169
KPO-7130	TTATGTAAATCGCTCCTTTTTAGGTG	pKV169
KPO-7131	CACCTAAAAAGGAGCGATTTACATAAATACCCGTTTTTTGGGCTAACAG	pKV169
KPO-7132	GTGGTGGTGGTGCTCGAGAAAACGAAAGGCCCAGTCTTTC	pKV169
KPO-7413	TCGTTTTATTTGATGCCTCTAGATTAACTGCGTTCATGCAAGGG	pKV174

KPO-7414	GCTAACAGGAGGAATTAACCATGGTTATGGCGGATAAGAAAGC	pKV174
KPO-7415	CTGCGTTCTGATTTAATCTAGAAAAAATTAACTGCGTTCATGCAAGGG	pKV175
KPO-7416	CGCAACTCTCTACTGTTTCTCCATGGTTATGGCGGATAAGAAAGC	pKV175
KPO-7417	GCTGCTCTGACTTCATTCCTAA	*ptsG* qRT-PCR
KPO-7418	CGTAAGCCAGACCTGCTAATAC	*ptsG* qRT-PCR
KPO-7419	GTGGTGCCTTCTTACCTCTATC	*treB* qRT-PCR
KPO-7420	CCAATCACGCGACCAAATG	*treB* qRT-PCR
KPO-7421	GTGATTGGTCTGGCGTTCTT	*nagE* qRT-PCR
KPO-7422	AGCGGAGTCTGTTTCTGTTTC	*nagE* qRT-PCR
KPO-7423	ACCTCTACAACCCACTCTCTC	*ptsH* qRT-PCR
KPO-7424	CCCGCTAATTCACCACACATA	*ptsH* qRT-PCR
KPO-7425	GCAAACGCTAGGTCTGGTAA	*ptsI* qRT-PCR
KPO-7426	GTTGGTCCATCAGAGCAACTA	*ptsI* qRT-PCR
KPO-IGRF	GGCTGCGAATTCCACAGATGACTCCTTACTGTAATTTTTCC	EMSA of *vcdRP*
KPO-IGRR	GCCCGAAGCTTCATGTTCGAATCCTTTAGTGGT	EMSA of *vcdRP*
pBAD-ATGrev	GGTTAATTCCTCCTGTTAGC	pEVS143 and pBAD1K-derivates

6.6 General methods

Bacterial strains and growth conditions

A complete list of strains used in this study is provided in Table 6.5. *V. cholerae* C6706 was used as the wild-type strain and all the mutant strains were generated as described previously (Papenfort *et al*, 2015). Briefly, RK2/RP4-based conjugal transfer was used to introduce pKAS32-derived plasmids from *E. coli* S17λpir plasmid donor strains. Subsequently, trans-conjugants were selected for ampicillin (200mg/ml) resistance. Polymyxin B (6.25 mg/ml) was used to specifically inhibit *E. coli* growth. Subsequently, single colonies were transferred to fresh plates and selected for streptomycin (5000 mg/ml) resistance. Finally, mutants were confirmed by PCR and sequencing. *B. subtilis* 168 and *V. natriegens* ATCC14048 served as the wild-type hosts for cloning and plasmid transformation as described previously (Brockmeier *et al*, 2006; Weinstock *et al*, 2016). All bacterial strains were grown at 37°C in LB broth or M9 minimal medium (supplemented with 0.4% casaminoacids and 0.4% glucose or glycerol) or in AKI medium (0.5% sodium chloride, 0.4% yeast extract and 1.5% bacto-peptone, 0.3% sodium bicarbonate). *V. natriegens* was grown in LB broth supplemented with 5% sodium chloride.

Cloning using polymerase chain reaction (PCR) and transformation into *E. coli*

DNA fragments of interest were amplified by PCR using Q5® Hi-Fidelity DNA polymerase and the DNA oligonucleotides used are listed in Table 6.12. For screening of bacterial transformants, cells were picked from plates and streaked into tubes to serve as template in colony PCR using GoTaq® polymerase. PCR products were purified using the Zymo PCR Purification kit according to the manufacturer's instructions. Site-directed mutagenesis, the deletion or addition of sequence stretches were introduced by amplification of the original plasmids and self-ligation of purified PCR products. Multiple fragments were joined by Gibson assembly using manufacturer's instructions. PCR fragments amplified to obtain plasmid variants were incubated with DpnI for 1h at 37°C to digest template DNA. All other restriction enzyme digests were performed in the buffers and under the conditions suggested by the manufacturer. Digested DNA fragments and linearized vectors were ligated by T4 DNA ligase prior to transformation into chemically competent *E. coli*: 1 µl of plasmid DNA or 5 µl of ligation reactions was mixed with 50 µl of chemically competent *E. coli* TOP10 or S17λpir cells. Upon incubation on ice for 20 min, cells were subjected to heat shock at 42°C for 1min 30 sec. Cells were chilled on ice for 2 min and resuspended in 600 µl LB medium. Recovery was carried out for 60 min at 37°C, 200 rpm.

Plasmid construction

The plasmids and DNA oligonucleotides used in this study are listed in Table 6.11 and Table 6.12, respectively. The plasmids for the over-expression sRNA library to screen for CT repression were cloned into pEVS143 (Dunn *et al*, 2006) linearized with KPO-0092/1397 followed by digestion with XbaI enzyme and subsequent ligation via T4 ligase with the corresponding sRNA inserts. For the amplification of the sRNAs, *V. cholerae* wild-type (KPS-0014) chromosomal DNA served as a template with the following combinations of oligonucleotides: KPO-1076/1077 (pAS001), KPO-1072/1073 (pAS002), KPO-1070/1071 (pAS003), KPO-0456/0465 (pKP333), KPO-1003/1004 (pNP001), KPO-0999/1000 (pNP002), KPO-1024/1025 (pNP003), KPO-1005/1006 (pNP004), KPO-1015/1016 (pNP005), KPO-1021/1022 (pNP006), KPO-1011/1012 (pNP007), KPO-1009/1010 (pNP008), KPO-1219/1220 (pNP009), KPO-1001/1002 (pNP010), KPO-1007/1008 (pNP011), KPO-1026/1027 (pNP012), KPO-1013/1014 (pNP015), KPO-1082/1083 (pRH001), KPO-1092/1093 (pRH002), KPO-1090/1091 (pRH003), KPO-1860/1861 (pSG002) and KPO-1448/1449 (pYH002). The remaining plasmids for the sRNA library screen were cloned into pEVS143 (Dunn *et al*, 2006) linearized with KPO-0092/1397, with the corresponding sRNA inserts amplified from KPS-0014 chromosomal DNA, followed by Gibson assembly. These sRNAs were amplified using KPS-0014 chromosomal DNA as a template with the following combinations of oligonucleotides: KPO-2568/2553 (pMD090), KPO-2570/2565 (pMD104), KPO-2571/2572 (pMD105), KPO-1375/1376 (pRH0009) and KPO-1383/1384 (pRH010). The inserts for *vcdRP* truncation plasmids were amplified with KPO-2082/2083 (pSG006), KPO-2082/2085 (pSG008), KPO-2082/2109 (pMD062) and KPO-2082/2110 (pMD063) and cloned into pEVS143 linearized using KPO-1949/0092 by Gibson assembly. Site-directed mutagenesis of pNP009 using KPO-2090/2091 and KPO-6476/6477 yielded pMD055 and pMD389, respectively. pMD083 was generated from pNP009 by amplifying the plasmid in two fragments with KPO-1529/2323 and KPO-1525/2322 respectively, and subsequently joined by Gibson assembly. For pMD004, the *rrnB* terminator from pKP8-35 (Papenfort *et al*, 2006) was amplified with KPO-1484/1485 and cloned by Gibson assembly into pKP-331 (Papenfort *et al*, 2015) linearized with KPO-0196/1397. pMD072 was generated by amplifying the insert from pNP009 using KPO-2100/2101 and cloned via Gibson assembly into pMD004 linearized with KPO-0196/1397. Site-directed mutagenesis of pMD072 using KPO-2090/2091 yielded pMD077. For pMD080, an artificial 5'UTR, MCS and T1 terminator were inserted via overlapping primers using KPO-2259/2260/2261/2329/1492 and cloned into pEVS143 (Dunn *et al*, 2006) linearized with KPO-0092/1949. For pMD087, the codon modified VcdP sequence was custom synthesized by GeneArt and cloned using Gibson assembly into pMD004 linearized with KPO-2391/2392. The insert fragment for pMD111 was amplified from pMD087 with KPO-2330/1484 and cloned into pMD080 linearized with pBAD-ATGrev/KPO-1397. The SPA tag was inserted using KPO-1715/4716 into linearized pMD065 with KPO-4168/2410 to yield pKV114. The translational GFP fusions were cloned as previously described (Corcoran *et al*, 2012) by employing already determined transcriptional start site annotations (Papenfort *et al*, 2015). Briefly, *treB* (pNP058), *ptsG* (pMD161), *nagE* (pMD0162) and *ptsHI* (pMD164) inserts were amplified from KPS-0014 chromosomal DNA using the primers indicated in Table 6.11 and introduced into the pXG-10SF backbone using NheI and NsiI restriction sites or via Gibson assembly. The M2* mutation was introduced via site-directed mutagenesis into the translational GFP fusions. To this end, amplification with KPO-6560/6561, KPO-6480/6481, KPO-6562/6563 and KPO-6558/6559 yielded pMD401, pMD402, pMD403 and pMD405, respectively. VcdRP transcriptional mKate2-based reporter (pMD064) was generated by amplifying KPS-0014 chromosomal DNA using KPO-2111/2112 and ligating into pYH010 digested with SphI and SalI enzymes, via T4 ligase. Site-directed mutagenesis of pMD064 with KPO-6954/6955 yielded pKV164. All pKAS32-derived plasmids (Skorupski & Taylor, 1996) were constructed by linearizing

the backbone using KPO-0267/0268 and cloning the corresponding 'up' and 'down' homologous fragments by Gibson assembly. An additional flank containing the SPA tag sequence where appropriate was also included. For the amplification of the 'up' and 'down' flanking fragments, KPS-0014 chromosomal DNA served as a template with the following combinations of oligonucleotides for each of the flanking fragments: pASp2 (KPO-0167/0273 and KPO-0170/0274), pASp17 (KPO-1872/1873 and KPO-1874/1875), pASp18 (KPO-1876/1877 and KPO-1878/1879), pMD054 (KPO-1282/1283 and KPO-1284/1285), pRH023 (KPO-2243/2244 and KPO-2245/2246), pRH024 (KPO-2249/2250 and KPO-2251/2252), pKV150 (KPO-6306/6307 and KPO-6308/6309), pKV154 (KPO-6524/6525 and KPO-6526/6527), pKV155 (KPO-6306/6670 and KPO-6309/6669), and pKV157 (KPO-6472/6473 and KPO-6744/6745 along with KPO-4814/6178 for SPA tag amplified from pKV114). Site-directed mutagenesis of pKV150 using KPO-6484/6485 and KPO-6486/6487 yielded pKV152 and pKV153, respectively. Plasmid pMD408 was constructed by cloning *gltA* amplified from KPS-0014 chromosomal DNA using KPO-6588/6589 into pET15b (Novagen) linearized with KPO-4202/4203. Site-directed mutagenesis of pMD408 using KPO-6669/6670 yielded pKV156. pKV169 was generated by Gibson assembly of linearized pBSmul2 (pKV168) using KPO-7127/7130 and insert amplified from pMD111 using KPO-7131/7132. The insert fragment for pKV175 and pKV176 were amplified from KPS-0014 chromosomal DNA using KPO-7413/7414 and 7415/7416, respectively, and cloned via Gibson assembly into linearized pMD080 (with KPO-3236/pBAD-ATGrev) and pMD004 (with KPO-0196/1488), respectively. 3xFlag tag was amplified from pASp2 and cloned into pMD004 that was linearized with KPO-0196/1379 via Gibson assembly to generate pASp6. Site-directed mutagenesis of pASp6 using KPO-0100/0101, KPO-0095/0096 and KPO-0098/0099 yielded pKV41, pKV42 and pKV43 respectively. The resistance cassette of pASp6 was replaced with chloramphenicol using KPO-1518/1519 amplified from pYH010 to generate pASp7. pASp10 was generated by linearizing pASp7 with KPO-0196/1379 and amplifying the insert using KPO-1544/1545. Site-directed mutagenesis of pASp10 using KPO-0100/0101 followed by DpnI digestion yielded pASp11. pKV44 was generated by cloning the full-length *mbrA* gene amplified from KPS-0014 using KPO-2471/2472 into linearized pMH29 (Huber *et al*, 2020) using KPO-2294/2295 by Gibson assembly. Site-directed mutagenesis of pKV44 using KPO-0095/0100 yielded pKV45. pKV72 was cloned by linearizing the backbone using KPO-1952/1953 and combining the insert amplified from KPS-0014 using KPO-3010/3011 via Gibson assembly. Site-directed mutagenesis of pKV52 yielded the R1, R2 and R3 mutants using KPO-3435/3436, KPO-3437/ 3438 and KPO-3439/3440, respectively. pKV70 was generated by cloning the insert amplified from KPS-0014 using KPO-4563/4564 into linearized pKV69. pKV70 was linearized using KPO-0101/1423 and joined via Gibson assembly with the insert amplified from pASp2 using KPO-5522/5523 to generate pKV79. Spot®-tagged variants of MbrA were generated by site-directed mutagenesis of pKV70 using KPO-5587/5588 (pKV99), KPO-5589/5590 (pKV100), KPO-5591/5592 (pKV101), KPO-5593/5594 (pKV102), KPO-5595/5596 (pKV103), KPO-5597/5598 (pKV104), KPO-5599/5600 (pKV105), KPO-5601/5602 (pKV106), KPO-5603/5604 (pKV107), KPO-5605/5606 (pKV108) and KPO-5607/5608 (pKV109).

Agarose gel electrophoresis of DNA

DNA fragments of different sizes were separated using 1-2% (w/v) agarose gels in 1xTAE buffer as described by (Green *et al*, 2012). Prior to loading, samples were mixed with DNA loading buffer (ratio 6:1) and separated at 130 V for 20-45 min. 1kB DNA ladder from New England Biolabs served as size standards. DNA fragments were visualized by the addition of RedSafe (0.02% (v/v)) to agarose gel solutions. If desired, DNA fragments were excised from gels under UV light and recovered using the Zymo Gel Extraction Kit.

TransFLP method to induce natural competence

The TransFLP method was carried out as previously described (Blokesch, 2012). Briefly, *V. cholerae* wild-type cells were grown aerobically in LB medium at 30 °C until mid-exponential phase. Bacterial cell pellets corresponding to 1 OD_{600} units were harvested and washed twice in defined artificial medium (DASW) before resuspending the cells in 2 volumes of DASW. 1 ml of this culture was added to to 50-80 mg of sterile chitin flakes and incubated at 30 °C for 16-24 hours without movement. The transforming PCR was generated and ≥ 200 ng of this DNA was mixed carefully to the chitin flakes without extensively detaching the bacteria from the chitin surfaces. The cultures were further incubated at 30 °C for 24 hours without movement. The cultures were vortexed extensively for ≥ 30 s and 100-300 µl was spread on selective LB medium plates containing kanamycin. The plates were incubated at 30 °C for 16-24 hours or until colonies are visible. Single transformants were isolated and the desired mutations were confirmed by Sanger sequencing.

RNA isolation and northern blotting

Total RNA was prepared and blotted as described previously (Peschek *et al*, 2019). Briefly, total RNA samples corresponding to 4 OD_{600} units were harvested and mixed with 0.2 volumes of stop-mix. Cell pellets were obtained by centrifugation, thoroughly resuspended in 1 ml of EXTRAzol, and the samples were incubated at room temperature for 5 min. Subsequently, they were transferred to phase lock gel™ tubes, 200µl of chloroform was added, mixed by inversion and incubated for 5 min at room temperature to allow phase separation. After the samples were centrifuged at maximal speed for 15 min at 12°C, the upper phase was transferred to a new tube and 400 µl of isopropanol was added. After incubating at room temperature for 30 min, the samples were centrifuged at maximal speed for 30 min at 4°C. Following two washes with 75% ethanol, the pellets were resuspended in 20-30µl of nuclease-free water and quantified using a NanoDrop.

Alternatively, RNA was prepared by the hot phenol method. To this end, bacterial cell pellets corresponding to 4 OD_{600} units were resuspended in 600 µl lysozyme solution (0.5 mg/ml lysozyme in TE buffer, pH 8.0) and 60 µl of 10% (w/v) SDS were added. The suspension was mixed by inversion and incubated at 64 °C in a water bath for 2 min. The pH was equilibrated by the addition of 66 µl of sodium acetate (pH 5.2), and samples were mixed with 750 µl phenol. Tubes were incubated at 64 °C for 6 minutes with frequent mixing. Upon 1 min chilling on ice, samples were centrifuged for 10 min at 13,000 rpm and 4°C to ensure phase separation. The aqueous layer was transferred to a phase-lock tube, mixed with 750 µl chloroform and centrifuged again for 10 min at 13,000 rpm at 4°C. RNA was precipitated from the aqueous layer by the addition of 1.4 ml of a 30:1 ethanol:sodium acetate (pH 6.5) mix. RNA pellets were washed with 75% ethanol and air-dried. The pellets were resuspended in 75-100µl of nuclease-free water and quantified using a NanoDrop.

5-10 µg RNA was then loaded on a denaturing 6% polyacrylamide 5M Urea gel along with (γ-^{32}P)-ATP labelled 50bp size marker and subsequently blotted on Amersham Hybond XL membranes (1h, 50V, 4°C) in 1xTBE. The membranes were then UV-crosslinked and hybridized in Roti-Hybri-Quick buffer at 42°C with (γ-^{32}P)-ATP end-labeled DNA oligonucleotides or 63°C with (α-^{32}P)-UTP labeled riboprobes listed in Table 6.12. Following three washes in SSC buffers, the membranes were sealed and exposed to phosphor-imaging plates for 6-14 h. Signals were visualized using a Typhoon Phosphorimager.

RNA purification using SV total RNA isolation system

RNA purified using the SV Total RNA Isolation System was employed in qRT-PCR. RNA isolated according to the manufacturer's instructions. Briefly, bacterial cell pellets corresponding to 2 OD_{600}

units resuspended in 100 µl lysozyme solution and samples were incubated for 4 min at room temperature. The samples were mixed with 75 µl of lysis reagent and 350 µl RNA dilution buffer was added. The lysates were incubated for 3 min at 70°C and subsequently centrifuged for 10 min at 13,000 rpm at room temperature. The supernatant was mixed with 200 µl 95% ethanol and loaded on a spin column provided with the kit. After centrifugation at room temperature for 1 min at 13,000 rpm, the eluate was discarded, and the column was washed with 600 µl wash buffer. After an additional centrifugation step, 50 µl of a DNase I mix (5 µl 90 mM $MnCl_2$, 40 µl DNase I core buffer and 5 µl DNase I - all provided with the kit) was applied to the membrane and samples were incubated for 30 min at room temperature. Digestion was stopped by the addition of 200 µl DNase I stop-mix (different from the standard stop-mix used for RNA isolation) and the columns were centrifuged at room temperature for 1 min at 13,000 rpm. Following two wash steps with 600 µl and 250 µl wash buffer, respectively, the column was transferred to a sterile tube and 75-100 µl RNase-free water were added. After incubation for 1 min at room temperature the RNA was eluted by centrifugation for 2 min at 13,000 rpm.

Generation of radiolabelled DNA oligonucleotides and riboprobes

For labelling, 1 pmol of the oligo was incubated with 25 µCi of (γ-^{32}P)-ATP in the presence of 1 U polynucleotide kinase (PNK) and 1x PNK buffer for 1 h at 37°C in a 20 µl reaction. Unincorporated nucleotides were removed using Microspin G-25 Columns. DNA templates for T7 *in vitro* transcription of riboprobes were amplified by PCR from template plasmids or genomic DNA using gene specific primer sets. *In vitro* transcription was performed with the MAXIscript kit using 200 ng of template DNA in the presence of 25 µCi (α-^{32}P)-UTP at 37°C for 1 h. Following DNase I digestion (1 U) for 15 min at 37°C, the riboprobes were purified over a MicroSpin G50 column.

Determination of RNA stability

To determine RNA stability, cultures were grown in triplicates to appropriate growth phases when transcription was inhibited by the addition of rifampicin (final concentration: 500 µg/ml). Prior to, and at 2, 4, 8, 16 as well as 32 min post rifampicin treatment, culture aliquots were mixed with 0.2 volumes of stop-mix and samples were frozen in liquid nitrogen. RNA was prepared, and either quantified on Northern blots or by qRT-PCR.

Quantitative real-time PCR (qRT-PCR)

The qRT-PCR was performed on independent biological triplicates using Luna Universal One-Step RT qPCR kit in a MyiQ™ Single-Color Real-Time PCR Detection System. *recA* served as the reference house-keeping gene and the oligonucleotides used for all qRT-PCR analyses are provided in Table 6.12.

Electrophoretic mobility shift assays (EMSAs)

EMSAs were carried out as previously described (Manneh-Roussel *et al*, 2018). Briefly, DNA fragments for the EMSA were generated by PCR using KPO-IGRF and KPO-IGRR, using genomic DNA from *V. cholerae* strain N16961 as a template. The resulting PCR products were purified and end-labelled with (γ-^{32}P)-ATP using T4 PNK. The resulting radiolabelled fragments were incubated with different concentrations of CRP in buffer containing 40 mM Tris acetate (pH 7.9), 1 mM $MgCl_2$, 100 mM KCl and 0.2 mM cAMP. Herring sperm DNA was added as a non-specific competitor at a final concentration of 12.5 µg/ml. Reactions were incubated at 37°C for 20 min, before being loaded onto a 7.5 % non-denaturing polyacrylamide gel run in 0.5x TBE buffer.

SDS-PAGE and Western blotting

Total protein samples corresponding to 1 OD_{600} units were collected at the desired OD_{600} and cell pellets were re-suspended in 1x Laemmli buffer to a final concentration of 0.01 OD per µl. The samples were immunoblotted as previously described (Papenfort *et al*, 2015). Signals were visualized using a Fusion FX EDGE imager and band intensities were quantified using the BIO-1D software. 3xFlag and SPA-tagged proteins were detected using mouse anti-Flag antibody. The SPA epitope contains the 3×FLAG and the calmodulin binding peptide sequences separated by a TEV protease cleavage site (Zeghouf *et al*, 2004). HA and 6xHis-tagged samples were detected using mouse monoclonal anti-HA and rabbit monoclonal anti-6xHis antibodies, respectively. RNAP served as a loading control and was detected using rabbit anti-RNAP antibody. The corresponding secondary antibodies used were goat anti-mouse HRP-conjugated IgG antibody and goat anti-rabbit HRP-conjugated IgG antibody.

Genetic screen for cholera toxin repression

A sRNA over-expression library was screened in a *V. cholerae* C6706 Δ*hapR* background for repression of cholera toxin (CT). To this end, secreted protein fractions were prepared in independent biological triplicates as previously described (Herzog *et al*, 2019). Specifically, the strains were cultured in AKI medium (0.5 % sodium chloride, 0.4 % yeast extract and 1.5 % bacto-peptone) supplemented with 0.3 % sodium bicarbonate to stimulate the production of CT. Growth under AKI conditions involve biphasic cultures. In the first phase, the cultures were grown in a still tube for 4h at 37°C. Subsequently, in the second phase, the cultures were poured into a flask to continue growing with shaking. 2 ml of sample was harvested after 16 h of continuous shaking followed by trichloroacetic acid (TCA) preparation of proteins. Pellets were re-suspended in volumes of 1x Laemmli buffer relative to the OD_{600} measurements of the respective culture. The same amount of each sample was analyzed simultaneously on two SDS-PAGEs: the first set of gels served as a loading control and were stained using Coomassie brilliant blue stain (0.1% Coomassie R-250 in 40% ethanol, 10% acetic acid). Additionally, the second set of gels were subjected to immunoblotting using rabbit polyclonal anti-cholera toxin antibody. The corresponding secondary antibody used was goat anti-rabbit HRP-conjugated IgG antibody. Signals were visualized using Fusion FX EDGE imager and the band intensities from both sets of gels were quantified using the BIO-1D software. Each secreted fraction from the immunoblot was normalized to its corresponding loading control from the Coomassie stained gel. CT levels were calculated relative to an empty control plasmid.

Fluorescence reporter measurements

Fluorescence assays of bacterial reporters were performed as previously described (Corcoran *et al*, 2012). Briefly, cell pellets corresponding to 1 OD_{600} units were collected by centrifugation. Subsequently, the pellets were washed twice in equal volume of 1x PBS (pH 7.4) and fluorescence intensity was quantified using a Spark 10M plate reader (Tecan, Männedorf, Switzerland). Control strains not expressing fluorescent proteins were used to subtract background autofluorescence.

RNA-seq analysis: identification of VcdRP and MbrA targets

V. cholerae C6706 Δ*vcdRP* strains harboring pBAD1K-ctrl, pBAD1K-*vcdRP*, pBAD1K-*vcdR*, or pBAD1K-*vcdP* were cultivated to early exponential phase (OD_{600} of 0.1) and treated with L-arabinose (0.2 % final concentration). After 15 minutes of treatment, 4.0 OD_{600} units of cells were harvested with 0.2 volumes of stop-mix. *V. cholerae* C6706 wild-type and Δ*mbrA* strains harboring pCtrl or Δ*mbrA* strain harboring pMbrA were cultivated in LB medium to early stationary phase (OD_{600} of 1.0). Total

RNA was isolated with EXTRAzol as described above, and DNase digested with TURBO DNase. Depletion of ribosomal RNA was performed using the Ribo-Zero rRNA removal kit for Gram-negative bacteria. Integrity of the prepared RNA was tested using an Agilent 2100 Bioanalyzer. cDNA libraries for the VcdRP and MbrA transcriptome analyses were prepared using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina and NEB Small RNA library Kit, respectively, according to the manufacturer's instructions. The libraries were then 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 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 the sRNAs Vcr001-Vcr230 (Huber *et al*, 2020; Papenfort *et al*, 2015) using the 'RNA-Seq Analysis' tool with standard parameters. Reads mapping in CDS were counted, and genes with a total count cut-off >10 in all samples were considered for analysis. Read counts were normalized (CPM) and transformed (log_2). Differential expression was tested using the built-in tool corresponding to edgeR in exact mode with tag-wise dispersions (empirical analysis of DGE). Genes with an absolute fold change ≥2.0 and an FDR-adjusted p-value ≤0.05 were considered as differentially expressed. Enrichment of GO (gene ontology) terms were analyzed using the DAVID tool v6.8 (Huang *et al*, 2009a, 2009b).

CLIP-seq sample preparation

For each biological replicate, 200 ml bacterial culture was grown to early stationary phase (OD_{600} of 1.0). Half of the culture was directly placed in a 22 × 22 cm plastic tray and irradiated with UV light at 800 mJ/cm². Cells were pelleted in 50 ml fractions by centrifugation for 40 min at 6,000 *g* and 4°C, resuspended in 800 µl NP-T buffer and mixed with 1 ml glass beads (0.1 mm radius). Cells were lysed using a bead-ruptor and centrifuged for 15 min at 16,000 *g* and 4°C. Cell lysates were transferred to new tubes and centrifuged for 15 min at 16,000 *g* and 4°C. The cleared lysates were mixed with one volume of NP-T buffer with 8 M urea, incubated for 5 min at 65°C in a thermomixer with shaking at 900 rpm and diluted 10× in ice-cold NP-T buffer. Anti-FLAG magnetic beads were washed three times in NP-T buffer (30 µl 50% bead suspension was used for a lysate from 100 ml bacterial culture), added to the lysate, and the mixture was rotated for one hour at 4°C. Beads were collected by centrifugation at 800 *g*, resuspended in 1 ml NP-T buffer, transferred to new tubes, and washed twice with high-salt buffer and twice with NP-T buffer. Beads were resuspended in 100 µl NP-T buffer containing 1 mM $MgCl_2$ and 2.5 U benzonase nuclease and incubated for 10 min at 37°C in a thermomixer with shaking at 800 rpm, followed by a 2 min incubation on ice. After one wash with high-salt buffer and two washes with CIP buffer, the beads were resuspended in 100 µl CIP buffer with 10 units of calf intestinal alkaline phosphatase and incubated for 30 min at 37°C in a thermomixer with shaking at 800 rpm. After one wash with high-salt buffer and two washes with PNK buffer, one-tenth of the beads was removed for subsequent Western blot analysis. The remaining beads were resuspended in 100 µl PNK buffer with 10 U of T4 polynucleotide kinase and 10 µCi (γ-^{32}P)-ATP and incubated for 30 min at 37°C. After three washes with NP-T buffer, the beads were resuspended in 20 µl protein loading buffer and incubated for 3 min at 95°C. The magnetic beads were collected on a magnetic separator, and the supernatant was loaded and separated on a 15% SDS–polyacrylamide gel. RNA–protein complexes were transferred to a nitrocellulose membrane; the protein marker was highlighted with a radioactively labeled marker pen and exposed to a phosphor screen for 30 min. The autoradiogram was used as a template to cut out the labeled RNA–protein complexes from the membrane. Each membrane piece was further cut into smaller pieces, which were incubated for 30 min in a thermomixer at 37°C with shaking at 1,000 rpm in 400 µl PK solution whereafter 100 µl 9 M urea was added and the incubation was continued for additional

30 min. About 450 µl of the PK solution/urea was mixed with 450 µl phenol:chloroform:isoamyl alcohol in a phase-lock tube and incubated for 5 min in a thermomixer at 30°C with shaking at 1,000 rpm followed by centrifugation for 12 min at 16,000 *g* and 4°C. The aqueous phase was precipitated with 3 volumes of ice-cold ethanol, 1/10 volume of 3 M sodium acetate pH 5.2, and 1 µl of GlycoBlue in LoBind tubes. The precipitate was pelleted by centrifugation (30 min, 16,000 *g*, 4°C), washed with 80% ethanol, centrifuged again for 15 min at 16,000 *g*, 4°C, dried 2 min at room temperature, and resuspended in 10 µl sterile water. cDNA libraries were prepared using the NEBNext Multiplex Small RNA Library Prep Set for Illumina according to the manufacturer's instructions

VcdP co-immunoprecipitation and LC-MS analysis

VcdP co-immunoprecipitations were performed as previously described for tagged *V. cholerae* cells (Huber *et al*, 2020). Briefly, *V. cholerae* wild-type cells carrying either an empty vector control or a SPA-tagged over-expression plasmid of *vcdP* were grown in LB medium to mid-log phase. The SPA epitope contains the 3×FLAG and the calmodulin binding peptide sequences separated by a TEV protease cleavage site (Zeghouf *et al*, 2004). Hence, the lysates corresponding to 50 OD_{600} units were subjected to immunoprecipitation using monoclonal anti-Flag antibody and Protein G Sepharose. Protein samples were analyzed on immunoblots using anti-Flag antibodies. Subsequently, these samples were processed using the single-pot solid-phase-enhanced sample preparation (SP3) protocol (Hughes *et al*, 2019) to identify potential protein partners binding to VcdP. The samples were reduced with Tris (2-carboxyethyl) phosphine, alkylated with chloroacetamide and subsequently precipitated onto magnetic beads (SpeedBeads™ magnetic carboxylate) using ethanol. Following two wash steps, the bead-associated precipitated proteins were digested in solution with trypsin. The peptides were separated by chromatography on a Dionex U3000 nanoHPLC system equipped with an Acclaim PepMap 100 C18 column (2 µm, 75 µm × 500 mm) coupled to a QExactive Plus Orbitrap MS. The raw MS data were processed using the Proteome Discoverer software package v2.4.0.305. The files were searched individually using SequestHT algorithm node against a protein database containing genome-derived proteins of *V. cholerae* serotype O1 (strain ATCC 39315 / El Tor N16961- access:2018.06.26), with VcdP::SPA and mouse IgG appended along with the cRAP list of common laboratory contaminants. Searches were performed with semi-trypsin specificity with a maximum of 4 missed cleavages. The results were filtered using the following criteria: protein level FDR <5%, at least two high confidence peptides (FDR <1%) of which at least one is unique and peptide assigned as high confidence with an FDR ≤ 1%. The proteins that were identified are shown in Table 2. To validate the findings from the LC-MS data, cells expressing chromosomally-tagged *gltA*::HA or *gltA*::6xHis either an empty vector control (pCtrl) or a SPA-tagged over-expression plasmid of *vcdP* were subjected to reciprocal co-immunoprecipitations using monoclonal anti-Flag antibody and monoclonal anti-HA or monoclonal anti-6xHis antibodies, respectively.

Purification of proteins

V. cholerae citrate synthase GltA and its variant that is no longer inhibited by NADH, GltA F383A, were expressed with a N-terminal 6xHis in the pET15b vector and transformed into BL21 (DE3) *E. coli* cells. The proteins were purified as previously described for tagged VqmR protein (Papenfort *et al*, 2017). Briefly, 6xHis::GltA (or GltA F383A) expressing cells were grown at 37°C with shaking to mid-log phase. IPTG (final concentration of 1mM) was added, and the cultures were grown for an additional 2.5 h. Pellets were collected and re-suspended in lysis buffer (20 mM imidazole pH 8.0, 450 mM potassium acetate, 50 mM HEPES pH 7.4 and 10 mM BME in PBS supplemented with 1x cOmplete EDTA-free protease inhibitor cocktail. The cells were lysed via sonication for 1 min 45 s at

an amplitude of 20% with 5 s pulse and 15s pause. The cleared lysates were applied to Ni-NTA resin and rotated end-to-end for 2 h at 4°C, following which the resin-lysate mixture was washed two times with wash buffer (40 mM imidazole pH 8.0, 450mM potassium acetate, 50mM HEPES pH 7.4, 5% (v/v) glycerol and 5mM BME). The washed mixture was then re-suspended in 10ml lysis buffer (per ml beads) and loaded on a polypropylene column. After washing with 4x bed volume of resin, on-column cleavage was induced using elution buffer (250mM imidazole pH 8.0, 50mM potassium chloride, 50mM HEPES pH 7.4 and 5mM BME). Protein purification was verified by SDS-PAGE analysis and the concentrations were measured using BCA assay.

The full-length MbrA protein as well as the mutated protein without its transmembrane domains was expressed from the pTYB11 expression vector in *E. coli* ER2566 cells and purified following the Impact Kit protocol according to the manufacturer's instructions. Briefly, cells were grown to mid-log phase and induced with IPTG (final concentration of 0.5mM) for 15 h at 20°C. Cells were harvested and resuspended in column buffer (20 mM Tris pH 8.5, 500mM naCl, 1 mM EDTA) and lysed by sonication for 1 min 45 s at an amplitude of 20% with 5 s pulse and 15s pause. Cleared lysates were loaded on a column containing chitin binding domain. After 40 h of incubation at room temperature, on-column cleavage was induced using column buffer. Protein purification was verified by SDS-PAGE analysis and the concentrations were measured using BCA assay.

MbrA crystallization and structure determination

Purified version of mutated MbrA was concentrated to 43 mg/ml and crystallization screens were performed as previously described (Kowalinski *et al*, 2007). Briefly, the sitting-drop vapor-diffusion method at 291K was employed upon mixing equal volumes (0.5 µl) of protein solution and crystallization buffer with a reservoir volume of 100 µl. The protein was crystallized in a condition containing 0.2 M calcium chloride, 0.1 M HEPES pH 7.5, 28% PEG 400. The crystal appeared after 7 months and were flash-frozen in liquid nitrogen after cryo-protection by transfer into cryo-solution containing mother liquor and 20% (v/v) glycerol. Diffraction data were measured under cryogenic conditions (100 K) at the European Synchrotron Radiation Facility (ESRF) beamline ID29. The structure was solved by molecular replacement with *Phaser* (McCoy, 2006)using PDB-ID 4ED5 as a search model.

Citrate synthase activity assay

Cell extracts were obtained from *V. cholerae* wild-type, Δ*vcdRP*, Δ*gltA* and F383A *gltA* cells carrying either an empty control vector, the VcdP plasmid, or VcdP* plasmid. The cells were grown in LB medium to late stationary phase and cell pellets were lysed using bead ruptor in a buffer containing 1 M Tris and 0.5 mM EDTA (TE) at pH 8.0. The resulting cell lysates were used to measure the citrate synthase activity in accordance with the protocols described earlier (Anderson & Duckworth, 1988; Duckworth & Tong, 1976; Pereira *et al*, 1994). Specifically, protein concentrations were determined using Pierce™ BCA protein assay kit. 1-50 µl of the lysate was used as input for the assay using a Spark 10M plate reader. Equimolar concentration of the substrates acetyl coenzyme A and oxaloacetate were added (final concentration. 0.1 mM each). The coupled enzymatic reaction led to the formation of citrate, with the release of thiols. These thiol groups were detected using 0.1 mM of Ellman's reagent (5,5′-dithiobis-(2-nitrobenzoic acid) or DTNB). The resulting colorimetric product at 412 nm was proportional to the enzymatic activity of citrate synthase present. For *in vitro* citrate synthase activity measurements, a pre-determined concentration of purified GltA and GltA F383A protein variants were mixed with a specified amount of synthetic VcdP or VcdP* peptides (calculated per monomer of the protein) and used as input for the assay. The enzyme kinetics were assayed at

15 s intervals for at least 30 min and the activity was determined from the initial velocities. The increase in product was exponential over a short period of time (in the order of 2-3 min) before saturation. The change in absorbance per minute was determined within the linear range. The citrate synthase activity was calculated as change in absorbance per minute per milligram of protein input using the extinction coefficient value of $13.6mM^{-1}$ for TNB at 412 nm.

Metabolite measurements using mass-spectrometry

Glycolysis and citric acid cycle metabolites were measured for *V. cholerae* WT, Δ*vcdRP and* Δ*gltA* each harboring an empty control plasmid (pCtrl) and Δ*vcdRP* with pVcdRP expression plasmids (pVcdRP and pVcdP) as well as Δ*gltA* harboring pVcdP. A modified LC-MS method as described previously was used for the measurement and analysis (Buescher *et al*, 2010). Briefly, cells were grown in LB medium to exponential and stationary phase and were subsequently quenched in ice-cold methanol. The quenched samples were centrifuged at 3000 *g* for 10 min at -9°C. The frozen cell pellets were thawed on ice and re-suspended in 500 µl methanol. After three freeze (liquid nitrogen) and thaw cycles, the sample was centrifuged at 16000 *g* and -4 °C for 5 min. The supernatant was transferred into a new tube and stored on ice. The extraction step was repeated, and the supernatants were combined prior to evaporation. The mass spectral data was acquired on a QTRAP 6500+® system coupled on-line with an Agilent 1290 II infinity UPLC system. Each sample was re-suspended in 100 µl Milli-Q water and 10 µl were injected onto a XSelect HSS T3 XP column (2.1 x 150 mm, 2.5 µm, 100 Å). Metabolites were eluted at a flow rate ranging from 0.4 ml/min to 0.15 ml/min with a non-linear gradient. Mobile phase A and B were 10 mM tributylamine, 10 mM acetic acid, 5% methanol and 2% 2-propanol (pH 7.1) in water and 100% 2-propanol, respectively. The autosampler was kept at 5°C and the temperature of the column oven was set to 40°C. Identification and relative quantification were based on specific MRM transitions measured in negative mode electrospray

Sequence alignment and gene synteny analysis

The *vcdRP* gene and *mbrA* gene and their respective genomic loci among various bacterial species were aligned using the MultAlin webtool (Corpet, 1988). Vch: *Vibrio cholerae* (NCBI: txid243277), Vfu: *Vibrio furnissii* (NCBI: txid29494), Vha: *Vibrio harveyi* (NCBI: txid33843), Vpa: *Vibrio parahaemolyticus* (NCBI: txid670), Vvu: *Vibrio vulnificus* (NCBI: txid672), Van: *Vibrio anguillarum* (NCBI: txid55601), Vsp: *Vibrio splendidus* (NCBI: txid575788), Pgb: *Photobacterium gaetbulicola* (NCBI: txid16723), Saly: *Salinivibrio* sp. YCSC6 (NCBI: txid24061), Son: *Shewanella oneidensis* MR-1 (NCBI: txid1082), Shp: *Shewanella putrefaciens* (NCBI: txid435), Pdj: *Pseudoalteromonas donghaensis* (NCBI: txid72335) and Avr: *Aeromonas veronii* (NCBI: txid3586). Gene synteny analysis of the genomic loci encoding *vcdRP* and *mbrA* in various bacterial strains was performed using SynTax (Oberto, 2013).

Quantification and statistical analyses

The statistical parameters including the number of independent biological replicates is indicated in the legend corresponding to each figure. Normality test was performed using Shapiro-Wilk whereas equal variance was determined using Brown-Forsythe test. The data were tested for significant differences using ANOVA and post-hoc Tukey or Dunnett's tests, where appropriate, and the p-values have been indicated against each figure. Statistical analysis was performed using Prism 9.0.0. No blinding or randomization was used in the experiments. Also, no estimation of power analysis was calculated before performing the experiments. Northern and Western blots were visualized using Fiji (Schindelin *et al*, 2012), and the latter were quantified using the BIO-1D software.

Chapter 7

References

Acosta N, Pukatzki S & Raivio TL (2015) The *Vibrio cholerae* Cpx Envelope Stress Response Senses and Mediates Adaptation to Low Iron. *J Bacteriol* 197: 262–276

Adhya S (2003) Suboperonic Regulatory Signals. *Sci Signal* 2003: pe22–pe22

Almagro-Moreno S & Boyd EF (2009) Sialic Acid Catabolism Confers a Competitive Advantage to Pathogenic *Vibrio cholerae* in the Mouse Intestine. *Infection and Immunity* 77: 3807–3816

Anderson DH & Duckworth HW (1988) *In vitro* mutagenesis of *Escherichia coli* citrate synthase to clarify the locations of ligand binding sites. *Journal of Biological Chemistry* 263: 2163–2169

Andresen L & Holmqvist E (2018) CLIP-Seq in Bacteria: Global Recognition Patterns of Bacterial RNA-Binding Proteins. *Methods Enzymol* 612: 127–145

Araújo-Bazán L, Huecas S, Valle J, Andreu D & Andreu JM (2019) Synthetic developmental regulator MciZ targets FtsZ across *Bacillus* species and inhibits bacterial division. *Molecular Microbiology* 111: 965–980

Azam MS & Vanderpool CK (2018) Translational regulation by bacterial small RNAs via an unusual Hfq-dependent mechanism. *Nucleic Acids Res* 46: 2585–2599

Azam MS & Vanderpool CK (2020) Translation inhibition from a distance: The small RNA SgrS silences a ribosomal protein S1-dependent enhancer. *Molecular Microbiology* 114: 391–408

Baek J, Lee J, Yoon K & Lee H (2017) Identification of Unannotated Small Genes in *Salmonella*. *G3 (Bethesda)* 7: 983–989

Bækkedal C & Haugen P (2015) The Spot 42 RNA: A regulatory small RNA with roles in the central metabolism. *RNA Biol* 12: 1071–1077

Baharoglu Z, Krin E & Mazel D (2012) Connecting Environment and Genome Plasticity in the Characterization of Transformation-Induced SOS Regulation and Carbon Catabolite Control of the *Vibrio cholerae* Integron Integrase. *Journal of Bacteriology* 194: 1659–1667

Baker CS, Eöry LA, Yakhnin H, Mercante J, Romeo T & Babitzke P (2007) CsrA Inhibits Translation Initiation of *Escherichia coli hfq* by Binding to a Single Site Overlapping the Shine-Dalgarno Sequence. *Journal of Bacteriology* 189: 5472–5481

Balaban N & Novick RP (1995) Translation of RNAIII, the *Staphylococcus aureus agr* regulatory RNA molecule, can be activated by a 3′-end deletion. *FEMS Microbiol Lett* 133: 155–161

Ball AS, Chaparian RR & van Kessel JC (2017) Quorum Sensing Gene Regulation by LuxR/HapR Master Regulators in *Vibrios*. *J Bacteriol* 199: e00105-17

Balsalobre C, Silván JM, Berglund S, Mizunoe Y, Uhlin BE & Wai SN (2006) Release of the type I secreted α-haemolysin via outer membrane vesicles from *Escherichia coli*. *Molecular Microbiology* 59: 99–112

Banki MR & Wood DW (2005) Inteins and affinity resin substitutes for protein purification and scale up. *Microbial Cell Factories* 4: 32

Bardill JP & Hammer B (2012) Non-coding sRNAs regulate virulence in the bacterial pathogen *Vibrio cholerae*. *RNA Biol* 9: 392–401

Bartholomäus A, Kolte B, Mustafayeva A, Goebel I, Fuchs S, Benndorf D, Engelmann S & Ignatova Z (2021) smORFer: a modular algorithm to detect small ORFs in prokaryotes. *Nucleic Acids Res*: gkab477

Bassler BL (2002) Small Talk: Cell-to-Cell Communication in Bacteria. *Cell* 109: 421–424

Bateman A, Birney E, Cerruti L, Durbin R, Etwiller L, Eddy SR, Griffiths-Jones S, Howe KL, Marshall M & Sonnhammer ELL (2002) The Pfam Protein Families Database. *Nucleic Acids Res* 30: 276–280

Baumann P, Baumann L, Bang SS & Woolkalis MJ (1980) Reevaluation of the taxonomy of *Vibrio*, *Beneckea*, and *Photobacterium*: Abolition of the genus *Beneckea*. *Current Microbiology* 4: 127–132

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

Beisel CL & Storz G (2011) The base pairing RNA Spot 42 participates in a multi-output feedforward loop to help enact catabolite repression in *Escherichia coli*. *Mol Cell* 41: 286–297

Bendezú FO, Hale CA, Bernhardt TG & de Boer PAJ (2009) RodZ (YfgA) is required for proper assembly of the MreB actin cytoskeleton and cell shape in *E. coli*. *EMBO J* 28: 193–204

Benito Y, Kolb FA, Romby P, Lina G, Etienne J & Vandenesch F (2000) Probing the structure of RNAIII, the *Staphylococcus aureus agr* regulatory RNA, and identification of the RNA domain involved in repression of protein A expression. *RNA* 6: 668–679

Beveridge TJ (1999) Structures of gram-negative cell walls and their derived membrane vesicles. *J Bacteriol* 181: 4725–4733

Beyhan S, Tischler AD, Camilli A & Yildiz FH (2006) Transcriptome and Phenotypic Responses of

Vibrio cholerae to Increased Cyclic di-GMP Level. *Journal of Bacteriology* 188: 3600–3613

Bilecen K, Fong JCN, Cheng A, Jones CJ, Zamorano-Sánchez D & Yildiz FH (2015) Polymyxin B Resistance and Biofilm Formation in *Vibrio cholerae* Are Controlled by the Response Regulator CarR. *Infect Immun* 83: 1199–1209

Bina J, Zhu J, Dziejman M, Faruque S, Calderwood S & Mekalanos J (2003) ToxR regulon of *Vibrio cholerae* and its expression in vibrios shed by cholera patients. *PNAS* 100: 2801–2806

Blokesch M (2012) TransFLP — A Method to Genetically Modify *Vibrio cholerae* Based on Natural Transformation and FLP-recombination. *J Vis Exp*: 3761

Blount ZD, Barrick JE, Davidson CJ & Lenski RE (2012) Genomic Analysis of a Key Innovation in an Experimental *E. coli* Population. *Nature* 489: 513–518

Blount ZD, Borland CZ & Lenski RE (2008) Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. *Proc Natl Acad Sci U S A* 105: 7899–7906

Blount ZD, Lenski RE & Losos JB (2018) Contingency and determinism in evolution: Replaying life's tape. *Science* 362: eaam5979

Bobrovskyy M, Vanderpool CK & Richards GR (2015) Small RNAs Regulate Primary and Secondary Metabolism in Gram-negative Bacteria. *Microbiol Spectr* 3

Bohn C, Rigoulay C & Bouloc P (2007) No detectable effect of RNA-binding protein Hfq absence in *Staphylococcus aureus*. *BMC Microbiol* 7: 10

Boin MA, Austin MJ & Häse CC (2004) Chemotaxis in *Vibrio cholerae*. *FEMS Microbiology Letters* 239: 1–8

Bradley ES, Bodi K, Ismail AM & Camilli A (2011) A Genome-Wide Approach to Discovery of Small RNAs Involved in Regulation of Virulence in *Vibrio cholerae*. *PLoS Pathog* 7

Brantl S (2002) Antisense-RNA regulation and RNA interference. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* 1575: 15–25

Brantl S (2007) Regulatory mechanisms employed by *cis*-encoded antisense RNAs. *Current Opinion in Microbiology* 10: 102–109

Brauer MJ, Yuan J, Bennett BD, Lu W, Kimball E, Botstein D & Rabinowitz JD (2006) Conservation of the metabolomic response to starvation across two divergent microbes. *Proc Natl Acad Sci U S A* 103: 19302–19307

Brielle R, Pinel-Marie M-L & Felden B (2016) Linking bacterial type I toxins with their actions. *Current Opinion in Microbiology* 30: 114–121

Brockmeier U, Wendorff M & Eggert T (2006) Versatile Expression and Secretion Vectors for *Bacillus subtilis*. *Curr Microbiol* 52: 143–148

Brown L & Elliott T (1996) Efficient translation of the RpoS sigma factor in *Salmonella typhimurium* requires host factor I, an RNA-binding protein encoded by the *hfq* gene. *Journal of Bacteriology* 178: 3763–3770

Brown RC & Taylor RK (1995) Organization of *tcp*, *acf*, and *toxT* genes within a ToxT-dependent

operon. *Molecular Microbiology* 16: 425–439

Brückner R & Titgemeyer F (2002) Carbon catabolite repression in bacteria: choice of the carbon source and autoregulatory limitation of sugar utilization. *FEMS Microbiology Letters* 209: 141–148

Bryson K, Loux V, Bossy R, Nicolas P, Chaillou S, van de Guchte M, Penaud S, Maguin E, Hoebeke M, Bessières P & J-F Gibrat (2006) AGMIAL: implementing an annotation strategy for prokaryote genomes as a distributed system. *Nucleic Acids Res* 34: 3533–3545

Buescher JM, Moco S, Sauer U & Zamboni N (2010) Ultrahigh Performance Liquid Chromatography−Tandem Mass Spectrometry Method for Fast and Robust Quantification of Anionic and Aromatic Metabolites. *Anal Chem* 82: 4403–4412

Bouloc P & Repoila F (2016) Fresh layers of RNA-mediated regulation in Gram-positive bacteria. *Current Opinion in Microbiology* 30: 30–35

Burkholder PR & Giles NH (1947) Induced biochemical mutations in *Bacillus subtilis*. *Am J Bot* 34: 345–348

Butz HA, Mey AR, Ciosek AL & Payne SM (2019) *Vibrio cholerae* CsrA Directly Regulates varA To Increase Expression of the Three Nonredundant Csr Small RNAs. *mBio* 10

Bzowska A, Kulikowska E & Shugar D (2000) Purine nucleoside phosphorylases: properties, functions, and clinical aspects. *Pharmacol Ther* 88: 349–425

Camacho A, Bouhenia M, Alyusfi R, Alkohlani A, Naji MAM, Radiguès X de, Abubakar AM, Almoalmi A, Seguin C, Sagrado MJ, *et al* (2018) Cholera epidemic in Yemen, 2016–18: an analysis of surveillance data. *The Lancet Global Health* 6: e680–e690

Cameron DE, Urbach JM & Mekalanos JJ (2008) A defined transposon mutant library and its use in identifying motility genes in *Vibrio cholerae*. *Proc Natl Acad Sci U S A* 105: 8736–8741

Carmichael G, Weber K, Niveleau A & Wahba A (1975) The host factor required for RNA phage Qbeta RNA replication *in vitro*. Intracellular location, quantitation, and purification by polyadenylate-cellulose chromatography. *Journal of Biological Chemistry* 250: 3607–3612

Cases I, Velázquez F & de Lorenzo V (2007) The ancestral role of the phosphoenolpyruvate–carbohydrate phosphotransferase system (PTS) as exposed by comparative genomics. *Research in Microbiology* 158: 666–670

Chakrabarti SR, Chaudhuri K, Sen K & Das J (1996) Porins of *Vibrio cholerae*: purification and characterization of OmpU. *J Bacteriol* 178: 524–530

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

Chen J & Gottesman S (2017) Hfq links translation repression to stress-induced mutagenesis in *E. coli*. *Genes Dev* 31: 1382–1395

Chen J, Morita T & Gottesman S (2019) Regulation of Transcription Termination of Small RNAs and by Small RNAs: Molecular Mechanisms and Biological Functions. *Front Cell Infect Microbiol* 9

Childers BM & Klose KE (2007) Regulation of virulence in *Vibrio cholerae*: the ToxR regulon. *Future*

Microbiol 2: 335–344

Chou C-H, Bennett GN & San K-Y (1994) Effect of Modulated Glucose Uptake on High-Level Recombinant Protein Production in a Dense *Escherichia coli* Culture. *Biotechnology Progress* 10: 644–647

Commichau FM, Forchhammer K & Stülke J (2006) Regulatory links between carbon and nitrogen metabolism. *Current Opinion in Microbiology* 9: 167–172

Conner JG, Teschler JK, Jones CJ & Yildiz FH (2016) Staying alive: *Vibrio cholerae*'s cycle of environmental survival, transmission, and dissemination. *Microbiol Spectr* 4

Contreras I, Toro CS, Troncoso G & Mora GCY 1997 (1997) *Salmonella typhi* mutants defective in anaerobic respiration are impaired in their ability to replicate within epithelial cells. *Microbiology* 143: 2665–2672

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

Corpet F (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res* 16: 10881–10890

Cosson P, Perrin J & Bonifacino JS (2013) Anchors aweigh: protein localization and transport mediated by transmembrane domains. *Trends Cell Biol* 23: 511–517

Craig SA, Carpenter CD, Mey AR, Wyckoff EE & Payne SM (2011) Positive Regulation of the *Vibrio cholerae* Porin OmpT by Iron and Fur. *J Bacteriol* 193: 6505–6511

Crawford JA, Kaper JB & DiRita VJ (1998) Analysis of ToxR-dependent transcription activation of *ompU*, the gene encoding a major envelope protein in *Vibrio cholerae*. *Molecular Microbiology* 29: 235–246

Dandekar T, Huynen M, Regula JT, Ueberle B, Zimmermann CU, Andrade MA, Doerks T, Sánchez-Pulido L, Snel B, Suyama M, Yuan YP, Herrmann R & Bork P (2000) Re-annotating the *Mycoplasma pneumoniae* genome sequence: adding value, function and reading frames. *Nucleic Acids Res* 28: 3278–3288

Darty K, Denise A & Ponty Y (2009) VARNA: Interactive drawing and editing of the RNA secondary structure. *Bioinformatics* 25: 1974–1975

Davis AR, Gohara DW & Yap M-NF (2014) Sequence selectivity of macrolide-induced translational attenuation. *Proc Natl Acad Sci U S A* 111: 15379–15384

De Mets F, Van Melderen L & Gottesman S (2019) Regulation of acetate metabolism and coordination with the TCA cycle via a processed small RNA. *Proc Natl Acad Sci USA* 116: 1043–1052

Delmar JA, Su C-C & Yu EW (2014) Bacterial multi-drug efflux transporters. *Annu Rev Biophys* 43: 93–117

Desnoyers G, Morissette A, Prévost K & Massé E (2009) Small RNA-induced differential degradation of the polycistronic mRNA *iscRSUA*. *EMBO J* 28: 1551–1561

Deutscher J, Francke C & Postma PW (2006) How Phosphotransferase System-Related Protein Phosphorylation Regulates Carbohydrate Metabolism in Bacteria. *Microbiol Mol Biol Rev* 70:

939–1031

D'Lima NG, Khitun A, Rosenbloom AD, Yuan P, Gassaway BM, Barber KW, Rinehart J & Slavoff SA (2017) Comparative Proteomics Enables Identification of Nonannotated Cold Shock Proteins in *E. coli*. *J Proteome Res* 16: 3722–3731

Donald LJ, Crane BR, Anderson DH & Duckworth HW (1991) The role of cysteine 206 in allosteric inhibition of *Escherichia coli* citrate synthase. Studies by chemical modification, site-directed mutagenesis, and 19F NMR. *Journal of Biological Chemistry* 266: 20709–20713

Doucette CD, Schwab DJ, Wingreen NS & Rabinowitz JD (2011) α-ketoglutarate coordinates carbon and nitrogen utilization via Enzyme I inhibition. *Nat Chem Biol* 7: 894–901

Dubey AK, Baker CS, Romeo T & Babitzke P (2005) RNA sequence and secondary structure participate in high-affinity CsrA-RNA interaction. *RNA* 11: 1579–1587

Duckworth HW & Tong EK (1976) The binding of reduced nicotinamide adenine dinucleotide to citrate synthase of *Escherichia coli* K12. *Biochemistry* 15: 108–114

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*. *AEM* 72: 802–810

Duperthuy M, Sjöström AE, Sabharwal D, Damghani F, Uhlin BE & Wai SN (2013) Role of the *Vibrio cholerae* Matrix Protein Bap1 in Cross-Resistance to Antimicrobial Peptides. *PLoS Pathog* 9

Durand S, Braun F, Lioliou E, Romilly C, Helfer A-C, Kuhn L, Quittot N, Nicolas P, Romby P & Condon C (2015) A Nitric Oxide Regulated Small RNA Controls Expression of Genes Involved in Redox Homeostasis in *Bacillus subtilis*. *PLOS Genetics* 11: e1004957

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

Durica-Mitic S, Göpel Y, Amman F & Görke B (2020) Adaptor protein RapZ activates endoribonuclease RNase E by protein–protein interaction to cleave a small regulatory RNA. *RNA* 26: 1198–1215

Durica-Mitic S, Göpel Y & Görke B (2018) Carbohydrate Utilization in Bacteria: Making the Most Out of Sugars with the Help of Small Regulatory RNAs. *Microbiology Spectrum* 6: 6.2.07

Duss O, Michel E, Diarra dit Konté N, Schubert M & Allain FH-T (2014) Molecular basis for the wide range of affinity found in Csr/Rsm protein-RNA recognition. *Nucleic Acids Res* 42: 5332–5346

Dutta T & Srivastava S (2018) Small RNA-mediated regulation in bacteria: A growing palette of diverse mechanisms. *Gene* 656: 60–72

Echazarreta MA & Klose KE (2019) *Vibrio* Flagellar Synthesis. *Front Cell Infect Microbiol* 0

Eddy SR (2002) Computational Genomics of Noncoding RNA Genes. *Cell* 109: 137–140

ENCODE Project Consortium (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* 489: 57–74

Engel F, Ossipova E, Jakobsson P-J, Vockenhuber M-P & Suess B (2020) sRNA scr5239 Involved in Feedback Loop Regulation of *Streptomyces coelicolor* Central Metabolism. *Front Microbiol* 10

Ensign SA (2006) Revisiting the glyoxylate cycle: alternate pathways for microbial acetate assimilation. *Molecular Microbiology* 61: 274–276

Escolar L, Pérez-Martín J & de Lorenzo V (1999) Opening the iron box: transcriptional metalloregulation by the Fur protein. *J Bacteriol* 181: 6223–6229

Evans DJ & Richardson SH (1968) In Vitro Production of Choleragen and Vascular Permeability Factor by *Vibrio cholerae*. *Journal of Bacteriology* 96: 126–130

Federle MJ & Bassler BL (2003) Interspecies communication in bacteria. *J Clin Invest* 112: 1291–1299

Fiebig A, Herrou J, Willett J & Crosson S (2015) General Stress Signaling in the Alphaproteobacteria. *Annu Rev Genet* 49: 603–625

Field M (1978) Cholera Toxin, Adenylate Cyclase, and the Process of Active Secretion in the Small Intestine: The Pathogenesis of Diarrhea in Cholera. In *Physiology of Membrane Disorders*, Andreoli TE Hoffman JF & Fanestil DD (eds) pp 877–899. Boston, MA: Springer US

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

Finkelstein RA, Atthasampunna P, Chulasamaya M & Charunmethee P (1966) Pathogenesis of Experimental Cholera: Biologic Activities of Purified Procholeragen A. *The Journal of Immunology* 96: 440–449

Floyd KA, Lee CK, Xian W, Nametalla M, Valentine A, Crair B, Zhu S, Hughes HQ, Chlebek JL, Wu DC, Hwan Park J, Farhat AM, Lomba CJ, Ellison CK, Brun YV, Campos-Gomez J, Dalia AB, Liu J, Biais N, Wong GCL & Yildiz FH (2020) c-di-GMP modulates type IV MSHA pilus retraction and surface attachment in *Vibrio cholerae*. *Nature Communications* 11: 1549

Fong JCN, Karplus K, Schoolnik GK & Yildiz FH (2006) Identification and Characterization of RbmA, a Novel Protein Required for the Development of Rugose Colony Morphology and Biofilm Structure in *Vibrio cholerae*. *Journal of Bacteriology* 188: 1049–1059

Franze de Fernandez MT, Eoyang L & August JT (1968) Factor Fraction required for the Synthesis of Bacteriophage Qβ-RNA. *Nature* 219: 588–590

Friedman RC, Kalkhof S, Doppelt-Azeroual O, Mueller SA, Chovancová M, von Bergen M & Schwikowski B (2017) Common and phylogenetically widespread coding for peptides by bacterial small RNAs. *BMC Genomics* 18: 553

Fröhlich KS, Förstner KU & Gitai Z (2018) Post-transcriptional gene regulation by an Hfq-independent small RNA in *Caulobacter crescentus*. *Nucleic Acids Research* 46: 10969–10982

Fröhlich KS & Papenfort K (2020) Regulation outside the box: New mechanisms for small RNAs. *Molecular Microbiology* 114: 363–366

Fröhlich KS, Papenfort K, Berger AA & Vogel J (2012) A conserved RpoS-dependent small RNA controls the synthesis of major porin OmpD. *Nucleic Acids Res* 40: 3623–3640

Fröhlich KS & Vogel J (2009) Activation of gene expression by small RNA. *Curr Opin Microbiol* 12: 674–682

Fu H, Elena RC & Marquez PH (2019) The roles of small RNAs: insights from bacterial quorum sensing. *ExRNA* 1: 32

Gallego-Hernandez AL, DePas WH, Park JH, Teschler JK, Hartmann R, Jeckel H, Drescher K, Beyhan S, Newman DK & Yildiz FH (2020) Upregulation of virulence genes promotes *Vibrio cholerae* biofilm hyperinfectivity. *PNAS* 117: 11010–11017

Gangarosa EJ, Bennett JV & Boring JR (1967) Differentiation between *Vibrio cholerae* and *Vibrio cholerae* biotype El Tor by the polymyxin B disc test: comparative results with TCBS, Monsur's, Mueller-Hinton and nutrient agar media. *Bull World Health Organ* 36: 987–990

Gao H, Ma L, Qin Q, Qiu Y, Zhang J, Li J, Lou J, Diao B, Zhao H, Shi Q, Zhang Y &Kan B (2020) Fur Represses *Vibrio cholerae* Biofilm Formation via Direct Regulation of *vieSAB, cdgD, vpsU*, and *vpsA-K* Transcription. *Front Microbiol* 11

Gao H, Xu J, Lu X, Li J, Lou J, Zhao H, Diao B, Shi Q, Zhang Y & Kan B (2018) Expression of Hemolysin Is Regulated Under the Collective Actions of HapR, Fur, and HlyU in *Vibrio cholerae* El Tor Serogroup O1. *Front Microbiol* 9

Garai P & Blanc-Potard A (2020) Uncovering small membrane proteins in pathogenic bacteria: Regulatory functions and therapeutic potential. *Molecular Microbiology* 114: 710–720

Gassel M, Möllenkamp T, Puppe W & Altendorf K (1999) The KdpF subunit is part of the K(+)-translocating Kdp complex of *Escherichia coli* and is responsible for stabilization of the complex *in vitro*. *J Biol Chem* 274: 37901–37907

Geissmann TA & Touati D (2004) Hfq, a new chaperoning role: binding to messenger RNA determines access for small RNA regulator. *The EMBO Journal* 23: 396–405

Germain-Amiot N, Augagneur Y, Camberlein E, Nicolas I, Lecureur V, Rouillon A & Felden B (2019) A novel *Staphylococcus aureus cis–trans* type I toxin–antitoxin module with dual effects on bacteria and host cells. *Nucleic Acids Research* 47: 1759–1773

Gerovac M, Wicke L, Chihara K, Schneider C, Lavigne R & Vogel J (2020) A Grad-seq view of RNA and protein complexes in *Pseudomonas aeruginosa* under standard and bacteriophage predation conditions. *bioRxiv*: 2020.12.06.403469

Giglio KM, Fong JC, Yildiz FH & Sondermann H (2013) Structural Basis for Biofilm Formation via the *Vibrio cholerae* Matrix Protein RbmA. *Journal of Bacteriology* 195: 3277–3286

Gimpel M, Preis H, Barth E, Gramzow L & Brantl S (2012) SR1—a small RNA with two remarkably conserved functions. *Nucleic Acids Research* 40: 11659–11672

Goli B & Nair AS (2012) The elusive short gene – an ensemble method for recognition for prokaryotic genome. *Biochemical and Biophysical Research Communications* 422: 36–41

Göpel Y & Görke B (2014) Lies and deception in bacterial gene regulation: the roles of nucleic acid decoys. *Molecular Microbiology* 92: 641–647

Göpel Y, Lüttmann D, Heroven AK, Reichenbach B, Dersch P & Görke B (2011) Common and divergent features in transcriptional control of the homologous small RNAs GlmY and GlmZ in *Enterobacteriaceae*. *Nucleic Acids Res* 39: 1294–1309

Gorke B & Vogel J (2008) Noncoding RNA control of the making and breaking of sugars. *Genes & Development* 22: 2914–2925

Gottesman S & Storz G (2011) Bacterial Small RNA Regulators: Versatile Roles and Rapidly Evolving Variations. *Cold Spring Harb Perspect Biol* 3: a003798

Goyal A, Belardinelli R & Rodnina MV (2017) Non-canonical Binding Site for Bacterial Initiation Factor 3 on the Large Ribosomal Subunit. *Cell Reports* 20: 3113–3122

Green J, Stapleton MR, Smith LJ, Artymiuk PJ, Kahramanoglou C, Hunt DM & Buxton RS (2014) Cyclic-AMP and bacterial cyclic-AMP receptor proteins revisited: adaptation for different ecological niches. *Curr Opin Microbiol* 18: 1–7

Green MR, Sambrook J & Sambrook J (2012) Molecular cloning: a laboratory manual 4th ed. Cold Spring Harbor, N.Y: Cold Spring Harbor Laboratory Press

Gregory GJ, Morreale DP & Boyd EF (2020) CosR Is a Global Regulator of the Osmotic Stress Response with Widespread Distribution among Bacteria. *Applied and Environmental Microbiology* 86: e00120-20

Groisman EA, Kayser J & Soncini FC (1997) Regulation of polymyxin resistance and adaptation to low-Mg^{2+} environments. *J Bacteriol* 179: 7040–7045

Gruber CC & Sperandio V (2015) Global Analysis of Posttranscriptional Regulation by GlmY and GlmZ in Enterohemorrhagic *Escherichia coli* O157:H7. *Infection and Immunity* 83: 1286–1295

Gupta RK, Luong TT & Lee CY (2015) RNAIII of the *Staphylococcus aureus agr* system activates global regulator MgrA by stabilizing mRNA. *Proc Natl Acad Sci U S A* 112: 14036–14041

Hacker J, Blum-Oehler G, Mühldorfer I & Tschäpe H (1997) Pathogenicity islands of virulent bacteria: structure, function and impact on microbial evolution. *Molecular Microbiology* 23: 1089–1097

Hall BG (1982) Chromosomal mutation for citrate utilization by *Escherichia coli* K-12. *J Bacteriol* 151: 269–273

Hamann K, Zimmann P & Altendorf K (2008) Reduction of turgor is not the stimulus for the sensor kinase KdpD of *Escherichia coli*. *J Bacteriol* 190: 2360–2367

Handford PA, Ner SS, Bloxham DP & Wilton DC (1988) Site-directed mutagenesis of citrate synthase; the role of the active-site aspartate in the binding of acetyl-CoA but not oxaloacetate. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* 953: 232–240

Handler AA, Lim JE & Losick R (2008) Peptide inhibitor of cytokinesis during sporulation in *Bacillus subtilis*. *Molecular Microbiology* 68: 588–599

Hankins JV, Madsen JA, Giles DK, Brodbelt JS & Trent MS (2012) Amino acid addition to *Vibrio cholerae* LPS establishes a link between surface remodeling in Gram-positive and Gram-negative bacteria. *Proc Natl Acad Sci U S A* 109: 8722–8727

Hankins JV, Madsen JA, Giles DK, Childers BM, Klose KE, Brodbelt JS & Trent MS (2011) Elucidation of a novel *Vibrio cholerae* lipid A secondary hydroxy-acyltransferase and its role in innate immune recognition. *Mol Microbiol* 81: 1313–1329

Hassan H & Troxell B (2013) Transcriptional regulation by Ferric Uptake Regulator (Fur) in pathogenic bacteria. *Frontiers in Cellular and Infection Microbiology* 3: 59

Hayes CA, Dalia TN & Dalia AB (2017) Systematic genetic dissection of PTS in *Vibrio cholerae* uncovers a novel glucose transporter and a limited role for PTS during infection of a mammalian host. *Molecular Microbiology* 104: 568–579

Hébrard M, Kröger C, Srikumar S, Colgan A, Händler K & Hinton J (2012) sRNAs and the virulence

of *Salmonella enterica* serovar Typhimurium. *RNA Biol* 9: 437–445

Heidelberg JF, Eisen 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, Qin H, Dragoi I, Sellers P, McDonald L, Utterback T, Fleishmann RD, Nierman WC, White O, Salzberg SL, Smith HO, Colwell RR, Mekalanos JJ, Venter JC & Fraser CM (2000) DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature* 406: 477–483

Heidrich N, Chinali A, Gerth U & Brantl S (2006) The small untranslated RNA SR1 from the *Bacillus subtilis* genome is involved in the regulation of arginine catabolism. *Molecular Microbiology* 62: 520–536

Heidrich N, Moll I & Brantl S (2007) *In vitro* analysis of the interaction between the small RNA SR1 and its primary target ahrC mRNA. *Nucleic Acids Res* 35: 4331–4346

Hemm MR, Paul BJ, Miranda-Ríos J, Zhang A, Soltanzad N & Storz G (2010) Small Stress Response Proteins in *Escherichia coli*: Proteins Missed by Classical Proteomic Studies. *J Bacteriol* 192: 46–58

Hemm MR, Paul BJ, Schneider TD, Storz G & Rudd KE (2008) Small membrane proteins found by comparative genomics and ribosome binding site models. *Mol Microbiol* 70: 1487–1501

Henderson JC, Fage CD, Cannon JR, Brodbelt JS, Keatinge-Clay AT & Trent MS (2014) Antimicrobial Peptide Resistance of *Vibrio cholerae* Results from an LPS Modification Pathway Related to Nonribosomal Peptide Synthetases. *ACS Chem Biol* 9: 2382–2392

Henke JM & Bassler BL (2004) Three Parallel Quorum-Sensing Systems Regulate Gene Expression in *Vibrio harveyi*. *J Bacteriol* 186: 6902–6914

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

Hogema BM, Arents JC, Bader R, Eijkemans K, Yoshida H, Takahashi H, Aiba H & Postma PW (1998) Inducer exclusion in *Escherichia coli* by non-PTS substrates: the role of the PEP to pyruvate ratio in determining the phosphorylation state of enzyme IIAGlc. *Molecular Microbiology* 30: 487–498

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

Holmqvist E & Vogel J (2018) RNA-binding proteins in bacteria. *Nature Reviews Microbiology* 16: 601–615

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

Hopper DJ & Cooper RA (1971) The regulation of *Escherichia coli* methylglyoxal synthase; a new control site in glycolysis? *FEBS Letters* 13: 213–216

Hör J, Di Giorgio S, Gerovac M, Venturini E, Förstner KU & Vogel J (2020a) Grad-seq shines light on unrecognized RNA and protein complexes in the model bacterium *Escherichia coli*. *Nucleic*

Acids Research 48: 9301–9319

Hör J, Garriss G, Di Giorgio S, Hack L-M, Vanselow JT, Förstner KU, Schlosser A, Henriques-Normark B & Vogel J (2020b) Grad-seq in a Gram-positive bacterium reveals exonucleolytic sRNA activation in competence control. *The EMBO Journal* 39: e103852

Hör J, Gorski SA & Vogel J (2018) Bacterial RNA Biology on a Genome Scale. *Molecular Cell* 70: 785–799

Hori K & Yanazaki Y (1974) Nucleotide sequence specific interaction of host factor I with bacteriophage Qβ RNA. *FEBS Letters* 43: 20–22

Horler RSP & Vanderpool CK (2009) Homologs of the small RNA SgrS are broadly distributed in enteric bacteria but have diverged in size and sequence. *Nucleic Acids Res* 37: 5465–5476

Houot L, Chang S, Pickering BS, Absalon C & Watnick PI (2010) The Phosphoenolpyruvate Phosphotransferase System Regulates *Vibrio cholerae* Biofilm Formation through Multiple Independent Pathways. *JB* 192: 3055–3067

Houot L & Watnick PI (2008) A Novel Role for Enzyme I of the *Vibrio cholerae* Phosphoenolpyruvate Phosphotransferase System in Regulation of Growth in a Biofilm. *J Bacteriol* 190: 311–320

Hoyos M, Huber M, Förstner KU & Papenfort K (2020) Gene autoregulation by 3′ UTR-derived bacterial small RNAs. *eLife* 9: e58836

Hu D, Liu B, Feng L, Ding P, Guo X, Wang M, Cao B, Reeves PR & Wang L (2016) Origins of the current seventh cholera pandemic. *PNAS* 113: E7730–E7739

Huang DW, Sherman BT & Lempicki RA (2009a) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4: 44–57

Huang DW, Sherman BT & Lempicki RA (2009b) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* 37: 1–13

Huber M, Fröhlich KS, Radmer J & Papenfort K (2020) Switching fatty acid metabolism by an RNA-controlled feed forward loop. *Proc Natl Acad Sci USA* 117: 8044–8054

Hücker SM, Ardern Z, Goldberg T, Schafferhans A, Bernhofer M, Vestergaard G, Nelson CW, Schloter M, Rost B, Scherer S & Neuhaus K (2017) Discovery of numerous novel small genes in the intergenic regions of the *Escherichia coli* O157:H7 Sakai genome. *PLOS ONE* 12: e0184119

Huergo LF & Dixon R (2015) The Emergence of 2-Oxoglutarate as a Master Regulator Metabolite. *Microbiol Mol Biol Rev* 79: 419–435

Hughes CS, Moggridge S, Müller T, Sorensen P, Morin G & Krijgsveld J (2019) Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nature Protocols* 14(1):68-85.

Huq A, Sack RB, Nizam A, Longini IM, Nair GB, Ali A, Morris JG, Khan MNH, Siddique AK, Yunus M, Albert MJ, Sack DA & Colwell RR (2005) Critical factors influencing the occurrence of *Vibrio cholerae* in the environment of Bangladesh. *Appl Environ Microbiol* 71: 4645–4654

Ingolia NT, Ghaemmaghami S, Newman JRS & Weissman JS (2009) Genome-Wide Analysis *in Vivo* of Translation with Nucleotide Resolution Using Ribosome Profiling. *Science* 324: 218–223

Ingolia NT, Lareau LF & Weissman JS (2011) Ribosome Profiling of Mouse Embryonic Stem Cells Reveals the Complexity and Dynamics of Mammalian Proteomes. *Cell* 147: 789–802

Irani MH & Maitra PK (1977) Properties of *Escherichia coli* Mutants Deficient in Enzymes of Glycolysis. *Journal of Bacteriology* 132: 398–410

Irastortza-Olaziregi M & Amster-Choder O (2021) RNA localization in prokaryotes: Where, when, how, and why. *WIREs RNA* 12: e1615

Islam MS, Zaman MH, Islam MS, Ahmed N & Clemens JD (2020) Environmental reservoirs of *Vibrio cholerae*. *Vaccine* 38: A52–A62

Iwanaga M & Kuyyakanond T (1987) Large production of cholera toxin by *Vibrio cholerae* O1 in yeast extract peptone water. *J Clin Microbiol* 25: 2314–2316

Iwanaga M, Yamamoto K, Higa N, Ichinose Y, Nakasone N & Tanabe M (1986) Culture Conditions for Stimulating Cholera Toxin Production by *Vibrio cholerae* O1 El Tor. *Microbiology and Immunology* 30: 1075–1083

Jaffe JD, Berg HC & Church GM (2004) Proteogenomic mapping as a complementary method to perform genome annotation. *Proteomics* 4: 59–77

Jakes KS & Finkelstein A (2010) The colicin Ia receptor, Cir, is also the translocator for colicin Ia. *Molecular Microbiology* 75: 567–578

Jan AT (2017) Outer Membrane Vesicles (OMVs) of Gram-negative Bacteria: A Perspective Update. *Front Microbiol* 8

Jeong Y, Kim J-N, Kim MW, Bucca G, Cho S, Yoon YJ, Kim B-G, Roe J-H, Kim SC, Smith CP & Cho BK (2016) The dynamic transcriptional and translational landscape of the model antibiotic producer *Streptomyces coelicolor* A3(2). *Nat Commun* 7

Jermyn WS & Boyd EF 2002 (2002) Characterization of a novel *Vibrio* pathogenicity island (VPI-2) encoding neuraminidase (*nanH*) among toxigenic *Vibrio cholerae* isolates. *Microbiology* 148: 3681–3693

Jones CJ, Utada A, Davis KR, Thongsomboon W, Sanchez DZ, Banakar V, Cegelski L, Wong GCL & Yildiz FH (2015) C-di-GMP Regulates Motile to Sessile Transition by Modulating MshA Pili Biogenesis and Near-Surface Motility Behavior in *Vibrio cholerae*. *PLOS Pathogens* 11: e1005068

Jordan PA, Thomson AJ, Ralph ET, Guest JR & Green J (1997) FNR is a direct oxygen sensor having a biphasic response curve. *FEBS Letters* 416: 349–352

Jørgensen MG, Thomason MK, Havelund J, Valentin-Hansen P & Storz G (2013) Dual function of the McaS small RNA in controlling biofilm formation. *Genes Dev* 27: 1132–1145

Kadner RJ, Murphy GP & Stephens CMY 1992 (1992) Two mechanisms for growth inhibition by elevated transport of sugar phosphates in *Escherichia coli*. *Microbiology* 138: 2007–2014

Kaito C, Saito Y, Ikuo M, Omae Y, Mao H, Nagano G, Fujiyuki T, Numata S, Han X, Obata K, Hasegawa S, Yamaguchi H, Inokuchi K, Ito T, Hiramatsu K & Sekimizu K (2013) Mobile genetic element SCCmec-encoded *psm-mec* RNA suppresses translation of *agrA* and attenuates MRSA virulence. *PLoS Pathog* 9: e1003269

Kaito C, Saito Y, Nagano G, Ikuo M, Omae Y, Hanada Y, Han X, Kuwahara-Arai K, Hishinuma T, Baba T, Ito T, Hiramatsu K & Sekimizu K (2011) Transcription and translation products of the cytolysin gene *psm-mec* on the mobile genetic element SCCmec regulate *Staphylococcus aureus* virulence. *PLoS Pathog* 7: e1001267

Kalamorz F, Reichenbach B, März W, Rak B & Görke B (2007) Feedback control of glucosamine-6-phosphate synthase GlmS expression depends on the small RNA GlmZ and involves the novel protein YhbJ in *Escherichia coli*. *Molecular Microbiology* 65: 1518–1533

Kamp HD & Higgins DE (2011) A protein thermometer controls temperature-dependent transcription of flagellar motility genes in *Listeria monocytogenes*. *PLoS Pathog* 7: e1002153

Kamp HD, Patimalla-Dipali B, Lazinski DW, Wallace-Gadsden F & Camilli A (2013) Gene Fitness Landscapes of *Vibrio cholerae* at Important Stages of Its Life Cycle. *PLOS Pathogens* 9: e1003800

Kanehisa M & Sato Y (2020) KEGG Mapper for inferring cellular functions from protein sequences. *Protein Sci* 29: 28–35

Kanjilal S, Citorik R, LaRocque RC, Ramoni MF & Calderwood SB (2010) A Systems Biology Approach To Modeling *Vibrio cholerae* Gene Expression under Virulence-Inducing Conditions. *Journal of Bacteriology* 192: 4300–4310

Kaper JB, Morris JG & Levine MM (1995) Cholera. *Clin Microbiol Rev* 8: 48–86

Karaolis DKR, Somara S, Maneval DR, Johnson JA & Kaper JB (1999) A bacteriophage encoding a pathogenicity island, a type-IV pilus and a phage receptor in cholera bacteria. *Nature* 399: 375–379

Kazi MI, Conrado AR, Mey AR, Payne SM & Davies BW (2016) ToxR Antagonizes H-NS Regulation of Horizontally Acquired Genes to Drive Host Colonization. *PLOS Pathogens* 12: e1005570

Keiler KC (2011) RNA localization in bacteria. *Current Opinion in Microbiology* 14: 155–159

Kelley JT & Parker CD (1981) Identification and preliminary characterization of *Vibrio cholerae* outer membrane proteins. *J Bacteriol* 145: 1018–1024

Kempf B & Bremer E (1998) Uptake and synthesis of compatible solutes as microbial stress responses to high-osmolality environments. *Archives of Microbiology* 170: 319–330

Khan MA, Durica-Mitic S, Göpel Y, Heermann R & Görke B (2020) Small RNA-binding protein RapZ mediates cell envelope precursor sensing and signaling in *Escherichia coli*. *The EMBO Journal* 39: e103848

Kimata K, Takahashi H, Inada T, Postma P & Aiba H (1997) cAMP receptor protein–cAMP plays a crucial role in glucose–lactose diauxie by activating the major glucose transporter gene in *Escherichia coli*. *Proc Natl Acad Sci U S A* 94: 12914–12919

Klose KE (2001) Regulation of virulence in *Vibrio cholerae*. *International Journal of Medical Microbiology* 291: 81–88

Koendjbiharie JG, van Kranenburg R & Kengen SWM (2021) The PEP-pyruvate-oxaloacetate node: variation at the heart of metabolism. *FEMS Microbiology Reviews* 45

Korla K & Mitra CK (2014) Modelling the Krebs cycle and oxidative phosphorylation. *Journal of*

Biomolecular Structure and Dynamics 32: 242–256

Kouokam JC, Wai SN, Fällman M, Dobrindt U, Hacker J & Uhlin BE (2006) Active Cytotoxic Necrotizing Factor 1 Associated with Outer Membrane Vesicles from Uropathogenic *Escherichia coli*. *Infect Immun* 74: 2022–2030

Kovacikova G & Skorupski K (1999) A *Vibrio cholerae* LysR Homolog, AphB, Cooperates with AphA at the *tcpPH* Promoter To Activate Expression of the ToxR Virulence Cascade. *J Bacteriol* 181: 4250–4256

Kovacikova G & Skorupski K (2002) Regulation of virulence gene expression in *Vibrio cholerae* by quorum sensing: HapR functions at the *aphA* promoter. *Molecular Microbiology* 46: 1135–1147

Kowalinski E, Bange G, Wild K & Sinning I (2007) Expression, purification, crystallization and preliminary crystallographic analysis of the proliferation-associated protein Ebp1. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 63: 768–770

Krebs HA (1937) The Intermediate metabolism of Carbohydrates. *The Lancet* 230: 736–738

Kulp A & Kuehn MJ (2010) Biological Functions and Biogenesis of Secreted Bacterial Outer Membrane Vesicles. *Annu Rev Microbiol* 64: 163–184

Kumar K, Desai V, Cheng L, Khitrov M, Grover D, Satya RV, Yu C, Zavaljevski N & Reifman J (2011) AGeS: A Software System for Microbial Genome Sequence Annotation. *PLoS One* 6

Küper C & Jung K (2005) CadC-Mediated Activation of the *cadBA* Promoter in *Escherichia coli*. *MIP* 10: 26–39

Kusama H & Craig JP (1970) Production of Biologically Active Substances by Two Strains of *Vibrio cholerae*. *Infect Immun* 1: 80–87

Lalaouna D, Carrier M-C, Semsey S, Brouard J-S, Wang J, Wade JT & Massé E (2015) A 3′ External Transcribed Spacer in a tRNA Transcript Acts as a Sponge for Small RNAs to Prevent Transcriptional Noise. *Molecular Cell* 58: 393–405

LaNoue KF, Bryla J & Williamson JR (1972) Feedback Interactions in the Control of Citric Acid Cycle Activity in Rat Heart Mitochondria. *Journal of Biological Chemistry* 247: 667–679

LaRocque RC, Harris JB, Dziejman M, Li X, Khan AI, Faruque ASG, Faruque SM, Nair GB, Ryan ET, Qadri F, Mekalanos JJ & Calderwood SB (2005) Transcriptional Profiling of *Vibrio cholerae* Recovered Directly from Patient Specimens during Early and Late Stages of Human Infection. *Infect Immun* 73: 4488–4493

Lee AT & Cerami A (1987) Elevated glucose 6-phosphate levels are associated with plasmid mutations *in vivo*. *PNAS* 84: 8311–8314

Lee C-R, Park Y-H, Kim M, Kim Y-R, Park S, Peterkofsky A & Seok Y-J (2013) Reciprocal regulation of the Autophosphorylation of Enzyme I^{Ntr} by Glutamine and α-Ketoglutarate in *Escherichia coli*. *Mol Microbiol* 88: 473–485

Lee S, Liu B, Lee S, Huang S-X, Shen B & Qian S-B (2012) Global mapping of translation initiation sites in mammalian cells at single-nucleotide resolution. *PNAS* 109: E2424–E2432

Lee SH, Hava DL, Waldor MK & Camilli A (1999) Regulation and Temporal Expression Patterns of *Vibrio cholerae* Virulence Genes during Infection. *Cell* 99: 625–634

Lekshmi N, Joseph I, Ramamurthy T & Thomas S (2018) Changing facades of *Vibrio cholerae*: An enigma in the epidemiology of cholera. *Indian J Med Res* 147: 133–141

Lenski RE (2017) Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations. *ISME J* 11: 2181–2194

Lenski RE, Rose MR, Simpson SC & Tadler SC (1991) Long-Term Experimental Evolution in *Escherichia coli*. I. Adaptation and Divergence During 2,000 Generations. *The American Naturalist* 138: 1315–1341

Lenz DH, Miller MB, Zhu J, Kulkarni RV & Bassler BL (2005) CsrA and three redundant small RNAs regulate quorum sensing in *Vibrio cholerae*. *Molecular Microbiology* 58: 1186–1202

Lenz DH, Mok KC, Lilley BN, Kulkarni RV, Wingreen NS & Bassler BL (2004) The Small RNA Chaperone Hfq and Multiple Small RNAs Control Quorum Sensing in *Vibrio harveyi* and *Vibrio cholerae*. *Cell* 118: 69–82

Li CC, Crawford JA, DiRita VJ & Kaper JB (2000) Molecular cloning and transcriptional regulation of *ompT*, a ToxR-repressed gene in *Vibrio cholerae*. *Molecular Microbiology* 35: 189–203

Li L & Chao Y (2020) sPepFinder expedites genome-wide identification of small proteins in bacteria. *bioRxiv*: 2020.05.05.079178

Liao C-H, Yao L & Ye B-C (2014) Three genes encoding citrate synthases in *Saccharopolyspora erythraea* are regulated by the global nutrient-sensing regulators GlnR, DasR, and CRP. *Molecular Microbiology* 94: 1065–1084

Licht A, Preis S & Brantl S (2005) Implication of CcpN in the regulation of a novel untranslated RNA (SR1) in *Bacillus subtilis*. *Mol Microbiol* 58: 189–206

Lichty JJ, Malecki JL, Agnew HD, Michelson-Horowitz DJ & Tan S (2005) Comparison of affinity tags for protein purification. *Protein Expression and Purification* 41: 98–105

Liu X, Jiang H, Gu Z & Roberts JW (2013) High-resolution view of bacteriophage *lambda* gene expression by ribosome profiling. *Proc Natl Acad Sci U S A* 110: 11928–11933

Liu Z, Hsiao A, Joelsson A & Zhu J (2006) The Transcriptional Regulator VqmA Increases Expression of the Quorum-Sensing Activator HapR in *Vibrio cholerae*. *J Bacteriol* 188: 2446–2453

Liu Z, Miyashiro T, Tsou A, Hsiao A, Goulian M & Zhu J (2008) Mucosal penetration primes *Vibrio cholerae* for host colonization by repressing quorum sensing. *Proc Natl Acad Sci U S A* 105: 9769–9774

Lloyd CR, Park S, Fei J & Vanderpool CK (2017) The Small Protein SgrT Controls Transport Activity of the Glucose-Specific Phosphotransferase System. *Journal of Bacteriology* 199

Lo Scrudato M & Blokesch M (2013) A transcriptional regulator linking quorum sensing and chitin induction to render *Vibrio cholerae* naturally transformable. *Nucleic Acids Res* 41: 3644–3658

Loh E, Dussurget O, Gripenland J, Vaitkevicius K, Tiensuu T, Mandin P, Repoila F, Buchrieser C, Cossart P & Johansson J (2009) A *trans*-Acting Riboswitch Controls Expression of the Virulence Regulator PrfA in *Listeria monocytogenes*. *Cell* 139: 770–779

Loui C, Chang AC & Lu S (2009) Role of the ArcAB two-component system in the resistance of *Escherichia coli* to reactive oxygen stress. *BMC Microbiol* 9: 183

Louis VR, Russek-Cohen E, Choopun N, Rivera ING, Gangle B, Jiang SC, Rubin A, Patz JA, Huq A & Colwell RR (2003) Predictability of *Vibrio cholerae* in Chesapeake Bay. *Appl Environ Microbiol* 69: 2773–2785

Love MI, Huber W & Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15: 550

Mahlapuu M, Håkansson J, Ringstad L & Björn C (2016) Antimicrobial Peptides: An Emerging Category of Therapeutic Agents. *Frontiers in Cellular and Infection Microbiology* 6: 194

Majdalani N, Chen S, Murrow J, John KS & Gottesman S (2001) Regulation of RpoS by a novel small RNA: the characterization of RprA. *Molecular Microbiology* 39: 1382–1394

Maloney PC (1994) Bacterial transporters. *Current Opinion in Cell Biology* 6: 571–582

Mandlik A, Livny J, Robins WP, Ritchie JM, Mekalanos JJ & Waldor MK (2011) RNA-seq-based monitoring of infection-linked changes in *Vibrio cholerae* gene expression. *Cell Host Microbe* 10: 165–174

Mangold M, Siller M, Roppenser B, Vlaminckx BJM, Penfound TA, Klein R, Novak R, Novick RP & Charpentier E (2004) Synthesis of group A streptococcal virulence factors is controlled by a regulatory RNA molecule. *Mol Microbiol* 53: 1515–1527

Manneh-Roussel J, Haycocks JRJ, Magán A, Perez-Soto N, Voelz K, Camilli A, Krachler A-M & Grainger DC (2018) cAMP Receptor Protein Controls *Vibrio cholerae* Gene Expression in Response to Host Colonization. *mBio* 9

Marden JN, Diaz MR, Walton WG, Gode CJ, Betts L, Urbanowski ML, Redinbo MR, Yahr TL & Wolfgang MC (2013) An unusual CsrA family member operates in series with RsmA to amplify posttranscriptional responses in *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A* 110: 15055–15060

Maris C, Dominguez C & Allain FH-T (2005) The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *The FEBS Journal* 272: 2118–2131

Mark Glover JN, Chaulk SG, Edwards RA, Arthur D, Lu J & Frost LS (2015) The FinO family of bacterial RNA chaperones. *Plasmid* 78: 79–87

Mars RAT, Mendonça K, Denham EL & van Dijl JM (2015) The reduction in small ribosomal subunit abundance in ethanol-stressed cells of *Bacillus subtilis* is mediated by a SigB-dependent antisense RNA. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1853: 2553–2559

Martin JE, Waters LS, Storz G & Imlay JA (2015) The *Escherichia coli* Small Protein MntS and Exporter MntP Optimize the Intracellular Concentration of Manganese. *PLOS Genetics* 11: e1004977

Martin WF (2020) Older Than Genes: The Acetyl CoA Pathway and Origins. *Front Microbiol* 11

Maruyama K, Sato N & Ohta N (1999) Conservation of structure and cold-regulation of RNA-binding proteins in cyanobacteria: Probable convergent evolution with eukaryotic glycine-rich RNA-binding proteins. *Nucleic Acids Research* 27: 2029–2036

Massé E & Gottesman S (2002) A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*. *Proc Natl Acad Sci U S A* 99: 4620–4625

Massé E, Vanderpool CK & Gottesman S (2005a) Effect of RyhB Small RNA on Global Iron Use in

Escherichia coli. *J Bacteriol* 187: 6962–6971

Massé E, Vanderpool CK & Gottesman S (2005b) Effect of RyhB Small RNA on Global Iron Use in *Escherichia coli*. *J Bacteriol* 187: 6962–6971

Matson JS, Livny J & DiRita VJ (2017) A putative *Vibrio cholerae* two-component system controls a conserved periplasmic protein in response to the antimicrobial peptide polymyxin B. *PLOS ONE* 12: e0186199

Matson JS, Yoo HJ, Hakansson K & DiRita VJ (2010) Polymyxin B Resistance in El Tor *Vibrio cholerae* Requires Lipid Acylation Catalyzed by MsbB. *Journal of Bacteriology* 192: 2044–2052

Matthey N & Blokesch M (2016) The DNA-Uptake Process of Naturally Competent *Vibrio cholerae*. *Trends in Microbiology* 24: 98–110

Matz C, McDougald D, Moreno AM, Yung PY, Yildiz FH & Kjelleberg S (2005) Biofilm formation and phenotypic variation enhance predation-driven persistence of *Vibrio cholerae*. *PNAS* 102: 16819–16824

Maurus R, Nguyen NT, Stokell DJ, Ayed A, Hultin PG, Duckworth HW & Brayer GD (2003) Insights into the Evolution of Allosteric Properties. The NADH Binding Site of Hexameric Type II Citrate Synthases. *Biochemistry* 42: 5555–5565

McCoy AJ (2006) Solving structures of protein complexes by molecular replacement with *Phaser*. *Acta Crystallogr D Biol Crystallogr* 63: 32–41

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

Meibom KL, Blokesch M, Dolganov NA, Wu C-Y & Schoolnik GK (2005) Chitin Induces Natural Competence in *Vibrio cholerae*. *Science* 310: 1824–1827

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.e7

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

Mellin JR, McClure R, Lopez D, Green O, Reinhard B & Genco C (2010) Role of Hfq in iron-dependent and -independent gene regulation in *Neisseria meningitidis*. *Microbiology (Reading)* 156: 2316–2326

Merrell DS, Butler SM, Qadri F, Dolganov NA, Alam A, Cohen MB, Calderwood SB, Schoolnik GK & Camilli A (2002) Host-induced epidemic spread of the cholera bacterium. *Nature* 417: 642–645

Mey AR, Wyckoff EE, Kanukurthy V, Fisher CR & Payne SM (2005) Iron and Fur Regulation in *Vibrio cholerae* and the Role of Fur in Virulence. *Infect Immun* 73: 8167–8178

Meydan S, Marks J, Klepacki D, Sharma V, Baranov PV, Firth AE, Margus T, Kefi A, Vázquez-Laslop N & Mankin AS (2019) Retapamulin-Assisted Ribosome Profiling Reveals the Alternative Bacterial Proteome. *Molecular Cell* 74: 481-493.e6

Meydan S, Vázquez-Laslop N & Mankin AS (2018) Genes within Genes in Bacterial Genomes. *Microbiology Spectrum* 6

Michaux C, Holmqvist E, Vasicek E, Sharan M, Barquist L, Westermann AJ, Gunn JS & Vogel J (2017) RNA target profiles direct the discovery of virulence functions for the cold-shock proteins CspC and CspE. *PNAS* 114: 6824–6829

Miller MB, Skorupski K, Lenz DH, Taylor RK & Bassler BL (2002) Parallel Quorum Sensing Systems Converge to Regulate Virulence in *Vibrio cholerae*. *Cell* 110: 303–314

Miller SL & Smith-Magowan D (1990) The Thermodynamics of the Krebs Cycle and Related Compounds. *Journal of Physical and Chemical Reference Data* 19: 1049–1073

Milner JL & Wood JM (1989) Insertion *proQ220*::Tn5 alters regulation of proline porter II, a transporter of proline and glycine betaine in *Escherichia coli*. *J Bacteriol* 171: 947–951

Miravet-Verde S, Ferrar T, Espadas-García G, Mazzolini R, Gharrab A, Sabido E, Serrano L & Lluch-Senar M (2019) Unraveling the hidden universe of small proteins in bacterial genomes. *Mol Syst Biol* 15

Mitarai N, Benjamin J-AM, Krishna S, Semsey S, Csiszovszki Z, Massé E & Sneppen K (2009) Dynamic features of gene expression control by small regulatory RNAs. *PNAS* 106: 10655–10659

Mitchell WJ, Misko TP & Roseman S (1982) Sugar transport by the bacterial phosphotransferase system. Regulation of other transport systems (lactose and melibiose). *Journal of Biological Chemistry* 257: 14553–14564

Mitobe J, Yanagihara I, Ohnishi K, Yamamoto S, Ohnishi M, Ishihama A & Watanabe H (2011) RodZ regulates the post-transcriptional processing of the *Shigella sonnei* type III secretion system. *EMBO Rep* 12: 911–916

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

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

Modell JW, Hopkins AC & Laub MT (2011) A DNA damage checkpoint in *Caulobacter crescentus* inhibits cell division through a direct interaction with FtsW. *Genes Dev* 25: 1328–1343

Modell JW, Kambara TK, Perchuk BS & Laub MT (2014) A DNA Damage-Induced, SOS-Independent Checkpoint Regulates Cell Division in *Caulobacter crescentus*. *PLOS Biology* 12: e1001977

Møller T, Franch T, Udesen C, Gerdes K & Valentin-Hansen P (2002) Spot 42 RNA mediates discoordinate expression of the *E. coli* galactose operon. *Genes Dev* 16: 1696–1706

Moreno R & Rojo F (2008) The Target for the *Pseudomonas putida* Crc Global Regulator in the Benzoate Degradation Pathway Is the BenR Transcriptional Regulator. *Journal of Bacteriology* 190: 1539–1545

Morfeldt E, Taylor D, von Gabain A & Arvidson S (1995) Activation of alpha-toxin translation in *Staphylococcus aureus* by the *trans*-encoded antisense RNA, RNAIII. *EMBO J* 14: 4569–4577

Morris JG (2011) Cholera—Modern Pandemic Disease of Ancient Lineage. *Emerg Infect Dis* 17: 2099–2104

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

Mukherjee A, Cao C & Lutkenhaus J (1998) Inhibition of FtsZ polymerization by SulA, an inhibitor of septation in *Escherichia coli*. *Proc Natl Acad Sci U S A* 95: 2885–2890

Mukherjee S & Bassler BL (2019) Bacterial quorum sensing in complex and dynamically changing environments. *Nature Reviews Microbiology* 17: 371–382

Mukhopadhyay AK, Chakraborty S, Takeda Y, Nair GB & Berg DE (2001) Characterization of VPI Pathogenicity Island and CTXφ Prophage in Environmental Strains of *Vibrio cholerae*. *J Bacteriol* 183: 4737–4746

Müller SA, Kohajda T, Findeiß S, Stadler PF, Washietl S, Kellis M, von Bergen M & Kalkhof S (2010) Optimization of parameters for coverage of low molecular weight proteins. *Anal Bioanal Chem* 398: 2867–2881

Nackerdien ZE, Keynan A, Bassler BL, Lederberg J & Thaler DS (2008) Quorum Sensing Influences *Vibrio harveyi* Growth Rates in a Manner Not Fully Accounted For by the Marker Effect of Bioluminescence. *PLOS ONE* 3: e1671

Nakahigashi K, Takai Y, Kimura M, Abe N, Nakayashiki T, Shiwa Y, Yoshikawa H, Wanner BL, Ishihama Y & Mori H (2016) Comprehensive identification of translation start sites by tetracycline-inhibited ribosome profiling. *DNA Research* 23: 193–201

Neiditch MB, Federle MJ, Miller ST, Bassler BL & Hughson FM (2005) Regulation of LuxPQ receptor activity by the quorum-sensing signal autoinducer-2. *Mol Cell* 18: 507–518

Ner SS, Bhayana V, Bell AW, Giles IG, Duckworth HW & Bloxham DP (1983) Complete sequence of the *gltA* gene encoding citrate synthase in *Escherichia coli*. *Biochemistry* 22: 5243–5249

Nesvizhskii AI (2014) Proteogenomics: concepts, applications and computational strategies. *Nature Methods* 11: 1114–1125

Nguyen NT, Maurus R, Stokell DJ, Ayed A, Duckworth HW & Brayer GD (2001) Comparative Analysis of Folding and Substrate Binding Sites between Regulated Hexameric Type II Citrate Synthases and Unregulated Dimeric Type I Enzymes,. *Biochemistry* 40: 13177–13187

Niederhoffer EC, Naranjo CM, Bradley KL & Fee JA (1990) Control of *Escherichia coli* superoxide dismutase (*sodA* and *sodB*) genes by the ferric uptake regulation (*fur*) locus. *Journal of Bacteriology* 172: 1930–1938

Nielsen AT, Dolganov NA, Rasmussen T, Otto G, Miller MC, Felt SA, Torreilles S & Schoolnik GK (2010) A Bistable Switch and Anatomical Site Control *Vibrio cholerae* Virulence Gene Expression in the Intestine. *PLoS Pathog* 6: e1001102

Nizet V, Beall B, Bast DJ, Datta V, Kilburn L, Low DE & De Azavedo JC (2000) Genetic locus for streptolysin S production by group A streptococcus. *Infect Immun* 68: 4245–4254

Notley-McRobb L, Death A & Ferenci T 1997 (1997) The relationship between external glucose concentration and cAMP levels inside *Escherichia coli*: implications for models of phosphotransferase-mediated regulation of adenylate cyclase. *Microbiology* 143: 1909–1918

Novick RP, Ross HF, Projan SJ, Kornblum J, Kreiswirth B & Moghazeh S (1993) Synthesis of

staphylococcal virulence factors is controlled by a regulatory RNA molecule. *EMBO J* 12: 3967–3975

Nuss AM, Beckstette M, Pimenova M, Schmühl C, Opitz W, Pisano F, Heroven AK & Dersch P (2017) Tissue dual RNA-seq allows fast discovery of infection-specific functions and riboregulators shaping host–pathogen transcriptomes. *PNAS* 114: E791–E800

Nye MB, Pfau JD, Skorupski K & Taylor RK (2000) Vibrio cholerae H-NS Silences Virulence Gene Expression at Multiple Steps in the ToxR Regulatory Cascade. *J Bacteriol* 182: 4295–4303

Oberto J (2013) SyntTax: a web server linking synteny to prokaryotic taxonomy. *BMC Bioinformatics* 14: 4

ÓhÉigeartaigh SS, Armisén D, Byrne KP & Wolfe KH (2011) Systematic discovery of unannotated genes in 11 yeast species using a database of orthologous genomic segments. *BMC Genomics* 12: 377

ÓhÉigeartaigh SS, Armisén D, Byrne KP & Wolfe KH (2014) SearchDOGS Bacteria, Software That Provides Automated Identification of Potentially Missed Genes in Annotated Bacterial Genomes. *Journal of Bacteriology* 196: 2030

Orr MW, Mao Y, Storz G & Qian S-B (2020) Alternative ORFs and small ORFs: shedding light on the dark proteome. *Nucleic Acids Res* 48: 1029–1042

Palacios IM & Johnston DSt (2001) Getting the Message Across: The Intracellular Localization of mRNAs in Higher Eukaryotes. *Annu Rev Cell Dev Biol* 17: 569–614

Papenfort K, Förstner KU, Cong J-P, Sharma CM & Bassler BL (2015) Differential RNA-seq of *Vibrio cholerae* identifies the VqmR small RNA as a regulator of biofilm formation. *PNAS* 112: E766–E775

Papenfort K, Podkaminski D, Hinton JCD & Vogel J (2012) The ancestral SgrS RNA discriminates horizontally acquired *Salmonella* mRNAs through a single G-U wobble pair. *PNAS* 109: E757–E764

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

Papenfort K, Sun Y, Miyakoshi M, Vanderpool CK & Vogel J (2013) Regulation of glucose homeostasis by small RNA mediated activation of sugar phosphatase mRNA. *Cell* 153: 426–437

Papenfort K & Vanderpool CK (2015) Target activation by regulatory RNAs in bacteria. *FEMS Microbiol Rev* 39: 362–378

Park SJ, McCabe J, Turna J & Gunsalus RP (1994) Regulation of the citrate synthase (*gltA*) gene of *Escherichia coli* in response to anaerobiosis and carbon supply: role of the *arcA* gene product. *J Bacteriol* 176: 5086–5092

Patra T, Koley H, Ramamurthy T, Ghose AC & Nandy RK (2012) The Entner-Doudoroff Pathway Is Obligatory for Gluconate Utilization and Contributes to the Pathogenicity of *Vibrio cholerae*. *J Bacteriol* 194: 3377–3385

Peekhaus N & Conway T (1998) What's for dinner?: Entner-Doudoroff metabolism in *Escherichia coli*. *J Bacteriol* 180: 3495–3502

Pereira DS, Donald LJ, Hosfield DJ & Duckworth HW (1994) Active site mutants of *Escherichia coli* citrate synthase. Effects of mutations on catalytic and allosteric properties. *Journal of Biological Chemistry* 269: 412–417

Perrenoud A & Sauer U (2005) Impact of Global Transcriptional Regulation by ArcA, ArcB, Cra, Crp, Cya, Fnr, and Mlc on Glucose Catabolism in *Escherichia coli*. *J Bacteriol* 187: 3171–3179

Peschek N, Herzog R, Singh PK, Sprenger M, Meyer F, Fröhlich KS, Schröger L, Bramkamp M, Drescher K & Papenfort K (2020) RNA-mediated control of cell shape modulates antibiotic resistance in *Vibrio cholerae*. *Nature Communications* 11: 6067

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

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

Pflüger-Grau K & Görke B (2010) Regulatory roles of the bacterial nitrogen-related phosphotransferase system. *Trends in Microbiology* 18: 205–214

Pilonieta MC, Erickson KD, Ernst RK & Detweiler CS (2009) A Protein Important for Antimicrobial Peptide Resistance, YdeI/OmdA, Is in the Periplasm and Interacts with OmpD/NmpC. *Journal of Bacteriology* 191: 7243–7252

Polayes DA, Rice PW & Dahlberg JE (1988) DNA polymerase I activity in *Escherichia coli* is influenced by Spot 42 RNA. *Journal of Bacteriology* 170: 2083–2088

Prévost K, Salvail H, Desnoyers G, Jacques J-F, Phaneuf É & Massé E (2007) The small RNA RyhB activates the translation of *shiA* mRNA encoding a permease of shikimate, a compound involved in siderophore synthesis. *Molecular Microbiology* 64: 1260–1273

Price NL & Raivio TL (2009) Characterization of the Cpx Regulon in *Escherichia coli* Strain MC4100.

Provenzano D & Klose KE (2000) Altered expression of the ToxR-regulated porins OmpU and OmpT diminishes *Vibrio cholerae* bile resistance, virulence factor expression, and intestinal colonization. *PNAS* 97: 10220–10224

Pulvermacher SC, Stauffer LT & 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: 42–50

Quandt EM, Deatherage DE, Ellington AD, Georgiou G & Barrick JE (2014) Recursive genome wide recombination and sequencing reveals a key refinement step in the evolution of a metabolic innovation in *Escherichia coli*. *PNAS* 111: 2217–2222

Quandt EM, Gollihar J, Blount ZD, Ellington AD, Georgiou G & Barrick JE (2015) Fine-tuning citrate synthase flux potentiates and refines metabolic innovation in the Lenski evolution experiment. *eLife* 4: e09696

Queck SY, Khan BA, Wang R, Bach T-HL, Kretschmer D, Chen L, Kreiswirth BN, Peschel A, Deleo FR & Otto M (2009) Mobile genetic element-encoded cytolysin connects virulence to methicillin resistance in MRSA. *PLoS Pathog* 5: e1000533

Quendera AP, Seixas AF, dos Santos RF, Santos I, Silva JPN, Arraiano CM & Andrade JM (2020) RNA-Binding Proteins Driving the Regulatory Activity of Small Non-coding RNAs in Bacteria. *Front Mol Biosci* 7

R. Cerqueira F & Vasconcelos ATR (2020) OCCAM: prediction of small ORFs in bacterial genomes by means of a target-decoy database approach and machine learning techniques. *Database* 2020

Raina M & Storz G (2017) SgrT, a Small Protein That Packs a Sweet Punch. *Journal of Bacteriology* 199

Ramamurthy T, Nandy RK, Mukhopadhyay AK, Dutta S, Mutreja A, Okamoto K, Miyoshi S-I, Nair GB & Ghosh A (2020) Virulence Regulation and Innate Host Response in the Pathogenicity of *Vibrio cholerae*. *Frontiers in Cellular and Infection Microbiology* 10: 520

Régnier P & Hajnsdorf E (2013) The interplay of Hfq, poly(A) polymerase I and exoribonucleases at the 3′ ends of RNAs resulting from Rho-independent termination. *RNA Biology* 10: 602–609

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

Remington S, Wiegand G & Huber R (1982) Crystallographic refinement and atomic models of two different forms of citrate synthase at 2 ·7 and 1 ·7 Å resolution. *Journal of Molecular Biology* 158: 111–152

Remington SJ (1992) Structure and Mechanism of Citrate Synthase. In *Current Topics in Cellular Regulation*, Stadtman ER & Chock PB (eds) pp 209–229. Academic Press

Rice JB, Balasubramanian D & Vanderpool CK (2012) Small RNA binding-site multiplicity involved in translational regulation of a polycistronic mRNA. *PNAS* 109: E2691–E2698

Rice JB & Vanderpool CK (2011) The small RNA SgrS controls sugar–phosphate accumulation by regulating multiple PTS genes. *Nucleic Acids Research* 39: 3806–3819

Rice PW & Dahlberg JE (1982) A gene between *polA* and *glnA* retards growth of *Escherichia coli* when present in multiple copies: physiological effects of the gene for Spot 42 RNA. *J Bacteriol* 152: 1196–1210

Richard AL, Withey JH, Beyhan S, Yildiz F & DiRita VJ (2010) The *Vibrio cholerae* virulence regulatory cascade controls glucose uptake through activation of TarA, a small regulatory RNA. *Molecular Microbiology* 78: 1171–1181

Richards GR, Patel MV, Lloyd CR & Vanderpool CK (2013) Depletion of glycolytic intermediates plays a key role in glucose-phosphate stress in *Escherichia coli*. *J Bacteriol* 195: 4816–4825

Roberts SA & Scott JR (2007) RivR and the small RNA RivX: the missing links between the CovR regulatory cascade and the Mga regulon. *Mol Microbiol* 66: 1506–1522

Robertson GT & Roop RM (1999) The *Brucella abortus* host factor I (HF-I) protein contributes to stress resistance during stationary phase and is a major determinant of virulence in mice. *Molecular Microbiology* 34: 690–700

Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G & Mesirov JP (2011) Integrative Genomics Viewer. *Nat Biotechnol* 29: 24–26

Roggentin P, Schauer R, Hoyer LL & Vimr ER (1993) The sialidase superfamily and its spread by

horizontal gene transfer. *Molecular Microbiology* 9: 915–921

Rohwer JM, Bader R, Westerhoff HV & Postma PW (1998) Limits to inducer exclusion: inhibition of the bacterial phosphotransferase system by glycerol kinase. *Molecular Microbiology* 29: 641–652

Roier S, Zingl FG, Cakar F, Durakovic S, Kohl P, Eichmann TO, Klug L, Gadermaier B, Weinzerl K, Prassl R, Lass A, Daum G, Reidl J, Feldman MF & Schild S (2016) A novel mechanism for the biogenesis of outer membrane vesicles in Gram-negative bacteria. *Nat Commun* 7: 10515

Romeo T & Babitzke P (2018) Global regulation by CsrA and its RNA antagonists. *Microbiol Spectr* 6

Romeo T, Gong M, Liu MY & Brun-Zinkernagel AM (1993) Identification and molecular characterization of *csrA*, a pleiotropic gene from *Escherichia coli* that affects glycogen biosynthesis, gluconeogenesis, cell size, and surface properties. *J Bacteriol* 175: 4744–4755

Rueden CT, Schindelin J, Hiner MC, DeZonia BE, Walter AE, Arena ET & Eliceiri KW (2017) ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics* 18: 529

Russell RJ, Ferguson JM, Hough DW, Danson MJ & Taylor GL (1997) The crystal structure of citrate synthase from the hyperthermophilic archaeon pyrococcus furiosus at 1.9 A resolution. *Biochemistry* 36: 9983–9994

Russell RJ, Hough DW, Danson MJ & Taylor GL (1994) The crystal structure of citrate synthase from the thermophilic Archaeon, *Thermoplasma acidophilum*. *Structure* 2: 1157–1167

Rutherford ST & Bassler BL (2012) Bacterial Quorum Sensing: Its Role in Virulence and Possibilities for Its Control. *Cold Spring Harb Perspect Med* 2

Rutherford ST, Kessel JC van, Shao Y & Bassler BL (2011) AphA and LuxR/HapR reciprocally control quorum sensing in *Vibrios*. *Genes Dev* 25: 397–408

Sáenz-Lahoya S, Bitarte N, García B, Burgui S, Vergara-Irigaray M, Valle J, Solano C, Toledo-Arana A & Lasa I (2019) Noncontiguous operon is a genetic organization for coordinating bacterial gene expression. *PNAS* 116: 1733–1738

Sahyoun N & Cuatrecasas P (1975) Mechanism of activation of adenylate cyclase by cholera toxin. *PNAS* 72: 3438–3442

Salerno C, Ovadi J, Keleti T & Fasella P (1982) Kinetics of Coupled Reactions Catalyzed by Aspartate Aminotransferase and Glutamate Dehydrogenase. *Eur J Biochem* 121: 511–517

Saliba A-E, C Santos S & Vogel J (2017) New RNA-seq approaches for the study of bacterial pathogens. *Curr Opin Microbiol* 35: 78–87

Salvail H, Caron M-P, Bélanger J & Massé E (2013) Antagonistic functions between the RNA chaperone Hfq and an sRNA regulate sensitivity to the antibiotic colicin. *EMBO J* 32: 2764–2778

Sauer U & Eikmanns BJ (2005) The PEP—pyruvate—oxaloacetate node as the switch point for carbon flux distribution in bacteria. *FEMS Microbiology Reviews* 29: 765–794

Saul-McBeth J & Matson JS (2019) A Periplasmic Antimicrobial Peptide-Binding Protein Is Required for Stress Survival in *Vibrio cholerae*. *Frontiers in Microbiology* 10: 161

Sayed N, Nonin-Lecomte S, Réty S & Felden B (2012) Functional and Structural Insights of a *Staphylococcus aureus* Apoptotic-like Membrane Peptide from a Toxin-Antitoxin Module. *J Biol Chem* 287: 43454–43463

Scheuerl T, Hopkins M, Nowell RW, Rivett DW, Barraclough TG & Bell T (2020) Bacterial adaptation is constrained in complex communities. *Nat Commun* 11: 754

Schild S, Tamayo R, Nelson EJ, Qadri F, Calderwood SB & Camilli A (2007) Genes induced late in infection increase fitness of *Vibrio cholerae* after release into the environment. *Cell Host Microbe* 2: 264–277

Schleicher L, Muras V, Claussen B, Pfannstiel J, Blombach B, Dibrov P, Fritz G & Steuber J (2018) *Vibrio natriegens* as Host for Expression of Multisubunit Membrane Protein Complexes. *Front Microbiol* 9

Schrader JM, Zhou B, Li G-W, Lasker K, Childers WS, Williams B, Long T, Crosson S, McAdams HH, Weissman JL & Shapiro L (2014) The Coding and Noncoding Architecture of the *Caulobacter crescentus* Genome. *PLoS Genet* 10

Scrudato ML & Blokesch M (2012) The Regulatory Network of Natural Competence and Transformation of *Vibrio cholerae*. *PLOS Genetics* 8: e1002778

Seitz P & Blokesch M (2013) DNA-uptake machinery of naturally competent *Vibrio cholerae*. *Proc Natl Acad Sci U S A* 110: 17987–17992

Sengupta N, Paul K & Chowdhury R (2003) The Global Regulator ArcA Modulates Expression of Virulence Factors in *Vibrio cholerae*. *Infect Immun* 71: 5583–5589

Sharma CM, Papenfort K, Pernitzsch SR, Mollenkopf H-J, Hinton JCD & Vogel J (2011) Pervasive post-transcriptional control of genes involved in amino acid metabolism by the Hfq-dependent GcvB small RNA. *Mol Microbiol* 81: 1144–1165

Sharma SS, Chong S & Harcum SW (2006) Intein-mediated protein purification of fusion proteins expressed under high-cell density conditions in *E. coli*. *J Biotechnol* 125: 48–56

Shikuma NJ, Davis KR, Fong JNC & Yildiz FH (2013) The Transcriptional Regulator, CosR, Controls Compatible Solute Biosynthesis and Transport, Motility and Biofilm Formation in *Vibrio cholerae*. *Environ Microbiol* 15: 1387–1399

Shikuma NJ & Yildiz FH (2009) Identification and characterization of OscR, a transcriptional regulator involved in osmolarity adaptation in *Vibrio cholerae*. *J Bacteriol* 191: 4082–4096

Shimada T, Fujita N, Yamamoto K & Ishihama A (2011) Novel Roles of cAMP Receptor Protein (CRP) in Regulation of Transport and Metabolism of Carbon Sources. *PLoS One* 6

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

Shiomi D, Sakai M & Niki H (2008) Determination of bacterial rod shape by a novel cytoskeletal membrane protein. *EMBO J* 27: 3081–3091

Siebold C, Flükiger K, Beutler R & Erni B (2001a) Carbohydrate transporters of the bacterial phosphoenolpyruvate: sugar phosphotransferase system (PTS). *FEBS Letters* 504: 104–111

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

Singh S & Mittal A (2016) Transmembrane Domain Lengths Serve as Signatures of Organismal Complexity and Viral Transport Mechanisms. *Sci Rep* 6: 22352

Singleton FL, Attwell RW, Jangi MS & Colwell RR (1982) Influence of salinity and organic nutrient concentration on survival and growth of *Vibrio cholerae* in aquatic microcosms. *Appl Environ Microbiol* 43: 1080–1085

Skorupski K & Taylor RK (1996) Positive selection vectors for allelic exchange. *Gene* 169: 47–52

Skorupski K & Taylor RK (1997) Cyclic AMP and its receptor protein negatively regulate the coordinate expression of cholera toxin and toxin-coregulated pilus in *Vibrio cholerae*. *PNAS* 94: 265–270

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

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

Song M, Kim H-J, Kim EY, Shin M, Lee HC, Hong Y, Rhee JH, Yoon H, Ryu S, Lim S & Choy HE (2004) ppGpp-dependent Stationary Phase Induction of Genes on *Salmonella* Pathogenicity Island 1. *Journal of Biological Chemistry* 279: 34183–34190

Song T, Mika F, Lindmark B, Liu Z, Schild S, Bishop A, Zhu J, Camilli A, Johansson J, Vogel J & Wai SN (2008) A new *Vibrio cholerae* sRNA modulates colonization and affects release of outer membrane vesicles. *Mol Microbiol* 70: 100–111

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

Spangler BD (1992) Structure and function of cholera toxin and the related Escherichia coli heat-labile enterotoxin. *Microbiol Rev* 56: 622–647

Srere PA (1969) Citrate synthase: [EC 4.1.3.7. Citrate oxaloacetate-lyase (CoA-acetylating)]. In *Methods in Enzymology* pp 3–11. Academic Press

Stalmach ME, Grothe S & Wood JM (1983) Two proline porters in *Escherichia coli* K-12. *J Bacteriol* 156: 481–486

Storz G, Wolf YI & Ramamurthi KS (2014) Small Proteins Can No Longer Be Ignored. *Annu Rev Biochem* 83: 753–777

Sun X, Zhulin I & Wartell RM (2002) Predicted structure and phyletic distribution of the RNA-binding protein Hfq. *Nucleic Acids Research* 30: 3662–3671

Svenningsen SL, Tu KC & Bassler BL (2009) Gene dosage compensation calibrates four regulatory RNAs to control *Vibrio cholerae* quorum sensing. *EMBO J* 28: 429–439

Svenningsen SL, Waters CM & Bassler BL (2008) A negative feedback loop involving small RNAs accelerates *Vibrio cholerae*'s transition out of quorum-sensing mode. *Genes Dev* 22: 226–238

Syed KA, Beyhan S, Correa N, Queen J, Liu J, Peng F, Satchell KJF, Yildiz F & Klose KE (2009) The Vibrio cholerae Flagellar Regulatory Hierarchy Controls Expression of Virulence Factors. *Journal of Bacteriology* 191: 6555–6570

Tamayo R, Patimalla B & Camilli A (2010) Growth in a Biofilm Induces a Hyperinfectious Phenotype in *Vibrio cholerae*. *Infect Immun* 78: 3560–3569

Tamplin ML & Colwell RR (1986) Effects of microcosm salinity and organic substrate concentration on production of *Vibrio cholerae* enterotoxin. *Appl Environ Microbiol* 52: 297–301

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

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: 2853–2856

Timmermans J & Van Melderen L (2010) Post-transcriptional global regulation by CsrA in bacteria. *Cell Mol Life Sci* 67: 2897–2908

Tischler AD & Camilli A (2004) Cyclic diguanylate (c-di-GMP) regulates *Vibrio cholerae* biofilm formation. *Molecular Microbiology* 53: 857–869

Tong EK & Duckworth HW (1975) The quaternary structure of citrate synthase from *Escherichia coli* K12. *Biochemistry* 14: 235–241

Tree JJ, Granneman S, McAteer SP, Tollervey D & Gally DL (2014) Identification of Bacteriophage-Encoded Anti-sRNAs in Pathogenic *Escherichia coli*. *Mol Cell* 55: 199–213

Tsai M-J, Wang J-R, Yang C-D, Kao K-C, Huang W-L, Huang H-Y, Tseng C-P, Huang H-D & Ho S-Y (2018) PredCRP: predicting and analysing the regulatory roles of CRP from its binding sites in *Escherichia coli*. *Scientific Reports* 8: 951

Tsou AM & Zhu J (2010) Quorum Sensing Negatively Regulates Hemolysin Transcriptionally and Posttranslationally in *Vibrio cholerae*. *Infection and Immunity* 78: 461–467

Tsui H-CT, Leung H-CE & Winkler ME (1994) Characterization of broadly pleiotropic phenotypes caused by an *hfq* insertion mutation in *Escherichia coli* K-12. *Molecular Microbiology* 13: 35–49

Tu KC & Bassler BL (2007) Multiple small RNAs act additively to integrate sensory information and control quorum sensing in *Vibrio harveyi*. *Genes Dev* 21: 221–233

Udden SMN, Zahid MSH, Biswas K, Ahmad QS, Cravioto A, Nair GB, Mekalanos JJ & Faruque SM (2008) Acquisition of classical CTX prophage from *Vibrio cholerae* O141 by El Tor strains aided by lytic phages and chitin-induced competence. *Proc Natl Acad Sci U S A* 105: 11951–11956

Ul Haq I, Müller P & Brantl S (2021) SR7 – a dual-function antisense RNA from *Bacillus subtilis*. *RNA Biology* 18: 104–117

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

Urban JH & Vogel J (2008) Two Seemingly Homologous Noncoding RNAs Act Hierarchically to Activate *glmS* mRNA Translation. *PLoS Biol* 6

Utada AS, Bennett RR, Fong JCN, Gibiansky ML, Yildiz FH, Golestanian R & Wong GCL (2014) *Vibrio cholerae* use pili and flagella synergistically to effect motility switching and conditional surface attachment. *Nat Commun* 5: 4913

Vadeboncoeur C & Pelletier M (1997) The phosphoenolpyruvate:sugar phosphotransferase system of oral *Sstreptococci* and its role in the control of sugar metabolism. *FEMS Microbiology Reviews* 19: 187–207

Vaitkevicius K, Lindmark B, Ou G, Song T, Toma C, Iwanaga M, Zhu J, Andersson A, Hammarström M-L, Tuck S & Wai SN (2006) A *Vibrio cholerae* protease needed for killing of *Caenorhabditis elegans* has a role in protection from natural predator grazing. *Proc Natl Acad Sci U S A* 103: 9280–9285

Vallenet D, Belda E, Calteau A, Cruveiller S, Engelen S, Lajus A, Le Fèvre F, Longin C, Mornico D, Roche D, ouy Z, Salvignol G, Scarpelli C, Thil Smith AA, Weiman M & Médigue C (2013) MicroScope—an integrated microbial resource for the curation and comparative analysis of genomic and metabolic data. *Nucleic Acids Res* 41: D636–D647

Van Assche E, Van Puyvelde S, Vanderleyden J & Steenackers HP (2015) RNA-binding proteins involved in post-transcriptional regulation in bacteria. *Front Microbiol* 6

Vanderpool CK & Gottesman S (2004) Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system. *Molecular Microbiology* 54: 1076–1089

VanOrsdel CE, Kelly JP, Burke BN, Lein CD, Oufiero CE, Sanchez JF, Wimmers LE, Hearn DJ, Abuikhdair FJ, Barnhart KR, Duley ML, Ernst SEG, Kenerson BA, Serafin AJ & Hemm MR (2018) Identifying New Small Proteins in *Escherichia coli*. *Proteomics* 18

Venditti V, Ghirlando R & Clore GM (2013) Structural basis for enzyme I inhibition by α-ketoglutarate. *ACS Chem Biol* 8: 1232–1240

Venturini E, Svensson SL, Maaß S, Gelhausen R, Eggenhofer F, Li L, Cain AK, Parkhill J, Becher D, Backofen R, Barquist L, Sharma CM, Westermann AJ, Vogel J (2020) A global data-driven census of *Salmonella* small proteins and their potential functions in bacterial virulence. *microLife* 1

Verdon J, Girardin N, Lacombe C, Berjeaud J-M & Héchard Y (2009) delta-hemolysin, an update on a membrane-interacting peptide. *Peptides* 30: 817–823

Virant D, Traenkle B, Maier J, Kaiser PD, Bodenhöfer M, Schmees C, Vojnovic I, Pisak-Lukáts B, Endesfelder U & Rothbauer U (2018) A peptide tag-specific nanobody enables high-quality labeling for dSTORM imaging. *Nat Commun* 9: 930

Vogel J (2009) A rough guide to the non-coding RNA world of *Salmonella*. *Molecular Microbiology* 71: 1–11

Vogel J, Argaman L, Wagner EGH & Altuvia S (2004) The small RNA IstR inhibits synthesis of an SOS-induced toxic peptide. *Curr Biol* 14: 2271–2276

Vogel J & Luisi BF (2011) Hfq and its constellation of RNA. *Nature Reviews Microbiology* 9: 578–589

Vornhagen J, Sun Y, Breen P, Forsyth V, Zhao L, Mobley HLT & Bachman MA (2019) The *Klebsiella pneumoniae* citrate synthase gene, *gltA*, influences site specific fitness during infection. *PLOS Pathogens* 15: e1008010

Wadler CS & Vanderpool CK (2007) A dual function for a bacterial small RNA: SgrS performs base pairing-dependent regulation and encodes a functional polypeptide. *PNAS* 104: 20454–20459

Wai SN, Lindmark B, Söderblom T, Takade A, Westermark M, Oscarsson J, Jass J, Richter-Dahlfors A, Mizunoe Y & Uhlin BE (2003) Vesicle-Mediated Export and Assembly of Pore-Forming Oligomers of the Enterobacterial ClyA Cytotoxin. *Cell* 115: 25–35

Waldor MK & Mekalanos JJ (1996) Lysogenic Conversion by a Filamentous Phage Encoding Cholera Toxin. *Science* 272: 1910–1914

Wang H, Yin X, Orr MW, Dambach M, Curtis R & Storz G (2017) Increasing intracellular magnesium levels with the 31-amino acid MgtS protein. *PNAS* 114: 5689–5694

Wang H, Zeng F, Liu Q, Liu H, Liu Z, Niu L, Teng M & Li X (2013) The structure of the ARE-binding domains of Hu antigen R (HuR) undergoes conformational changes during RNA binding. *Acta Cryst D* 69: 373–380

Wang Q, Millet YA, Chao MC, Sasabe J, Davis BM & Waldor MK (2015) A Genome-Wide Screen Reveals that the *Vibrio cholerae* Phosphoenolpyruvate Phosphotransferase System Modulates Virulence Gene Expression. *Infection and Immunity* 83: 3381

Wang S & Huang T (2021) Applications of Deep Learning in Biomedicine. In *Systems Medicine*, Wolkenhauer O (ed) pp 29–39. Oxford: Academic Press

Wang X & Quinn PJ (2010) Endotoxins: Lipopolysaccharides of Gram-Negative Bacteria. (ed) pp 3–25. Dordrecht: Springer Netherlands

Wanner BL & Wilmes-Riesenberg MR (1992) Involvement of phosphotransacetylase, acetate kinase, and acetyl phosphate synthesis in control of the phosphate regulon in *Escherichia coli*. *J Bacteriol* 174: 2124–2130

Warren AS, Archuleta J, Feng W & Setubal JC (2010) Missing genes in the annotation of prokaryotic genomes. *BMC Bioinformatics* 11: 131

Washietl S, Findeiß S, Müller SA, Kalkhof S, von Bergen M, Hofacker IL, Stadler PF & Goldman N (2011) RNAcode: Robust discrimination of coding and noncoding regions in comparative sequence data. *RNA* 17: 578–594

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

Waters CM & Bassler BL (2006) The *Vibrio harveyi* quorum-sensing system uses shared regulatory components to discriminate between multiple autoinducers. *Genes Dev* 20: 2754–2767

Waters CM, Lu W, Rabinowitz JD & Bassler BL (2008) Quorum Sensing Controls Biofilm Formation in *Vibrio cholerae* through Modulation of Cyclic Di-GMP Levels and Repression of *vpsT*. *Journal of Bacteriology* 190: 2527–2536

Waters LS, Sandoval M & Storz G (2011) The *Escherichia coli* MntR Miniregulon Includes Genes Encoding a Small Protein and an Efflux Pump Required for Manganese Homeostasis. *J Bacteriol* 193: 5887–5897

Waters LS & Storz G (2009) Regulatory RNAs in Bacteria. *Cell* 136: 615–628

Watnick PI & Kolter R (1999) Steps in the development of a *Vibrio cholerae* El Tor biofilm. *Molecular*

Microbiology 34: 586–595

Wayne Albers RRW (2012) Cell Membrane Structures and Functions. In Basic Neurochemistry (ed) pp 26–39. New York: Academic Press

Weaver J, Mohammad F, Buskirk AR & Storz G (2019) Identifying Small Proteins by Ribosome Profiling with Stalled Initiation Complexes. *mBio* 10

Weinstock MT, Hesek ED, Wilson CM & Gibson DG (2016) *Vibrio natriegens* as a fast-growing host for molecular biology. *Nature Methods* 13: 849–851

Weitzman PDJ (1966) Regulation of citrate synthase activity in *Escherichia coli*. *Biochimica et Biophysica Acta (BBA) - Enzymology and Biological Oxidation* 128: 213–215

Weitzman PDJ & Jones D 1975 (1975) The Mode of Regulation of Bacterial Citrate Synthase as a Taxonomic Tool. *Microbiology* 89: 187–190

Westermann AJ, Venturini E, Sellin ME, Förstner KU, Hardt W-D & Vogel J (2019) The Major RNA-Binding Protein ProQ Impacts Virulence Gene Expression in *Salmonella enterica* Serovar Typhimurium. *mBio* 10

Whipp MJ, Camakaris H & Pittard AJ (1998) Cloning and analysis of the *shiA* gene, which encodes the shikimate transport system of *Escherichia coli* K-12. *Gene* 209: 185–192

Wiegand G & Remington SJ (1986) Citrate synthase: structure, control, and mechanism. *Annu Rev Biophys Biophys Chem* 15: 97–117

Wielgoss S, Barrick JE, Tenaillon O, Cruveiller S, Chane-Woon-Ming B, Médigue C, Lenski RE & Schneider D (2011) Mutation Rate Inferred From Synonymous Substitutions in a Long-Term Evolution Experiment With *Escherichia coli*. *G3 (Bethesda)* 1: 183–186

Winans SC & Bassler BL (2002) Mob Psychology. *J Bacteriol* 184: 873–883

Withey JH & DiRita VJ (2006) The toxbox: specific DNA sequence requirements for activation of *Vibrio cholerae* virulence genes by ToxT. *Molecular Microbiology* 59: 1779–1789

Wong SM, Carroll PA, Rahme LG, Ausubel FM & Calderwood SB (1998) Modulation of Expression of the ToxR Regulon in *Vibrio cholerae* by a Member of the Two-Component Family of Response Regulators. *Infect Immun* 66: 5854–5861

Wu R, Zhao M, Li J, Gao H, Kan B & Liang W (2015) Direct regulation of the natural competence regulator gene *tfoX* by cyclic AMP (cAMP) and cAMP receptor protein (CRP) in *Vibrios*. *Scientific Reports* 5: 14921

Xavier KB & Bassler BL (2005) Interference with AI-2-Mediated Bacterial Cell-Cell Communication. *Nature* 437: 750–753

Xu Q, Dziejman M & Mekalanos JJ (2003) Determination of the transcriptome of *Vibrio cholerae* during intraintestinal growth and midexponential phase *in vitro*. *PNAS* 100: 1286–1291

Yakhnin AV, Baker CS, Vakulskas CA, Yakhnin H, Berezin I, Romeo T & Babitzke P (2013) CsrA activates *flhDC* expression by protecting *flhDC* mRNA from RNase E-mediated cleavage. *Mol Microbiol* 87: 851–866

Yan D, Lenz P & Hwa T (2011) Overcoming fluctuation and leakage problems in the quantification

of intracellular 2-oxoglutarate levels in *Escherichia coli*. *Appl Environ Microbiol* 77: 6763–6771

Yan J, Sharo AG, Stone HA, Wingreen NS & Bassler BL (2016) *Vibrio cholerae* biofilm growth program and architecture revealed by single-cell live imaging. *PNAS* 113: E5337–E5343

Yang C, Ko B, Hensley CT, Jiang L, Wasti AT, Kim J, Sudderth J, Calvaruso MA, Lumata L, Mitsche M, Rutter J, Merritt ME & DeBerardinis RJ (2014) Glutamine Oxidation Maintains the TCA Cycle and Cell Survival during Impaired Mitochondrial Pyruvate Transport. *Molecular Cell* 56: 414–424

Yeh JI, Kettering R, Saxl R, Bourand A, Darbon E, Joly N, Briozzo P & Deutscher J (2009) Structural Characterizations of Glycerol Kinase: unraveling phosphorylation-induced long-range activation. *Biochemistry* 48: 346–356

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

Yu RR & DiRita VJ (2002) Regulation of gene expression in *Vibrio cholerae* by ToxT involves both antirepression and RNA polymerase stimulation. *Molecular Microbiology* 43: 119–134

Zeghouf M, Li J, Butland G, Borkowska A, Canadien V, Richards D, Beattie B, Emili A & Greenblatt JF (2004) Sequential Peptide Affinity (SPA) System for the Identification of Mammalian and Bacterial Protein Complexes. *J Proteome Res* 3: 463–468

Zhang A, Wassarman KM, Ortega J, Steven AC & Storz G (2002) The Sm-like Hfq Protein Increases OxyS RNA Interaction with Target mRNAs. *Molecular Cell* 9: 11–22

Zhang C, Wei Z-H & Ye B-C (2013a) Quantitative monitoring of 2-oxoglutarate in *Escherichia coli* cells by a fluorescence resonance energy transfer-based biosensor. *Appl Microbiol Biotechnol* 97: 8307–8316

Zhang D, Abovich N & Rosbash M (2001) A Biochemical Function for the Sm Complex. *Molecular Cell* 7: 319–329

Zhang Y, Fonslow BR, Shan B, Baek M-C & Yates JR (2013b) Protein Analysis by Shotgun/Bottom-up Proteomics. *Chem Rev* 113: 2343–2394

Zhu J, Miller MB, Vance RE, Dziejman M, Bassler BL & Mekalanos JJ (2002) Quorum-sensing regulators control virulence gene expression in *Vibrio cholerae*. *Proc Natl Acad Sci U S A* 99: 3129–3134

Zingl FG, Kohl P, Cakar F, Leitner DR, Mitterer F, Bonnington KE, Rechberger GN, Kuehn MJ, Guan Z, Reidl J, *et al* (2020) Outer membrane vesiculation facilitates surface exchange and *in vivo* adaptation of *Vibrio cholerae*. *Cell Host Microbe* 27: 225-237.e8

Manuscripts generated from this study

A dual-function RNA balances carbon uptake and central metabolism in *Vibrio cholerae*

Kavyaa Venkat*, Mona Hoyos*, James R. J. Haycocks, Liam Cassidy, Beatrice Engelmann, Ulrike Rolle-Kampczyk, Martin von Bergen, Andreas Tholey, David C. Grainger, and Kai Papenfort

*Shared first-authorship

The EMBO Journal (2021) 40:e108542 ; https://doi.org/10.15252/embj.2021108542

List of figures

Acknowledgements

The Papenfort lab has been a significant part of my 20's and as much as I would not miss Jena, I would certainly miss being part of the lab. I am sincerely grateful to Kai for his relentless support and endless optimism over the years. I would like to thank Kathi and Andreas for their perpetual source of laughter in the ''family lab'' as well as Mona and Daniel for being fantastic office-mates. Thank you to all the current and former members of the lab for all the insightful discussions and mindful conversations about science and life! I am glad to have met some wonderful people including Micha, Alex, Sara and Yannik; and hope to catch up with you again over coffee and chai lattes in the future! I am especially thankful for the Jena tribe: Anne, Malte, Manuel, Matthias, George and Yvonne for making my time here feel less of a drab and wish you guys the very best.

I would also like to extend my gratitude to Prof. Dr. Marc Bramkamp, Prof. Dr. Andreas Klingl and Prof. Dr. Kai Thormann for their encouraging feedback and guidance during the inter-city / country-wide TAC meetings. Thank you, Prof. Dr. Simon Ringgaard for readily agreeing to be my second supervisor as well as Dr. Arne Weiberg, Dr. Noémie Becker, Prof. Dr. Nicolas Gompel and Prof. Dr. Andreas Klingl for your interest and time to be a part of my dissertation committee. I am also grateful for the constant support and encouragement over the years from both Nadine and Francesca as well as the Life Science Munich Graduate School.

The last few years would be incomplete without the 3,650 cups of coffee and 35,616 km of travelling between Munich and Jena. Beyond science, I am grateful to my dear friends for being a constant source of joy and for their support through thick and thin: Swathi for the endless phone calls from different continents (and I am so sorry I missed your wedding…), Krishna for all the fun times and (scary) hikes, Akash and Prateek for the lovely chai and all the experimental food; as well as Sharada, Sonal, Priyamvadha, Ashwath, Shankar, Swarna, Arpi, Sunil and Puja for all the good times.

Thank you Amma for supporting the choices I made and always being a pillar of strength.

I strive to have the mental strength that you did, Shobha maa. You will always be dearly missed.

Words fall short to describe how fortunate I am to have you my dear Pushkar. Thank you for always sticking around in the high of highs and the low of lows.

To all the 'love in the time of *Vibrio cholerae*!'

Am Schluß, wie sichs gehört, Friede, Freude und Eierkuchen!

Curriculum Vitae