GENOME-WIDE INVESTIGATION OF *VIBRIO PARAHAEMOLYTICUS* TYPE III SECRETION SYSTEM-1 REGULATION AND CHITIN METABOLISM

by

Landon J. Getz

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

at

Dalhousie University
Halifax, Nova Scotia
July 2022

Dalhousie University is located in Mi'kma'ki, the
ancestral and unceded territory of the Mi'kmaq.
We are all Treaty people.

Dedicated to Jordon and Susan

Table of Contents

List of Tables

List of Figures

Abstract

Vibrio parahaemolyticus is the leading cause of seafood-borne gastroenteritis in humans following the consumption of contaminated seafood, and is a consequential organism for aquaculture and human health. The pandemic strain of *V. parahaemolyticus* encodes toxins and hemolysins, two type III secretion systems, and two type VI secretion systems. Primarily, two type III secretion systems, T3SS-1 and T3SS-2, drive rapid host-cell cytotoxicity and enterotoxicity, respectively. How *V. parahaemolyticus* uses these secretion systems to coordinate virulence is not understood. It is clear that T3SS-1 is expressed in calcium-depleted, magnesium-supplemented rich media. However, it is not clear how this signal is transduced into the cell, or through what mechanism the expression of the T3SS-1 master regulator ExsA is induced. Using a transposon mutagenesis screen, we identified HlyU, a SmtB/ArsR family transcriptional regulator as essential for *exsA* expression. Next, we explored the mechanism of HlyU-dependent expression of *exsA*, and identified a non-β-DNA superstructure – a DNA cruciform – involved in the regulation of *exsA*. Through a currently unknown mechanism, HlyU and the cruciform structure appear to interact to de-repress *exsA* expression. Taken together, these studies contribute to our understanding of *exsA* expression, and T3SS coordination, in *V. parahaemolyticus*. In addition to virulence, *Vibrio* spp., including *V. parahaemolyticus*, catabolize and assimilate chitin in the environment. Importantly, many *Vibrio* spp. interactions with seafood and aquacultural organisms are driven by their ability to sense, move towards, and degrade this common and rich carbon and nitrogen source. To better understand the connection between their environmental life and their pathogenic one, I characterized two important sugar catabolism regulators, and their role in cellular functions including chitinase secretion, biofilm formation, and T3SS-1 expression. Then, I employed a transposon-sequencing approach, which combines high-efficiency transposon mutagenesis with high-throughput DNA sequencing, to explore the genes necessary for chitin catabolism genome-wide. Here, I identified uncharacterized sugar transport and transcriptional regulator genes as essential for growth on chitin. *Vibrio* spp. are a growing global burden. Understanding how they survive in the environment, and how they coordinate virulence, is vital to predicting and managing emerging disease.

List of Abbreviations and Symbols Used

AHPND	acute-hepatopancreatic necrosis disease
BSA	bovine serum albumin
CAA	chloroacetylaldehyde
cAMP	Cyclic-AMP
ChIP-Seq	chromatin-immunoprecipitation sequencing
ChiRP	chitin regulated pilus
CPS	counts per second
cryo-EM	cryo-electron microscopy
cryoET	cryoelectron tomography
DMEM	Dulbecco's modified Eagle's medium
dsDNA	double-stranded DNA
EMS	early mortality syndrome
EPEC	enteropathogenic *E. coli*
FAO	Food and Agriculture Organization of the United Nations
GbpA	GlcNAc-binding protein A
GlcNAc	*N*-acetyl-D-glucosamine

HMM	hidden Markov model
HRP	horseradish peroxidase
IP	immunoprecipitation
IP6	inositol hexakisphosphate
LDH	lactate dehydrogenase
LEE	locus of enterocyte effacement
LPS	lipopolysaccharide
MAM	multivalent adhesion molecule
MAPK	mitogen-activate protein kinase
MARTX	multifunctional-autoprocessing repeats-in-toxin
MC	membrane complex
mce	mammalian cell entry
MM9	Minimal M9
MSHA	mannose sensitive hemagglutinin
MurNAc	N-acetylmuramic acid
Neu5Ac	N-acetylneuroaminic acid
PBS	phosphate-buffered saline
PEP	phosphoenolpyruvate
PTS	phosphoenolpyruvate:sugar phosphotransferase system

RNA	Ribonucleic Acid
ROK	repressor orf kinase
rSAP	recombinant shrimp alkaline phosphatase
Sct	secretion and cellular translocation
SPA	single-particle analysis
SPI-1	Salmonella pathogenecity island-1
ssDNA	single-stranded DNA
ST36	multilocus sequence type 36
T2SS	type II secretion system
T3SS	type III secretion system
T3SS-1	type III secretion system - 1
T3SS-2	type III secretion system - 2
T6SS	Type VI Secretion System
T6SS-1	type VI secretion system - 2
T6SS-2	type VI secretion system - 2
TDH	thermostable direct hemolysin
TnSeq	transposon sequencing
TRH	TDH-related hemolysin
TSS	transcriptional start site
VPS	*Vibrio* polysaccharide
wHTH	winged helix-turn-helix

ZINB zero-inflated negative binomial

ZOC zone-of-clearing

Acknowledgements

It is difficult to consider exactly where to begin in recognizing folks for the contributions they've made to my growth (both personal and professional) during my PhD studies and the generation of this thesis. These contributions are surely innumerable in any real sense.

I want to start by acknowledging Mi'kma'ki, the traditional and unceded territories of the Mi'kmaq people which I have called home since the beginning of my post-secondary education. My education as a young person did not prepare me well to understand the complexities and the collective trauma that surrounds the agreements that were made, and subsequently broken, by settlers on Turtle Island. I am thankful for the opportunity to study in Mi'kma'ki, and to appreciate the connection that all living things have with the land here, and everywhere. Particularly important to my work, I also appreciate, am grateful, and would like to acknowledge the oceans which border Mi'kma'ki, where my study organism lives and thrives in collaboration with the land, the sea, and the animals that call it home. I endeavour to approach science, and my scientific pursuits with the understanding that western ways of knowing do not stand alone in a vacuum, and that approaching problems through multiple lenses – in the words of Elder Dr. Albert Marshall, two-eyed seeing – stands only to strengthen our resolve in what we know, and what we can know.

The greatest of thanks go out to my partner Jordon (for whom this thesis is dedicated) for driving my growth as a person and as a proud member of the 2SLGBTQIA+ community. It is his constant search for the truth which drives my curiosity and my ambition. Susan, our cat, has also provided endless companionship during my PhD.

I'de also like to thank my supervisor, Dr. Nikhil Thomas for allowing me the freedom to explore. Not only was I able to make contributions to science and the *Vibrio* community, but through Dr. Thomas' mentorship I was also able to explore other areas of interest including science communication, bioethics, and scientific governance and policy. Dr. Thomas' mentorship has made me into the scientist I am today, and for that I am grateful.

During my thesis project, we initiated collaborations with a number of research groups, including with Dr. Jacqueline Upham at the Canadian Food Inspection Agency, and with Dr. John Archibald, Dr. Morgan Langille, and the clinical microbiology lab at the Nova Scotia Health Authority, led by Dr. Jason LeBlanc. These folks set out the terms for what would be successful collaborations and have taught me important skills in communication, team-work, and scientific group-work. Particularly, Dr. Upham performed genome sequencing for us as part of the collaboration, some of which is found in this thesis.

My thesis committee, Dr. Julie LaRoche, Dr. Morgan Langille, and Dr. Song Lee's continual support and engagement with my thesis project helped drive it into the document you see here. Regular meetings and support from this core group contributed significantly to my success. These folks, along with Dr. Jason LeBlanc, made up my comprehensive examination committee, and I appreciate both their patience and guiding hands during this stressful time in my degree. Dr. John Rohde, Dr. Craig McCormick, and Dr. Zhenyu Cheng all supported my scientific pursuits by donating time, energy, and the occasional reagent to help make my experiments a success. Dr. Lois Murray is an excellent experimentalist and was always around to lend a helping hand with complicated experiments. She was a regular confidant. Dr. Francoise Baylis and Dr. Graham Dellaire helped me explore my interests in bioethics, scientific governance, and science

communication, and significantly broadened my horizons as a scientist. Dr. Jacqueline Gahagan, Dr. Heather Andreas, and Dr. Phillip Joy supported my efforts regarding Equity, Diversity, and Inclusion for 2SLGBTQIA+ folks in STEM in the Atlantic Provinces and I would not have been able to found and organize Queer Atlantic Canadian Science, Technology, Engineering, and Mathematics without them. Although this contribution is not directly related to my thesis, it should be noted that scientific pursuits don't stand in isolation of identity, and having folks to support my personal growth alongside my scientific growth has contributed to my success in no small part.

A significant number of folks contributed to my work from outside of the Dalhousie community, including Dr. James Gauld, Dr. Andrew Camilli, Dr. Alexander Bond, Dr. Angela Crane, and Dr. Jennifer MacDonald provided me with opportunities and support that I am incredibly grateful for. In particular, Dr. Gauld and Dr. Bond provided much needed personal support during tough times, and for that I am grateful.

Of course, a number of funding agencies supported this work including the Natural Sciences and Engineering Research Council of Canada (NSERC), Killam Trusts, the Vanier Canadian Graduate Scholarship Program, and Research Nova Scotia (formerly Nova Scotia Health Research Foundation). Without so much support for and from Canadian science funding, I would not have been able to pursue this PhD.

I am certain that I have forgotten to acknowledge folks here, and if I have I hope that you will take my greatest apologies knowing that I could not have succeeded without your help, and without the help of every person who has walked with me on this journey.

Chapter 1

Introduction

Prokaryotic organisms make up approximately 13% of the globe's biomass measured by carbon (70 Gt C) and are the second most abundant organisms on earth, second only to plants which make up approximately 80% of the globe's biomass (450 Gt of C) [1]. These remarkably abundant organisms are essential for a variety of nutrient cycles – including but not limited to nitrogen, carbon, phosphorous, and sulfur cycles – which provide energy and vital nutrients to other forms of life. This is also true of marine prokaryotes, which provide the foundations of marine food webs and nutrient cycles [2]. Marine microbes are responsible for half of the primary production on Earth, converting CO_2 from the atmosphere into oxygen and sugar [3, 4].

These global cycles are at risk of disruption, with anthropogenic climate change significantly impacting global ecosystems. This impact will not spare prokaryotes [5]. While dysbiosis in a variety of marine ecosystems will occur through climate change induced modifications in ocean temperature, acidity, and salinity, a particularly troubling concern surrounds microbial pathogens and the impacts that changes in these marine attributes will have on them. For example, both ocean acidification and thermal stress drive significant changes in species composition of the coral holobiont – the symbiotic host-microbe interactions that allow the coral to survive – shifting towards opportunistic organisms like the *Vibrionacea* and *Alteromonadaceae* [6–9]. Studies have shown that thermal stress can increase the presence of *Vibrio* spp. virulence factor genes, including flagella, chitinases, and pili [7].

Vibrio spp. are some of the most broadly studied marine bacteria due to their remarkable diversity and relative ease of isolation in laboratory culture. In fact, the causative agent of cholera, *Vibrio cholerae*, was the one of the first pathogens to be described, first by Filippo Pacini in 1854 and later isolated in pure culture by Robert Koch in 1884 [10–12]. The *Vibrio* genus is a gram-negative, rod-shaped, halophilic group of bacteria which contains over 120 individual species [13]. *Vibrio* spp. exist in a variety of niches including free-living in the water column, associated with other organisms like bivalves and molluscs [14], corals [15], shrimp [16, 17], fishes [18], zooplankton (including copepods) [19], squid [20], sponges [21], and seagrasses [22–24], as-well as associated with abiotic particles and in soil sediment [25] (Figure 1A).

Much of this ecology is driven by a strongly conserved chitin metabolism pathway which makes *Vibrio* spp. important in the cycling of marine carbon [26, 27]. For example, colonization of chitinous marine plankton like copepods is considered to be a significant environmental reservoir for *V. cholerae* [28]. Notably, although some of these interactions are beneficial – such as nitrogen fixation in the coral mucus of *Mussismilia hispida* by *V. harveyi* and *V. alginolyticus* – many of these interactions are pathogenic in nature [29, 30]. These myriad interactions in the environment make *Vibrio* spp. a microbial group of concern, especially in the world's changing oceans.

Vibrio spp. drive pathogenicity in a number of animals including mammals, coral, fish, and crustaceans. Primary culprits include *V. parahaemolyticus*, *V. alginolyticus*, *V. harveyi*, *V. cholerae*, and *V. vulnificus*. Importantly, some of these species, like *V. parahaemolyticus*, *V. alginolyticus*, and *V. harveyi*, infect agriculturally important organisms and cause significant economic damages [31–33]. Four major human pathogens belong in the *Vibrio* genus as well including *V. cholerae*, *V. parahaemolyticus*, *V. vulnificus*, and *V. alginolyticus* [30].

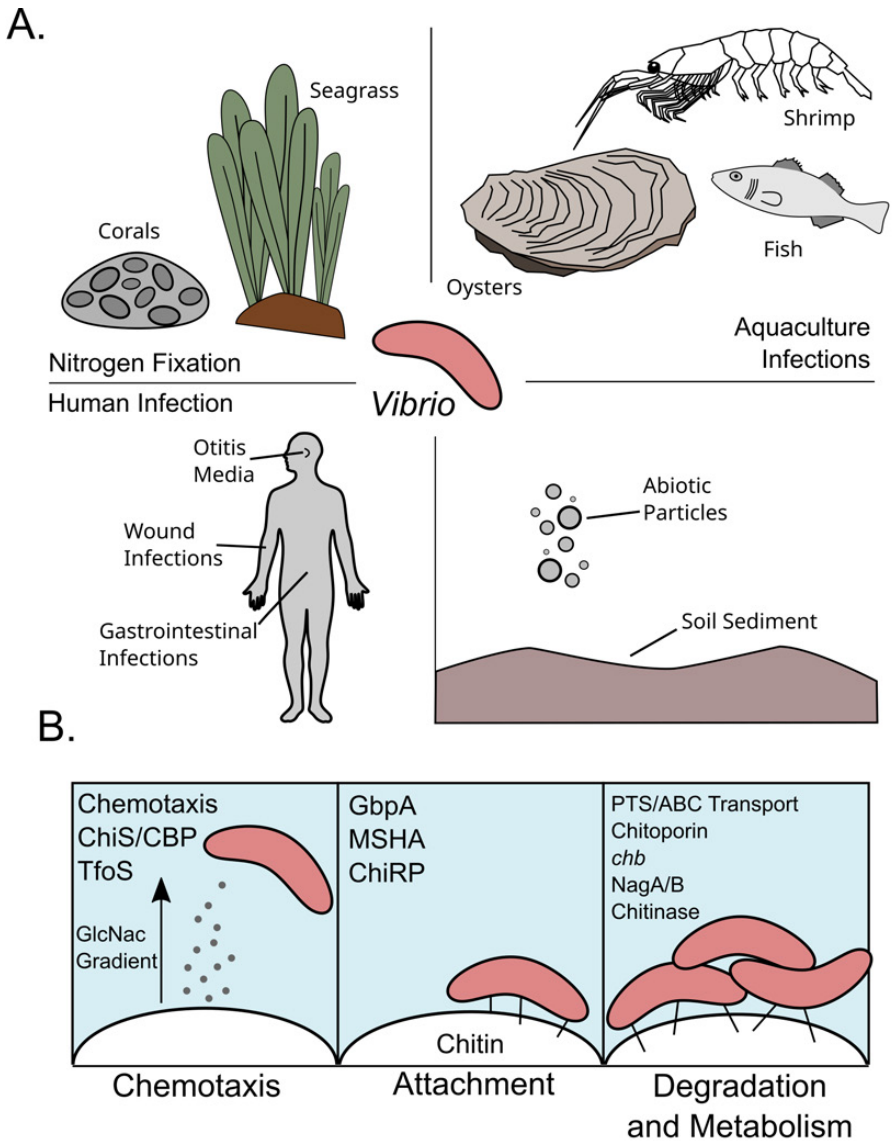

Figure 1.1: *Vibrio* Environmental Interactions. A. A broad visual representation of *Vibrio* spp. environmental interactions. *Vibrio* spp. interact with a variety of substrates in the environment to cause disease, fix nitrogen, and colonize surfaces such as abiotic particles and soil sediment. B. An overview of chitin sensing and degradation by marine *Vibrio* spp. The primary proteins/genes involved in each process are labelled. The black lines between the *Vibrio* cell and the chitin molecule represent attachment to the molecule.

Connecting the many lifestyles of the *Vibrionacea* – regarding basic biology, environmental ecology and survival, as well as host-pathogen interactions – will advance and clarify knowledge in the field of prokaryotic biology and physiology alone. Perhaps more practically, considering the broader picture of *Vibrio* biology will be vital in preventing future pathogen emergence, enhancing pathogen surveillance, and dealing with infections as they occur in both agriculturally important animals and in humans. Recent studies have indicated that there is a strong link between climate change and *Vibrio* infections [34, 35]. Perhaps most obvious is the direct correlation between increased sea surface temperature and the number of *Vibrio* infections in humans [34]. This raises concerns about how well we understand *Vibrio* pathogenesis, and if we are prepared for future *Vibrio* spp. epidemics or pandemics.

The work in this thesis will explore aspects of *V. parahaemolyticus* host-pathogen interactions and environmental survival, through exploration of type III secretion and chitin metabolism. In preparation of this exploration, I will introduce *Vibrio* biology by first examining a broad overview of their ecology and diversity, with a particular focus on the molecular mechanisms of chitin metabolism and degradation. I will then explore *Vibrio* pathogenesis broadly including a discussion of both agriculturally important pathogens and human pathogens. Following, considering this work is primarily focused on molecular mechanisms of *V. parahaemolyticus* pathogenesis, I will explore *V. parahaemolyticus* host-pathogen interactions in detail, with a particular focus on Type III secretion system biology and regulation. I will conclude this chapter by providing a rationale for the four linked but distinct studies presented in this thesis.

1.1 The Ecology of *Vibrio* Species

1.1.1 General *Vibrio* Biology

Vibrio species have a number of shared traits which influence their ecology and niche selection. *Vibrio* carry two chromosomes, both of which encode essential genes, and use a complex partitioning mechanism to maintain both chromosomes [36–38]. *Vibrio* spp. generally prefer higher temperatures (>18°C) and lower salinity conditions (<2.5% NaCl), although some *Vibrio* spp. have differential salt requirements, like *V. cholerae* which can survive in fresh water [39–41]. Importantly, the tolerance of *V. cholerae* for low salt conditions allowed it to spread and survive in drinking water, contributing to a number of distinct pandemics. Also, *Vibrio* spp. are known to be some of the fastest growing prokaryotic organisms on record, with *V. parahaemolyticus* and *V. vulnificus* doubling every 12-14 minutes and 20-25 minutes, respectively, and others species like *V. natriegens* having a doubling-time less than 10 minutes [42, 43]. For this reason, *V. natriegens* has been explored as a new tool for recombinant DNA technologies and synthetic biology [44, 45]. This growth rate impacts their ecology, as introduction of the appropriate nutrients to a marine environment allows rapid response and massive increase in *Vibrio* population. For example, the arrival of sub-saharan dust to tropical north Atlantic water – dust which often contains important nutrients for biological life, including iron – generates *Vibrio* blooms [46, 47].

The *Vibrio* genus contains remarkable diversity with over 120 known species and a pan-genome which contains over 17,000 gene families [13, 48]. These species represent symbiotic organisms, like *V. harveyi* and *V. alginolyticus*, which contributes to nitrogen cycling in *Mussismilia hispida*, or *V. fischeri*, who colonizes the light organ of the Hawaiian Bobtail Squid (*Euprymna scolopes*) and protects them from predators [20, 49, 50]. Others represent organisms of consequence

for both agricultural health and human health, such as *V. parahaemolyticus*, *V. cholerae*, *V. vulnificus*, and *V. coralliilyticus*, to name only a few [29, 51–56]. The broadness of hosts, ecosystems, and functions of *Vibrio* spp. make them particularly important organisms to study and understand (Figure 1.1).

One mechanism which drives the diversity of *Vibrio* spp. includes their genome plasticity and predisposition for horizontal gene transfer [13, 57–59]. Natural competence in *Vibrios* – the ability to uptake DNA from the environment and incorporate it into the genome – is linked to the colonization of chitin, where both expression of Type VI secretion systems (T6SS), as well as natural competence machinery is induced [59–62]. Importantly, the two systems appear to be linked, in that T6SS killing of neighbouring cells (see Section 1.2.4 Type VI Secretion Systems) releases large pieces of DNA which can be taken up by the *Vibrio* cell [63, 64]. Indeed, studies have indicated that *Vibrio* can capture very large pieces of DNA – many pieces larger that 100 kbp – from T6SS-targeted prey [65]. Horizontal gene transfer can dramatically shape the genome of *Vibrio* spp., as seen by Efimov *et al.*, where the acquisition of *Shewanella* genes by *V. vulnificus* biotype 1 led to the generation of wound-infection causing *V. vulnificus* biotype 3 in aquaculture ponds in Israel [29, 66, 67].

DNA uptake by *Vibrio* spp. is initiated by a regulatory cascade that begins at TfoS, a chitin oligosaccharide sensor which activates the expression of a small RNA, *tfoR* [62, 68]. TfoR is Hfq-dependent, and interacts with the 5' untranslated region of the mRNA for *tfoX* [62]. TfoX is a master transcriptional regulator of natural competence, driving the expression of many genes involved in the natural competence regulon [59, 69]. The presence of Type IV pili allows *V. cholerae* to pick up environmental DNA, which then enters the cell through the Type IV pilus machine [70]. ComEA unwinds the double-stranded DNA molecule, while ComEC carries a ssDNA strand through the inner membrane to be incorporated into the chromosome [59, 71–73]. Natural competence is also under the control of quorum

sensing, where at low-cell density a DNase is expressed which degrades dsDNA as it enters the cell. When quorum sensing regulators are present at high cell density, this gene is repressed and natural competence can proceed [59, 74].

1.1.2 *Vibrio* Interactions with Chitin

Chitin provides the foundation for many of the ecological relationships of *Vibrio* species. Chitin is the second most abundant carbon polymer on Earth, second only to cellulose, and the most abundant carbon polymer in the world's oceans. Chitin makes up the exoskeletons of a variety of marine phytoplankton, including microscopic crustaceans like copepods [75]. Macroscopic organisms, like crabs, molluscs, and shrimp also contain chitinous shells. Chitin is a β-1,4-N-linked polymer of N-acetyl-D-glucosamine (GlcNAc) and is a stable polymer, requiring specialized enzymes and organisms to recycle the carbon and nitrogen it contains. The *Vibrionacea* are one such group of organisms, as they encode a conserved chitin catabolism pathway and are capable of surviving on chitin as their sole carbon source [27].

Metabolism and degradation of chitin in the environment is a complex operation, involving: 1) chemotaxis and sensing, 2) attachment, and 3) degradation and assimilation (Figure 1.1B) [19] (Figure 1.1B). *Vibrio* spp. use chemotaxis to move towards the source of the GlcNAc oligosaccharides, and membrane-bound regulators like ChiS and TfoS to control the expression of proteins necessary for interaction with and degradation of the chitin molecule, including through the expression of chitinases and specialized catabolism enzymes (Figure 1.1B) [27, 59, 76]. *V. cholerae* encodes 68 open-reading frames annotated as putative chemotaxis genes, and studies have shown that GlcNAc oligosaccharides, GlcNAc monomers, and the heterodisaccharide 4-O-(N-acetyl-β-d-glucosaminyl)-d-glucosamine (GlcNAc-GlcN) drive chemotaxis in members of the *Vibrio* spp. [77–81]. Next, *Vibrio* must attach

to chitin, and does so with a variety of secreted factors, namely GlcNAc-binding protein A (GbpA) and Type IV pili (Figure 1.1B). Following attachment, the secretion of a number of chitinase enzymes in conjunction with specialized sugar catabolism proteins allow for the successful degradation and assimilation of the carbon and nitrogen found in chitin. (Figure 1.1B.)

Three primary chitin attachment mechanisms have been described, including two Type IV pili and an attachment molecule GbpA. The **M**annose **S**ensitive **H**em**a**gglutinin (MSHA) pilus has been shown as essential for the initial attachment of *V. cholerae* during biofilm formation, and is constitutively expressed in *V. cholerae* [76, 82, 83]. Alternatively, the **Chi**tin **R**egulated **P**ilus (ChiRP) discovered by Meibom *et al.* is activated by ChiS and is only expressed in media containing $(GlcNAc)_{2\text{-}6}$ oligosaccharides. ChiRP confers an advantage on chitinous surfaces, and is involved in bacterial agglutination during biofilm formation [84]. *Vibrio* also secretes an attachment molecule GbpA via the general secretion pathway (the general secretion pathway is also referred to as Type II Secretion, see Section 1.2.3 The Type II Secretion System), which is highly conserved in *V. cholerae* environmental and clinical strains, and aids in the colonization of biotic chitinous surfaces [85]. GbpA is a 4-domain protein where domain 1 interacts with chitinous surfaces, and domain 2 and 3 interact with the *V. cholerae* cell surface [86]. Domain 4 is so far unresolved, although sequence similarity to a chitinase chitin-binding domain suggests that it also has chitin binding activity. GbpA is proposed to be secreted by *Vibrio* spp., coating chitinous surfaces and enhancing colonization.

Chitin catabolism requires a complex array of enzymes and proteins needed to import, modify, and assimilate chitin from its large polymeric form, into smaller oligosaccharides, and eventually to GlcNAc monomers. Upon the sensing of GlcNAc oligomers by TfoS as well as ChiS, the expression of a number of chitinases is activated to degrade the chitin polymer [87]. After oligosaccharides are released

from the chitin polymer, these oligosaccharides can be imported into the cell through a chitoporin in the outer membrane, which is also regulated by ChiS [76, 88]. Following import, three seperate genetic regulons are activated, for degradation of nonacetylated chitin residues $(GlcN)_2$, $(GlcNAc)_2$, and GlcNAc monomers. The assimilation of $(GlcNAc)_2$ is coordinated by the sensor histidine kinase ChiS which directly activates the *chb* operon from the inner membrane [76, 89, 90]. Importantly, a GlcNAc-oligosaccharide periplasmic binding protein (CBP) interacts with ChiS in the absence of GlcNAc polymers, and prevents ChiS activation of *chb*. However, upon binding to GlcNAc polymers, CBP can no longer interact with ChiS, and ChiS is free to activate *chb* [89, 90]. The *chb* operon encodes an ATP-binding cassette (ABC) transporter necessary for the import of $(GlcNAc)_2$ across the inner membrane, and a number of enzymes which convert $GlcNAc_2$ into GlcNAc-6-Phosphate (GlcNAC-6-P). Further, the *nagE-nagBC* and *nagA* genes are required for catabolism of GlcNAc monomers and are regulated by NagC. NagC is a ROK (repressor orf kinase) family transcriptional repressor that regulates a number of genes involved in aminosugar catabolism. However, NagC is under allosteric regulation by GlcNAc-6-P, where when bound to GlcNAc-6-P, NagC cannot bind DNA and therefore does not repress its regulon. In turn, the NagE-NagBC divergent operon, as well as NagA are expressed. NagE encodes a phosphoenolpyruvate-dependent phosphotransferase system (PTS) which moves GlcNAc across the inner membrane and converts it to GlcNAc-6-P. NagA encodes a N-acetylglucosamine-6-phosphate deacetylase which converts GlcNAc-6-P to glucosamine-6-phosphate (GlcN-6-P), while NagB encodes a glucosamine-6-phosphate deaminase which converts GlcN-6-P to fructose-6-phosphate and NH_3. Alternatively, microarray analysis of *V. cholerae* identified seven genes differentially regulated when grown on $(GlcN)_2$, including a cellobiose PTS gene cluster [76]. In the absence of the IIC component of this PTS system, *V. cholerae* exhibited a growth defect on $(GlcN)_2$ indicating that *V.*

cholerae can grow on $(GlcN)_2$, and requires the PTS IIC component to do so [76]. All three of these pathways for the catabolism of chitin oligosaccharides provide fructose-6-phosphate which can feed into the glycolysis and pentose-phosphate central metabolic pathways.

As noted, chitin and GlcNAc have significant impacts on *Vibrio* spp. ecology. *Vibrio* interaction with chitin in the environment serves as the foundation for their colonization of chitinous organisms [91]. Chitin induces a number of cellular-responses, driving biofilm formation and protozoan-grazing persistence in *V. cholerae* [83, 92]. Further, *V. fischeri* identifies and chemotaxes towards the *E. scolopes* light organ through the sensing of $(GlcNAc)_{\geq 2}$ oligosaccharides [80]. *V. cholerae* adheres to the planktonic crustacean *Daphnia pulex* and *Daphnia magna*, and interacts with many other chitinous organisms such as the copepod *Tigriopus fulvus*, shrimp, and crabs using a variety of extracellular chitin-binding proteins and secretion systems [28, 76, 91, 93–95] (Figure 1.1A). These interactions provide *Vibrio* spp. with a valuable source of carbon and nitrogen, but can also concentrate pathogenic *Vibrio* spp., which can significantly increase risk of infection in humans (see Section 1.1.5 Human Pathogenic *Vibrio* spp.) [96–98].

1.1.3 From Environmental Organisms to Pathogens: *Vibrio* spp. and Emergent Virulence

Pathogenic bacteria are only pathogenic given the alignment of a variety of factors that give rise to something larger than the sum of its parts: pathogenesis. This process is known as emergent virulence, the process through which synergy between bacteria and host drives symptoms and disease [99]. For bacteria which are found environmentally and as human pathogens, this is a particularly important feature of virulence, as often the mechanisms used to invade, colonize, and drive disease in a host are tools these bacteria can use for environmental

survival as well. The understanding of how pathogenesis emerges is vital to predicting and managing future infections, and requires broad exploration and study of the ecology of bacterial pathogens. For example, *V. cholerae* uses GbpA to attach to chitinous surfaces in the environment, allowing for the efficient degradation and uptake of nutrients [85]. However, GbpA is also necessary for efficient attachment and colonization of host-cell surfaces, as GlcNAc decorates the outside of human epithelial cells and is found in intestinal mucin [28, 86]. While I have thus far discussed broad *Vibrio* ecology, I will now turn to their pathogenesis by exploring infectious organisms in aquaculture (primarily shrimp) and in humans, followed by a particular focus on *V. parahaemolyticus* host-pathogen interactions to provide appropriate context for this thesis.

1.1.4 *Vibrio* Pathogenesis in Aquaculture and Acute Hepatopancreatic Necrosis Disease

In 2020, the Food and Agriculture Organization of the United Nations (FAO) reported that world fisheries produced 179 million tonnes in 2018, while aquaculture alone contributed 82 million tonnes. Global fish production is valued at 410$ billion USD, while aquaculture alone is valued at 250$ USD billion [100]. The FAO also reports that while known pathogens contribute significantly to economic loss in the aquaculture setting, previously unknown emerging pathogens cause production losses roughly every three to five years [101]. Many of these established and emerging infections are attributed to *Vibrio* spp. and vibriosis remains the most common disease in aquaculture organisms worldwide [102, 103].

A number of well-described *Vibrio* spp. cause disease in agriculturally important species. *Vibrio* in these contexts are primarily opportunistic pathogens that cause a variety of diseases in crustaceans including acute hepatopancreatic necrosis disease (AHPND), a variety of shell diseases including luminescent vibriosis shell disease syndrome, and *Vibrio*-caused bacteremia [33, 104–109]. *V.*

harveyi, *V. parahaemolyticus*, *V. alginolyticus*, and *V. anguillarum* are important pathogens in a variety of aquacultured fish species, causing septicemia following infection [31]. Opportunistic *Vibrio* spp. are linked to infections in the Pacific oyster (*Crassostreae gigas*), particularly following a viral infection from *Ostreif herpesvirus*, leading to Pacific Oyster Mortality Syndrome [110]. Primarily, these diseases are seen in the summer months where water temperatures are more conducive to *Vibrio* spp. growth and spread. As well, in warmer waters, along with the over-crowded conditions common in aquaculture, fish can become stressed and immunocompromised, leading to further spread of the pathogenic *Vibrio* spp. as shown by Cheng and colleagues in the Orange-spotted grouper (*Epinephelus coidoides*) challenged with pathogenic *V. alginolyticus* [111].

An emerging and significant problem for aquaculture is the discovery of a strain of *V. parahaemolyticus* which causes early mortality syndrome (EMS, more recently known as acute-hepatopancreatic necrosis disease; AHPND) in agriculturally important shrimp species [112]. AHPND causes mass mortality events at infected aquaculture farms, through the destruction of the moribund shrimp hepatopancreas, an essential organ for digestion [16]. AHPND was first described as EMS in 2009 in China, and later spread to southeast Asia and Mexico [113]. AHPND was more recently identified in the United States in 2019 [114]. AHPND is estimated to have caused USD 43$ billion in losses across Asia and in Mexico [112]. Primarily, AHPND is caused by strains of *V. parahaemolyticus* carrying a set of *Photorhabdus* insect-related toxins $PirA^{Vp}$ and $PirB^{Vp}$ (See section 1.2.2 Toxins and Hemolysins) [113].

While *Vibrio* spp. can drive pathogenesis in agriculturally important organisms, disease to aquaculture animals is not the only disease concern caused by *Vibrio* spp. Species like *V. vulnificus* and *V. parahaemolyticus* can colonize and persist in molluscan shellfish such as oysters, leading to human disease [115, 116]. While not a direct disease or mortality in the affected seafood, infections from the

consumption of contaminated seafood cause significant economic losses [117]

1.1.5 Human Pathogenic *Vibrio* species

There are 4 primary human pathogenic species of *Vibrio*: *V. cholerae*, *V. parahaemolyticus*, *V. vulnificus*, and *V. alginolyticus*. However, a number of other *Vibrio* spp., can cause disease in humans, including *V. fluvialis*, *V. hollisae*, *V. mimicus*, and *V. metschnikovii* [30]. These *Vibrio* spp. largely drive infection through the consumption of contaminated water or seafood and cause gastroenteritis. Some species, like *V. cholerae*, *V. parahaemolyticus* and *V. vulnificus* frequently cause wound infections through contact with contaminated seawater or seafood [118]. Importantly, while *V. cholerae* and *V. parahaemolyticus* infections rarely cause septicemia, *V. vulnificus* infections are often associated with sepsis through either wound infections or the consumption of contaminated seafood, and have a high mortality (~25% for wound infections, ~50% for seafood consumption) [30, 119]. In the United States, *V. vulnificus* is responsible for >95% of contaminated seafood deaths [120].

The most prevalent and serious *Vibrio* spp. infectious disease is cholera, caused by the aptly named *V. cholerae*. *V. cholerae* causes upwards of 5 million cases and 100,000 deaths annually, and is endemic in India and Bangladesh [121–123]. Many serogroups of *V. cholerae* have been identified, largely characterized by the O-antigen of the lipopolysaccharide (LPS). The O1 and O139 serogroup of *V. cholerae* often carry cholera toxin, which is the key virulence factor driving severe dehydrating diarrhea. Cholera toxin is an AB_5 type toxin which interacts with the ganglioside GM1 on epithelial cells in the small intestine and ultimately drives significant water and salt loss from intestinal cells [30, 124, 125]. The diarrhea associated with cholera toxin is often referred to as rice-water stool, as it is associated with massive amounts of water loss, sometimes reaching >1L per hour [121]. While the diarrheal disease is thought to aid in the dissemination of

the bacteria from the host into the environment, it also aids in the acquisition of key nutrients [126]. Non-O1 and non-O139 strains of *V. cholerae* – which typically do not carry cholera toxin – can still drive serious gastrointestinal disease as well, without the characteristic rice-water stool [30].

V. vulnificus infection can occur from the consumption of contaminated seafood, or through wound infections, and is often characterized by a rapidly progressing and often fatal septicemia. A number of virulence factors, including secreted toxins such as multifunctional-autoprocessing repeats-in-toxin (MARTX), are essential for bacterial spreading beyond the intestine and for the widespread tissue damage seen in both intestinal and wound infections [127]. Importantly, *V. vulnificus* is an opportunistic pathogen, generally only causing disease in people with increased serum iron, often seen in liver disease patients [128, 129]. Understanding *V. vulnificus* pathogenesis remains an important priority, as the disease progression is often too rapid for treatment prior to severe septicemia.

While some *Vibrio* spp. are more common, others like *V. alginolyticus* can cause disease in humans although do so at a low rate and respond well to antibiotic treatment [30]. *V. alginolyticus* is often associated with wound infections, as well as eye and ear infections, although these infections rarely progress to sepsis [130]. However, the incidence of even the low-rate *Vibrio* infections is increasing [131]. As noted previously, corollary evidence from the Baltic Sea shows that increase in sea surface temperature associated with global climate change causes significant increases in *Vibrio* spp. infections [34]. In fact, long term ocean warming has increased *Vibrio* prevalence in northern oceans over time, according to retrospective molecular analyses of North Sea water samples [132]. While non-cholera *Vibrio* infections remain comparatively sparse, their increasing prevalence warrants considerable research into both their ecology and their host-pathogen interactions.

1.1.6 The *Vibrio parahaemolyticus* O3:K6 Pandemic Strain

V. parahaemolyticus, while not as prevalent as cholera or as fatal as *V. vulnificus*, also causes gastrointestinal disease in human hosts following the consumption of contaminated seafood, primarily undercooked molluscan seafood like oysters [55]. In fact, *Vibrio parahaemolyticus* is the leading cause of seafood-borne gastroenteritis worldwide. While a variety of toxins and secretion systems have been identified in *V. parahaemolyticus*, how it coordinates these secretion systems and toxins to cause disease remains somewhat unclear, in part due to a lack of effective animal models which replicate human infection (explored in detail in Section 1.2 *Vibrio parahaemolyticus* host-pathogen interactions).

V. parahaemolyticus was a known human pathogen prior to the emergence of the pandemic O3:K6 strain, causing occasional outbreaks from seafood consumption in the United States, while being prevalent in Japan [133, 134]. Surveillance of *V. parahaemolyticus* infections began in Kolkata, India began in 1994, and this surveillance identified a significant increase in infections in 1996. When compared with strains from before and after 1996 isolated from travellers in Japan, it became clear that the increase in infections was due to a new, unique O3:K6 serotyped strain isolated in Japan in 1995 [135]. Following identification of a similar O3:K6 strain in the United States [136], Chile [52], and Spain [137], it was labelled as the first human-pathogenic pandemic strain of *V. parahaemolyticus* [138]. This strain, labelled RIMD2210633, was sequenced in 2003 and remains the model strain for laboratory studies of *V. parahaemolyticus* host-pathogen interactions with humans [139]. Following this pandemic outbreak, O3:K6 infections were – and are – reported regularly from contaminated, under-cooked seafood [140–144]. The British Columbia Centre for Disease Control (BCCDC) indicates that *V. parahaemolyticus* infections are increasing in prevalence in Canada, with the largest outbreak recorded in British Columbia with 67 people

falling ill in 2015 [145, 146]. *V. parahaemolyticus* was estimated to cause roughly 30,000 infections annually in the United States in 2011 [147].

More recently, an endemic strain of *V. parahaemolyticus* in the United Stated Pacific Northwest, labelled ST36 (multilocus sequence type 36) was identified in Spain and the eastern coast of the United States [148, 149]. This particular strain has been identified to have a significantly lower infectious dose than the pandemic O3:K6 strain, at 10^3 - 10^4 cells compared to 10^7 - 10^8 cells for the O3:K6 strain [30, 150]. While not yet found worldwide, the spread of ST36 to other continents indicates that it has likely established environmental reservoirs and remains an emerging risk for the development of a new pandemic strain [151].

1.2 *Vibrio parahaemolyticus* Host-Pathogen Interactions

As noted above, *V. parahaemolyticus* O3:K6 is the leading-cause of seafood-borne gastroenteritis globally. Given its importance as a human pathogen, and the nature of its disease linked to agriculturally important animals, *V. parahaemolyticus* pathogenesis has been a topic of considerable research. However, while we understand that *V. parahaemolyticus* encodes a number of secretion systems and toxins, exactly how these mechanisms are coordinated to cause disease remains unclear. Here, I will explore *V. parahaemolyticus* host-pathogen interactions, including host attachment, Type II, III, VI, and VI secretion systems, as well as the relevant hemolysins and toxins. For the purposes of this thesis, particular attention will be paid to the Type III secretion systems encoded by *V. parahaemolyticus* and their regulation.

1.2.1 Host Attachment and Colonization: Extracellular Factors and Type IV Secretion

At the onset of infection, gastrointestinal bacterial pathogens must first establish a niche within a host, out-competing the native microbiome for limited space and adhere to host surfaces. As such, bacterial pathogens have evolved a variety of mechanisms to attach. The mechanisms used by *V. parahaemolyticus* for the attachment and colonization to the human intestine include a Type IV pili (MSHA), GbpA, cell-surface proteins (MAM7) and the two T6SSs [86, 152–154].

Many gram-negative bacteria encode Type IV pili machinery, often but not exclusively used for surface adherence. The type IV pili found in many *Vibrio* spp. including *V. cholerae* and *V. parahaemolyticus*, are composed of 4 primary components: one or more cytoplasmic ATPases and inner membrane platform, an outer-membrane pore, the pilus filament itself, and an alignment complex connects the inner membrane and outer membrane components [155–158]. Type IV pili form two subclasses that vary based on their amino acid sequence and length. Type IVa pili tend to have shorter mature pili sequences after peptidase cleavage and shorter leader sequences, while Type IVb pili have longer mature pilins, and longer leader sequences. The Type IV pili found in *V. parahaemolyticus* resemble the well-studied *P. aeruginosa* Type IVa pilus, and its structural proteins are presented here. The outer membrane complex is composed of PilQ subunits which form an outer membrane pore, along with the PilF pilotin protein [159, 160]. The inner membrane platform protein is composed of four proteins, PilC which forms the integral membrane platform, and PilB, PilT, and PilU which form a cytoplasmic ATPase [161]. PilMNOP form an alignment complex which connects the outer membrane pore with the inner membrane platform [162]. Finally, the pilus itself is generated from the major pilin subunit PilA, along with five minor pilins [157, 163]. This occurs through the polymerization of membrane localized

pilin subunits into a pilus structure, which is extended by the ATPase. Retraction ATPases are also common in Type IVa pili, depolymerizing the pilus to retract it back into the cell [164]. Type IV pili are involved in extracellular DNA uptake, inter-bacterial interactions and biofilm formation, as well as twitching motility and surface attachment [155, 165, 166].

V. parahaemolyticus encodes two Type IV pili, MSHA and ChiRP [84, 167]. ChiRP was discovered by Meibom *et al.* – first in *V. cholerae* – by use of a microarray with or without exposure to GlcNAc [76]. ChiRP presents a significant growth advantage to *V. cholerae* in the presence of chitin [76]. The MSHA pilus was also first discovered in *V. cholerae* [168]. MHSA is expressed during growth of *V. parahaemolyticus* in laboratory growth conditions, while ChiRP is regulated in *V. cholerae* by the presence of GlcNAc in these same conditions [84]. Decreased attachment is observed in mutants unable to produce both ChiRP and the MSHA pilus, implicating both pili in the colonization of chitinous surfaces in the environment and for biofilm formation [84]. Because the MSHA pilus is expressed in the absence of chitin, Shime-Hatorri and colleagues posit that MSHA may be involved in the initial attachment of *V. parahaemolyticus* to chitinous surfaces, followed by the upregulation and biosynthesis of ChiRP, allowing for more robust attachment to chitinous surfaces [84].

While these structures clearly aid in the adherence to surfaces in the environment, the MSHA pilus has also been implicated in colonization of the host intestine in *V. parahaemolyticus*. *V. parahaemolyticus* generates MSHA-dependent aggregates on glass coverslips, and an *mshA* pilin gene null mutant is reduced for the ability to attach to glass coverslips indicating that it is involved in surface attachment [84]. In agreement with Shime-Hatorri *et al.*, O'Boyle and colleagues show that deletion of essential genes involved in the formation of the MSHA pilus significantly reduces virulence towards Caco-2 cells *in vitro*, and drives a 60% decrease in cellular adherence [152]. In the same study,

O'Boyle *et al.* show that the MSHA pilus has lectin like features, binding to a number of cell associated molecules like blood groups A and B, oligosaccharide antigens Lewis A and X, as well as lacto-N-fucopentaose I, asialo-GM1 ganglioside, and lacto-N-difucohexaose I [152].

Another example of environmental chitin attachment facilitating colonization in mammalian hosts is the secreted protein GbpA, which interacts with GlcNAc and intestinal mucin *in vitro* [85, 86]. A *gbpA* null mutant is deficient for colonization and virulence in a mouse model in *V. cholerae*, and shares 70% sequence identity with the *V. paraheamolyticus* GbpA protein, suggesting that it may be involved in *V. parahaemolyticus* infection as well [28, 169]. Indeed, GbpA is activated upon surface-sensing in *V. parahaemolyticus* [170, 171].

Other molecules generated by *V. parahaemolyticus* also aid in adherence, such as multivalent adhesion molecules (MAM), which are used directly for adherence to mammalian cells [172]. MAM proteins have either six mammalian cell entry (mce) domains (MAM6) or seven mce domains (MAM7) and are found exclusively in gram-negative pathogens. *V. parahaemolyticus* encodes a MAM7 protein which is involved in attachment to HeLa and Caco-2 epithelial cells, RAW264.7 macrophages, and 3T3 fibroblasts *in vitro* [153]. MAM7 null mutants were also unable to generate infection phenotypes comparable to the RIMD2210633 wildtype strain in *C. elegans* infections [172]. MAM7 in *V. parahaemolyticus* facilitates both protein-protein adhesion (MAM7-fibronectin) and protein-lipid adhesion (MAM7-phosphatidic acid) as well as tight adherence to host cells.

V. parahaemolyticus also encodes two T6SSs which are located each on their respective chromosomes – T6SS-1 on chromosome 1 and T6SS-2 on chromosome 2 (See Section 1.2.4 Type VI Secretion Systems). Both T6SS-1 and T6SS-2 have been implicated in adherence to HeLa cell monolayers by the deletion of T6SS-1 and T6SS-2 core genes *icmF1*, *icmF2*, *hcp1*, and *hcp2* and can be complemented by reintroduction of these genes *in trans* [154]. While both T6SS-1 and T6SS-2

seem to be important for adhesion to HeLa cells, only T6SS-2 was necessary for adherence to Caco-2 cells, where deletion of *icmF1* and *hcp1* did not impact cell adherence. The mechanism through which T6SS-1 and T6SS-2 adhere to cells *in vitro* remains unclear [173], although T6SS-2 has been linked to autophagy in macrophages and therefore clearly interacts with host cells [174].

Many of the systems that *V. parahaemolyticus* uses for adherence and colonization in hosts are regulated by quorum sensing. Indeed, deletions of important quorum sensing regulators *luxO* and *aphA*, as well as overexpression of a quorum sensing master regulator *opaR*, significantly reduced *V. parahaemolyticus* colonization in animal models [175]. While this study indicates direct changes in *V. parahaemolyticus* metabolism in these null mutants which likely contribute to decreased fitness in a host, many other studies have shown more directly that quorum sensing is involved in the regulation of the MSHA pilus, capsular polysaccharide production, GpbA expression, and T6SS regulation [176–180]. For examples, GbpA protein expression is modulated by quorum sensing, where GbpA expression and secretion can be detected during low-cell density, but not at high-cell density [177]. A Δ*hapR* strain of *V. cholerae*, which cannot integrate the quorum-sensing signals and act to regulate quorum-sensing regulated genes, maintains high levels of secreted GbpA in the culture supernatant [177, 181]. While the expression of GbpA is constant in *Vibrio cholerae* in laboratory conditions, the decrease of secreted GbpA in the culture supernatant is due to the expression of a HapR regulated protease which degrades GbpA at high cell density [169].

1.2.2 Toxins and Hemolysins

V. parahaemolyticus encodes a number of toxin and hemolysin proteins which contribute to its pathogenesis in both humans and shrimp. While toxins PirA/PirB, thermostable direct hemolysin (TDH) and TDH-related hemolysin

(TRH) are probably the most relevant and important toxins for infection, *V. parahaemolyticus* also encodes a thermolabile hemolysin (TLH), as well as a number of collagenases and proteases [182–184]. Further, *in vitro* studies showed that *V. parahaemolyticus* cells can invade and survive intracellularly, and require a secreted toxin/lipase VPA0226 for egress from host cells [185].

Acute Hepatopancreatic Necrosis Disease (AHPND) occurs in moribund shrimp and causes significant economic loss after infection spreads in an aquaculture environment [112]. Early accounts showed that while sloughing of hepatopancreas cells was associated with AHPND, bacteria were not found in the lesions, suggesting that a secreted toxin was the culprit. *V. parahaemolyticus* strains that contain a virulence plasmid named pVA1 were eventually discovered to cause AHPND [112]. The plasmid is 63-70kb in size, and encodes homologs of the *Photorhabdus* insect-related toxins $PirA^{Vp}$ and $PirB^{Vp}$ [113]. Injection of moribund shrimp with recombinant $PirB^{Vp}$ alone was not sufficient to cause AHPND-like histology, however introduction of $PirB^{Vp}$ and $PirAB^{Vp}$ heterodimers induced histological changes consistent with AHPND, with $PirAB^{Vp}$ heterodimers driving the most severe change in histology. When fed *per os*, neither protein alone drove significant histological changes, but in combination induced AHPND-like changes in histology, implicating both $PirA^{Vp}$ and $PirB^{Vp}$ as necessary for naturally occuring AHPND [113]. Curing of pVA1 in infectious strains of *V. parahaemoylticus* abolished their ability to cause AHPND, further implicating the pVA1 plasmid and $PirAB^{Vp}$ as the virulence factors necessary for AHPND. Crystal structures of $PirA^{Vp}$ and $PirB^{Vp}$ resemble those of the *Bacillus thuringiensis* crystal insecticidal (Cry) toxin, and therefore likely function in a similar fashion where $PirA^{Vp}$ forms pores in membranes, while $PirB^{Vp}$ is involved in receptor binding [113]. Further, a more recent study identified two aminopeptidases called *Lv*APN1 and *Lv*APN2 which are expressed by *V. parahaemolyticus*-challenged *L. vannamei* hemocytes. Importantly, $PirAB^{Vp}$ appears to interact with *Lv*APN1 to gain access

to hemocytes and drive hemocyte destruction [186]. While previous work with PirAB^{Vp} identified that it damaged stomach and hepatopancreatic cells, it also appears that it may have a direct affect on hemocytes via *Lv*APN1 [113, 187].

Human pathogenic strains of *V. parahaemolyticus* carrying the VpPAI pathogenicity island encode two copies of the TDH gene (*tdhA* and *tdhS*), as well as a non-VpPAI encoded TDH-related hemolysin (TRH) [139, 183]. TDH is primarily responsible for a hemolytic ring on blood agar plates in particular conditions, known as the Kanagawa phenomenon [184, 188]. TDH forms tetramers and generates 2nm pores in the membranes of erythrocytes in a mechanism that is not dependent on lipid rafts, although the cytotoxicity seen in TDH against nucleated cells appears to be dependent on sphingomylin and lipid rafts [183, 189, 190]. TRH shares 63% sequence similarity with TDH and forms tetramers like TDH [191]. TRH also has hemolytic activity *in vitro*, although its role during infection remains unclear. Importantly, while these toxins contribute to fluid accumulation and enterotoxicity, and are expressed during infection, strains with genetic deletions of these toxin genes still generate significant cytotoxicity against HeLa cells and drive enterotoxicity in a rabbit ileal loop model of infection [192–195]. This suggests that while they contribute to pathogenesis, other factors are essential for cytotoxicity and enterotoxicity.

1.2.3 The Type II Secretion System

Type II Secretion systems resemble the structure of Type IV pili, using a similar secretion apparatus to push target proteins into the extracellular spaces. Type II secretion systems are generally used by bacteria to secrete large exoproteins like chitinases, cellulases, collagenases, and proteases that contribute to biofilm formation and extracellular polymer degradation [196, 197]. Exotoxins, such as cholera toxin from *V. cholerae* [198], and hemolysins are also often targeted for Type II secretion.

Type II Secretion systems (T2SS) can be divided into three complexes. The outer membrane complex is comprised of GspD multimeric pores and the GspS pilotin, while the inner membrane complex is comprised of GspC, GspF, GspL, and GspM proteins, as well as the cytoplasmic ATPase GspE [199, 200]. The inner membrane complex is also known as the assembly platform (AP). The pseudopilus structure comprises the third complex and is comprised of the major pseudopilin GspG, and a number of minor psuedopilins (GspH, GspI, GspJ, and GspK) [200]. A peptidase, GspO, processes the psuedopilins into mature form to be incorporated into the growing pseudopilus structure by GspE [201, 202]. Generally T2SS genes are structured into a single operon from *gspA-O*, although variations are common and not all *gspA-O* genes are found in all bacteria. Indeed the number of essential genes for T2SS expression varies from 12 - 15 [196]. *gspS* may be co-located with the other T2SS genes on the chromosome, or elsewhere [203].

The T2SS spans the periplasm and both membranes to push proteins into the extracellular space. The Lol machinery is involved in secretion of GspD and GspS subunits into the periplasm, where the GspS pilotin aids in the insertion of GspD subunits into the outer membrane [200, 204, 205]. The AP, the psuedopilus, and a cytoplasmic ATPase make up the rest of the components and localize at the inner membrane [200]. GspCFLM span the inner membrane, and generate a secretion platform that interacts with the ATPase GspE in the cytoplasm. It is unclear how this complex forms in the inner membrane, but it is suggested that the outer membrane complex aids in this insertion of GspC and GspM, as they form florescent foci in *V. cholerae* only in the presence of the outer membrane complex [206]. Finally the ATPase GspE, in combination with pseudopilins (GspGHIJK) and the prepilin protease (GspO) assemble the pseudopilus at the inner membrane complex.

Proteins targeted to be secreted by the Type II Secretion system carry an N-terminal signal sequence that targets them to be secreted into the periplasm

by the SecYEG or Tat complexes [207–209]. How exactly substrates are loaded into the secretion system remains unclear, however it appears that interactions with GspC and GspD are necessary for substrate loading [209, 210]. Finally, the ATPase is stimulated to assemble the pseudopilus and push the secreted protein out of the cell through the outer membrane complex [211, 212].

In *V. parahaemolyticus*, TDH carries an N-terminal secretion signal for secretion using the Sec translocon followed by type II secretion [184, 213]. Further, the lipase (VPA0226) necessary for intracellular egress *in vitro* is also secreted via the T2SS and Sec pathways. Finally, GbpA is also targeted for Type II Secretion [86]. Presumably, a number of other unidentified/uncharacterized proteases and mucin degrading enzymes are also secreted through the T2SS in *V. parahaemolyticus* [197].

1.2.4 The Type VI Secretion Systems

Type VI secretion systems are unique bacterial protein secretion machines which resemble molecular spears and deliver effector proteins directly into target cells, most notably other bacteria [214, 215]. These secretion systems contribute to virulence by aiding colonization of host intestinal tracts, by competing out the host microbiome, or by directly injecting effector proteins into eukaryotic cells to change host cell biology [214]. In some cases, this change in host cell biology can actually impact the microbiome, as seen in *V. cholerae*, where T6SS expressing cells drive an increase in peristalsis in zebrafish, functioning to expel host gut symbionts to allow *V. cholerae* colonization [216].

The T6SS resembles a *Myoviridae* phage tail, which is inserted into the gram-negative inner membrane and acts like a microbial spear, contracting to push effector proteins through the membrane of neighbouring cells [217]. The contractile machinery is made up of a minimum set of 13 proteins, a membrane complex (MC), a baseplate (BP), and the tail-tip complex (TTC) [214]. The

membrane complex consists of three proteins, two integral membrane protein TssM and TssL, and an outer membrane lipoprotein TssJ [218, 219]. TssJ aids in the assembly of the inner membrane complex [220]. Together, TssJLM form a pore in the inner and outer membrane, span the periplasm, and provide a channel for the rest of the contractile machinery. Following, the MC recruits a number of baseplate proteins TssA, TssE, TssF, TssG, TssK, and Valine-glycine repeat protein G (VgrG) to act as a platform for the assembly of the TTC [221, 222]. The TTC is primarily made up of Hemolysin-coregulated protein (Hcp) which form stacked hexameric rings, followed by the assembly of the contractile sheath made up of TssB and TssC, which enclose the Hcp hexamers [221, 223]. However, how Hcp contraction and secretion of the VgrG tip protein occurs is still unclear, although TssB and TssC inter-subunit interactions appear to be essential in this contraction process [224]. Upon contraction, a AAA^+ ATPase ClpV recognition motif is accessible in TssC, permitting disassembly and sheath recycling by ClpV [225, 226].

VgrG is a trimeric protein which forms the tip of the contractile machinery. The VgrG trimer is sometimes capped by a Pro-Ala-Ala-Arg (PAAR)-repeat containing protein which sharpens the tip and presumably enhances the ability of VgrG, and the contractile machinery, to pierce the target cell [227]. Effectors can either be extensions of the structural components of the T6SS inner tube, or may interact directly with the VgrG or Hcp, occasionally with the help of accessory proteins [222]. These effectors belong to a few major classes: Type six amidase effectors (Tae), Type six glycoside effectors (Tge), Type six DNase effectors (Tde) and Type six lipase effectors (Tle) [228–231]. These effectors target peptidoglycan (Tae and Tge), have nuclease activity (Tde), and disrupt membrane stability (Tle). Bacteria which encode these toxic effectors also encode immunity proteins to protect themselves from damage prior to the secretion of these effectors [222]. Importantly, structural proteins can have their own effector functions, for example

some VgrG proteins cross-link actin [232].

V. parahaemolyticus encodes two T6SS's each found on a different chromosome. T6SS-1 is encoded on chromosome 1 and is primarily found in clinical isolates, while T6SS-2 is encoded on chromosome 2 and is found in many strains of *V. parahaemolyticus* – suggesting T6SS-2 has an ancestral origin, while T6SS-1 was likely more recently acquired [154]. T6SS-1 is activated in marine conditions (high salt, 23°C or 30°C), while T6SS-2 is located on chromosome 2 and is differentially regulated, expressed at 23°C, 30°C, and 37°C in low salt condition [233]. Here it is suggested that T6SS-2 functions inside the host, perhaps during infection, and T6SS-1 is involved in environmental fitness, especially in warmer waters [233]. Indeed VgrG2 of T6SS-2 in *V. parahaemolyticus* induces autophagy in macrophages and both secretion systems appear to promote cell-adhesion to monolayers *in vitro* [154, 174]. T6SS-1 is also found in all strains that cause AHPND in shrimp, suggesting that T6SS-1 confers a fitness advantage [234].

1.2.5 The Two Type III Secretion Systems

Type III Secretion Systems (T3SS) are highly sophisticated and well-studied protein secretion systems [235, 236]. T3SS structure resembles a molecular needle which transverses three biological membranes, both bacterial membranes and the host cell membrane to directly deliver effector proteins into host cells, to various effect [237]. These nanomachines are found widely-distributed in gram-negative bacteria – including a variety of human pathogenic *E. coli*, *Salmonella*, *Shigella*, *Yersinia*, *Pseudomonas*, and *Vibrio* species – and generally target conserved cellular processes like cytoskeletal structure, host immune responses, signal transduction, and vesicle transport [238]. Some intracellular pathogens like *Salmonella* and *Shigella*, which invade epithelial cells, use T3SS needles to gain entry to host cells, modifying cytoskeletal structure to induce membrane ruffling

and internalization, followed by the secretion of factors to escape the vacuole and enter the cytosol, as well as evade host immune responses [239]. Similarly, primarily extracellular pathogens like *Vibrio*, *Yersinia*, *Psuedomonas* and *E. coli* pathogens deliver effector proteins to alter host cell biology, generally by modifying host cell cytoskeletal structure and interrupting host immune responses [240].

T3SSs are composed of greater than 20 proteins, which comprise a inner and outer membrane pore, a cytoplasmic ring and ATPase export complex, a needle filament, and a needle tip and translocon pore (Table 1.1) [241]. Secreted proteins are called effectors as they cause changes to host cell biology to aid in the survival, invasion, and persistence in a host [240]. Here, when describing the structure of the T3SS, I will use the unified secretion and cellular translocation (Sct) nomenclature first proposed in 1998 and later expanded to provide an overview of the common T3SS structure [242, 243]. Fortunately, *Vibrio* nomenclature follows the proposed universal nomenclature closely, and so both are presented here (Table 1.1, Figure 1.2) [244]. The basal body structure of the needle apparatus spans both bacterial membranes and the periplasm, and is composed of outer membrane (SctC), inner membrane (SctJ) and periplasmic (SctD) rings (Figure 1.2) [238]. A pilotin protein is involved in membrane insertion and assembly of the ring components, and lytic transglycosylase enzymes are encoded to disrupt peptidoglycan to aid assembly of the structure [245]. The basal body also includes the needle filament composed of the inner rod (SctI) protein and the needle protein (SctF), both of which form polymeric helical structures. The export apparatus, which is located at the base of the basal body at the inner membrane and acts as a platform for substrate entry, is composed of inner-membrane components SctS, SctT, and SctR, as well as the autoprotease SctU and the export gate SctV [243]. The C-ring and ATPase complex, which provides energy for secretion and acts as a hub for substrate sorting and loading, is composed of the cytoplasmic ring structure (SctQ) and ATPase components SctN, SctL, SctO, and SctK. SctN is the

Universal Nomenclature	Function	T3SS-1 Nomenclature	T3SS-2 Nomenclature
SctC	OM Ring	VscCI (VP1696)	VscC2 (VPA1339)
SctD	Periplasmic Ring	VscD1 (VP1695)	–
SctJ	IM Ring	VscJ1 (VP1690)	VscJ2 (VPA1367)
-	Pilotin	ExsB$^{?}$ (VP1700)	–
SctS	Export Component	VscS1 (VP1673)	VscS2 (VPA1335)
SctT	Export Component	VscT1 (VP1674)	VscT2 (VPA1341)
SctR	Export Component	VscR1 (VP1672)	VscR2 (VPA1342)
SctU	Autoprotease	VscU1 (VP1675)	VscU2 (VPA1354)
SctV	Export Gate	VscV1 (VP1662)	VscV2 (VPA1355)
SctQ	Cytoplasmic Ring	VscQ1 (VP1671)	VscQ2 (VPA1349)
SctN	ATPase	VscN1 (VP1668)	VscN2 (VPA1338)
SctK	ATPase Co-Factor	VscK1 (Vp1698)	–
SctL	Stator	VscL1 (VP1688)	–
SctO	Stalk	VscO1 (VP1669)	–
SctP	Needle-length Regulator	VscP1 (Vp1670)	–
SctI	Inner Rod	VscI1 (VP1691)	–
SctF	Needle	VscF1 (VP1696)	–
SctW	Switch Regulator	VscW1 (VP1667)	–
SctE	Translocon Pore	VopD1 (VP1656)	VopD2 (VPA1361)
SctB	Translocon Pore	VopB1 (VP1657)	VopB2 (VPA1362)
SctA	Needle Tip	VscA1 (VP1659)	VopW (VPA1345)

Table 1.1: T3SS-1 and T3SS-2 Structural Components. Dashes indicate proteins yet to be identified, $^{?}$ indicates an unclear role in type III secretion. Table modified from Matsuda *et al.* [244].

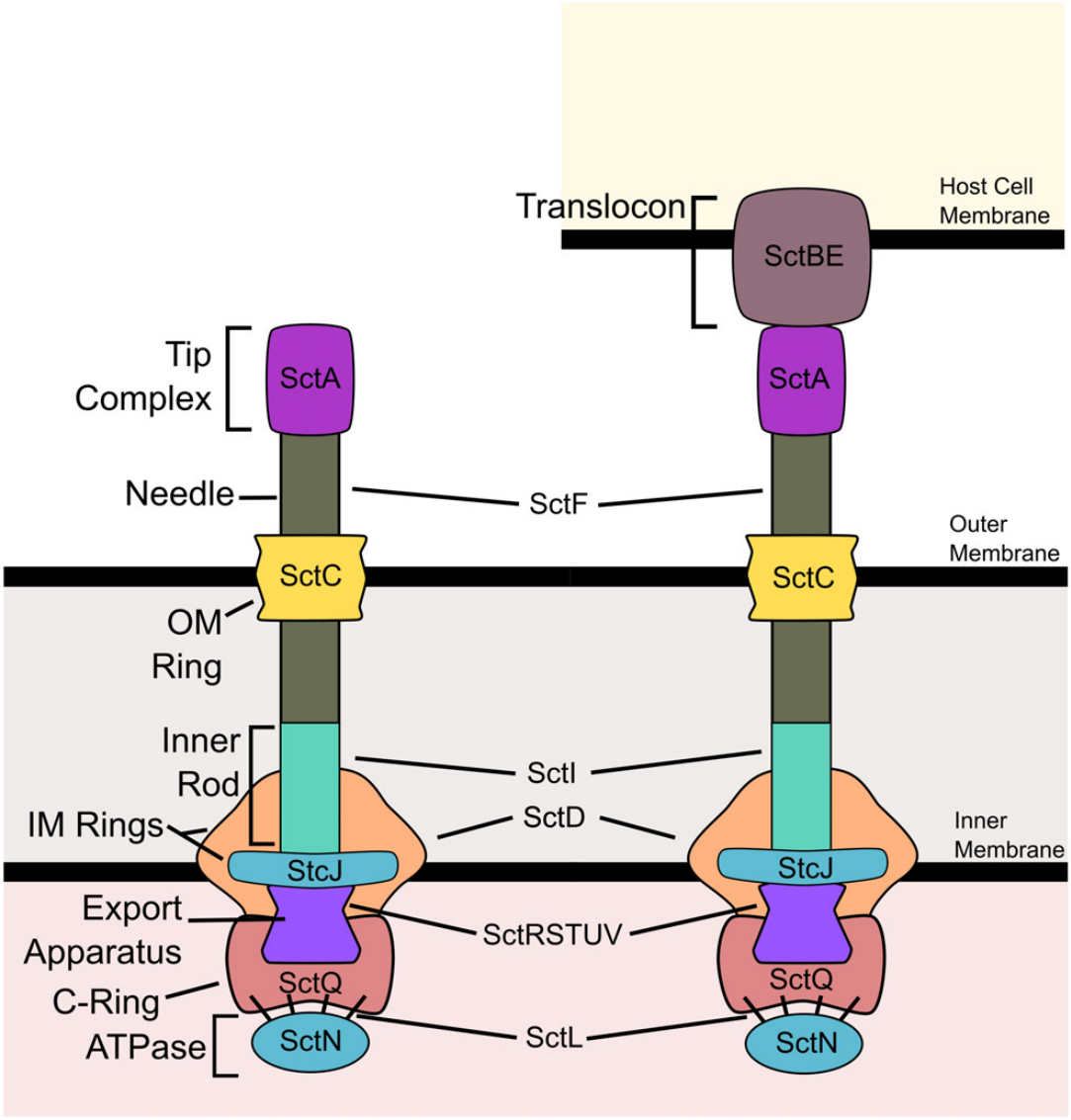

Figure 1.2: Universal T3SS Core Structure. Universal nomenclature is used to show the core structure of the Type III Secretion System. Labels are shown for membranes, and protein components are labelled with their universal nomenclature proteins as well as their function [242, 243].

ATPase, SctL is the stator protein, SctO is a stalk protein, and SctK is an ATPase co-factor which together make up the ATPase complex. Finally, the needle tip protein (SctA) and translocator proteins (SctB and SctE) bridge the needle apparatus and the host cell (Figure 1.2). Two additional protein regulators, SctP and SctW regulate the needle length and substrate switching between middle (translocators) and late (effector) secretion, respectively [238, 246, 247]. Taken together, these proteins form a multi-megadalton nanomachine which capture effector proteins at the cytoplasmic ring, secrete them via the ATPase complex into the needle lumen contained by multiple membrane rings, to directly secrete effectors into host cells through the needle tip and translocon proteins (Figure 1.2) [248]. Cryoelectron microscopy (cryo-EM) – in particular two techniques called single-particle analysis (SPA) and cryoelectron tomography (cryoET) has generated significant advances in the resolution and visualization of these structures – and has resolved much of the needle apparatus in bacterial membranes [237, 249, 250].

Most *Vibrio* spp. encode a T3SS, while *V. parahaemolyticus* encodes two T3SSs, one on each chromosome 1 and chromosome 2. These two secretion systems, T3SS-1 and T3SS-2 named for the chromosome they are found on, are implicated in different functions, namely host cell cytotoxicity and intestinal enterotoxicity respectively [251, 252]. The structural proteins of T3SS-1 are conserved across a number of different *Vibrio* spp., and are well characterized, while the structural components of T3SS-2 remains only partially characterized (see Table 1.1). Notably, the structures of the needle (SctF/VscF2), IM ring (SctD/VscD2), inner rod (SctI/VscI2), and the gate keeper protein (SctW/VscW2), as well as 3 proteins in the ATPase complex (SctLOP/VscLOP2), remain elusive [244]. Further, T3SS-2 is located on a pathogenecity island (VpPAI), where a number of genes – more than half – remain uncharacterized.

As noted, T3SS-2 causes enterotoxicity in hosts, and has been linked to invasion

of host cells *in vitro* [251–253]. Importantly, studying T3SS-2 against human cell culture *in vitro* requires the deactivation of T3SS-1, as the rapid host cell cytotoxicity caused by this strain makes it difficult to study T3SS-2 function alone [194, 254]. Identified secreted proteins from T3SS-2 include VopA/P, VopL, VopT, VopV, VopC, VopZ, VopO and VPA1380 (Table 1.2) [252, 254–262]. VopA/P inhibits the mitogen-activated protein kinase (MAPK) pathway in host cells to evade immune responses [256]. This is unique from the similar effector YopJ in *Yersinia* spp., which inhibits both MAPK and NF_kB signalling pathways [139]. VopA/P functions as an acetylase which acetylates MAPK kinases and IKKβ in essential residues for phosphorylation, which in turn prevent their activation from upstream kinases [256]. VopV, VopC, VopL, and VopO all impact cytoskeletal structure through different mechanisms. First, VopV binds to the polymer form of actin (F-actin) and is the primary driver of enterotoxicity and fluid accumulation in the rabbit ileal loop model of infection [263]. Further, VopC appears to be the effector essential for the internalization of *V. parahaemolyticus* cells as seen in cell culture [254]. VopC deamidates glutamine 61 in CDC42 and Rac – two Rho-GTPases essential for cytoskeletal structuring – generating membrane ruffles which internalize *V. parahaemolyticus* cells [254, 264, 265]. Also, VopL modulates cytoskeletal structure by binding to actin monomers and encouraging the formation of polymerized actin fibers [258]. Wang *et al.* suggest that VopL may also enhance the uptake and invasion of host cells by *V. parahaemolyticus* by generating a favourable microenvironment for growth [255]. However, the role of intracellular infection by *V. parahaemolyticus in vivo* remains unclear. VopL is also involved in inhibiting the host ROS response, again aiding in the survival of intracellular *V. parahaemolyticus* [266]. Finally, VopO interacts with GEF-H1 (a RhoA guanine nucleotide exchange factor or GEF) which, through mechanisms which are still not completely resolved, activates the RhoA-ROCK pathways and induces stress fibre formation in host cells. Two additional effectors, VopZ and

Effector	Locus Tag	Function	Effect on Host Cells
T3SS-1			
VopQ	*vp1680*	Binds V_0 subunit of H+-ATPase	Activates Autophagy, cell lysis, MAPK activation, IL-8 secretion
VopS	*vp1686*	Inhibition of Rho GTPases by AMPylation	Cell rounding, phagocyte invasion
VPA0450 (VepA)	*vpa0450*	Phosphatidylinositol phophatase	membrane blebbing
VopR	*vpa1683*	Binds PIP_2	Promoters effector refolding
T3SS-2			
VopA/P	*vpa1346*	Inhibition of MAPK by acetylation	Blocks phosphorylation cascade and ATP binding
VopT	*vpa1327*	Ras ADP-ribosylation	Cytotoxicity
VopL	*vpa1370*	Actin nucleation	Stress fibre formation, cell shape change
VopC	*vpa1321*	Activation of Rac and CDC42 by deamidation	Invasion of non-phagocytic cells
VopV	*vpa1357*	Actin binding and bundling	Enterotoxcity and blunting of villi
VopZ	*vpa1336*	Inhibition of TAK1 and downstream MAPK and NF_kB	Enterotoxicity and colonization
VopO	*vpa1329*	Binds to GEF-H1 to activate Rho-ROCK pathway	Stress fibre formation.
VPA1380	*vpa1380*	Cysteine catalysis dependent on IP6	Toxicity in Yeast

Table 1.2: T3SS-1 and T3SS-2 Secreted Effectors in *V. parahaemolyticus*. Effector names and gene locus tags for *V. parahaemolyticus* RIMD2210633 are provided. Both the molecular function of the effectors, as well as their effect on host cells are presented. Table modified from Wang *et al.* [255].

VPA1380, have attributable phenotypes, but their mechanisms remain unclear. VopZ causes fluid accumulation, and epithelial damage and shedding by inhibiting the activation of the TAK1 kinase, and by extension both the MAPK and NF_kB pathways [262]. VPA1380 is an inositol hexakisphosphate (IP6) cysteine protease that drives toxicity in yeast cells, but is not toxic in cells which lack IP6 [253, 260, 263]. Importantly, other large *Vibrio* toxins often contain IP6 domains, although the exact function of VPA1380 remains unknown [267]. Taken together, T3SS-2 effectors drive enterotoxicity, actin rearrangement, fluid accumulation, and dampen innate immune responses.

Conversely, T3SS-1 causes host cell cytotoxicity in cell culture, and does so rapidly in 4 hours at a low multiplicity of infection (MOI) [251, 268, 269]. Functional protein secretion studies from our lab and others have identified a significant portion of the T3SS-1 secreted proteins including VscA1 (also known as VcrV), VopQ, VopB, VopS, VopD, VscW1 (also known as VopN), and VPA0450 [268, 270]. VscA1 is the needle tip protein, VscW1 is the switch regulator, while VopB and VopD are translocators [244, 270, 271]. Two particularly potent effectors, VopQ (also known as VepA) and VPA0450, drive oncosis and membrane blebbing, respectively (Table 1.2) [270, 272–277]. VopQ causes oncosis of the host cell through interaction with the V_o subunit of the lysosomal H+ V-ATPase, and causes the lysis of lysosomes and the spilling of the acidic and proteolytic contents into the cytosol, leading to rapid cell death [273]. VopQ also increases autophagy and activates JNK, ERK, p38 innate immune molecules [274, 275, 278]. Further, VPA0450 has inositol polyphosphate 5-phosphatase activity and cleaves phospholipid phosphatidylinositol (4,5)-bisphosphate (PIP_2) to D5 phosphate at the plasma membrane, and disrupts binding of the cytoskeleton to the plasma membrane, leading to membrane blebbing [276]. Other accessory effector proteins include VopS and VopR. VopS encodes an AMPylase which AMPylates host Rho GTPases and mediates actin rearrangement, which causes cell rounding

and allows the bacteria to escape phagocytosis of infected cells by macrophages [279, 280]. VopR contains a phosphoinositide-binding domain (BPD) similar to VPA0450, and therefore localizes to the plasma membrane where its binding to PIP_2 induces its folding in the host cell. Although the function of VopR is less clear, it may help refold effectors after their secretion into host cells [255].

T3SS-2 is considered to have been acquired more recently, while T3SS-1 is an ancestral secretion system as determined by GC% content from sequence analysis [139, 251]. The VpPAI that contains T3SS-2 contains the secretion system, as well as two copies of a thermodirect hemolysin (TDH), *tdhA* and *tdhS*, two hemolytic toxins commonly associated with virulence in *V. parahaemolyticus* [251]. Genetic comparisons of the *V. parahaemolyticus* environmental and clinical isolates indicate that T3SS-2 is commonly present in clinical isolates [281]. T3SS-2 is also linked to colonization of the mammalian intestine by a number of studies, including a whole genome survey using transposon sequencing [282, 283]. Intriguingly, T3SS-2 has also been implicated in the invasion of cells in tissue culture by *V. parahaemolyticus*, although the importance of this in animal models has not been established [239, 254, 266]. While T3SS-2 is generally considered a necessary pathogenecity determinant for *V. parahaemolyticus* gastrointestinal disease, T3SS-2 has also been implicated in the defense against predation by a number of marine protists [284]. Conversely, the function of T3SS-1 remains somewhat elusive although it is quite clear that T3SS-1 encodes potent effector proteins that target conserved cellular processes and can lead to rapid cell death *in vitro*, in many cell lines [251, 268]. A coordinated infectious effort, first by T3SS-2 to promote colonization and enterotoxicity, followed by further invasion and expression of T3SS-1, has been proposed, but remains understudied, in part due to the lack of mammalian models which adequately replicate human infection [193].

1.3 T3SS Gene Regulation in *V. parahaemolyticus*

T3SS-1 and T3SS-2 expression paradigms are divergent, with different regulatory circuits acting to drive expression of the machinery. However, both of these systems are repressed by a common gene repressor, histone-like nucleoid structuring protein or H-NS.

T3SS-2 is primarily activated by two ToxR-like transcriptional regulators VtrA (VPA1332) and VtrB (VPA1348) [285]. In line with the known role for T3SS-2 in enterotoxicity, VtrA, along with another protein VtrC (VPA1333), drive this activation of VpPAI genes through the sensing of bile acid in the environment and in turn activate the VpPAI master-transcriptional regulator VtrB [192, 286]. Bile acids are important mediators of digestion and are readily found in the small intestine where *V. parahaemolyticus* deploys T3SS-2 to generate enterotoxicity. VtrA and VtrC form an inner membrane-bound heterodimer which binds taurodeoxycholate and glycodeoxycholate bile acid components, but not chenodeoxycholate or cholate, in a primarily hydrophobic cleft of the VtrC β-barrel [286]. A model has been proposed where the DNA-binding domain of VtrA is activated by the binding of bile salt components to VtrAC, which in turn allows VtrA to activate the expression of *vtrB*, which then drives the expression of VpPAI genes including T3SS-2 [285, 286]. Interestingly, the VtrABC system has similarities with other regulatory circuits in *Vibrio* spp., where the dimerization of ToxR/ToxS and TcpP/TcpH is enhanced in the presence of bile acid by disulfide-bond formation in *Vibrio cholerae* [287, 288]. Finally, a screen of transposon mutants in mammalian intestinal colonization identified that ToxR augments the activity of VtrA, allowing for the expression of *vtrB*, through an as yet unknown mechanism [283].

T3SS-1 follows a divergent regulatory paradigm. Decreased calcium concentration and increased magnesium concentrations contribute to the

activation of this secretion system in a variety of studies (Figure 1.3A) [268, 269, 289–291]. In our lab, we simulate these condition through the addition of Mg^{2+} ions, and remove calcium from the growth media by chelation with EGTA [268]. Host-cell contact has also been shown to induce T3SS activity in *P. aeruoginosa* and *V. parahaemolytiucs*, although the mechanism through which this occurs is unknown (Figure 1.3A) [171, 292–294]. Through a signal transduction cascade which is also not resolved, these conditions drive the expression of an AraC-family master transcription regulator for T3SS-1 called ExsA [289, 290, 295, 296]. Importantly, the ExsA protein drives the expression of over 20 genes involved in the biosynthesis of T3SS-1 [297]. ExsA in *V. parahaemolyticus* is homologous to ExsA found in *P. aeruginosa* which controls an analogous T3SS [298–301].

Following genetic activation of *exsA*, the gene is auto-regulatory – a common feature of AraC-type regulators – and when expressed, two ExsA proteins bind to a tri-partite DNA motif which contains the following elements: TTTAGN4TT, a downstream A-box, and a downstream GC-box (Figure 1.3A) [290]. ExsA from *V. parahaemolyticus* can complement *P. aeruginosa exsA* null mutants, indicating that the binding motifs must have similar elements [302]. Indeed, previous studies have identified the TTTAGN4TT, as well as the importance of the A-box and GC-box in both *V. parahaemolyticus* and *P. aeruginosa*, respectively [296, 303]. This tri-partite binding motif is also found at regulatory sites for the ExsA regulon and serves as a conserved binding site for this protein [290]. While *exsA* is auto-regulatory, it is also repressed via quorum-sensing through the binding of the quorum-sensing master regulator OpaR (Figure 1.3A) [304]. At high cell density, OpaR is highly expressed, and in turn, *exsA* is repressed by OpaR binding to sites upstream of the *exsA* gene.

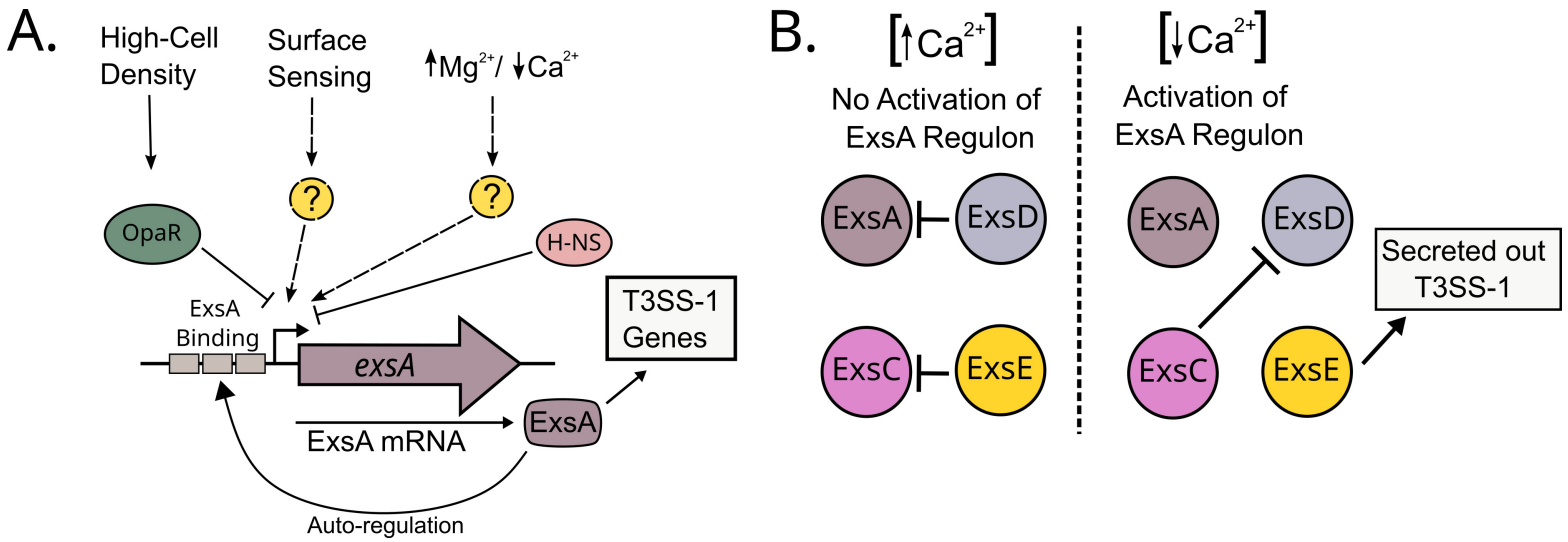

Figure 1.3: T3SS-1 Master-Regulator *exsA* Regulation in *V. parahaemolyticus*. A. The *exsA* gene is presented, including its promoter and the autoregulatory site for ExsA binding. H-NS and OpaR proteins are known to repress T3SS-1 expression through repression of *exsA*. Increased Mg^{2+} and decreased Ca^{2+}, as well as surface sensing, are known to activate *exsA* expression, through unknown mechanisms. Once ExsA is expressed, it can activate its own regulon, containing over 20 genes involved in the biogenesis of the T3SS-1. B. ExsA itself is regulated by a protein-binding cascade. In high Ca^{2+} conditions, ExsE is present in the cell and binds ExsC, and ExsD is free to bind ExsA, acting as an anti-activator and preventing ExsA activation of its regulon. Alternatively, in low Ca^{2+} conditions, ExsE is secreted out of the cell, and ExsC can sequester the anti-activator ExsD. Here, ExsA can go on to activate T3SS-1 genes.

Following ExsA expression, additional protein factors ExsDCE interact with ExsA either directly or indirectly to control its activation of T3SS-1 (Figure 1.3B) [296, 305]. Here, ExsD acts as an anti-activator and binds ExsA to prevent it from binding DNA and activating gene operons involved in T3SS-1 biosynthesis. ExsC is an anti-anti-activator protein which binds to ExsD and prevent's ExsD interaction with ExsA, and in turn allows ExsA to interact with DNA and activate T3SS-1 genes. Further, ExsE interacts with ExsC, preventing the anti-anti-activator interaction with ExsD, allowing for anti-activation of ExsA. When T3SS-1 machinery is generated upon inducing conditions, ExsE is secreted out of the cell through the T3SS, allowing ExsC to sequester ExsD, thereby freeing ExsA to activate its regulon (Figure 1.3B) [301, 305, 306]. This complex protein-protein regulatory system is regulated by similar conditions as the T3SS itself, as ExsE is secreted from the cell by low extracellular calcium [301, 306].

Many studies have identified H-NS independently as involved in the repression of virulence genes in *V. parahaemolyticus*, including T3SS-1, T3SS-2 and *tdh*, T6SS-1, T6SS-2 [269, 307, 308]. H-NS is referred to as a genome sentinel, as it has a known function of binding to and preventing the expression of horizontally acquired DNA, of which virulence genes often are [307, 309–312]. Evolutionarily, this is an important function for bacteria to have, as horizontally acquired DNA may contain toxins or incompatible proteins which could disrupt fitness or kill the cell.

H-NS binds to AT-rich high-affinity binding sites, and nucleates from those regions creating large regions of H-NS-DNA complexes [313]. These nucleated regions of DNA can bridge with one another, condensing the nucleoid and supercoiling the DNA [314]. H-NS is thought to prevent transcription in these regions by either trapping RNA polymerase in bridged or folded DNA or preventing RNA polymerase from binding to its recognition sequences by steric hinderance [315]. In H-NS-bridged DNA regions, both transcription elongation and initiation

would be blocked, while in nucleated regions that are unbridged, RNA polymerase may be able to continue transcription elongation that began elsewhere, but could not initiate transcription at sites of promoter coverage by H-NS [315, 316]. Further, H-NS pausing or slowing of RNA polymerase at bridged regions may aid in Rho-dependent termination [316]. Importantly, Mg^{2+} concentrations appear to disrupt H-NS bridging, and promote H-NS filamentation on the DNA helix [317]. Interestingly, increased Mg^{2+} concentration appears to enhance T3SS-1 expression, consistent with the observation that Mg^{2+} concentration can modify H-NS filamentation on DNA [268].

A common paradigm for removing H-NS repression is the expression of de-repressor proteins. Bacteria appear to have evolved de-repression mechanisms to remove H-NS repression of genes that confer fitness advantages [318] Global studies of the LuxR quorum-sensing master regulator in *V. harveyi*, for example, identified that LuxR is only necessary in the presence of H-NS, and that it is sufficient to remove H-NS from promoter DNA *in vitro* [319]. Similarly, H-NS silences many genes in the ToxR regulon of *V. cholerae*, and ToxR antagonizes this interaction, removing H-NS from DNA [320, 321]. Importantly, ToxR is also unnecessary for activation of virulence genes in *V. cholerae* in the absence of H-NS, highlighting its role as a de-repressor like LuxR [321]. HlyU also de-represses a number of hemolysin genes in a variety of *Vibrio* spp., presumably by disrupting H-NS binding and repression [322]. How exactly these de-repressing regulators remove H-NS from DNA remains unclear. While H-NS binding sites sometimes overlaps directly with the de-repressor binding site, this is not consistently the case, and the distance between some H-NS binding sites and de-repressor DNA binding sites suggests that it may be more complicated than simply competitive binding or steric-hinderance interaction between H-NS and the de-repressor protein [323, 324]. For example, while HlyU regulation of the *vvhBA* toxin genes in *V. vulnificus* overlaps the H-NS binding site, suggesting a competitive binding

interaction between repressor and de-repressor, the HlyU binding site at the *vah1* toxin gene does not overlap [322, 325–328] Further, changes in DNA topology and superstructure have been linked to removal of gene repression by H-NS in the case of *virF* and *proU* expression in *S. flexneri* [329–331]. In the absence of competitive binding or steric hinderance, gene de-repressor regulators could modify DNA topology to de-repress H-NS, although this remains underexplored [318].

1.4 Rationale and Hypotheses

V. parahaemolyticus infects and causes disease in agriculturally important organisms and is also the cause of a pandemic diarrheal disease. *V. parahaemolyticus*, and in fact most *Vibrio* spp. serve important ecological roles in the environment. Understanding how these organisms interact with the environment, translate those environmental cues into changes in cell structure and function, and cause disease in their hosts is important to mitigate damage and predict future outcomes. With the dramatic changes expected to occur to the world's oceans due to anthropogenic climate change, enabling the prediction of future outcomes is ever prescient. In this thesis, I explore the biology of *V. parahaemolyticus* from a few perspectives that are central to many of its environmental interactions, and its interactions with human hosts. I begin by exploring T3SS-1 activation by previously established conditions, in an effort to connect environmental sensing to *exsA* activation. Following, I explore the sugar catabolism regulator NagC, and its related regulator Mlc, for connections between environmental sugar sensing and other *Vibrio* phenotypes like Type III and Type VI secretion, biofilm formation, and swimming motility. Finally, I deploy a whole-genome approach using transposon mutagenesis to identify the genomic content necessary for growth on chitin in the laboratory environment.

Given the lack of knowledge surrounding the function of T3SS-1 and how it aids *V. parahaemolyticus* survival and fitness, in Chapter 2 we explore how T3SS-1 is activated. Our lab previously found that T3SS-1 biogenesis can be activated in media containing additional Mg^{2+} and decreased Ca^{2+} when the cells are grown at 30°C [268]. In these conditions, ExsA is expressed and goes on to activate a number of genes involved in the biogenesis of the T3SS-1. While ExsA itself is controlled by a known protein cascade involving ExsCDE, as well as through auto-regulation and repression from H-NS and OpaR, the mechanism through which the Mg^{2+}/EGTA environmental signal confers *exsA* activation is unclear. Therefore, we hypothesized that a protein activator or de-repressor of H-NS can activate/de-repress *exsA* following exposure to inducing conditions. I address this by employing an *exsA-luxCDABE* transcriptional fusion on chromosome 2 of *V. parahaemolyticus* RIMD2210633, and subjecting this strain to transposon mutagenesis to identify genomic regions necessary for activation of *exsA* in the Mg^{2+} condition. In this study, we identify a known transcriptional regulator HlyU, involved in the expression of hemolysins in a variety of other *Vibrio* spp., as essential for the activation of T3SS-1 biogenesis through *exsA* expression. We also characterized HlyU binding to DNA at a region containing a DNA palindrome within the *exsA* promoter.

Following the identification of HlyU in Chapter 2, and its importance for expression of the T3SS-1 through *exsA* de-repression, we sought to better understand how HlyU interacts with the *exsA* promoter, and to characterize the activation paradigm of *exsA* (Chapter 3). HlyU should interact with the major groove of the DNA helix, presumably inducing a DNA bend, according to molecular models [332, 333]. However, a DNA-protein co-crystal has not been characterized to date. Further, HlyU is canonically thought to interact with DNA in such a way that it removes the repression generated by H-NS. However, it is unclear how this occurs, given that HlyU binding sites and H-NS binding sites, particularly at the

exsA promoter, do not overlap directly [269, 307]. In Chapter 3, we hypothesized that HlyU binding to DNA, at the palindromic region identified in Chapter 2 as well as based on previous literature, disrupts DNA superstructure which in turn leads to *exsA* de-repression.

While Chapters 2 and 3 deal with virulence factor activation directly, it remains clear that these virulence factors are influenced by the ecology of *Vibrio* spp, including their important ecological mediator, chitin catabolism. In Chapter 4, to explore how chitin catabolism might influence other cellular functions, including virulence, I sought to better understand the roles of two sugar catabolism regulators, NagC and Mlc, in *V. parahaemolyticus*. In particular, I assayed type VI secretion, biofilm formation, swimming motility, chitin degradation, and type III secretion as directed by previous literature on NagC and Mlc.

Finally in Chapter 5, I used our expertise in transposon mutagenesis and tied it to high-throughput DNA sequencing to characterize the genes genetic necessary for *V. parahaemolyticus* growth on colloidal chitin in minimal media. This study was also not driven by a particular hypothesis, but rather is a discovery-based study to identify novel factors involved in chitin catabolism in *V. parahaemolyticus*. Here, I developed a transposon mutagenesis tool for adequate selection of high-density transposon mutant libraries and generated a streptomycin resistance strain of *Vibrio parahaemolyticus*. Following selection of mutants and growth in both glucose and colloidal chitin containing minimal media, TnSeq was performed. This study overlaps with previous work on chitin catabolism and identified many of the previously identified essential genes. However, this study also identified novel porin and transcriptional regulator genes, as well as the T2SS as necessary for growth on colloidal chitin.

Chapter 2

The Transcriptional Regulator HlyU Positively Regulates Expression of *exsA*, Leading to Type III Secretion System 1 Activation in *Vibrio parahaemolyticus*

This work appears in:

Getz, L.J., Thomas, N.A. (2017) *The Transcriptional Regulator HlyU Positively Regulates Expression of exsA, Leading to Type III Secretion System 1 Activation in Vibrio parahaemolyticus.* **Journal of Bacteriology**, Vol. 200, No. 15. DOI: 10.1128/JB.00653-17.

2.1 Abstract

Vibrio parahaemolyticus is a marine bacterium that is globally recognized as the leading cause of seafood-borne gastroenteritis. *V. parahaemolyticus* uses various toxins and two type 3 secretion systems (T3SS-1 and T3SS-2) to subvert host cells during infection. We previously determined that *V. parahaemolyticus* T3SS-1 activity is upregulated by increasing the expression level of the master regulator ExsA under specific growth conditions. In this study, we set out to identify *V. parahaemolyticus* genes responsible for linking environmental and growth signals to *exsA* gene expression. Using transposon mutagenesis in combination with a sensitive and quantitative luminescence screen, we identify HlyU and H-NS as two antagonistic regulatory proteins controlling the expression of *exsA* and, hence, T3SS-1 in *V. parahaemolyticus*. Disruption of *hns* leads to constitutive unregulated *exsA* gene expression, consistent with its known role in repressing *exsA* transcription. In contrast, genetic disruption of *hlyU* completely abrogated *exsA* expression and T3SS-1 activity. A *V. parahaemolyticus hlyU* null mutant was significantly deficient for T3SS-1-mediated host cell death during *in vitro* infection. DNA footprinting studies with purified HlyU revealed a 56-bp protected DNA region within the *exsA* promoter that contains an inverted repeat sequence. Genetic evidence suggests that HlyU acts as a derepressor, likely by displacing H-NS from the *exsA* promoter, leading to *exsA* gene expression and appropriately regulated T3SS-1 activity. Overall, the data implicate HlyU as a critical positive regulator of *V. parahaemolyticus* T3SS-1-mediated pathogenesis.

2.2 Introduction

Vibrio parahaemolyticus is a ubiquitous marine bacterium and is a leading cause of seafood-borne gastroenteritis globally [334]. In brackish estuaries, these Gram-negative bacteria exist free-living in the water column, within

marine sediments, and in commensal relationships with many bivalve shellfish, including oysters [335]. After human consumption of contaminated shellfish, *V. parahaemolyticus* can cause vibriosis, leading to nausea, fever, diarrhea, and occasionally vomiting. Although generally self-limiting, *V. parahaemolyticus* causes significant morbidity during foodborne outbreaks [336]. *V. parahaemolyticus* environmental isolates frequently exhibit genetic profiles; however, a clinically relevant pandemic *V. parahaemolyticus* serotype (O3:K6) has emerged, being isolated in patients around the world [334, 337]. Additionally, specific *V. parahaemolyticus* strains that possess a toxin-antitoxin plasmid have emerged as significant shrimp pathogens, where major economic losses have occurred in aquaculture industries [113]. Thus, it appears these marine bacteria continue to evolve and are a serious threat to the seafood industry and human health.

V. parahaemolyticus virulence factors include hemolysins, toxins, and two type III secretion systems (T3SS), T3SS-1 and T3SS-2. The T3SS of pathogens acts as a molecular needle-like structure to deliver bacterial effector proteins into host cells [238]. Sequence analysis of various *V. parahaemolyticus* clinical isolates has revealed that many contain both T3SSs [139, 338]. T3SS-2 has primarily been linked to enterotoxicity [263, 339], whereas T3SS-1 has been linked to cellular disruption and rapid cytotoxicity [270, 276, 279]. Many effectors are delivered via T3SS-1 into host cells and have targeted actions. VopQ contributes to rapid cell death during *V. parahaemolyticus* infection by interacting with lysosomal H+ V-ATPases, causing lysosomal rupture and release of contents [273]. Additionally, T3SS-1 effector proteins VopS, VopR, and VPA0450 have been implicated in immune evasion and actin rearrangement [279, 280]. Other virulence mechanisms in *V. parahaemolyticus* include colonization factors (such as pili) and two type 6 secretion systems (T6SS) that likely aid in killing other bacteria or acting on macrophages during infection [84, 174, 255].

Over 40 genes within multiple genetic operons contribute to the formation of *V. parahaemolyticus* T3SS-1. The majority of these genes require the master regulator ExsA for their expression [296]. ExsA and its orthologues in other bacteria are members of the AraC/XylR family of transcriptional regulators, many of which are implicated in bacterial pathogenesis mechanisms [340]. In the cases of *Yersinia* species (LcrF), *Pseudomonas* species (ExsA), and enteropathogenic *Escherichia coli* (EPEC; GrlA), these regulators have been demonstrated to activate T3SS-related genes and have been linked to virulence phenotypes [302, 341–343].

ExsA transcriptional activator roles are well documented, although it is not known how *V. parahaemolyticus* interprets environmental signals to activate *exsA* expression and thus initiate T3SS-1 biogenesis. In this study, we used a genome-wide transposon mutagenesis approach to identify genetic regulators of *exsA* expression. Through the design of a sensitive and quantitative luminescence screen, we identified a cis-acting genetic switch and implicate HlyU as a critical *V. parahaemolyticus* virulence determinant.

2.3 Results

2.3.1 A Sensitive Functional Screen Identifies Genes Linked to *exsA* Promoter Activation

We initially generated a recombinant *V. parahaemolyticus* strain containing a promoterless bioluminescent reporter cassette (*luxCDABE*) transcriptionally fused to the *exsA* promoter region from *V. parahaemolyticus*. The $_{p}$*exsA-luxCDABE* fusion was integrated into the *tdhS* locus of the *V. parahaemolyticus* RIMD2210633 genome (see Section 2.5 Materials and Methods) (Figure 2.1A). We chose the *tdhS* locus, as TdhS is not required for T3SS-1 activity [270]. This strain (referred to as the Vp-lux strain) quantitatively

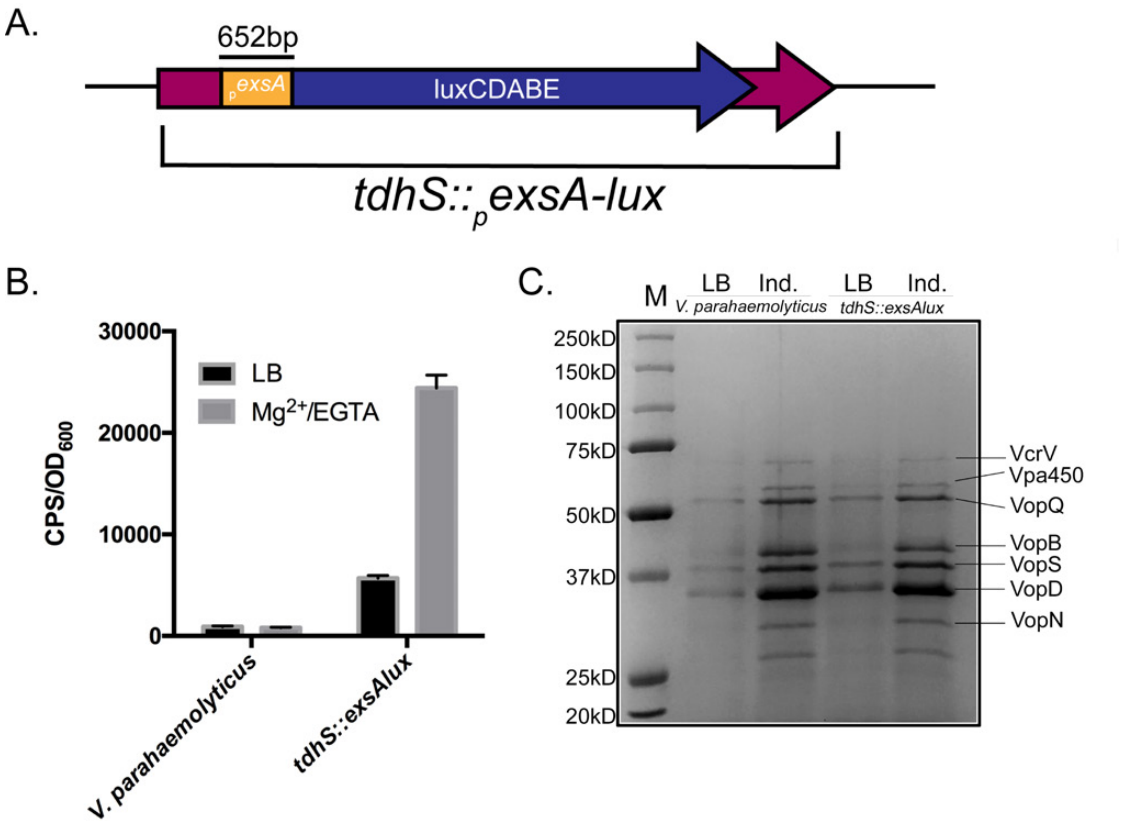

Figure 2.1: Integration of an *exsA* promoter-*luxCDABE* fusion into the *tdhS* locus (*tdhS*::*exsA*-lux) results in bioluminescence and a normal T3SS-1 secreted protein profile. (A) Schematic representation of the constructed *exsA* promoter-*luxCDABE* fusion inserted within the *tdhS* chromosomal locus. (B) Comparison of bioluminescence emitted by the indicated *V. parahaemolyticus* strains. Bacteria were cultured under T3SS-1-inducing conditions, sampled after 3 h of growth, and then assessed for light emission. cps/OD, counts per second divided by optical density (600 nm) of the cell suspension at time of collection. (C) Total secreted protein profiles of *V. parahaemolyticus* strains as determined by SDS-PAGE. Bacteria were grown in LB medium or LB supplemented with MgSO4 and EGTA (induced [Ind.] for T3SS-1 activity). Labeled protein species have been previously identified by mass spectrometry analyses in wild-type *V. parahaemolyticus* [268]. The smallest protein species visible was not previously identified, and is therefore unlabelled. M, protein standard.

reports the activity of the *exsA* promoter in real time by emitting bioluminescence, thereby allowing for sensitive in situ measurement of *exsA* promoter activation (Figure 2.1B). The strain responded to magnesium and EGTA (T3SS-1-inducing conditions) by increasing *exsA* promoter activity as expected (Figure 2.1B). A T3SS-1 protein secretion assay indicated that the Vp-lux strain secreted effectors similarly to wild-type *V. parahaemolyticus* (Figure 2.1C).

Consequently, we employed a conjugative plasmid encoding a hyperactive transposase and its associated transposon [344] to mutagenize the Vp-lux strain. An optimized triparental mating approach resulted in the isolation of 4,302 stable transposants at an average transposition frequency of 1×10^{-6}. Using a high-throughput 96-well plate formatted assay, the transposants were screened for *exsA* promoter activity via bioluminescence. The Vp-lux strain served as a reference for defined normal growth and light emission. The screen identified a spectrum of mutants that exhibited altered bioluminescence levels compared to those of the parent Vp-lux strain (Figure 2.2). Prior to investigating specific mutants, they were divided into 4 groups (or quadrants) to facilitate downstream analyses. This was achieved by establishing categories based on an optical density at 600 nm (OD_{600}) of 0.4 and a counts per second (cps) value of 1,000 and produced four quadrants (separate groups), including low-bioluminescence and normal growth, normal bioluminescence and normal growth, low bioluminescence and low growth, and high bioluminescence and any growth (Figure 2.2).

Mutants that fell into the low-luminescence and normal-growth quadrant were further subjected to statistical narrowing. Using a binning procedure, mutants were subdivided into three categories by luminescence emission (measured in counts per second; low bioluminescence, 0 to 100 cps; low-moderate bioluminescence, 101 to 300 cps; and moderate bioluminescence, 301 to 1,000 cps). The means and standard deviations for each group were calculated. The means and standard deviations were calculated as 63.30 ± 10.68, 155.1 ± 22.32,

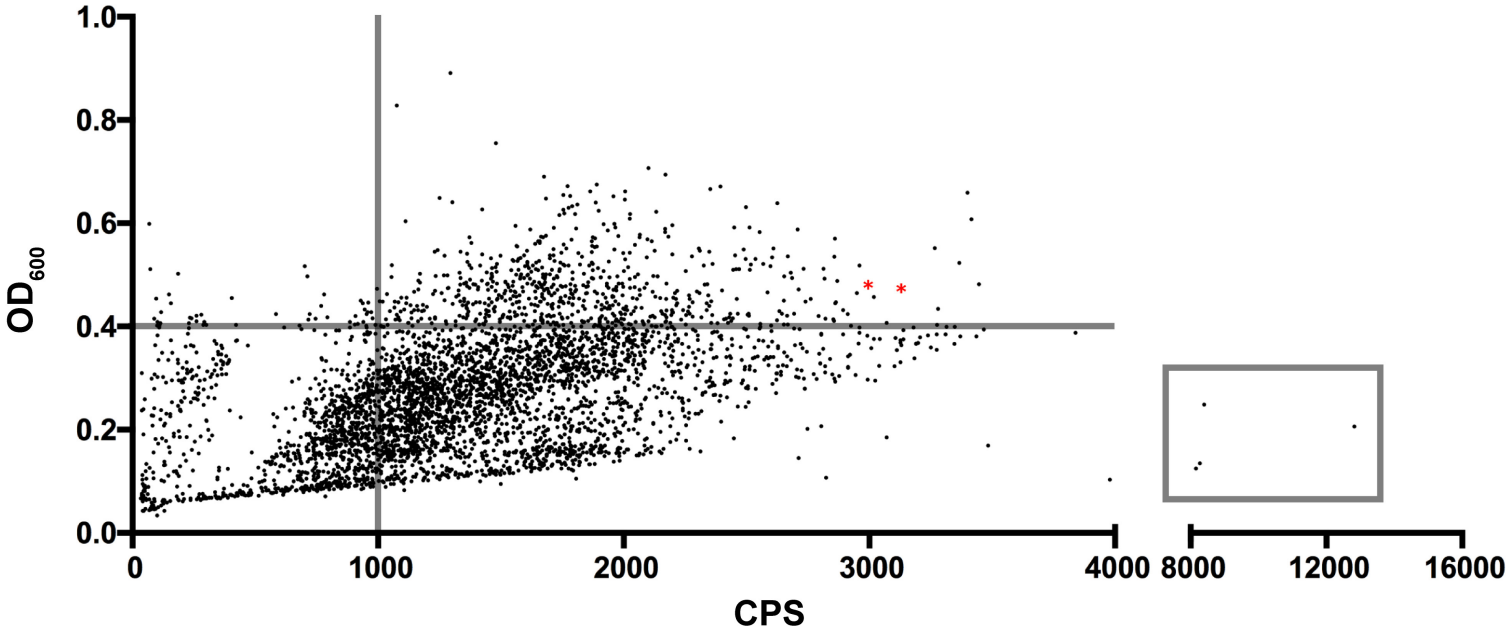

Figure 2.2: Bioluminescence scatterplot of Vp-lux Tn5 insertional mutants. Transposon insertion mutants were assayed for bioluminescence emission. Each dot represents an individual mutant, and two representative controls (Vp-lux strain) are indicated as red asterisks. The vertical and horizontal bars represent boundaries used to categorize mutants based on light emission (*exsA* promoter activity) and cell growth, respectively. The box encompasses high-light-emitting mutants selected for further study. CPS, counts per second.

and 542.7 ± 126.6, respectively. Six mutants whose cps reading was less than 1 standard deviation below the mean were then selected for further characterization in the low-bioluminescence group. Within the low-moderate and moderate bioluminescence groups, 3 and 4 mutants were selected for characterization, respectively. Lastly, 5 of the most bioluminescent mutants (Figure 2.2, gray box) were also selected for characterization. Details of the selected mutants and statistical data are listed in the supplemental material (see Table A.1 in the supplemental material).

2.3.2 Identification of Genes With Specific Insertional Transposons that Alter *exsA* Promoter Activity

Using a genomic DNA marker retrieval protocol, we successfully identified chromosomal locations of transposition within mutants from the low-moderate and moderate light-emitting groups. From the low-moderate luminescence group linked to deficient *exsA* promoter activity, we mapped transposon insertions exclusively within *vp0529* (Figure 2.3A). Notably, for *vp0529*, we identified three isolates with the exact same transposon insertion site. These likely represent a single clone that generationally expanded during the mutagenesis procedure or, alternatively, this chromosomal site is favored for transposon insertion. For the moderate bioluminescence group, we mapped two transposon insertion mutants, *vpa0179* (*phoX*) and *vp1473* (*merR* operon) (Figure A.1). Within the high-luminescence group linked to increased *exsA* promoter activity, 3 independent transposon insertions were mapped near or within *vp1133* (*hns*) (Figure 2.3B) and one insertion was mapped within *vp1633* (*rtx* toxin) (Figure A.1).

Our initial efforts focused on *vp0529* and *vp1133*, since multiple independent mutant insertions were isolated from the lux reporter screen for these genes (Fig. 3) and therefore strongly implied a role in *exsA* regulation. *in silico* analysis of *vp0529* using NCBI BLAST indicated sequence similarity to HlyU, a DNA

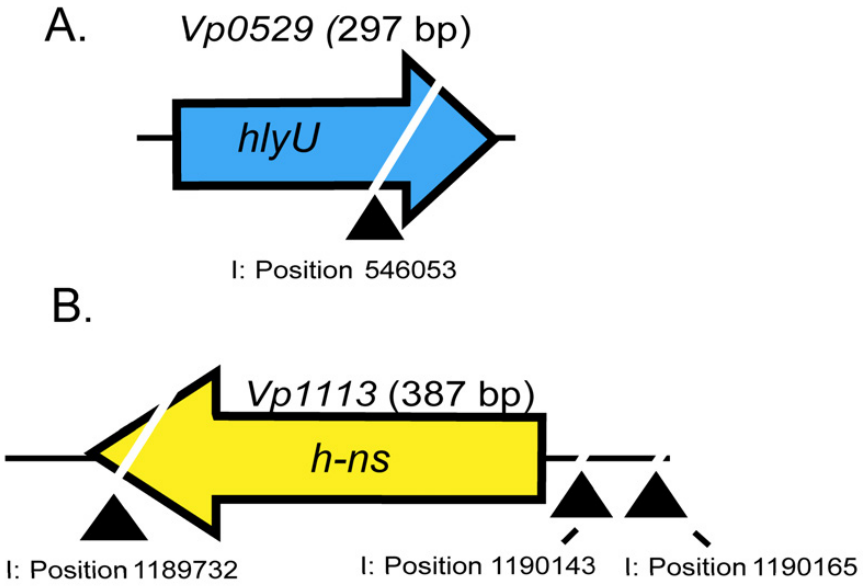

Figure 2.3: Schematic diagram of transposon insertion sites within selected genes of interest identified by luminescence screening of the Vp-lux Tn5 mutant library. The triangles represent approximate transposon insertion sites with the approximate DNA base location in the *V. parahaemolyticus* genome (NBCI accession numbers NC_004603 and NC_004605; I, chromosome I).

binding protein of many *Vibrio* species that acts as a positive regulator of virulence genes [345]. *vp1133* encodes a histone-like nucleoid-structuring protein (H-NS) and has previously been implicated in *exsA* repression [305]. To investigate this further, an assay that evaluates T3SS-1 protein secretion was performed. Critically, a marked decrease in many T3SS-1-dependent proteins was observed for the Vp0529 Tn5 insertion mutant compared to the Vp-lux parent strain (Figure 2.4A). In contrast, the *vp1133* mutant displayed secretion of many T3SS-1-dependent proteins, comparable to the parent strain.

HlyU has previously been implicated as an H-NS derepressor, leading to rtxA1 toxin expression in *V. vulnificus* [346]. To investigate if HlyU acts as a derepressor of H-NS in *V. parahaemolyticus* or, alternatively, as a *bona fide* transcriptional activator, we deleted *hlyU* in the context of the *hns*::Tn5 mutant (strain Vp-EL) and performed protein secretion and luciferase reporter assays. Vp-EL behaved similarly to the Δ*hns*::Tn5 mutant, exhibiting T3SS-1 effectors and translocators along with enhanced luciferase activity (driven by the *exsA* promoter) (Figure 2.4A and B). Therefore, in the absence of H-NS, HlyU is not required for T3SS-1 protein secretion or for *exsA* promoter activity. To confirm the role of HlyU for *exsA* promoter activity in cells expressing H-NS, we subjected the *hlyU*::Tn5 mutant to genetic complementation. Enhanced restoration of luminescence was observed for the *hlyU* complemented strain that was significantly above the level of the Vp-lux parent strain (Figure 2.4C).

2.3.3 *hlyU* Null Mutants are Deficient for T3SS-1-dependent Secretion and Cytotoxic Effects

To validate the phenotypes measured in the *hlyU*::Tn5 mutant and to rule out any effects due to transposon insertion, an *hlyU* null mutant derived from wild-type *V. parahaemolyticus* was generated (strain *hlyU*1) and tested for protein secretion. In complete agreement with the *hlyU*::Tn5 mutant, T3SS-1 protein secretion

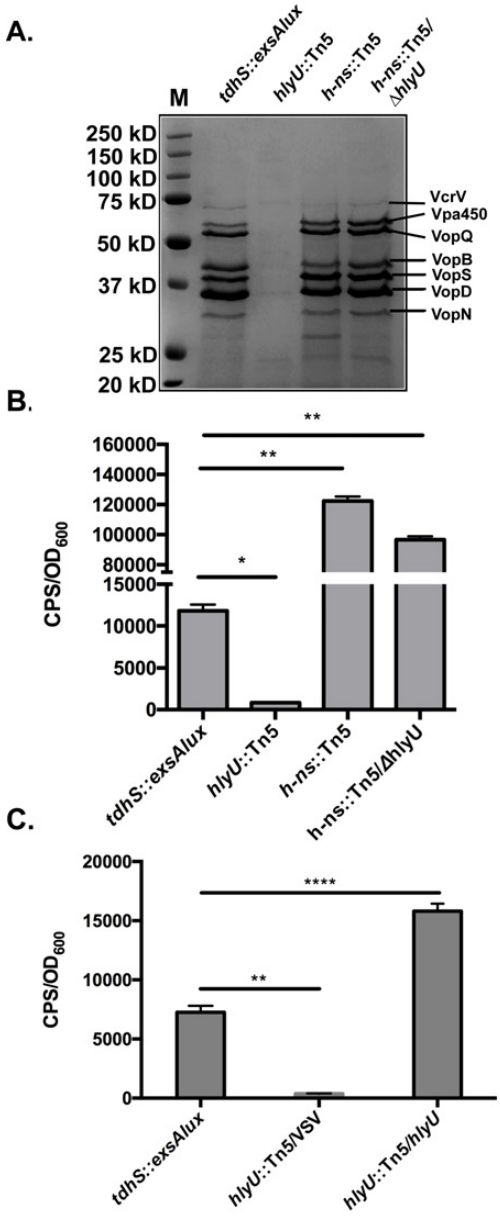

Figure 2.4: Protein secretion and *exsA* promoter activity assays with the Vp-lux (*tdhS*::*exsA*-lux) strain and specific transposon insertion mutants. (A) Total secreted proteins were precipitated from culture supernatants of the indicated Tn5 insertion mutants and subjected to SDS-PAGE, followed by Coomassie blue staining. A protein ladder was included (M), and previously identified proteins are labeled. (B) Luciferase assays were conducted with the indicated Tn5 insertion mutants and the *hns*::Tn5 Δ*hlyU* double mutant (strain Vp-EL) (n = 2). The error bars indicate standard deviations from the mean values. A two-tailed *t*-test was conducted to compare data sets. *, $P < 0.05$; **, $P < 0.005$. (C) Complementation of the *hlyU*::Tn5 mutant with pVSV105 encoding *hlyU* was performed, followed by a luciferase assay. A two-tailed t test was used to compare data sets. **, $P < 0.005$; ****, $P < 0.0001$

for *hlyU*1 was deficient compared to that of wild-type *V. parahaemolyticus* and appeared similar to that of a Δ*vscN1* T3SS-1-deficient strain (Figure 2.5A). Genetic complementation of strain *hlyU*1 with the wild-type *hlyU* coding region in trans fully restored T3SS-1 protein secretion.

To investigate the role of HlyU in the context of infection-like conditions, we conducted *in vitro* HeLa infections with *V. parahaemolyticus* strains followed by a lactate dehydrogenase (LDH) release cytotoxicity assay. As expected, wild-type *V. parahaemolyticus* resulted in rapid HeLa cytotoxicity in a *vscN1* (T3SS-1 ATPase)-dependent manner (Figure 2.5B), confirming previous reports [251, 268]. *hlyU*1 was significantly less cytotoxic than the parent strain (Figure 2.5B), comparing closely to the low cytotoxicity levels observed for the Δ*vscN1* strain. Genetic complementation of *hlyU*1 restored HeLa cytotoxicity to wild-type *V. parahaemolyticus* levels, indicating that HlyU contributed to *V. parahaemolyticus* cytotoxic activity during *in vitro* infection.

Taken together, the protein secretion, promoter activation, and cytotoxicity data all indicated that HlyU was required for *exsA* promoter activation that supported expression of T3SS-1-related genes, leading to HeLa cell cytotoxicity during infection.

2.3.4 Purified HlyU Binds Upstream of *exsA*

HlyU is a member of the SmtB/ArsR family of regulator proteins that bind to DNA using a winged helix-turn-helix (wHTH) domain structure [345]. Based on the *hlyU*1 deficiency for T3SS-1-associated phenotypes, we set out to investigate whether HlyU binds to *V. parahaemolyticus* DNA. We hypothesized that DNA near the *exsA* promoter (located within our lux reporter constructs) should contain sequences that direct HlyU binding to DNA. To test this hypothesis, we overexpressed and purified a HlyU-HIS fusion protein (see Materials and Methods). Native HlyU is 98 amino acids in length (NCBI accession no.

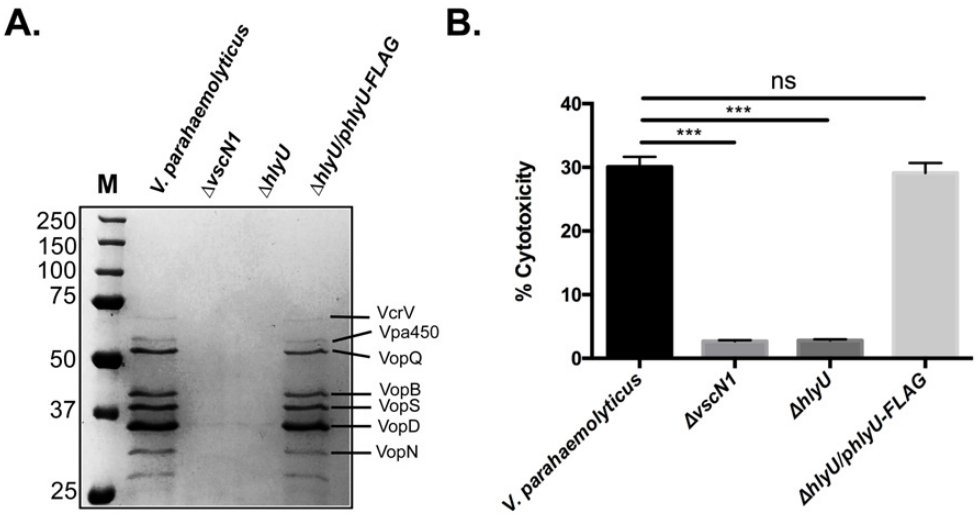

Figure 2.5: *V. parahaemolyticus* Δ*hlyU* mutant (strain *hlyU*1) is deficient for secretion of T3SS-1 proteins and exhibits reduced host cell cytotoxicity during infection. (A) Total secreted proteins were collected and subjected to SDS-PAGE, followed by Coomassie staining. A protein standard was included (M), and previously identified protein species are labeled. (B) Percent cytotoxicity was calculated after infection of HeLa cells with various strains by measuring released lactate dehydrogenase levels. All strains contained pVSV105 or p*hlyU*-FLAG as indicated. Statistical bars indicate standard deviations from the means, and statistical significance was determined by a paired two-tailed t test compared to the *V. parahaemolyticus* control. ***, $P < 0.001$. ns, not significant.

NP_796908.1) and has a predicted molecular mass of 11.2 kDa. Overexpressed HlyU-HIS appeared as an approximately 11-kDa protein by SDS-PAGE (Figure A.2). Electrophoresis mobility shift assays (EMSA) next were performed using an *exsA* promoter DNA fragment mixed with increasing amounts of purified HlyU protein. The *exsA* promoter DNA fragment demonstrated multiple shifts, with concentrated DNA species appearing with larger amounts of HlyU-HIS (Figure 2.6A). For these larger HlyU-HIS amounts, Sypro red and SYBR green staining revealed protein-DNA complexes which appeared as tight bands coinciding with reduced mobility through the acrylamide gel matrix (Figure A.3). In contrast, no mobility shift was apparent with increasing amounts of bovine serum albumin (BSA), indicating that the *exsA* promoter region did not support nonspecific protein binding (Figure 2.6A). Based on densitometric analyses of HlyU-bound *exsA* promoter DNA (Table A.2), an apparent binding affinity ranging from 5 to 6.3 μM (95% confidence interval) was calculated (Figure A.4). A nonspecific DNA fragment next was mixed with purified HlyU-HIS. While some minor shifts were observed at low HlyU levels (as seen by smearing), no major shifts or concentrated DNA species were apparent with increasing amounts of HlyU-HIS (Figure 2.6B). These results suggest that small amounts of HlyU-HIS interacted with DNA weakly, and larger HlyU amounts specifically bound the *exsA* promoter region DNA, leading to homogenous protein-DNA complexes.

2.3.5 HlyU Protects An Inverted Repeat Sequence within the *exsA* Promoter Region from DNase I Digestion

We set out to identify the DNA region bound by HlyU within the *exsA* promoter region. Fluorescently end-labeled *exsA* promoter DNA was mixed with purified HlyU-HIS and then treated with DNase I. The reaction mixtures were purified and then subjected to a DNA footprinting assay. As shown in Figure 2.6C, a 56-bp region repeatedly and specifically produced low fluorescence signals indicative of

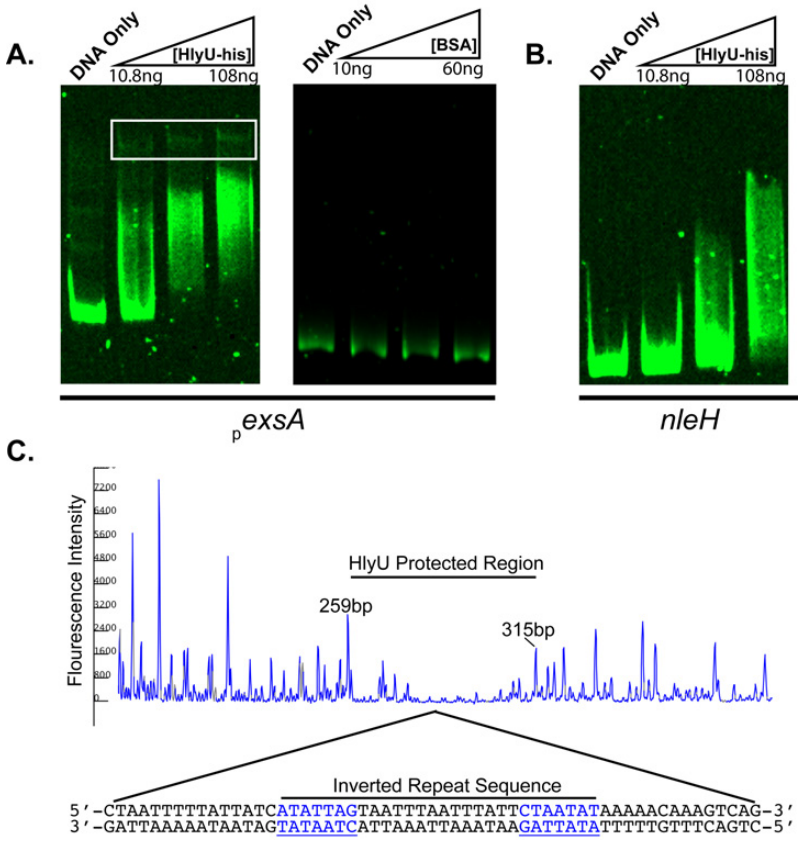

Figure 2.6: Electrophoresis mobility shift assay (EMSA) and DNA footprinting assays indicate HlyU binding to a region within the *exsA* promoter. (A) Six percent TBE-polyacrylamide gels were loaded with reaction mixtures containing *exsA* promoter DNA and increasing amounts of purified HlyU-HIS or bovine serum albumin (BSA). An *exsA* promoter DNA-only control was included as indicated. The white box indicates slow-migrating, concentrated DNA species. (B) A 6% TBE-polyacrylamide gel was loaded with reaction mixtures containing *nleH1* DNA and increasing amounts of purified HlyU-HIS. Following electrophoresis, gels were stained with SYBR green to specifically stain DNA. (C) A DNase I footprinting assay using purified HlyU-HIS was used to identify a DNA binding region for HlyU-HIS within the *exsA* promoter. The approximate base pair numbers of the HlyU-HIS footprint region are indicated (259 to 315 bp) and are based on capillary electrophoresis internal size standards. The DNA sequence associated with the identified HlyU protected region is displayed below the chromatogram. A 7-bp inverted repeat (labeled in blue) and separated by 14 bp is centrally located within the HlyU-HIS protected region.

an HlyU-protected region (undigested by DNase I). Sequence examination of this DNA region resulted in the identification of a perfect 7-base inverted repeat sequence, 5′-ATATTAG-3′, separated by an A/T tract of 14 bases. Furthermore, a 10-base direct-repeat palindromic sequence centered within the A/T tract (TAAATTAAAT) (Figure 2.6C). Bioinformatic analyses indicated that this region of DNA is highly conserved in all *V. parahaemolyticus* sequences in public databases and is located upstream of *exsA* in every case.

2.4 Discussion

We have identified a genetic regulatory switch that controls T3SS-1 gene expression using a transposon mutagenesis approach coupled to a sensitive luciferase reporter screen. We set out to identify genes that promote *exsA* gene expression, which encodes the master transcriptional regulator of T3SS-1 in *V. parahaemolyticus*. The data suggest that *V. parahaemolyticus* HlyU acts as a derepressor by binding to DNA and relieving H-NS-mediated repression of *exsA* gene expression, thus leading to T3SS-1 activation. Critically, we demonstrated that HlyU is strictly required for *V. parahaemolyticus* to secrete specific effectors and induce rapid T3SS-1-mediated host cell cytotoxicity.

Within *Vibrio* species, it is well established that HlyU proteins act as global regulators of virulence genes. In *V. vulnificus*, HlyU functions to derepress the *rtxA1* toxin gene antagonizing H-NS binding [324]. In this way HlyU acts as a derepressor, by removal of the H-NS-mediated repression, and not as a true RNA polymerase-recruiting transcriptional activator. In *V. cholerae*, HlyU acts to enhance promoter activity at genes associated with virulence. Insertional mutation of the *hlyU* gene in *V. cholerae* leads to a significant decrease in virulence within an infant mouse infection model, highlighting the importance of this gene for virulence [347]. In *V. anguillarum*, HlyU controls the expression of

the *rtxACHBDE* and *vahI* gene clusters, whose protein products mediate the *V. anguillarum* hemolysin and cytotoxicity activities in fish [328].

Our data provide evidence that *V. parahaemolyticus* utilizes HlyU to positively regulate T3SS-1 expression, which is, to our knowledge, the first report linking this regulator to type III secretion. Our transposon library approach identified *hlyU* and *hns* as encoding protein regulators of *exsA* expression. H-NS has previously been implicated in T3SS-1 regulation in *V. parahaemolyticus* [305]; therefore, we independently confirmed those findings using a different approach. Our data suggest that HlyU is critical to derepress the *exsA* promoter by disrupting H-NS activity. In the absence of H-NS, HlyU is not required for *exsA* promoter activity (Figure 2.4B). Instead, elevated *exsA* promoter activity was observed in the absence of H-NS, suggesting that HlyU is not strictly required to activate the *exsA* promoter and likely displaces a negative regulator. Collectively, the data suggest that HlyU and H-NS form a genetic regulatory switch that serves to tightly control T3SS-1 gene expression in *V. parahaemolyticus*. Different genetic regulatory factors likely contribute to auxiliary *exsA* regulation, as our screen did identify other genes (see Table A.1 in the supplemental material); however, the robustness of the phenotypes linked to *hlyU* and *hns* mutants implicate these respective genes in a central regulatory mechanism for *V. parahaemolyticus* T3SS-1 gene expression. Negative regulation mediated by H-NS is a common paradigm for virulence genes, especially those found within genomic pathogenicity islands [318].

H-NS is proposed to protect the genome from foreign DNA acquired from various genetic transfer events. H-NS propensity to binding A/T-rich sequences likely serves to silence deleterious gene expression, which could stabilize the acquisition of new DNA while limiting detrimental fitness costs [310]. Furthermore, pathogens capable of conditionally derepressing H-NS by expressing specialized DNA binding regulatory proteins would benefit from fine-tuning

of virulence gene expression. Indeed, there are multiple examples of H-NS derepression by DNA binding proteins, including HilD (*Salmonella*), Ler (*E. coli*), and VirB (*Shigella*) [348–350]. Many of these proteins bind at distinct DNA motifs and promote bending or alteration of DNA structure at specific genome sites. DNA binding locally displaces H-NS, thus supporting efficient promoter access and gene transcription [351]. In the case of *V. parahaemolyticus*, we demonstrate that HlyU binds to DNA downstream of a previously reported *exsA* transcriptional start site and adjacent to where H-NS binds the DNA (Figure 2.7) [307]. The current data are not able to identify the exact regulatory mechanism underlying the HlyU–H-NS regulation of *exsA* expression; however, we hypothesize that under certain environmental or infection conditions, particularly ones with elevated magnesium and depleted calcium, HlyU is crucial to trigger *exsA* promoter activity.

HlyU protein sequences are highly conserved among *Vibrio* species (ranging from 81 to 100% identity) [345]. Numerous structure-function studies have established that HlyU proteins form stable homodimers and possess a winged helix-turn-helix (wHTH) domain structure which is modeled to bind DNA within two major grooves [345, 352]. Furthermore, HlyU dimerization is necessary for DNA binding activity in *V. cholerae*, and specific amino acids within HlyU contribute to DNA binding [345]. These critical amino acids are conserved in all *V. parahaemolyticus* HlyU homologues. We demonstrate here that *V. parahaemolyticus* HlyU binds to the DNA upstream of *exsA* using EMSA. Interestingly, we did observe various degrees of DNA shifts that were dependent on the amount of HlyU (Figure 2.6A). A similar observation was previously reported for *V. cholerae* HlyU [345]. The current data cannot differentiate if the DNA shifts were due to (i) increased HlyU binding at multiple sites, creating larger DNA-protein complexes, or (ii) different physical conformations of DNA-protein complexes. DNase I footprinting assays repeatedly detected a single contiguous

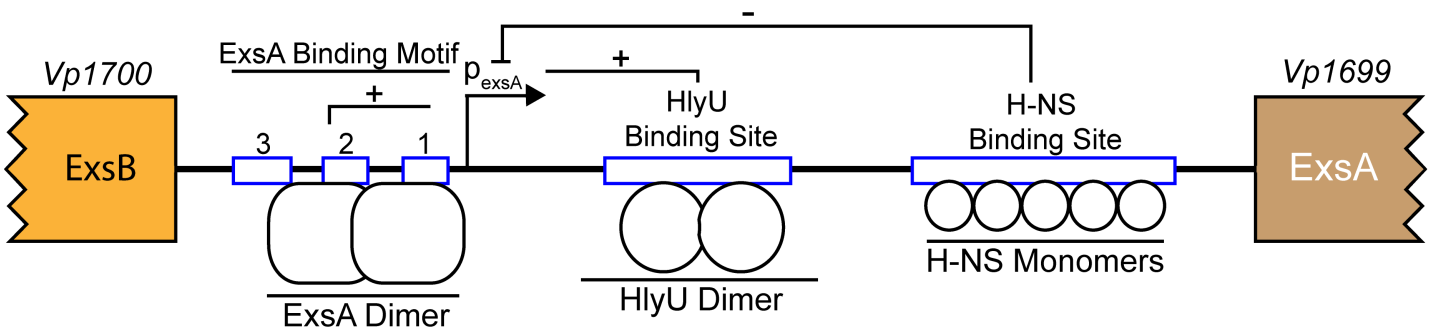

Figure 2.7: Schematic diagram of the intergenic region involved in *exsA* gene regulation. The ExsA binding motif and H-NS binding site have been previously identified. A putative HlyU binding site, as identified in this study, is shown. The ExsA binding motif has three box elements, each composed of conserved DNA sequences: 1, GC box; 2, A box; 3, TTAGN4TT. Protein-DNA interactions at specific sites within the *exsA* promoter region are shown. Plus and minus signs indicate positive and negative regulatory roles, respectively, on *exsA* promoter activity.

56-bp stretch of protected *exsA* promoter DNA, suggesting that HlyU binds at or near this site. We identified a perfect 7-base inverted repeat located within this putative HlyU binding site. This feature is in agreement with *V. cholerae* HlyU, which also binds at an inverted repeat near its *rtx* operon; however, the sequence motif is different in each case. HlyU proteins from other *Vibrio* species bind at different repeat-type sequence motifs [328, 345]; thus, it appears that some flexibility in sequence recognition exists. HlyU is considered a global virulence regulator, so it might be expected to recognize different sequences across the genome. Additional detailed studies will be required to determine sequence specificity requirements for HlyU binding to DNA.

In summary, we have identified a primary genetic switch composed of the DNA binding proteins HlyU and H-NS that serves to regulate T3SS-1 gene expression in *V. parahaemolyticus*. *V. parahaemolyticus* required HlyU to support the expression of ExsA, the master transcriptional regulator of multiple T3SS-1-associated genes. During infection-like conditions, HlyU was necessary for rapid host cell cytotoxicity mediated by T3SS-1. HlyU bound DNA near an inverted repeat located upstream of *exsA* which likely disrupts H-NS-mediated repression at this gene locus in *V. parahaemolyticus*.

2.5 Materials and Methods

2.5.1 Bacterial Strains, Growth Conditions, and Plasmids

Vibrio parahaemolyticus RIMD2210633 was grown in Luria broth (LB; L3522; Sigma) or LBS (10 g tryptone, 5 g yeast extract, 20 g NaCl, pH 8.0). All mutant derivatives described in this study were derived from *V. parahaemolyticus* RIMD2210633 (Table 2.1). All *Escherichia coli* strains were cultured in LB. The following antibiotics (Sigma) were used in growth medium as required: chloramphenicol (5 μg/ml and 30 μg/ml for *V. parahaemolyticus* and

Strain or plasmid	Description	Reference or source
RIMD2210633	Wild-type *V. parahaemolyticus*	8
Δ*vscN1* strain	*V. parahaemolyticus vscN1* null mutant	35
Vp-lux (*tdhS*::*exsA-lux*) strain	*V. parahaemolyticus* with the *exsA* promoter fused to *luxCDABE* gene cassette, integrated into chromosomal *tdhS* locus	This study
hlyU::Tn5 strain	Vp-lux strain with transposon-disrupting *hlyU*	This study
h-ns::Tn5 strain	Vp-lux strain with transposon-disrupting *hns*	This study
*hlyU*1	*V. parahaemolyticus* RIMD2210633 Δ*hlyU*	This study
Vp-EL	*hns*::Tn5-Δ*hlyU*	This study
DH5α	*Escherichia coli* cloning strain	Stratagene
DH5αλpir	*E. coli* host for oriR6K-dependent plasmid replication	
BL21(λDE3)	*E. coli* host for protein overexpression studies	Novagen
pRE112	Suicide plasmid, R6K origin of replication	[353]
pJW15	p15A-based plasmid with promoterless *luxCDABE* gene cassette	50
pΔ*tdhS*	Deletion construct for *tdhS* within pRE112	This study
p*tdhS-exsA*lux	Integration construct to place promoter *exsA-luxCDABE* onto the *V. parahaemolyticus* chromosome within the *tdhS* locus	This study
pEVS104	Conjugative helper plasmid with mobilization machinery, used in triparental matings	26
pEVS170	Mini-Tn5 transposon plasmid, erythromycin selection, R6K origin of replication	26
pVSV105	*Vibrio* shuttle vector with lac promoter and multiple cloning site, replication competent in *V. parahaemolyticus* and DH5αλpir	45
pFLAG-CTC	Cloning plasmid to create C-terminal FLAG sequence fusions	Sigma
pFLAG-CTC-*hlyU*	*hlyU* fused with FLAG sequence	This study
p*hlyU*-FLAG	Expresses *hlyU* under the control of the lac promoter in pVSV105	This study

Table 2.1: Bacterial strains and plasmids used in Chapter 2.

E. coli, respectively), erythromycin (10 μg/ml and 75 μg/ml for *V. parahaemolyticus* and *E. coli*, respectively), kanamycin (50 μg/ml), and ampicillin (100 μg/ml). *V. parahaemolyticus* was routinely grown at 30°C or 37°C aerobically with shaking at 200 rpm for 16 to 18 h. Agar (A5306; Sigma) was added to medium at 1.5% (wt/vol) for solid medium preparations. All plasmids used in this study are listed in Table 2.1.

2.5.2 Generation of Strain *hlyU*1 and Construction of a Plasmid for *hlyU*1 Genetic Complementation

Primers NT393 and NT390 (synthesized by Integrated DNA Technologies [IDT]) were used in a PCR (with Phusion DNA polymerase; M0535S; New England BioLabs) with *V. parahaemolyticus* genomic DNA as the template. This approach amplified a contiguous DNA fragment encompassing upstream flanking sequence of *tdhS* to a partial coding sequence of the gene. Similarly, primers NT391 and NT392 were used to amplify a distal *tdhS* coding sequence and downstream flanking DNA. These DNA fragments were digested with PacI, ligated together with T4 DNA ligase, and then used as a DNA template in a PCR with primers NT393 and NT392. The resulting DNA fragment was blunt-end cloned into the Eco53kI site of pRE112 to create pΔ*tdhS*. Primers NT337 and NT339 next were used with *V. parahaemolyticus* genomic DNA in a PCR to amplify the *V. parahaemolyticus exsA* promoter region. The resulting DNA product was subjected to restriction digestion with EcoRI and BamHI and then cloned upstream of the promoter-less *luxCDABE* gene cassette in pJW15. Finally, PacI digestion was used to excise the *exsA-luxCDABE* DNA fragment from pJW15, which was then cloned into the PacI site of pΔ*tdhS* to generate p*tdhS*::*exsA*lux. p*tdhS*::*exsA*lux then served as a chromosomal integration suicide construct via delivery into *V. parahaemolyticus* using triparental conjugal mating. Chromosomal integrants were selected with medium containing chloramphenicol and then streak purified. The integrants

Designation	Sequence (5'→ 3')	Description or purpose
NT337	CC**GAATTC**AATCGGTTACATTTAATTAGCGC	*exsA* promoter forward
NT339	CC**GGATCC**CCGTTTCTGTGTTTAGTTGGCCTG	*exsA* promoter reverse
AL390	CGC**TTAATTAA**TACATTGACCGGAGCTTG	*tdhS* mutant construction
AL391	CC**TTAATTAA**AGAGCGGTCATTCTGCTG	*tdhS* mutant construction
AL393	AGCTTACAGCTTGGTATGCCT	*tdhS* mutant construction
AL394	GTGGCTATGCACTGGCAGAT	*tdhS* mutant construction
NT394	AA**GAGCTC**ACAGTATCCACTTACGTTGTTACG	Δ*hlyU* mutant construction
NT395	AA**CTCGAG**TTCTTGTAGATTCATGTGTTGG	Δ*hlyU* mutant construction
NT396	AA**CTCGAG**TGCGCAAACTGATCGCACAGACTG	Δ*hlyU* mutant construction
NT397	AAGGTACAAGCACGAGCAATCAACTCGC	Δ*hlyU* mutant construction
NT398	AA**CTCGAG**TCAAGATACTTGGTTTAGTGAGC	*hlyU* genetic complementation
NT399	AA**GGTACC**GTTTGCGCAATAAAGACCGTGAAGG	*hlyU* genetic complementation
NT400	A**GAATTC**CATATGAATCTACAAGAAATGGAGAA	HlyU overexpression
NT401	AA**CTCGAG**GTTTGCGCAATAAAGACCGTGAAGG	HlyU overexpression
NT402	6-FAMN/CCGAATTCAATCGGTTACATTTAATTAGCGC	*exsA* promoter forward-6-FAM
M13F	GTAAAACGACGGCCAGT	Sequencing primer
nleH1-F	GCGGTACCATGCTATCACCATCTTCTGTAAA	*nleH1* gene fragment
nleH1-R	GCACAATTGCCAATTTTACTTAATACCACACTAATAAG	*nleH1* gene fragment

Table 2.2: List of oligonucleotides used in Chapter 2. Restriction enzyme recognition sites are bolded.

were then subjected to allelic exchange by SacB-mediated sucrose selection. Stable integrants within the *tdhS* allele were screened by PCR and then confirmed by Sanger sequencing. All oligonucleotides used in this study are listed in Table 2.2.

2.5.3 Mini-Tn5 Mutant Library Generation within the Vp-lux Reporter Strain

A mutagenesis procedure was followed, with minor modifications [344]. Briefly, a conjugal mating on LBS agar allowed for the delivery of the plasposon pEVS170 into the Vp-lux strain. The mating mixture was plated onto selective LBS agar medium (pH 8.0) containing 10 μg/ml erythromycin and incubated at 22°C for 36 h. Transposants were then individually picked and grown on M9 minimal medium overnight. Finally, the transposants were grown overnight in LB (pH 8.0) supplemented with erythromycin and then frozen in 20% (vol/vol) glycerol in 96-well plate format.

2.5.4 Luciferase Reporter Library Screen

Overnight cultures of the mini-Tn5 mutant library in 96-well plates were grown at 37°C in 5.0% CO_2. A volume of 200 μl of LB medium supplemented with 5 mM EGTA and 15 mM MgSO4 was accurately pipetted into each well of a sterile 96-well clear-bottom, white-walled plate (number 3632; Corning). A sterilized 96-metal-pin well replicator (V&P Scientific) was used to sample the overnight 96-well plate cultures, and then the replicator was used to inoculate each well of the 96-well plate supplemented with LB medium. These plates were incubated with shaking at 30°C and 250 rpm for 3.5 h. Luminometry (counts per second, read at 1 s per well) and OD_{600} endpoint readings were taken using a Victor X5 multilabel plate reader (PerkinElmer).

2.5.5 Statistical Binning to Categorize Transposon Mutants

To narrow mutants down to a reasonable number for genetic characterization, we undertook a statistical binning approach. All mutants that fell below 1,000 cps and above an OD_{600} of 0.4 were binned according to their counts per second into three categories: low glowers (less than 100 cps), low-moderate glowers (100 to 200 cps), and moderate glowers (200 to 1,000 cps). Statistical means and standard deviations were calculated for each group, and mutants that fell 1 standard deviation below the mean for each group were selected for characterization. These data are summarized in Table A.1 in the supplemental material.

2.5.6 Genetic Marker Retrieval

Genomic DNA from selected *V. parahaemolyticus* transposants was isolated and restriction enzyme digested to completion using HhaI, followed by a heat inactivation of HhaI. The digested DNA (final concentration of approximately 40 ng/μl) was then treated with T4 DNA ligase overnight. Finally, the ligated DNA was transformed into *E. coli* DH5αλpir using a standard procedure, and transformants containing self-replicating plasposons were selected on erythromycin LB agar medium. The plasposons were retrieved from the *E. coli* hosts using a standard miniprep procedure and were subjected to Sanger sequencing using the M13 forward sequencing primer. The sequencing data were compared to the RIMD2210633 reference *V. parahaemolyticus* genome to identify the genetic locus were transposition had occurred.

2.5.7 Protein Secretion Assays

T3SS-1 protein secretion assays were performed as previously described [268]. Culture conditions that support T3SS-1 expression were a starting OD_{600} of 0.025 in LB supplemented with 15 mM $MgSO_4$ and 5 mM EGTA and a 4-h incubation at

30°C (250 RPM).

2.5.8 Cytotoxicity Assays

HeLa cells (American Type Culture Collection) were cultured in Dulbecco's modified Eagle's medium (DMEM; Gibco 11995) supplemented with fetal bovine serum, seeded in a sterile 24-well plate (number 3526; Costar) at a density of 10^5 cells/mL, and incubated for 16 h at 37°C in 5.0% CO_2 prior to infection. The HeLa cells were washed twice with 1 mL of phosphate-buffered saline (PBS; 137 mM NaCl, 2.7 mM KCl, 8.1 mM Na2HPO4, 1.46 mM KH2PO4) before infection with selected *V. parahaemolyticus* strains. *V. parahaemolyticus* strains were cultured overnight in LB broth at 37°C and 200 rpm. The cultures were diluted in PBS and adjusted for cell number using OD_{600} measurement and then transferred to phenol red-free DMEM (Gibco 21063) (without serum), resulting in bacterial suspensions of ~5 × 10^5 cells/mL. One milliliter of the relevant suspension was added to the appropriate wells of a HeLa cell-seeded 24-well plate for a multiplicity of infection (MOI) of approximately 5 (verified by viable plate counts). Uninoculated DMEM was added to wells containing the uninfected HeLa cells and the maximal LDH release condition controls. The 24-well plate was incubated for 4 h at 37°C and 5.0% CO_2. A cytotoxicity kit (88954; Pierce) was used according to the manufacturer's instructions. The following formula was used to calculate percent cytotoxicity:

$$\frac{experimental OD_{490} - uninfected OD_{490}}{maximal release OD_{490}} \times 100$$

2.5.9 Construction of a Recombinant Plasmid to Overexpress and Purify HlyU-HIS

Primers NT400 and NT401 were used in a PCR with *V. parahaemolyticus* genomic DNA to amplify the *hlyU* open reading frame without its stop codon.

The resulting DNA fragment was digested with NdeI and XhoI and then cloned into the corresponding restriction sites within pET21a+ (Novagen), thus creating an in-frame fusion to a hexahistidine coding sequence (C-terminal HIS tag). The recombinant plasmid was initially transformed into DH5α and DNA sequence verified. Finally, the plasmid was moved into *E. coli* BL21(λDE3) for HlyU-HIS protein overexpression using a T7 inducible promoter system. Overexpressed HlyU-HIS was purified from the soluble fraction of bacterial lysates using nickel-nitrilotriacetic acid (Ni-NTA)–agarose (Qiagen) and column chromatography as previously described [354]. Extensively washed and purified HlyU-HIS was eluted from columns using an elution buffer (10 mM EDTA, 150 mM NaCl, 20 mM phosphate buffer).

2.5.10 EMSA

The electrophoresis mobility shift assay (EMSA) was performed as previously described [355], with a few minor modifications. Purified HlyU-HIS or BSA (B9000S; New England BioLabs) was mixed with a PCR-amplified *exsA* promoter DNA fragment in binding buffer (10 mM Tris [pH 7.5 at 20°C], 1 mM EDTA, 0.1 M KCl, 0.1 mM dithiothreitol, 5%, vol/vol, glycerol, 0.01 mg ml−1 BSA). A PCR-amplified *nleH1* gene fragment (derived from EPEC genomic DNA) served as an unrelated nonspecific DNA control. Six percent Tris-borate-EDTA (TBE)–polyacrylamide gels were prerun with 1.5 μl of 6× TBE loading dye (6 mM Tris, 0.6 mM EDTA, 30%, vol/vol, glycerol, 0.0006%, wt/vol, bromophenol blue, 0.0006%, wt/vol, xylene cyanol FF) and then loaded with the equilibrated protein-DNA samples. The gel was run for 4 h (100 V, 4°C) and then stained with SYBR green fluorescent DNA dye (Invitrogen) at a 1× concentration in TBE buffer and imaged using a VersaDoc MP5000 system (Bio-Rad). Protein staining and processing of TBE gels was performed using SYPRO Ruby Red (Bio-Rad) as previously described [356] and then imaged using a VersaDoc MP5000 system

(Bio-Rad).

2.5.11 6-FAM DNase I Footprinting Assay

A 6-carboxyfluorescein (6-FAM)-end-labeled *exsA* promoter fragment was amplified by PCR using *V. parahaemolyticus* genomic DNA as the template with primers NT337 and NT402. Using the same EMSA binding conditions, the 6-FAM-labeled *exsA* promoter PCR product was mixed with purified HlyU-HIS or BSA and allowed to equilibrate for 30 min. Various amounts of DNase I (0.5 to 2 U) were added to the protein-DNA mixture, followed by immediate incubation at 37°C for 20 min. To stop DNase I activity, reaction mixtures were rapidly heated to 75°C for 10 min and then purified using a PCR purification kit (Qiagen). The samples were then subjected to capillary electrophoresis on an ABI-3730XL DNA Analyzer (Genome Québec Innovation Center). The chromatogram from this analysis was matched with a Sanger DNA sequencing reaction of the same *exsA* promoter DNA fragment. The HlyU-HIS protected region was identified by searching the chromatogram for a region with decreased 6-FAM fluorescence output. This experiment was repeated four times and included independent binding and DNase I digestion reactions.

2.6 Acknowledgments

We acknowledge Aaron Liu, Courtney Nieforth, and Divya Thomas for technical assistance. Rosalie Fréchette, at the Genome Québec Innovation Center, assisted with troubleshooting the DNase I footprinting assay. This work was supported by an operating grant (RGPIN/342111-2013) from the Natural Sciences and Engineering Research Council of Canada (NSERC). The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Conceptualization: LJG, NAT

Methodology: LJG, NAT

Validation: LJG, NAT

Formal analysis: LJG, NAT

Investigation: LJG, NAT

Writing – original draft preparation: LJG, NAT

Writing – review and editing: LJG, NAT

Visualization: LJG

Supervision: NAT

Chapter 3

Attenuation of a DNA Cruciform by a Conserved Regulator Directs T3SS-1 Mediated Virulence in *Vibrio parahaemolyticus*

This work appears in:

Getz, L.J., Brown, J.M., Sobot, L., Chow, A., Mahendrarajah, J., Thomas, N.A. (2021) *Attenuation of a DNA Cruciform by a Conserved Regulator Directs T3SS-1 mediated virulence in Vibrio parahaemolyticus* **BioRxiv**, DOI: 10.1101/2022.03.07.483294.

3.1 Abstract

Pathogenic *Vibrio* species account for 3-5 million annual life-threatening human infections. Virulence is driven by bacterial hemolysin and toxin gene expression positively regulated by the winged helix-turn-helix (wHTH) HlyU transcriptional regulator family and silenced by Histone-like nucleoid-structuring protein (H-NS). In the case of *Vibrio parahaemolyticus*, HlyU is required for virulence gene expression associated with T3SS-1 although its mechanism of action is not understood. Here, we provide evidence for DNA cruciform attenuation mediated by HlyU binding to support concomitant virulence gene expression. Genetic and biochemical experiments revealed that upon HlyU mediated DNA cruciform attenuation, an intergenic cryptic promoter became accessible allowing for *exsA* mRNA expression and initiation of an ExsA autoactivation feedback loop at a separate ExsA-dependent promoter. Using a heterologous *E. coli* expression system, we reconstituted the dual promoter elements which revealed that HlyU binding and DNA cruciform attenuation were strictly required to initiate the ExsA autoactivation loop. The data indicate that HlyU acts to attenuate a transcriptional repressive DNA cruciform to support T3SS-1 virulence gene expression and reveals a non-canonical extricating gene regulation mechanism in pathogenic *Vibrio* species.

3.2 Introduction

Pathogenic *Vibrio* species cause millions of life-threatening human infections annually [30], as well as fatal infections in seafood organisms that contribute to major economic losses in aquaculture [25]. *Vibrio* spp. are responsible for severe diarrheal disease (in the case of *Vibrio parahaemolyticus* and *Vibrio cholerae*), as well as wound infections causing necrotizing fasciitis (*Vibrio vulnificus*). Both *V. parahaemolyticus* and *V. cholerae* employ a complex array of pathogenicity

factors, including pore forming toxins and secretion systems, to generate disease in a host [30].

These pathogens possess DNA binding proteins that either repress or activate virulence gene expression, many of which alter DNA conformation [192, 269, 286, 296, 357–359]. Specifically, the HlyU family of winged helix-turn-helix (wHTH) DNA-binding proteins positively regulate a subset of *Vibrio* spp. virulence genes [269, 328, 345, 360]. HlyU proteins dimerize and interact with consecutive major grooves of the double-stranded DNA helix and are modeled to induce variably angled DNA bends at intergenic regions [332, 352]. Small drug HlyU inhibitors have been discovered [361, 362] and some have been shown to limit disease in animal models of infection [362]. The exact mechanism underlying HlyU positive gene regulation remains unknown. In the cases of *V. vulnificus* and *Vibrio anguillarum*, competitive DNA binding studies have implicated HlyU in overcoming virulence gene silencing mediated by Histone-like Nucleoid-Structuring protein (H-NS) [324, 327]. Additional genetic evidence in support of this regulatory paradigm was later found in *V. parahaemolyticus*, where *hlyU* was no longer required to support T3SS-1 activity in the absence of *hns* [269]. Moreover in *V. cholerae*, a parallel example between the wHTH ToxT transcriptional activator and H-NS repression has been studied in context of *ctxAB* (cholera toxin) expression [323, 363]. How transcriptional regulators like HlyU and ToxT impact H-NS repression is not completely understood and various models for de-repression of H-NS have been proposed [322, 323, 364]. In many cases, H-NS DNA binding sites do not directly overlap with those of those of the transcriptional regulators raising the possibility of DNA topology or protein-DNA interactions that function at a distance to impact gene expression. For numerous pathogens, H-NS binding to DNA acts to constrain localized DNA supercoiling leading to dramatic changes in DNA topology such as looping, hairpins, and cruciforms [316, 365, 366]. In some cases, the altered DNA topology reduces

promoter accessibility for key transcriptional initiation events [367]. Pathogens must therefore possess a variety of mechanisms to overcome these repressive effects on virulence genes [310, 317, 318, 368, 369].

In *Vibrio parahaemolyticus*, we previously discovered the *exsBA* intergenic region was an HlyU binding site, with inverted repeat sequences separated by A/T rich DNA [269]. Such DNA nucleotide arrangements are known to form DNA cruciform (or 4-way) structures under certain energetic provisions [316, 370, 371]. DNA cruciform structures are dynamic non-β-DNA structures found in all domains of life and are implicated in repressing gene expression [316, 372], replication initiation in bacteriophages and plasmids [370, 373] and occasionally as DNA recombination intermediates [374]. The formation of DNA cruciform structures typically requires high superhelical density – as a mechanism of energy input into the DNA helix – constrained by negative supercoiling which is then relieved by contextual extrusion of DNA strands [375]. While DNA looping and bending mechanisms have been well characterized in bacterial genetic processes, functional DNA cruciform examples in bacterial chromosomes are rare and there are no examples linked to virulence gene regulation. Notably, protein crystallization studies have implicated HlyU binding to bent dsDNA [332, 333], however no HlyU-DNA co-crystal has been solved to date and the functional genetic outcomes have not been investigated.

In this study, we set out to investigate the mechanism of HlyU function in context of *exsA* gene expression and subsequent T3SS-1 gene regulation. *V. parahaemolyticus* Δ*hlyU* mutants are significantly impaired for *exsA* gene expression and exhibit reduced cytotoxicity during infection [269]. Conversely, *hns* null mutants exhibit significantly enhanced and de-regulated expression of *exsA* [269, 305, 307]. HlyU and H-NS binding sites within the *exsA* promoter region do not overlap and are separated by at least 100 base pairs [269, 307] making competitive DNA binding models challenging to reconcile.

Considering the nature and propensity of A/T rich supercoiled DNA to form cruciforms, we hypothesized that a DNA cruciform structure is involved in regulation of *exsA* expression. We present data that implicates HlyU in attenuating a transcriptionally repressive DNA cruciform leading to *V. parahaemolyticus* T3SS-1 mediated virulence.

3.3 Results

3.3.1 *in silico* Identification of Cruciform Forming Loci in *Vibrio* Species

Extensive genomic analyses revealed that potential cruciform forming DNA sequences are significantly enriched at the 3' end of genetic operons, within intergenic regions and near transcriptional promoters [376, 377]. To investigate DNA sequences capable of forming cruciform structures within the *V. parahaemolyticus exsBA* intergenic region, we undertook an *in silico* approach using Palindrome Analyser [376]. We allowed for a maximum of one sequence mismatch in the base-paired cruciform stem DNA and excluded cruciforms with stem sequences which contained less than six nucleotides. Furthermore, DNA cruciforms with looped-DNA which are spaced by less than ten nucleotides were also excluded. These criteria were used to identify DNA sequences which supported the annealing constraints for both DNA cruciform formation and HlyU binding [269, 328, 345, 360].

Twenty-one cruciform sequences which met our criteria for cruciform potential were identified within the ~650 nucleotide *exsBA* intergenic region (Table B.1). Critically, these putative cruciform structures all had positive predicted ΔG values, which is consistent with an energetic supercoiling input requirement for cruciform formation. Three cruciform structures are clustered around the HlyU binding site and have similar ΔG values between 12.94 to 14 (Table B.1).

We also investigated if HlyU binding sites from other *Vibrio* spp. have cruciform forming potential. For all the *Vibrio* spp. evaluated, the HlyU binding site (or sites) was in proximity or overlapped with a putative DNA cruciform structure (Tables B.2, B.3, B.4). These data indicate that DNA sequence features required for cruciform formation are present at a variety of intergenic HlyU binding sites in different *Vibrio* spp. The observations align with the known enrichment of cruciform forming DNA sequences at specific intergenic DNA locations [376]. In all cases, the respective DNA sequences exhibit A/T rich DNA (>90%) which aligns with the commonly described conditions that favor cruciform formation in vivo [290, 371, 373, 375].

3.3.2 Identification and Sequence Mapping of Cruciform Structures

To functionally evaluate the *Vibrio parahaemolyticus exsBA* intergenic region for cruciforms, we built on a previously developed T7 endonuclease assay with supercoiled plasmid DNA substrates [377]. Here, T7 endonuclease cleaves DNA cruciform structures in a sequential two-step process to produce a double-strand DNA cleavage and linearize plasmid DNA. A pUC(A/T) plasmid with an engineered stable cruciform served as a positive control for T7 endonuclease cleavage [377] and empty pBluescript acted as a negative control (Figure 3.1A). As expected, treatment of pUC(A/T) with T7 endonuclease linearized the plasmid, while treatment of pBluescript did not (Figure 3.1B, pBS and pUC(A/T), see lane 4 for each condition).

The *exsBA* intergenic region of *Vibrio parahaemolyticus* was cloned into a pBluescript vector (Figure 3.1A) and assayed by digestion with either T7 endonuclease alone or in combination with other restriction digests (Figure 3.1B). When recombinant pBluescript containing the *exsBA* intergenic region was treated with T7 endonuclease, dsDNA cleavage was consistently observed as evidenced by enrichment of DNA that migrated similarly to linearized DNA (Figure 3.1B, lane

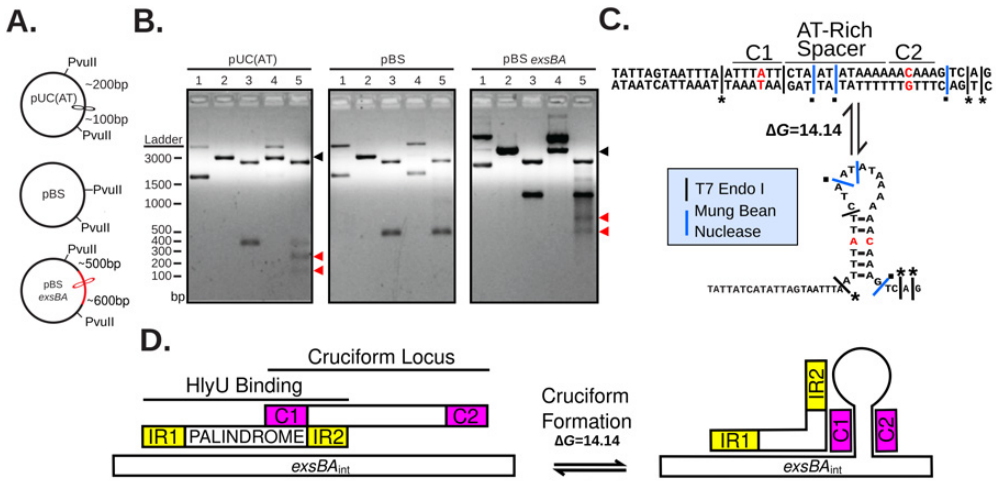

Figure 3.1: Identification of cruciform structures at the HlyU binding site in the *exsBA* locus of *V. parahaemolyticus*. (A) Schematic diagram of plasmid constructs used for T7 endonuclease I digestion analyses. Expected restriction fragment sizes and relative location of PvuII and cruciform sites are identified. The red line and cruciform structure indicate the cloned *exsBA* DNA region. (B) Restriction digestion of plasmid DNA visualized by agarose gel electrophoresis. 1-undigested plasmid, 2-linearized plasmid, 3-PvuII, 4-T7 Endonuclease, 5-T7 Endonuclease followed by PvuII. T7 Endonuclease targets cruciform structures to cause double-strand break in a two-step process. PvuII was used as it flanks the cloned DNA and allows for restriction mapping. The black arrowhead indicates linearized DNA for the respective cruciform forming plasmid constructs (compare lanes 2 and 4). Complete DNA cleavage by T7 endonuclease is rare due to the two step cut process (initial nicking) and dissipation of supercoiling in the intermediate stage, explaining the enriched slower migrating nicked DNA (lane 4, left and right panels). The red arrowheads in lane 5 of each panel indicate DNA cleavage fragments generated by initial cruciform cleavage by T7 endonuclease. pBluescript and pUC(AT) are negative and positive controls respectively. This experiment was repeated at least three times with representative data shown. (C) Schematic depicting cleavage results of the cruciform sequence mapping assay for the *exsBA* intergenic region in *V. parahaemolyticus*. Red nucleotides indicate stem mismatches in the cruciform sequence. The asterisks and black squares indicate cut sites that are consistent with the known specificities of T7 endonuclease and mung bean nuclease respectively. Cleavage data was obtained from multiple independent cloned DNA fragments (see methods). (D) Schematic diagram of the genomic locations of HlyU binding and the identified cruciform structure. Yellow rectangles depict inverted repeats (IR1 and IR2) that flank a palindromic DNA element that constitutes the HlyU binding site within the *exsBA* intergenic region. Cruciform stem DNA sequences (C1 and C2) are indicated by magenta rectangles. Only one 'side' of the stem-loop cruciform is shown in panels C and D for presentation purposes.

4). This suggested the existence of a DNA cruciform within the *exsBA* intergenic region.

Initial T7 endonuclease digestion of the pUC(A/T) and *exsBA* constructs followed by sequential PvuII digestion provided additional DNA cleavage products that allowed approximation of cruciform position by DNA fragment sizes (Figure 3.1B, lane 5, Figure B.1). As expected, uniform, cruciform-associated cleavage products of approximately 200 and 100 bp were consistently observed for pUC(A/T). For the *exsBA* construct DNA cleavage products ~600 and ~450bp in size were observed. Importantly, these *exsBA* associated DNA cleavage products overlap the HlyU binding site [269] and multiple predicted cruciform structures identified by Palindrome Analyser (Table B.1).

To further investigate the intergenic *exsBA* localized cruciform structure, we precisely sequence mapped the relevant T7 endonuclease cut sites (see methods). Here, we found compelling evidence for one of the cruciform structures identified by Palindrome Analyser (Figure 3.1C, Table B.1). Critically, three *exsBA* T7 endonuclease cut sites precisely mapped to the base of the cruciform stem, which is consistent with model synthetic cruciform cleavage studies (Figure 3.1C). Moreover, mung bean nuclease digestion assays efficiently localized single stranded DNA likely associated with 'loop' structures of the DNA cruciform (Figure 3.1C) [373]. These data indicate that a stable cruciform structure exists at a DNA locus that overlaps a HlyU binding site in *V. parahaemolyticus* (Figure 3.1D).

To further explore the cruciform forming capacity of HlyU binding sites in other *Vibrio* species, we subjected HlyU binding intergenic regions from *V. cholerae* (*tlh-hlyA* – 1104bp), *V. vulnificus* (*rtxA1* operon region – 362bp) and *V. anguillarum* (*plp-vah* – 493bp) to the T7 endonuclease and sequential cleavage experiments (Figure B.1). T7 endonuclease cleavage produced linearization of the intergenic regions from *V. cholerae* and *V. vulnificus*, and unique digestion products were observed on sequential digest with PvuII (Figure B.1D and E).

However, no significant cleavage from either the sequential digest or the T7 endonuclease digestion alone was observed for the intergenic fragment from *V. anguillarum* (Figure B.1F, see Discussion). A summary of cruciform location and cleavage data is presented in supplementary figure B.2. These experiments provide evidence that cruciform structures may form at the site of HlyU binding in a variety of *Vibrio* spp.

3.3.3 Evidence for *exsBA* Genetic Locus DNA Supercoiling and Cruciform Formation within *V. parahaemolyticus* Cells

Next, we set out to find evidence for DNA cruciform formation on the *V. parahaemolyticus* chromosome within the *exsBA* intergenic region. We considered the possibility that the *exsBA* genetic locus is subject to DNA supercoiling, as DNA cruciform formation generally results from excessive torsional forces. To investigate this possibility, we measured promoter activity for a chromosomally integrated *exsA* promoter fused to a *luxCDABE* reporter (Figure 3.2A) [269]. Bacteria were grown in the presence of increasing sublethal doses of novobiocin [0-0.4μM], which acts as a DNA gyrase inhibitor and potently reduces DNA supercoiling [378]. *exsA* promoter activity was reduced in a dose dependent manner with increasing concentrations of novobiocin (Figure 3.2B). These data indicate that the *exsBA* intergenic region is subject to DNA supercoiling which impacts on transcriptional activity.

Next, we set out to investigate DNA cruciform formation on chromosomal DNA. We employed a pulse-chase chemical treatment of cells with chloroacetaldehyde (CAA) which interacts with single-stranded DNA and unpaired bases to form DNA ethanoadducts [379]. We hypothesized that if a DNA cruciform exists at this locus in a proportion of cells, CAA treatment would reduce overall *exsA* promoter activity and thus negatively

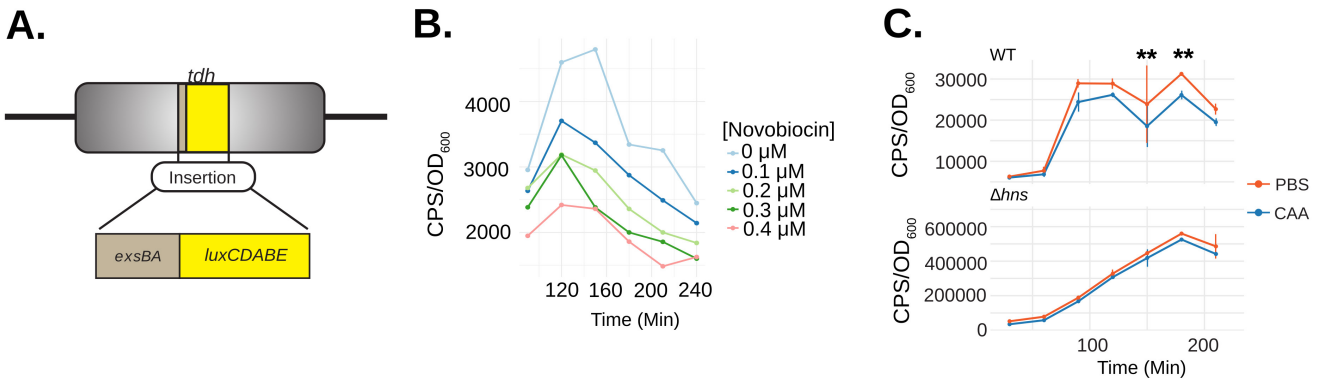

Figure 3.2: DNA supercoiling and cruciform formation within the *exsBA* genetic locus regulate *exsA* promoter activity. (A) Schematic depicting the *exsBA* intergenic-*luxCDABE* luciferase transcriptional reporter fusion integrated into the *V. parahaemolyticus tdh* chromosome locus to allow expression of the luciferase cassette [269]. (B) *V. parahaemolyticus* harboring the *exsBA-luxCDABE* reporter (depicted in panel A) was treated with increasing concentrations of novobiocin and *exsA* promoter activity was quantitatively measured based on in vivo real-time light emission. (C) Wildtype *V. parahaemolyticus* and Δ*hns* strains harbouring the integrated *exsBA-luxCDABE* transcriptional reporter were treated with either phosphate buffered saline (PBS) or 40nM chloroacetylaldehyde (CAA). Light emission was measured as counts per second (CPS) every 30 min, along with OD_{600}, for a total of 4 hrs. Statistical significance was determined by multiple unpaired *t*-tests, ** $p<0.01$, $n=2$ and data were visualized using R and the ggplot2 package.

impact the expression of a transcriptional luciferase reporter.

The experiments were performed using an initial pulse of 40nM CAA which did not inhibit growth (Figure B.3). DNA ethenoadduct formation via CAA modification was performed in wildtype and Δ*hns* strains containing an *exsA* promoter-*luxCDABE* reporter integrated into the chromosome of each strain to permit cis-regulation by DNA binding proteins, including HlyU and H-NS (Figure 3.2A). We reasoned that the Δ*hns* mutant, owing to its lack of H-NS containment of DNA negative supercoiling, should be deficient for DNA cruciform structures, and generally less affected for *exsA* promoter activity. *in situ* real time *exsA* promoter activity measurements revealed that the CAA-treated wildtype reporter strain had significantly reduced *exsA* promoter activity compared to untreated cells at 150 and 180 minutes (Figure 3.2C). In contrast, CAA treatment of the Δ*hns* reporter strain caused no a difference at these time points using a 2-way ANOVA with Bonferroni multiple test correction (Figure 3.2C). These data suggest that CAA mediated ethanoadduct modification of unpaired DNA bases reduced transcriptional activity of the luciferase reporter. Moreover, H-NS containment of DNA supercoiling may play a role in the formation of the unpaired DNA within a cruciform structure. Importantly, these data are consistent with the well-established requirement for DNA supercoiling in cruciform formation *in vivo* [371, 372]

3.3.4 Inverted Repeat and Palindrome Sequence Elements Contribute to HlyU Binding

The HlyU binding site upstream of *exsA* in *V. parahaemolyticus* consists of a 56-bp DNA stretch [269] that partially overlaps the cruciform-forming locus (Figure 3.1D). We sought to investigate the DNA sequence within this 56bp region that is necessary for HlyU binding using electrophoretic mobility shift assays (EMSA). For these experiments, dsDNA fragments spanning 56bp were formed by

annealing two single-stranded complementary oligos (see methods) and incubated with purified HIS-tagged HlyU (Figure B.4). HlyU bound to the wildtype DNA sequence producing two shifted DNA species in EMSA experiments (denoted as S1 and S2, (Figure 3.3A,C). Scrambling the palindrome sequence ATTTAATTTA between the perfect inverted repeats (ATATTAG and CTAATAT) to form mutant PAL1 abrogated HlyU binding to the DNA target (Figure 3.3A,B). This provided evidence that the palindrome sequence is necessary for HlyU interaction with DNA.

To further explore this specific *exsBA* region, we generated a series of mutations in the inverted repeat sequences adjacent to the palindrome, as well as the palindromic sequence itself (Figure 3.3B). Every modification to these sequence elements impacted HlyU interaction with DNA by EMSA (Figure 3.3C) with one caveat. One of the inverted repeat mutations distal to *exsA* (IR1) partially bound HlyU producing only S2, the low molecular-weight shift species (Figure 3.3C). This is the only mutation that is adjacent (does not overlap) to the cruciform stem sequences (C1 and C2) previously identified (Figure 3.1D) (see discussion). Taken together, these data suggest that both the inverted repeat sequences and the palindromic spacer are necessary elements for efficient HlyU binding and that the cruciform-forming locus within the *exsBA* intergenic region partially overlaps with the site of HlyU binding.

3.3.5 Mutations that Alter HlyU Binding Negatively Impact T3SS-1 Activity in *V. parahaemolyticus*

The genetic and biochemical data suggested that HlyU binding to DNA attenuates a transcriptionally repressive cruciform leading to *exsA* expression. To address this further, we investigated the *V. parahaemolyticus* chromosomal intergenic *exsBA* locus by mutational analysis. Mutations were introduced onto the chromosome by

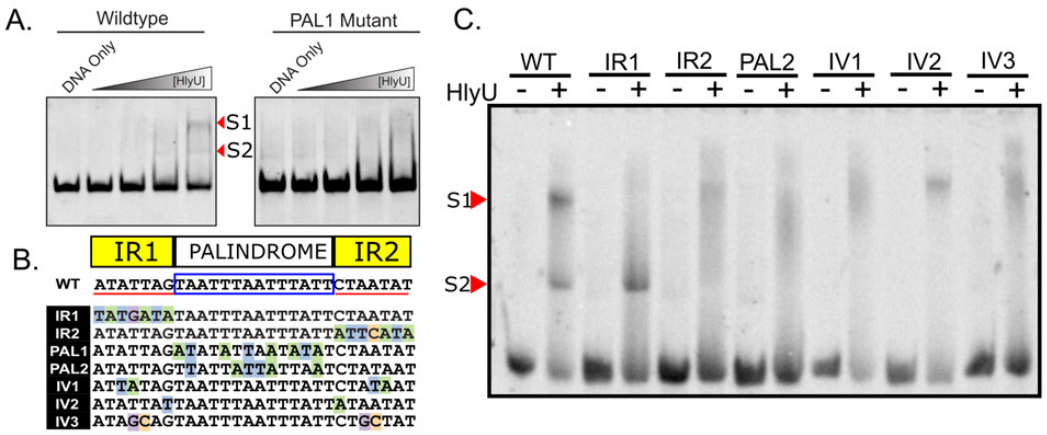

Figure 3.3: Mutational analysis identifies nucleotide elements important for HlyU interaction with DNA near the *exsBA* cruciform forming locus. (A) Oligonucleotides constituting a wildtype *exsBA* HlyU binding site, as well as a mutant in the palindromic region (PAL1), were annealed by sequential cooling and used as DNA targets for HlyU binding in electrophoretic mobilty shift assays (EMSA). Arrowheads refer to distinct bands that appear upon elevated amounts of purified HlyU. S1 and S2 indicate distinct shifted species in the EMSA assay. (B) Schematic of the specific mutant DNA sequences compared to wildtype (WT). Nucleotides with a coloured background identify differences from the wildtype sequence. Perfect inverted repeat DNA sequences are underlined in red separated by a boxed palindrome sequence. (C) Oligonucleotides of mutants described in (B) were annealed and subjected to EMSA with and without purified HlyU. Arrowheads refer to shifted HlyU-DNA complexes S1 and S2 All EMSA experiments were performed at least three times with representative data images shown.

allelic exchange and verified by DNA sequencing. The resulting mutant strains were compared to the parental strain, as well as T3SS-1 deficient Δ*hlyU* and Δ*vscN1* mutants [268, 269] using protein secretion and host cell infection assays. T3SS-1 specific protein secretion was markedly reduced in each mutant, except for mutant IR1 which appeared slightly reduced, as detected by SDS-PAGE analysis (Figure 3.4A). The ability of these mutants to cause cytotoxicity towards HeLa cells during infection was significantly impaired compared to the parental strain (Figure 3.4B). The IR1 and IV2 mutants exhibited intermediate cytotoxicity and therefore retain intermediate T3SS-1 activity. HeLa cytotoxicity was restored in all the mutants by complementation with *exsA in trans* (Table B.5), which bypasses the effect of the chromosomal mutations. These *in vivo* analyses show that mutations negatively impacting on HlyU binding produce a marked decrease in T3SS-1 activity as measured by secretion and infection (cytotoxicity) assays.

3.3.6 HlyU Binding Attenuates a DNA Cruciform in the *exsBA* Intergenic Region to Support Gene Expression

Next, we set out to determine if HlyU binding to DNA serves to attenuate cruciform formation to support efficient gene expression. Informed by the cruciform mapping data (Figure 3.1) and EMSA data (Figure 3.3), we aimed to generate modified *exsBA* DNA constructs that could still form cruciforms but were unable to bind HlyU, or were inefficient for HlyU binding. We reasoned that such DNA constructs could be generated given the intermediate phenotypes consistently observed for the IR1 and IV2 mutant bacterial strains (Figure 3.3). With this approach, we deleted each inverted repeat element independently (ΔIR1 and ΔIR2), both inverted repeat elements (ΔIR1ΔIR2), and the A/T rich palindrome which centers the inverted repeat sequences (ΔPAL).

The modified DNA was assessed with Palindrome Analyzer for cruciform forming potential. This analysis revealed that each DNA construct still retained

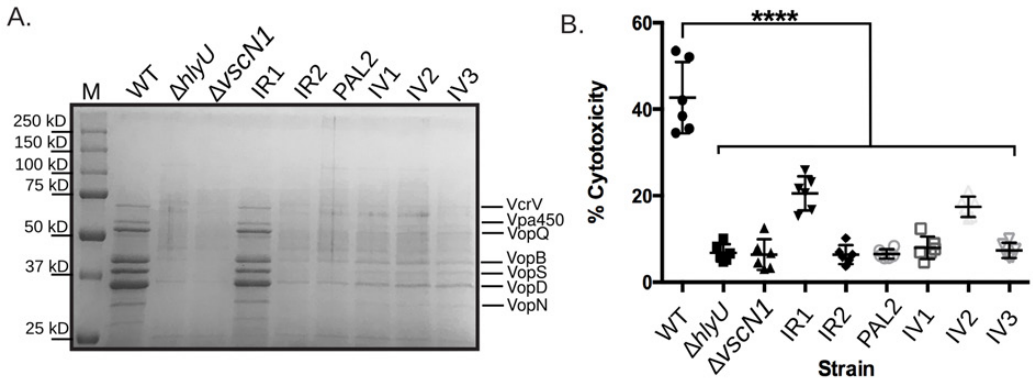

Figure 3.4: *V. parahaemolyticus* T3SS-1 protein secretion and host cell cytotoxicity are reduced in bacteria harbouring mutations at the HlyU binding site near the *exsBA* DNA cruciform forming locus. (A) Total secreted protein profiles derived from *V. parahaemolyticus* harbouring chromosomal mutations at the HlyU binding site and cruciform locus. The specific DNA mutations are listed in figure 3.3B and labeled accordingly for each gel lane. Proteins were visualized by Coomassie staining. M refers to protein standard. The dominant protein species have been previously identified from WT using mass spectrometry [268]. The experiment was performed at least three times with a representative stained protein gel shown. (B) A LDH release assay from cultured HeLa cells was used to determine host cytotoxicity upon infection with the indicated *V. parahaemolyticus* strains. Statistical significance was determined by two-way ANOVA (against WT) using a Bonferroni test correction. **** $p<0.001$, $n=3$. Δ*hlyU* and Δ*vscN1* are deficient for T3SS-1 activity and were included as comparative controls.

cruciform forming potential with comparable Δ*G* values ranging from 12.94 to 14.37 to the original cruciform (14.14) even with the nucleotide deletions (Figure B.5). Proof of cruciform formation for these constructs was obtained by demonstrating T7 endonuclease cleavage of the cloned supercoiled DNA sequences (Figure 3.5A). This further substantiates the propensity for A/T rich DNA (>90%) in intergenic regions to form cruciforms given appropriate base pairing and energetic provisions [366, 371]. Furthermore, the corresponding genetic deletions created different DNA juxtapositions for the *exsBA* intergenic DNA sequence which were predicted to alter HlyU binding. Indeed, purified HlyU bound inefficiently to the modified linear DNA fragments and required high amounts to initiate shifts in EMSA experiments (Figure 3.5B).

Next, the corresponding modified *exsBA* intergenic DNA fragments were fused to a promoter-less *luxCDABE* cassette in a plasmid reporter system and assessed for luciferase expression in wild type *V. parahaemolyticus*. In this system, chromosomally encoded ExsA proteins act *in trans* to activate the cloned *exsA* promoter in the plasmid construct to support light emission. Moreover, HlyU proteins acting *in trans* interact with appropriate *exsBA* DNA and inform on cruciform attenuation. We observed some HlyU-independent light emission from this plasmid system (i.e., within a Δ*hlyU* strain) however the level was significantly reduced compared to wild type *V. parahaemolyticus* (Figure B.6) thus demonstrating that maximal gene expression was HlyU-dependent. All the modified *exsBA* intergenic DNA constructs were introduced in wildtype *V. parahaemolyticus* and produced reduced light emission compared to the wildtype *exsBA* intergenic sequence (Figure 3.5C). Notably, the ΔPAL construct, which forms a cruciform but is missing the 14 palindromic nucleotides that make up the central core of the HlyU binding site was the most negatively impacted for light expression. The ΔIR2 and ΔIR1ΔIR2 constructs supported intermediate levels of light emission, below that of wildtype *exsBA* DNA but higher than ΔPAL and ΔIR1.

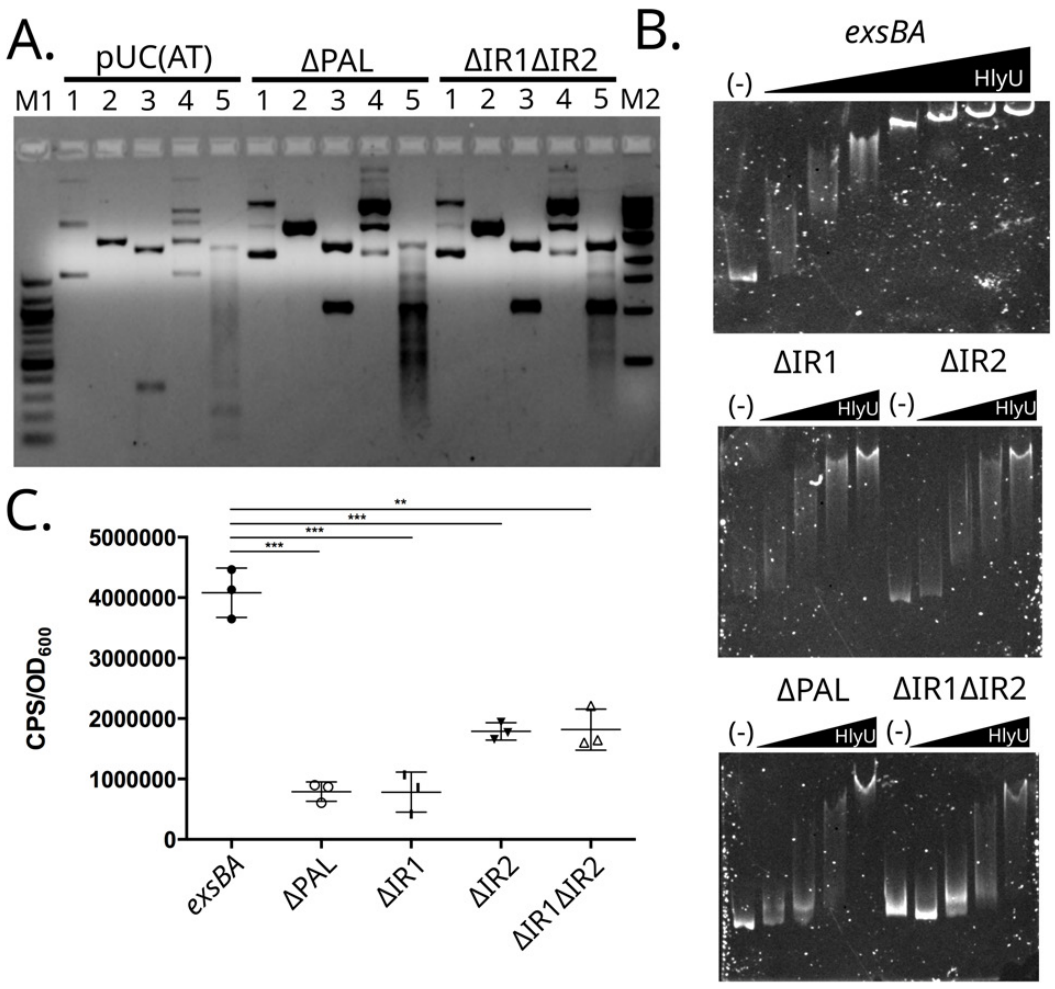

Figure 3.5: Genetic deletions of intergenic *exsBA* inverted repeat and palindrome sequences retain cruciform formation but are altered for HlyU binding efficiency. (A) Cruciform cleavage assays using T7 endonuclease and other restriction enzymes. 1-undigested plasmid, 2- linearized plasmid, 3-PvuII, 4-T7 Endonuclease, 5-T7 Endonuclease followed by PvuII. T7 Endonuclease targets cruciform structures to cause double-strand break in a two-step process. PvuII was used as it flanks the cloned DNA and allows for restriction mapping. Cruciform associated DNA fragments are indicated by arrowheads in lane 5 of the respective samples. M1 and M2 are DNA molecular size standards (B) EMSA with increasing amounts of purified HlyU protein mixed with the indicated *exsBA* genetic deletion fragments. The DNA species were stained with SYBR green. The (-) indicates no HlyU protein and therefore unshifted DNA template. (C) Luciferase activity of plasmid constructs with genetic transcriptional fusions to a *luxCDABE* cassette. The relevant *exsBA* intergenic deletions are indicated and '*exsBA*' represents the wildtype sequence. Multiple t-tests were performed to determine statistical significance, ***: p-value < 0.01, **:p-value<0.05. All experiments were repeated at least three times. Technical replicates are shown in panel C.

Combined with the T7 endonuclease cleavage results, these data indicate that different cruciforms produce repressive structures that impact on transcription activity. More importantly, the ability of HlyU to efficiently bind its target sequence has a positive effect on transcriptional activity, likely by competitively attenuating a cruciform structure.

3.3.7 HlyU is Required for ExsA Auto-activation

The HlyU binding site partially overlaps the cruciform forming locus and constitutes a co-localized cis-genetic element that is involved in *exsA* gene expression. However, the location of HlyU binding is approximately 70 base pairs downstream of the known auto-regulatory *exsA* promoter [269, 296]. Based on this positioning, it is unclear how HlyU positively impacts *exsA* promoter activity however alterations to local DNA topology induced by HlyU binding could be involved.

To better study this system and parse the role of HlyU that supports ExsA production, we reconstituted the *exsBA* intergenic region in a transcriptional luciferase reporter system within a heterologous *E. coli* strain (DH5αλpir) lacks native HlyU or ExsA proteins. Recombinant plasmids coding for HlyU and ExsA were generated, and we confirmed the expression of HlyU and ExsA in this system by immunoblotting cell lysates (Figure 3.6A).

As expected, expression of both HlyU and ExsA in *E. coli* supported robust transcriptional activity of the *exsBA* intergenic region as measured by the luciferase reporter (Figure 3.6B). By expressing HlyU alone in the heterologous system (without ExsA), we identified that HlyU was necessary and sufficient to drive low level luciferase expression (Figure 3.6B). This observation suggests the existence of an additional cryptic promoter downstream of the ExsA autoregulated promoter. Critically, ExsA alone was unable to generate luciferase expression from the *exsBA* intergenic region, revealing that HlyU is essential to support transcriptional

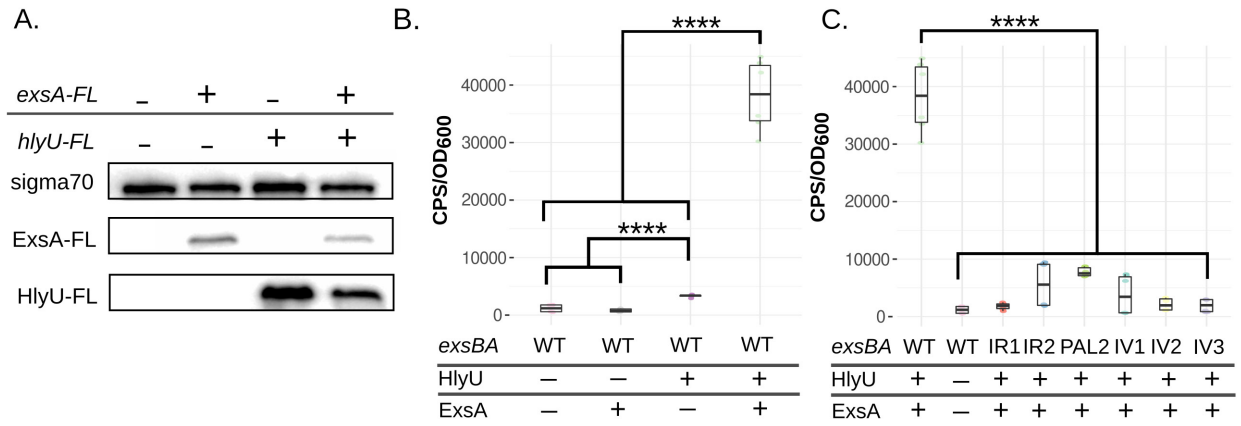

Figure 3.6: Reconstitution of the *exsBA* minimal regulon in *E. coli* provides evidence for a kick-start model of *exsA* regulation in *V. parahaemolyticus*. (A) *E. coli* DH5αλpir cell lysates from strains harboring plasmids as indicated were subjected to immunoblotting to detect ExsA and HlyU. (B) DH5αλpir containing an *exsBA* transcriptional reporter fusion, *hlyU*, and/or *exsA* were assessed for light emission after 3 hrs at 37°C. Photons per second (CPS) and cell density (OD_{600}) were measured. (C) Bacterial strains containing cruciform mutations in context of the *exsBA-luxCDABE* transcriptional reporter fusion were assessed for light emission as in panel B. Multiple t-tests were performed to determine statistical significance, ****: p-value < 0.0001. Data was plotted with R and the ggstatsplot package.

activity driven from the ExsA autoregulated promoter. This suggested that HlyU binding to DNA is involved in the removal of a repressive DNA cruciform. Indeed, mutations that alter HlyU binding yet maintain cruciform forming potential were significantly reduced for *exsA-luxCDABE* expression (Figure 3.6C). Moreover, these *E. coli* heterologous system data are in direct agreement with the inverted repeat deletion plasmid experiments performed in *V. parahaemolyticus* in Figure 3.5.

3.3.8 Transcriptional Start Site Mapping Identifies a Cryptic Promoter

The potential of a cryptic HlyU dependent promoter within the *exsBA* intergenic region was very interesting in that it would theoretically support initial ExsA production to autoregulate the ExsA-dependent *exsA* promoter (e.g. positive feedback loop). To address this possibility, we mapped transcriptional start sites (TSS) for mRNA species containing the *exsA* open reading frame using 5'-Rapid Amplification of cDNA Ends (5'RACE). As expected, we identified *exsA* mRNA originating from near the known auto-regulatory distal *exsA* promoter (Figure 3.7A,B). Notably, we discovered a shorter *exsA* mRNA transcript initiating downstream of the HlyU binding site and cruciform forming locus (Figure 3.7A,B). We then confirmed the presence of a cryptic proximal *exsA* promoter (P1) using a luciferase reporter fusion. The entire *exsBA* intergenic region generated significant luciferase expression consistent with the existence of the distal auto-regulatory promoter and with previous experiments (Figure 3.7A) [269]. However, when the auto-regulatory distal promoter P2 was removed (ΔP2), the newly identified P1 promoter was still able to drive luciferase expression (Figure 3.7B). Importantly, a *hlyU* null mutant generated significantly less luciferase expression from the same luciferase reporter, indicating that the P1 promoter requires HlyU for maximal

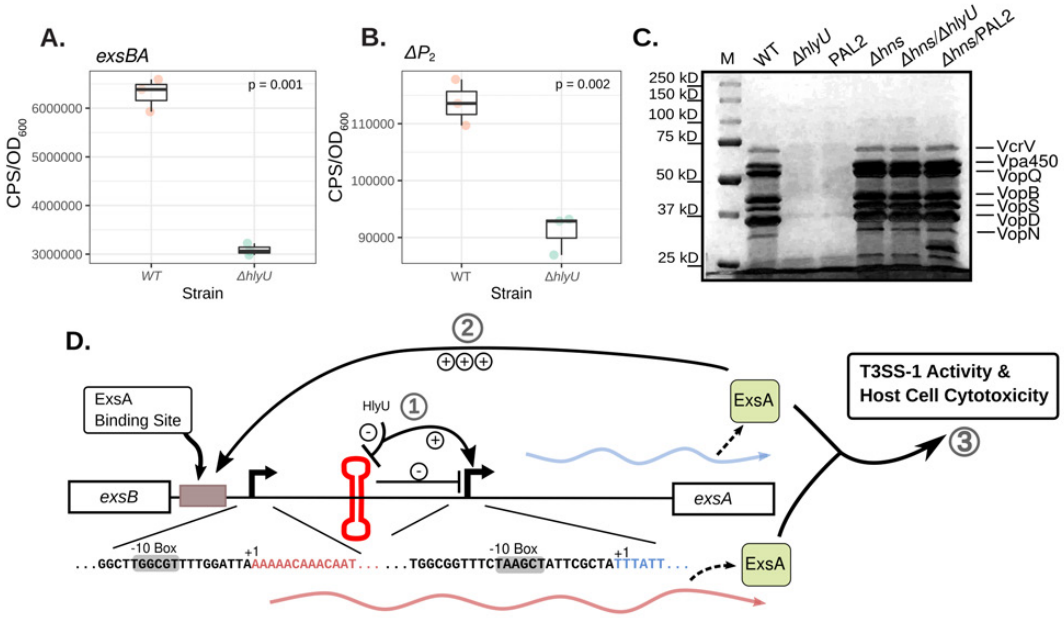

Figure 3.7: The *exsBA* intergenic region contains a autoregulatory and an internal promoter elements that require HlyU binding near a cruciform to initiate a positive transcriptional feedback loop. (A) The *exsBA* intergenic region drives high level light emission from lux reporter construct in an HlyU dependent manner in *V. parahaemolyticus*. (B) Deletion of the autoregulatory promoter [P2] reveals light emission from an internal promoter that requires HlyU for maximal activity. (C) HlyU binding to a palindromic sequence near the cruciform locus is no longer required in the absence of H-NS to support T3SS-1 associated gene expression. Total secreted proteins of the indicated strains were subjected to SDS-PAGE and then stained by Coomassie blue. The indicated T3SS-1 proteins have previously been identified by mass spectrometry analyses. (D) Schematic model of genetic regulation of *exsA* within the *exsBA* intergenic region. 1) HlyU binds to DNA at the *exsBA* intergenic region and attenuates a cruciform structure allowing the activation of an internal promoter for *exsA* expression. 2) Expressed ExsA can autoactivate a distal upstream promoter, driving a majority of ExsA expression (3). Expressed ExsA can then drive T3SS-1 activity and host-cell cytotoxicity. DNA sequences shown indicate the identified promoter and transcriptional start sites from the described 5'RACE experiment.

activity (Figure 3.7B). It is noteworthy that the HlyU-dependent P1 promoter is located within a DNA region that directly overlaps with H-NS binding (Figure B.7A) [307]. Collectively, these data identify at least two distinct mRNA species which support ExsA expression. These data also address how ExsA is first expressed through HlyU binding upstream of the P1 promoter, allowing ExsA to positively feedback and auto-regulate the distal promoter and therefore much of its own expression.

3.3.9 The Action of HlyU Binding at the Cruciform Locus is Dispensable in the Absence of H-NS

We and others have previously shown that *V. parahaemolyticus hns* mutants are deregulated for *exsA* expression and hyper-secrete T3SS-1 proteins [269, 305]. H-NS binds an approximately 200 bp span within the *exsBA* intergenic region [307]. It is possible that H-NS could mask the newly identified P1 promoter from cellular transcriptional machinery (Figure B.7A). The exact mechanism of H-NS gene silencing of *exsA* expression is unknown but based on the large section of DNA coverage, it likely involves H-NS nucleation that leads to nucleofilament mediated DNA 'stiffening' to silence transcription [331]. Moreover, H-NS contributes to constraining chromosomal DNA supercoiling [365] which could also impact on *exsA* expression. In support of this view, magnesium, which reduces H-NS associated nucleofilaments [317, 380] is a potent inducer of *exsA* expression and T3SS-1 activity in wild type *V. parahaemolyticus* [268]. Nonetheless, any influence of magnesium on H-NS must be synergistic with contextual HlyU action to support T3SS-1 activity as *hlyU* null mutants are defective for *exsA* expression and efficient T3SS-1 secretion (see Figure 3.4A) [269]. We attempted to rescue a *hlyU* mutant for T3SS-1 activity by culturing it in elevated amounts of magnesium (up to 55mM) which inhibits H-NS binding to DNA *in vitro* [380] but were not successful in restoring *exsA* expression (data not shown). This further highlighted

the critical role of HlyU for *exsA* expression and suggested synergistic actions of HlyU binding to DNA along with a separate undefined H-NS de-repression mechanism for efficient T3SS-1 expression.

We set out to investigate the T3SS-1 associated regulatory effect of HlyU binding to DNA in the presence and absence of H-NS. Notably, H-NS contributes to constraining DNA supercoiling which is a requisite condition for DNA cruciform formation [375]. We hypothesized that the important role of HlyU binding to *exsBA* intergenic DNA would be dispensable for T3SS-1 activity in the absence of H-NS. Indeed, a Δ*hns*Δ*hlyU* strain exhibited deregulated T3SS-1 activity revealed by elevated T3SS-1 proteins in a secretion assay (Figure 3.7C). Further, a Δ*hns*/PAL2 double mutant where HlyU is present but unable to efficiently bind *exsBA* intergenic DNA due to alteration of the target palindrome sequence, similarly exhibited high levels of T3SS-1 secreted proteins (Figure 3.7C). In contrast, the same PAL2 mutation in the presence of HlyU and H-NS was unable to support efficient T3SS-1 protein secretion (i.e. PAL2, Figure 3.7C), revealing a context specific H-NS associated phenotype for this mutation. This suggests that HlyU binding near the site of cruciform formation is a critical event for productive *exsA* expression in wild type *V. parahaemolyticus* with a condensed nucleoid under typical torsional stresses. In the absence of H-NS, the strict requirement for HlyU binding to *exsBA* intergenic DNA is eliminated. This suggests that in *V. parahaemolyticus*, a separate but synergistic mechanism of HlyU-mediated cruciform attenuation and H-NS activity contributes to *exsA* gene expression and T3SS-1 associated virulence.

3.4 Discussion

This study reports HlyU DNA-binding at the site of a DNA cruciform as the first step in an extensive regulatory cascade leading to T3SS-1 virulence gene

expression in *Vibrio parahaemolyticus*. We propose the following mechanism for DNA cruciform involvement in *exsA* expression and T3SS-1 activity in *Vibrio parahaemolyticus*: 1) HlyU binding to DNA attenuates a repressive DNA cruciform which supports internal promoter activity leading to initial low level *exsA* mRNA expression, 2) initial ExsA production (mediated by HlyU function) autoregulates the upstream *exsA* promoter to drive elevated *exsA* transcription, and 3) high cellular levels of ExsA drives the expression of multiple ExsA-dependent T3SS-1 gene operons [296] leading to host cell cytotoxicity (Figure 3.7D). While the mechanism with which H-NS is de-repressed using this model remains unclear, it is possible that changes in DNA topology by HlyU binding at the DNA cruciform disrupt H-NS binding [318, 329–331].

In many pathogenic *Vibrio* species, HlyU has been proposed to alleviate H-NS mediated virulence gene silencing by outcompeting and displacing H-NS from A/T rich DNA sequences [324, 327]. This view agrees with reports for *Vibrio* spp. DNA binding regulators such as LuxR and ToxT [319, 323] and aligns with the roles of DNA binding regulators in other pathogens (e.g., SsrB, Ler) [350, 381]. It is reasonable to consider that each DNA binding regulator displaces H-NS associated DNA in a contextual and localized manner. Moreover, H-NS nucleofilament DNA formation at A/T rich regions and DNA bridging mechanisms introduce negative DNA supercoiling and localized DNA structural changes [317]. Such conditions overcome the energetics of DNA duplex stability to generate unpaired DNA bases such as those found in cruciforms. We recognize our data for DNA cruciform formation always depended on DNA supercoiling whether it was *in vitro* (purified supercoiled plasmid DNA) or *in vivo* (CAA treated bacteria). Critically, these observations agree with previous DNA supercoiling mechanistic studies and A/T rich sequence requirements for DNA cruciform formation [366, 375, 382].

DNA cruciforms are very challenging to study due to the localized energy input and base pairing requirements that contribute to their formation. We used a

variety of traditional genetic and biochemical tools in addition to pharmacological (novobiocin) and chemical genetic approaches (CAA treatment) to investigate the DNA cruciform at the *V. parahaemolyticus exsBA* locus. Our T7 endonuclease cruciform cleavage data was in striking agreement with model oligonucleotide J-structure cruciform cleavage studies [383]. Specifically, we observed multiple T7 endonuclease cruciform cleavages 3-5 nucleotides from the cruciform base (Figure 3.1C), which is the established sequence independent location of cleavage for this enzyme. Additional independent evidence for single stranded 'loop' DNA associated with cruciforms was obtained with mung bean nuclease digestions. We also found T7 endonuclease cleavage evidence for cruciform forming DNA in proximity to HlyU binding sites in *V. cholerae* (*hlyA*) and *V. vulnificus* (*rtxA1* operon region), but not for *V. anguillarum* (*rtxH-rtxB*). Palindrome analyser indicated the putative *V. anguillarum* cruciform required more energy input for cruciform formation (ΔG=17.29) compared to other *Vibrio* spp. (ΔG=14.14 for *V. parahaemolyticus*, Figure 3.1C). This provides two possibilities: 1) that a cruciform structure doesn't exist at this site in *Vibrio anguillarum* or 2) that our plasmid system is incapable of creating the necessary conditions for cruciform formation outside of *V. anguillarum*. Nonetheless, DNA cruciforms are found near HlyU binding sites in multiple pathogenic *Vibrio* spp. In the case of the *V. parahaemolyticus exsBA* genetic locus, the DNA cruciform appears to operate as a transcriptionally repressive element that requires attenuation to permit *exsA* gene expression from a cryptic internal promoter that is masked by H-NS. As T3SS-1 biosynthesis is a major cellular investment in *V. parahaemolyticus* (expression of 40+ genes) [296], it follows that the entry master regulator ExsA is tightly repressed to prevent spurious expression. Furthermore, *exsA* expression is contextually de-repressed by HlyU to support coordinated T3SS-1 associated gene expression during host infection.

Our data expands knowledge relating to HlyU-DNA binding outcomes. The

DNA binding data along with HlyU crystal structures suggest that a core A/T rich palindrome forms DNA major grooves to accommodate α4 helical domains found within wHTH HlyU dimers [332, 333]. Furthermore, inverted repeat DNA elements interact with the HlyU 'wing' domains to support efficient binding [345]. Notably, the IR1 mutant *V. parahaemolyticus* strain presented in this study is particularly interesting as it produced an intermediate level of T3SS-1 activity and host cytotoxicity compared to wild type bacteria. The IR1 mutation alters the inverted DNA repeat sequence on the left side of the central core palindrome, while the right inverted DNA repeat element IR2 is unchanged (Figure 3.3B). Our data indicates that HlyU has reduced binding in context of IR1 altered DNA (Figure 3.3C). We speculate that this imperfect binding supported partial attenuation of cruciform formation, thus allowing for partial *exsA* expression. Critically, a similar outcome was not observed for the alterations of the DNA inverted repeat to the right of the core palindrome DNA (i.e., IR2 mutants). The key difference between IR1 and IR2 DNA elements is that IR1 is outside of the cruciform forming DNA stem and loop sequences, whereas IR2 directly overlaps and is within the cruciform DNA locus (Figure 3.1D). While additional studies will be required to unravel these experimental observations, it appears that HlyU binding in the immediate vicinity of the cruciform locus produces DNA topology changes that expose a cryptic promoter leading to *exsA* gene expression.

Our data supports two models for *V. parahaemolyticus* HlyU interaction with cruciform forming DNA. In the first model, HlyU interacts with double-stranded DNA and prevents the formation of a DNA cruciform by sterically hindering annealing of cruciform stem associated nucleotides. This model is best supported by our data, most notably by DNA-binding EMSA analysis using linear dsDNA. The requirement of HlyU for *exsA* promoter activation in situ (Figure 3.6B) is also consistent with this interpretation. An alternative possibility is that HlyU interacts with bent DNA found at the base of the cruciform structure and destabilizes the

cruciform directly by conformational changes induced by protein binding. Such an interaction is possible based on studies evaluating the formation and structure of DNA cruciforms [366, 371, 375, 383] and HlyU binding to a bent planar face of DNA [332, 345]. We explored this possibility with *in vitro* studies but were unsuccessful in showing a direct HlyU interaction with a synthetic cruciform DNA structure (data not shown). We cannot exclude the possibility that the synthetic DNA cruciform was improperly formed or assumed a conformation less favorable for HlyU binding. Therefore, the data has limitations in that it lacks a defined HlyU mechanism for direct cruciform binding. Instead the data suggest that cruciform attenuation is necessary for initial *exsA* gene expression and the process is strictly dependent on HlyU binding to DNA. Complex biophysical protein-DNA binding experiments, such as fluorescence-resonance energy transfer (FRET) or circular dichroism, beyond the scope of this study will be required to test the proposed models.

DNA cruciforms and other 4-way junctions are found in all living cells and some plasmids and viruses.The dynamic and temporal aspects of cruciform formation are incompletely understood however DNA supercoiling is thought to provide energy to facilitate DNA strand extrusion at A/T rich containing regions [375]. This outcome would likely occlude or prevent RNA polymerase access to specific DNA promoters [366]. Accordingly, specialized DNA binding proteins would be required to act near certain DNA cruciforms to permit transcription related activities. The data presented here newly implicate HlyU in destabilizing a transcriptionally repressive DNA cruciform thus contextually supporting access to a previously silenced genetic promoter. Considering that chromosomal DNA supercoiling and A/T rich DNA are common features of intergenic regions [370, 372, 375, 384, 385], we believe that cruciform attenuation, driven by specialized DNA binding proteins, may represent an overlooked extricating mechanism to de-repress gene expression.

3.5 Material and Methods

3.5.1 Bacterial Cultures and Growth Conditions

Vibrio parahaemolyticus RIMD2210633 was cultured in either LB-Miller (10g/L tryptone, 5g/L yeast extract, 10g/L NaCl) or LBS (10g/L tryptone, 5g/L yeast extract, 20g/L NaCl, 20mM Tris-HCl, pH 8.0). Antibiotic concentrations used for *V. parahaemolyticus* are as follows: chloramphenicol (2.5 μg/mL). *V. parahaemolyticus* was cultured at room temperature (22°C), 30°C, or 37°C, depending on a given experiment's requirements. *E. coli* strains were cultured in LB-Miller at 37°C containing the following antibiotics when necessary: ampicillin (100 μg/mL), chloramphenicol (30 μg/mL), and tetracycline (5 μg/mL).

3.5.2 Recombinant DNA Approaches

PCR and DNA cloning was performed using standard techniques. All DNA polymerases and restriction enzymes were purchased from New England Biolabs (NEB) unless stated otherwise. Control cloning and restriction digestion experiments were performed in parallel for interpretation purposes.

3.5.3 *in silico* Cruciform Analysis

Intergenic sequences known to contain HlyU binding sites by previous studies were selected from various *Vibrio* spp. and were used as input into Palindrome Analyser - an online bioinformatics tool which identifies cruciform forming DNA sequences and calculates the amount of energy required for cruciform formation [376]. Our cruciform identification required a cruciform stem of at least 6 base pairs, with a spacer/loop region of at least 10 bp. We allowed up to a single mismatch in the cruciform stem sequences. For each sequence, the 10 (or fewer) possible cruciforms requiring the smallest change in free energy to form are

detailed.

3.5.4 T7 Endonuclease and Cruciform Restriction Mapping Assays

HlyU binding sites within *Vibrio* spp. intergenic DNA regions were evaluated for cruciform structures by cloning PCR amplified DNA fragments or synthetic gBlocks (Integrated DNA Technologies (IDT), see Table B.6) into pBluescript and transforming the recombinant plasmid DNA into *E. coli* DH5α. Freshly prepared supercoiled plasmid DNA was then isolated from overnight cultures using a Monarch miniprep kit (NEB) and then immediately subjected to restriction digestion with T7 endonuclease (NEB) which cleaves at DNA cruciforms. To determine the approximate localization of DNA cruciform structures, a sequential digest with PvuII (following an initial T7 endonuclease digestion) was performed as previously described [377] followed by agarose gel electrophoresis to resolve digested DNA fragments.

To precisely detect cruciform cleavage sites, we designed a strategy using T7 endonuclease digestion or mung bean nuclease; the latter digests single stranded 'loop' DNA within DNA cruciforms [373]. Briefly, supercoiled plasmid DNA was treated with T7 endonuclease followed by reaction purification and treatment with mung bean nuclease to remove ssDNA overhangs and blunt the DNA to a T7 endonuclease cleavage site. Linearized dsDNA was then selectively extracted from an agarose gel, subjected to PvuII digestion (2 PvuII sites flank the DNA cruciform) yielding two blunt ended DNA fragments which were separately cloned into EcoRV treated pBluescript. A similar approach using only mung bean nuclease to detect ssDNA within cruciform 'loops' was also pursued. DNA sequencing of the resultant recombinant plasmids revealed DNA cruciform cleavage sites for either T7 endonuclease or mung bean nuclease as indicated by ligation to the EcoRV pBluescript site.

3.5.5 Transcriptional Reporter Assays Using *exsBA* DNA with Altered HlyU Binding Potential

Synthetic gene fragments (IDT) were designed with specific nucleotide deletions that were predicted to impact HlyU binding. The HlyU binding site is composed of two perfect inverted repeats (ATATTAG, CTAATAT) that flank a central (core) A/T rich palindromic sequence (TAATTTAATTTATT). Each of the altered *exsBA* DNA fragments along with a wild type *exsBA* fragment were separately cloned into a plasmid containing a promoter-less luxCDABE cassette to create transcriptional reporter constructs. The corresponding plasmids were incorporated into wild type *V. parahaemolyticus*. The bacteria were then cultured under T3SS-1 inducing conditions (magnesium supplementation and EGTA) [268] to induce *exsA* expression. Luciferase activity derived from the lux cassette was measured as light emission 2.5 hrs post induction using a Victor X5 luminometer as previously described [290].

3.5.6 *Vibrio parahaemolyticus* Chromosomal Mutant Generation

Δ*hlyU* and Δ*hns* null chromosomal mutants were generated using allelic exchange and sucrose selection as previously described [269]. Multiple DNA fragments (IDT, synthesized gBlocks) with specific nucleotide substitutions were designed to mutate HlyU binding site or cruciform forming locus within the *V. parahaemolyticus exsBA* intergenic region on the chromosome (see Table B.5, Table B.7 and Figure 3.3B). The DNA fragments were cloned into suicide plasmid pRE112 and used in allelic exchange experiments. All mutant strains were verified by DNA sequencing of PCR amplified chromosomal DNA to confirm genetic changes.

3.5.7 *in situ* Chromosomal Cruciform Modification by Chloroacetylaldehyde (CAA) Pulse-Chase Treatment

Cell permeable CAA was used to chemically modify in situ *V. parahaemolyticus* chromosomal DNA cruciforms by generating nucleotide base ethenoadducts at specific sites [379]. Specifically, CAA reacts with unpaired bases within DNA cruciforms resulting in localized DNA damage. Upon removal of CAA the damaged DNA can be assessed for effects on locus-specific gene expression. In this pulse-chase approach, 10^9 stationary phase cells from an overnight culture were harvested by centrifugation, washed in PBS (Phosphate Buffered Saline), and then treated with 40 nM CAA (or mock-treated with PBS) for 30 min. The cells were washed twice with PBS to remove CAA, and then immediately used to measure *exsA* promoter activity by an *in situ* real-time quantitative luciferase reporter assay. Cells within the population with CAA damaged DNA cruciform lesions were expected to emit less light than a paired mock-treated sample. We independently confirmed that a 30-min 40 nM CAA pulse treatment did not impair population-based cell growth metrics as determined by OD_{600} readings whereas higher concentrations were inhibitory and not pursued further (data not shown). Moreover, maintenance of CAA treatment during active growth conditions was not possible due to unpaired nucleotide bases associated with DNA replication events and various RNA species. Importantly, the assay is population based so cells that grow and repair CAA damaged DNA or suffer mutations are accounted for in the captured data.

3.5.8 *in vitro V. parahaemolyticus* T3SS-1 Protein Secretion Assay

T3SS-1 protein secretion assays were performed as previously described [268]. A characterized T3SS-1 defective strain Δ*vscN1* along with Δ*hlyU* [269] served as relevant controls allowing for secreted protein profile comparisons.

3.5.9 *V. parahaemolyticus* Host Cell Cytotoxicity Assays

HeLa cells (ATCC) were seeded in a 24-well dish at a density of 10^6 cells/mL and cultured for 16 hrs. Overnight cultures of *V. parahaemolyticus* strains were grown in LB-Miller at 37°C and diluted in DMEM (Dulbecco Modified Eagle's Medium, Invitrogen #11995) to generate a multiplicity of infection (MOI) of approximately 2. The cultured HeLa cells were rinsed twice with phosphate-buffered saline (pH 7.4, ThermoFisher; 10010023), followed by the addition of the relevant bacterial cells to initiate infection. The infection was incubated for 4 hrs at 37°C/5% CO2. Infection supernatants were collected and subjected to centrifugation (15000xg for 1 min) to remove bacteria and HeLa cells. The fluorescent CyQUANT™ LDH Cytotoxicity Assay (ThermoFisher; C20302) was used to measure lactate dehydrogenase (LDH) in the clarified supernatants as directed by the manufacturer. Percent cytotoxicity was calculated according to the following formula:

$$\%Cytotoxoicity = \frac{|Ab_{\text{infection}}| - |Ab_{\text{uninfected}}|}{|Ab_{\text{MaxRelease}}|} * 100$$

3.5.10 Recombinant HlyU-HIS Protein Purification

E. coli BL21(λDE3) containing a *V. parahaemolyticus hlyU*-HIS plasmid DNA construct [269] was cultured overnight. The following day, the bacteria were sub-cultured (1/50) and grown to an OD_{600} of 0.8 and then induced with 0.4 mM IPTG and cultured for an additional 3 hrs. The bacteria were harvested by centrifugation and the cell pellet was frozen at −20°C. Cell lysates were prepared, and nickel affinity chromatography was performed under soluble conditions as previously described [269]. Purified HlyU-HIS was subjected to Amicon ultrafiltration to enrich for dimeric HlyU (~22 kDa). Protein expression and purification steps were assessed by SDS-PAGE analysis (Figure B.5).

3.5.11 Electrophoretic Mobility Shift Assays

Electrophoretic mobility shift assays were performed as previously described [269]. Briefly, DNA oligonucleotides (sequences found in Table B.8) were mixed at equimolar concentrations and NaCl was added to a concentration of 50 mM to promote proper annealing of the oligonucleotides. The oligonucleotides were heated to 98°C followed by sequential cooling (1°C/5s) to 10°C thus generating annealed short dsDNA fragments.

A DNA master-mix in 1X EMSA buffer (1 mM Tris, 6 mM NaCl, 0.5 mM MgCl2, 0.01 mM EDTA, 0.1 mM CaCl2, 0.2% glycerol) was created for each dsDNA fragment and mixed with variable HlyU-HIS protein amounts or buffer alone to a final volume of 15µL. Reactions were allowed to equilibrate at room temperature for 30 min. Each reaction was subjected to TBE-PAGE for 1 hr at 100V at 4°C. Gels were stained with 1X SYBR Green (Invitrogen) for 30 min, rinsed in distilled H_2O, and visualized with the BioRad VersaDoc platform.

3.5.12 5' RACE

To determine mRNA transcriptional start sites in the *exsBA* intergenic region, a 5' RACE experiment was performed as previously described [386]. *Vibrio parahaemolyticus* RIMD2210633 was inoculated at a starting OD_{600} of 0.025 in Mg/EGTA containing LB and incubated at 30°C/250RPM for 3 hrs, prior to the isolation of total RNA. Total RNA was used for reverse transcription and PCR amplification. Primer AL400 (an *exsA* specific primer targeting the 3' end) was used for reverse transcription, and Gene Specific Primers (GSP) 1 and 2 were designed to target the *exsA* coding region of mRNA templates (Table B.8).

3.5.13 Reconstitution of *exsBA* Genetic Locus Transcriptional Activity in *E. coli*

Luciferase based reporter plasmids containing modified *exsBA* sequences were generated using standard cloning techniques (Table B.6) and were based on a verified *exsBA-luxCDABE*-pVSV105 transcriptional reporter plasmid [269]. pHlyU-FLAG which expresses C-terminal FLAG epitope tagged HlyU from a recombinant *tac* promoter was transformed into *E. coli* DH5αλpir. To generate a FLAG epitope tagged ExsA expression construct driven by the *lac* promoter, the *exsA* gene was PCR amplified from *V. parahaemolyticus* chromosomal DNA using oligonucleotides NT387 and NT388 followed by cloning into pFLAG-CTC. This plasmid served as template DNA in a PCR with primers NT472 and NT473 to generate *exsA*FL-pRK415 (for ExsA-FLAG expression). All plasmid constructs were verified by DNA sequencing. The expression plasmids or empty control plasmids were transformed into DH5αλpir in various combinations together with appropriate *exsBA* intergenic DNA luciferase reporter plasmids (Table B.6).

Luciferase assays were performed by inoculating fresh LB with OD_{600} normalized overnight *E. coli* cultures allowing cell growth to mid-log phase (~3 hrs). One hundred μL of culture was used to measure light emission measured as counts per second along with cell density readings (OD_{600}). Each culture was measured in triplicate and the experiment was repeated twice to attain statistical significance by multiple *t*-test. Graphs were plotted using R and the ggstatsplot package. An immunoblot was used to detect FLAG epitope tagged HlyU and ExsA from cell lysates. Anti-FLAG (Sigma) and anti-sigma70 (BioLegend) antibodies were used as primary antibodies, and goat anti-mouse HRP (Cell Signaling Technology) as secondary antibodies. Images were captured using a Bio-Rad VersaDoc system.

3.6 Acknowledgments

The authors would like to thank Ken Jarrell, Craig McCormick, John Rohde, John Archibald, and Andrew Roger for their insightful comments on the initial drafts of our manuscript. LJG is funded by a Vanier Canadian Graduate Scholarship, and a Killam PreDoctoral scholarship. LS and AC were funded by the National Science and Engineering Research Council (NSERC) Undergraduate Summer Research Award program. NAT holds an NSERC Discovery Grant-RGPIN 05807.

Conceptualization: LJG, JMB, NAT
Methodology: LJG, NAT
Validation: LJG, NAT
Formal analysis: LJG, LS, JMB, JM, AC, NAT
Investigation: LJG, JMB, LS, AC, JM, NAT
Writing – original draft preparation: LJG, NAT
Writing – review and editing: LJG, NAT
Visualization: LJG, LS
Supervision: NAT, LJG

Chapter 4

Exploring the Roles of Two Sugar Catabolism Regulators in *V. parahaemolyticus*

4.1 Introduction

Bacteria sense and integrate external environmental signals and restructure the cell to better survive and respond to the surrounding environment. This adaptation to environmental cues is vital to bacterial survival where nutrient sources are varied, and specialized proteins are required for nutrient transport and catabolism. Phosphoenolpyruvate:sugar phosphotransferase systems (PTS) are multicomponent sugar transport systems which contribute to this adaptation. The PTS is composed of three proteins, which can be encoded individually or fused: enzyme I (EI), histidine protein (HPr), and enzyme II (EII) [387]. EII is itself a multicomponent enzyme consisting of 3 subunits: EIIA, EIIB, and EIIC [388]. EII subunits can also be encoded individually or as a single protein. These proteins coordinate a phosphotransfer cascade which terminates in the phosphorylation of the sugar molecule as it is moved across the membrane. This phosphotransfer cascade is initiatied when phosphoenolpyruvate (PEP) is used to autophosphorylate EI, which in turn donates the phosphate to HPr. HPr then donates the phosphate to the EII protein, on subunit A (EIIA), and EIIA donates the phosphate to EII, subunit B (EIIB) [389]. EIIB, in coordination with EII, subunit C (EIIC), catalyzes the phosphorylation of a sugar molecule while simultaneously moving the molecule across the membrane [390].

While the PTS is important in sugar transport, these systems also act to sense

sugar signals and impact gene regulation. For example, the glucose PTS in *E. coli* is also used to monitor the presence and absence of glucose – the preferred carbon source – through the phosphorylation status of the EIIA enzyme [390, 391]. Here, $EIIA^{Glc}$ interacts with the membrane-bound enzyme adenlyate cyclase, only when $EIIA^{Glc}$ is phosphorylated in the absence of active transport of glucose by the Glucose PTS [392, 393]. Adenylate cyclase is then activated and produces the secondary messenger cyclic-AMP (cAMP) and drives catabolite repression [394–396]. Intriguingly, cAMP is linked to virulence in a number of bacterial pathogens [397–403].

ROK (repressor orf kinase) family transcriptional regulators contribute to the regulation of many PTSs [404]. These regulators contain an N-terminal helix-turn-helix (HTH) domain, necessary for DNA binding, a C-terminal cysteine-rich sugar-binding ROK motif, and a dimerization interface for the formation of homodimers [404, 405]. Protein crystal structures for a ROK family transcriptional regulator Mlc (**m**akes **l**arge **c**olonies) resolved 3 domains, where the HTH motif is found in domain 1 (aa 82-194), and the ROK motif as well as the homodimerization domain are found in domain 3 (aa 195-380) [405, 406]. Domain 2 and the HTH domain are involved in interaction with other protein partners, namely the $EIICB^{Glc}$ subunit of the glucose PTS [407]. The sequence specificity of ROK family transcriptional regulators is determined primarily by the amino-acid sequence in the HTH domain, however a linker region just downstream of the HTH domain also appears important for sequence recognition [406].

Mlc represses several PTS genes, including the glucose PTS genes *ptsG* and the *ptsHIcrr* operon, which constitute the $EIIBC^{Glc}$ subunit, HPr, EI enzyme, and soluble $EIIA^{Glc}$ subunit, respectively [408–412]. As well, Mlc represses *malT* and *manXYZ*, the regulator and components of the EII enzyme of the mannose PTS [413–415]. Mlc was first discovered in 1994, identified as a ROK family

protein, and was named after a phenotype observed on solid media: when *mlc* is expressed on a plasmid, it protected *E. coli* from acetate production and therefore made larger colonies [416]. Mlc repression is alleviated when glucose is present through the physical sequestering of Mlc to the membrane through interaction with dephosphorylated PtsG (also known as EIIBC^{Glc}) [409, 417]. When PtsG is donating its phosphoryl-group to glucose, it is no longer phosphorylated, and can sequester Mlc. In turn, when PtsG is inactive, it is in a phosphorylated state and cannot sequester Mlc allowing it to bind its DNA targets [417].

NagC is another ROK family transcriptional regulator involved in repression of the PTS responsible for the catabolism and transport of amino sugars, including N-acetylglucosamine. NagC represses the *nagE-nagBACD* operon in *E. coli* where NagE encodes the GlcNAc EIIBCA PTS components, and NagB and NagA are involved in GlcNAc-6-P conversion to GlcN-6-P and GlcN-6-P to fructose-6-phosphate, respectively [418–421]. NagC also activates the *glmUS* operon whose protein products drive biosynthesis of glucosamine and UDP-GlcNAc, respectively [422]. Importantly, UDP-GlcNAc is a precursor for a variety of cell wall components including peptidoglycan and the O-antigen, including lipopolysaccharide (LPS) [423, 424]. Unlike Mlc, DNA binding and therefore repression or activation by NagC can be relieved by allosteric regulation mediated by GlcNAc-6-phosphate, in both *E. coli* and *V. fischeri* [420, 421].

NagC and Mlc have similar functions in *E. coli* and some *Vibrio* species [425–427]. For example, the Mlc homolog in *V. cholerae* (VC2007) can restore repression of *ptsG* when expressed in *E. coli mlc* null mutants [425]. However, VC2007 is not sequestered to the membrane by the dephosphorylated glucose PTS IIBC^{Glc} subunit, but rather by dephosphorylated NagE, the EIIBCA component of the GlcNAc PTS [425]. VC2007 repression can also be partially alleviated by binding to GlcNAc-6-P, which is also not characteristic of *E. coli* Mlc [425].

While the roles of NagC and Mlc for sugar catabolism regulation are clear,

NagC and Mlc have also been implicated in regulation of a number of other cellular processes. In *V. cholerae*, *mlc* null mutants display disrupted biofilm formation compared to wildtype when grown in LB, which can be restored by complementation with either *V. cholerae* or *E. coli* Mlc [426]. In *V. parahaemolyticus* NagC is involved in biofilm formation, as well as capsule production, with a translucent colony morphology noted in a *nagC* mutant [423]. NagC represses GlcNAc catabolism genes in *V. fischeri* and *V. cholerae*, and facilitates colonization of the *Euprymna scolopes* (Bobtail Squid) light organ in *V. fischeri* [421, 427, 428]. While *nagC* null mutants in *V. fischeri* had reduced fitness in the presence of GlcNAc or $(GlcNAc)_2$, they had no defect in sensing or swimming towards GlcNAc [427]. Finally, regulation of T6SSs by various sugar molecules has been described in *V. fluvialis* and Mlc binding sites have been predicted in the promoter regions of T6SS-related genes in *Klebsiella pneumoniae* through a bioinformatic survey [429, 430]. However, it is currently not clear if these sugar catabolism regulators contribute to swimming or prey-killing via the T6SS in *V. parahaemolyticus*.

Mlc has also been implicated in virulence regulation. Mlc indirectly regulates genes found in the Salmonella pathogenicity island-1 (SPI-1) of *Salmonella enterica* serovar Typhimurium [431]. In a complicated transcriptional cascade, HilD de-represses the AraC/XylS transcriptional activator HilA, which activates expression of T3SS genes involved in cellular invasion and proliferation [432–436]. Mlc represses HilE, a protein which interacts with HilD and prevents HilD activation of SPI-1 genes [431]. In another virulence related example, T3SS expression in *P. aeruginosa* is activated by ExsA, where *exsA* itself is activated at the genetic level, in part, by Vfr [403, 437]. Vfr is a cyclic-nucleotide binding protein that binds to cAMP, which in turn allows it to interact with DNA and regulate the expression of *exsA* [403]. Considering sugar transport, specifically by the PTS, is linked to cellular levels of cAMP [390], it is likely that sugar catabolism

regulators like Mlc might impact T3SS expression in *V. parahaemolyticus* as well, although this has not been shown directly.

Altogether, little research has been conducted on NagC or Mlc in *V. parahaemolyticus*, with the work mostly limited to capsule and biofilm formation, or other phenotypes in two related *Vibrio* spp., *V. fischeri* and *V. cholerae* [421, 423, 426, 427]. RegPrecise – an online, manually curated transcriptional regulator database – identifies the NagC regulon in *Vibrio parahaemolyticus* to include *glmU*, *nagE1*, *ptsG*, *nagA*, *nagC*, *tfoX*, *chi*, *ptsHI*, *crr*, *chiA*, and *nagB*, among others [438]. These genes are involved in chitin degradation, GlcNAc transport across the membrane, GlcNAc catabolism, and natural competence. Importantly, the role of NagC identified by RegPrecise is inferred based on binding sites identified, and many of these regulatory roles have not been experimentally validated.

The purpose of this study was to characterize the roles of NagC and Mlc in biofilm formation, chitin utilization, prey-killing, swimming motility, and virulence gene expression using phenotypes in other species as a guide. I began this work by comparing the NagC and Mlc proteins from *E. coli* K12 with their homologues in *V. parahaemolyticus*, and generated *nagC* and *mlc* genetic knockouts, as well as their genetic complementation constructs. I performed a number of phenotypic assays designed to explore null *nagC* and *mlc* mutants in the aforementioned functions. I found that NagC is involved in biofilm formation and chitin utilization in *V. parahaemolyticus*, and identified an intriguing link between NagC and T3SS-1 expression.

4.2 Results

4.2.1 NagC and Mlc Sequences Comparisons with *E. coli*

Genes encoding homologues of the *E. coli* NagC and Mlc proteins can be found in *V. parahaemolyticus* and *V. cholerae* genomes. To better understand these proteins in *V. parahaemolyticus*, the predicted amino acid sequences of both *mlc* and *nagC* were compared between *E. coli* K12, *V. parahaemolyticus* RIMD2210633, and *V. cholerae* El Tor N16961. *V. cholerae* was included as previous literature on *V. cholerae mlc* can be used to compare more closely the *Vibrio* genus functions of these proteins with their *E. coli* homologues [425].

As expected, the *Vibrio* genus NagC (VC0993 and VP0828) and Mlc (VC2007 and VP2038) protein sequences are quite similar, with 83% identity and 89% similarity between NagC proteins and 88% identity and 94% similarity between Mlc proteins when compared via Needleman-Wunsch Global Align Protein Sequences using BLAST [439, 440]. *V. parahaemolyticus* NagC shared 60% identity and 64% similarity with the *E. coli* K12 NagC, and 46% identity and 70% similarity with *E. coli* Mlc. This is comparable to *V. cholerae* and *E. coli* with 60% identity and 74% similarity between NagC proteins, and 45% identity and 69% similarity between Mlc proteins. A protein sequence alignment was performed using Clustal Omega, and visualized using JalView (Figure 4.1) [441, 442]. T

Highly conserved amino acids found in the ROK motif of ROK family proteins include a conserved cysteine-rich region which coordinates a Zn^{2+} ion, corresponding to H247, C257, C259, and C264 in crystal structures of *E. coli* Mlc [405].

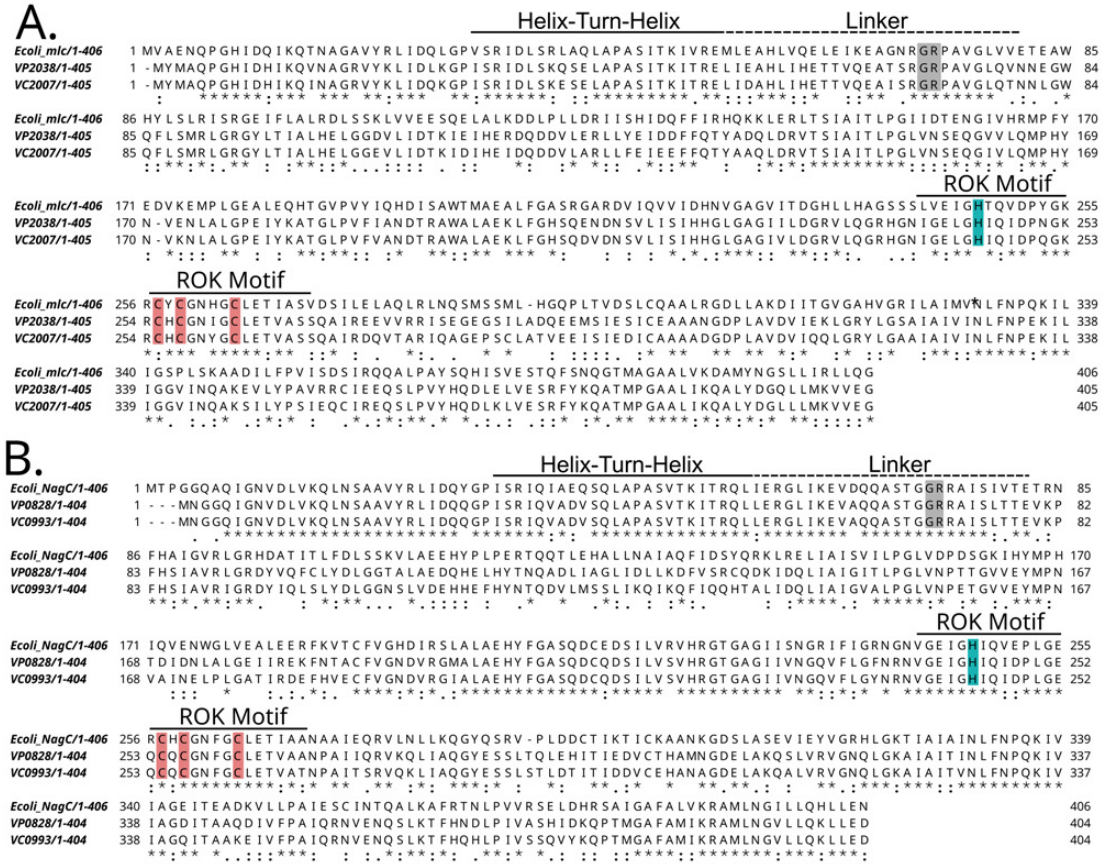

Figure 4.1: Comparison of *E. coli* K12, *V. parahaemolyticus* RIMD2210633, and *V. cholerae* El Tor N16961 Mlc and NagC Protein Sequences. Clustal Omega protein sequence alignments were viewed and exported on JalView [441, 442]. Labels for the helix-turn-helix motif, linker sequence, and ROK motif are presented. The grey box indicates an essential amino acid pair for DNA-binding [443]. Blue and red boxes indicate conserved ROK family H and C residues important for Zn^{2+} binding [405]. Symbols indicating degree of residue conservation are found below the sequence, and are as follows: * - fully conserved, : - conservation between groups with strongly similar properties, . - conservation between groups with weakly similar properties. A. Comparison of Mlc protein sequences. B. Comparison of NagC protein sequences.

These amino acids are also conserved in both *Vibrio* spp. Mlc and NagC proteins, and they are identified by the blue and red boxes in Figure 4.1. Further, conserved GR residues in the linker-region has been implicated in DNA-binding, which is also conserved among all of the protein sequences compared, indicated by the grey box in Figure 4.1. Overall, the similarity between *E. coli* K12 and *Vibrio* spp. NagC and Mlc protein sequences identify that the *Vibrio* spp. proteins likely have similar functions. Indeed, previous work has shown the *Vibrio cholerae mlc* can complement *E. coli mlc* null mutants [425].

4.2.2 NagC and Mlc are Not Responsible For Regulation of Type VI Secretion or Swimming Motility

To better understand the function of *nagC* and *mlc* in *V. parahaemolyticus*, gene deletions were generated in both *nagC* and *mlc*, and complemented strains were developed. First, the expression of the FLAG-tag encoding *nagC*FL and *mlc*FL complementation constructs were verified in *V. parahaemolyticus* by immunoblotting (Figure 4.2A). Growth was comparable to wildtype for all strains except the *mlc*FL complemented strain, which carried the *mlc* gene deletion and a copy of the *mlc*FL gene on a plasmid (Figure 4.2B). The Δ*mlc* complemented strain did reach a similar, albeit slightly reduced, final OD_{600} reading as wildtype, however the slope of the growth curve for the complemented Δ*mlc* strain was less steep than those for the other strains, taking much longer to reach the final OD_{600} reading. This observation indicates a growth defect in the logarithmic phase of growth (see Discussion).

No effect on swimming motility was noted for either of the null mutants, or the complemented *nagC* null mutant after inoculation of 0.2% agar LBS plates (Figure 4.2C), similar to previous results in *V. fischeri* [427]. However, the complemented *mlc* mutant shows reduced swimming motility compared to wildtype, which may be related to its growth defect (Figures 4.2C and B, see Discussion).

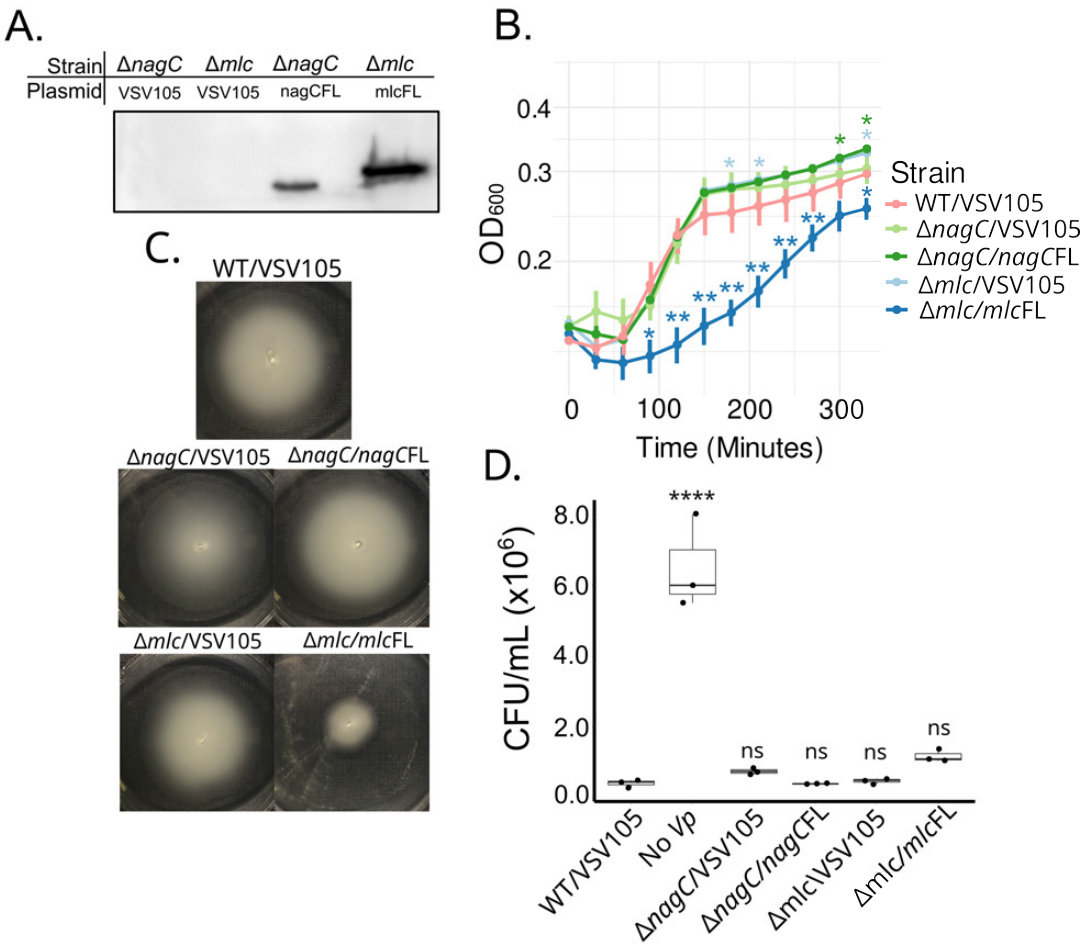

Figure 4.2: Growth, swimming, and prey-killing phenotypes for null mutants of *nagC* and *mlc*. A. An immunoblot was performed on cell lysates of the relevant complemented null mutant strains, as well as empty vector controls, using anti-FLAG antibodies. The immunoblot was visualized on the BioRad VersaDoc platform using BioRad ECL Western blot substrates. The strains and plasmids are labelled. B. A growth curve in LB media as measured by OD_{600} values, where the y-axis is transformed by $\log_{10}$. Error bars indicate standard deviation from mean values. *t*-tests were performed at each timepoint to determine statistical significance compared to wildtype, and statistical indicators are coloured according to the strain legend (n=3, *: $p<0.05$, **: $p<0.01$). C. Swimming assays were performed using 0.2% (w/v) agar LBS plates. Overnight cultures were normalized for OD_{600} followed by inoculation into a 0.2% (w/v) agar LBS plates. The experiment was performed 3 times and photos of representative experiments are shown. D. A prey-killing assay was performed using *E. coli* prey, and various *V. parahaemolyticus* predator strains. Cells were collected following 4-hr contact between predator and prey cells, and the prey cells were enumerated. DH5αλpir incubated without predator cells was used as a negative control. Statistical significance was obtained using an unpaired *t*-test for each strain compared to wildtype (n=3, ****: $p=<0.001$).

To determine if the *nagC* or *mlc* null mutants had deficiencies in *E. coli* prey-killing, the mutant and complemented strains were used for prey-killing assays on LBS at 30°C [233]. In brief, *V. parahaemolyticus* and *E. coli* DH5αλpir cells were mixed at a 4:1 ratio, plated on LBS agar plates, and incubated for 4 hrs. Following incubation, cells were selected on appropriate plates to determine colony forming units of the prey *E. coli* cells. If the T6SS is active, *E. coli* colony forming units should be significantly reduced, as shown previously by Salomon *et al.* [233]. As expected, *E. coli* cell counts were reduced after exposure to wildtype *V. parahaemolyticus*, while *E. coli* incubated alone form colonies at a significantly higher rate (Figure 4.2D). No significant difference was noted between wildtype *V. parahaemolyticus* and either the *nagC* or *mlc* null mutants, indicating that *nagC* and *mlc* null mutants have no significant defect in *E. coli* killing (Figure 4.2D). While this experiment is analogous to others done in *V. parahaemolyticus* which show definitive T6SS-dependent *E. coli* killing, the assay presented here lacks T6SS specific mutant controls, which limits the interpretation [233].

4.2.3 NagC is Responsible For Regulating Chitin Utilization, Biofilm Formation, and T3SS-1 Activity

NagC is responsible for aspects of biofilm formation and for an opaque colony phenotype in *V. parahaemolyticus* [423], as well as for chitin utilization in other *Vibrio* spp. [76, 427]. As such, the roles of *nagC* and *mlc* in biofilm formation and chitin utilization were explored. Indeed, a *nagC* null mutant was unable to generate biofilms comparable to wildtype, and can be complemented by the re-introduction of *nagC* on a plasmid (Figure 4.3A). The *mlc* mutant does not appear to have a defect in forming biofilms on glass tubes, however reintroduction of *mlc in trans*, slightly increased biofilm accumulation on glass tubes (Figure 4.1A). Conversely, in *V. cholerae*, Mlc regulates biofilm formation V. cholerae in similar growth conditions [426].

nagC null mutants appear de-repressed for chitinase secretion, as noted by the larger zone-of-clearing (ZOC) on chitin containing MM9 agar plates compared to wildtype (Figure 4.3B). This appears to be *nagC* specific, as the complemented strain (Δ*nagC/nagC*FL) resembles the ZOC seen in wildtype and apperas to restore the wildtype phenotype (Figure 4.3B). This is consistent with previous literature that NagC is a repressor of *chiA*, an important chitinase in *V. parahaemolyticus* [421]. *mlc* mutants did not appear to have an appreciable effect on chitinase secretion (Figure 4.3B).

In *P. aeruginosa*, an analogous T3SS to the T3SS-1 of *V. parahaemolyticus* is activated by the expression of ExsA, in which *exsA* is under the control of a cAMP-binding protein called Vfr [403, 437]. Sugar transport, specifically by the PTS, is linked to cellular levels of cAMP [390]. As well, Mlc is linked to *hilE* repression in *S. enterica* serovar Typhimurium, which in turn promotes T3SS activation. Therefore, I considered the possibility that NagC or Mlc in *V. parahaemolyticus* might regulate *exsA* and T3SS-1 expression as well. Indeed, when the *nagC* null mutant is generated in the *tdh*::*exsBA-luxCDABE V. parahaemolyticus* parent strain, *exsA* expression was reduced at 2 hours compared to wildtype in Mg^{2+}/EGTA containing media (Figure 2.3C). However, at 2.5 hours, this reduction was no longer present. Complementation of *nagC in trans* restored *exsA* expression to slightly above wildtype levels at the 2 hour timepoint (Figure 2.3C). Importantly, the difference between the Δ*nagC* and Δ*nagC*/*nagC*-FL complemented mutants was statistically significant at 2 hours, but not at 2.5 hours (Figure 2.3C). Additionally, *mlc* null mutants behave in opposition to Δ*nagC* displaying similar *exsA* expression to wildtype at 2 hours, and increased exsA expression at 2.5 hours (Figure 2.3C). Expression of *mlc in trans*, however, almost completely abrogated *exsA* expression, similar to the *hlyU*::Tn strain (Figure 2.3C). It is unclear why *mlc* complementation would influence *exsA* expression, considering no difference was observed in the *mlc* null background

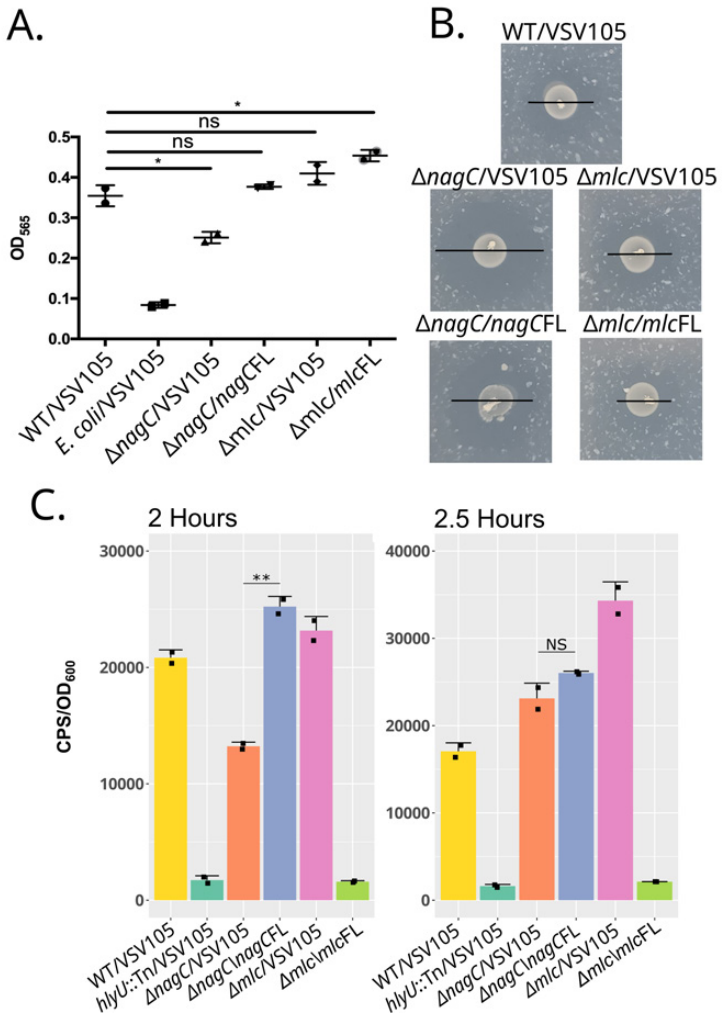

Figure 4.3: Biofilm formation, chitin utilization, and *exsA* expression phenotypes for the *nagC* and *mlc* null mutants. A. Biofilms were generated in glass tubes by inoculating *V. parahaemolyticus* strains or an *E. coli* control to an initial OD_{600} of 0.025 into LB. Cells were incubated at 37°C/200RPM for 16 hrs. Following incubation, biofilms were stained with crystal violet, washed, and destained with 95% ethanol. Measurements were taken at OD_{565} and statistical significance measured by unpaired *t*-test (n=2, *: $p<0.05$). B. Chitinase secretion and activity was assayed by zone-of-clearing. Overnight cultures were normalized by OD_{600}, and spotted onto MM9 agar plates containing colloidal chitin. Photos were taken after 7 days incubating at 30°C. Black lines indicate the circumference of the zone-of-clearing for each strain. C. *hlyU*::Tn, as well as *nagC* and *mlc* null and their complemented strains in the *tdh*::*exsA-luxCDABE* background were subjected to a luciferase assay in Mg^{2+}/EGTA containing LB and luminometry and cell density were measured at 2 and 2.5 hrs. Both CPS (counts-per-second) and OD_{600} were measured to allow for cell number normalization. Statistical significance was determined between the *nagC* null and *nagC* complemented strains using an unpaired *t*-test (n=2, **:$p<0.01$).

(see Discussion).

4.2.4 Deletion of *nagC* or *mlc* Does Not Affect Cytotoxicity Against Two Human Cell Lines

Provided the Δ*nagC* strains reduction in *exsA* activity was only observed at early expression timepoints, I characterized the ability of both *nagC* and *mlc* null mutants to generate functional T3SS-1 needles and drive cytotoxicity against human cells. To accomplish this, I used well established human cell cytotoxicity assays against two standard human cell lines: HeLa - a human epithelial cell line, HT29 - a human colonic cancer cell line (Figure 4.4). Cytotoxicity against both cell lines was observed, as expected, for the wildtype positive control. The Δ*hlyU* stain was used as a negative control as it is unable to express *exsA* and therefore does not cause cytotoxicity as observed (Figure 4.4) [269]. The Δ*nagC*, complemented Δ*nagC*, and the Δ*mlc* mutants showed no significant defect in cytoxicity when compared to wildtype for HeLa cell infection or HT29 cell infection (Figure 4.4). However, the complemented Δ*mlc* strain had reduced cytotoxicity against HeLa and HT29 cells (Figure 4.4; see Discussion).

The lack of significant differences in cytotoxicity for the null mutants of *nagC* and *mlc* are intriguing, given the previous finding that *exsA* expression is reduced in the *nagC* null mutant. This discrepancy may be explained by the feed-forward nature of *exsA*. Once enough *exsA* is expressed in the cell, it can auto-regulate its own promoter and drive sufficient *exsA* expression for T3SS-1 expression by the end of the 4 hour infection assay (see Discussion).

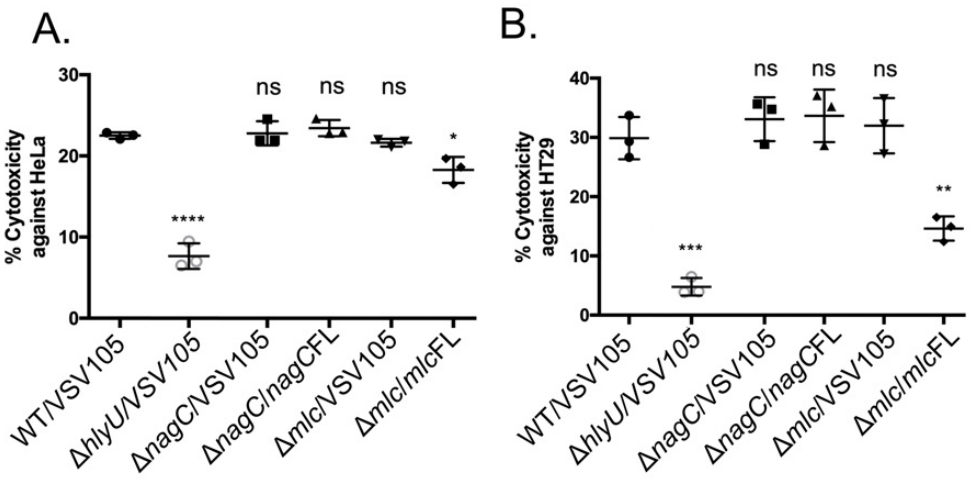

Figure 4.4: Deletion of *nagC* or *mlc* in *V. parahaemolyticus* does not affect T3SS-1 mediated cytotoxicity. *V. parahaemolyticus* strains were used to infect HeLa (A.) or HT29 (B.) human cell lines at an MOI of approximately 5. All strains carried pVSV105 or a complementation construct in the same background plasmid vector as labelled. Infection progressed for 4 hrs prior to the measurement of released LDH to measure host-cell cytoxicity. Δ*hlyU* was used as a negative control. Statistical significance was determined between wild-type and the null/complemented strains using an unpaired *t*-test. (n=3, *: $p<0.05$, **: $p<0.01$, ***: $p<0.001$, ****: $p<0.0001$)

4.2.5 *crp* Deletion Mutants Phenocopy the Δ*nagC* Mutants for *exsA* Expression

Given that the Δ*nagC* mutant appears to have an *exsA* expression defect, and that *nagC* is involved in the expression of the GlcNAc-specific PTS, I wanted to more directly explore the role of PTS-regulated secondary-messenger cAMP on *exsA* expression. To accomplish this, I generated a *crp* gene deletion in the *tdh*::*exsA-luxCDABE* background, as well as a complemented strain, and subjected them to growth and *exsA* expression assays. A growth curve comparing the growth of the *hlyU*::Tn, Δ*crp*, and Δ*crp*/*crp* strains to the wildtype background is shown in Figure 4.5A. Notably, the *crp* null mutant failed to reach the same final OD_{600} values as the wildtype background, and cell growth can be restored by supplying *crp in trans*. This is perhaps not unexpected as Δ*crp* mutants have been shown to have pleitropic effects on virulence, growth, capsule production, protease and hemolysin expression, and mouse lethality in *V. vulnificus* [444].

To determine the effects of the Δ*crp* mutant on *exsA* expression, a luciferase assay was performed using the chromosomal *tdh*::*exsA-luxCDABE* strains in *exsA* inducing conditions. CPS and OD_{600} was measured every 30 minutes for 270 minutes for the wildtype and *hlyU*::Tn mutant – serving as positive and negative controls respectively – as well as the Δ*crp* and Δ*crp*/*crp* complemented strains (Figure 4.5B). A sharp jump in *exsA* expression was identified beginning at around 30 minutes for the wildtype strain, which eventually levels out after 150 minutes as expected [269]. The *hlyU*::Tn mutant is significantly defective for *exsA* expression as previously identified (Figure 4.5C). The Δ*crp* mutant appeared to have an intermediate phenotype, even after correction of growth deficits by OD_{600} values, and was significantly reduced compared to wildtype (Figure 4.5B). Complementation with *crp in trans* restored *exsA* expression nearly to wildtype,

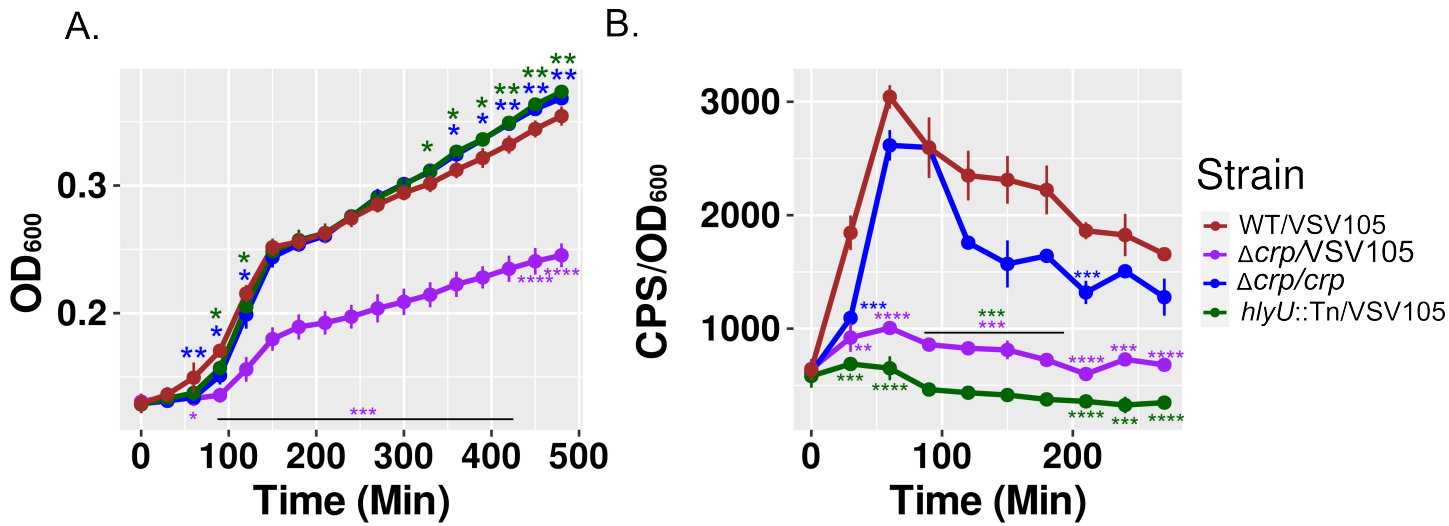

Figure 4.5: *crp* deletion mutants exhibit reduced *exsA* activity A. A growth curve in LB media was performed for various strains and wildtype in the *tdh*::*exsA-luxCDABE* backgrounds. Cells were normalized to an OD_{600} of 0.025, inoculated into triplicate wells of a 96-well plate, and incubated at 37°C in the VictorX5 Multilabel plate reader for 480 min. 96-well plate lids were treated to prevent condensation build-up prior to inoculation. Statistics were performed using t-tests against wildtype at each timepoint (n=3, *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$; ****: $p < 0.0001$). Statistical indicators are coloured as in the legend. Error bars indicate standard deviation from the mean values. B. Luciferase assays were performed in Mg^{2+} and EGTA containing LB using the same strains as A. Cells were normalized to a starting OD of 0.025, and pipetted into 96-well plates with treated lids to prevent condensation build-up. Cells were incubated at 30°C in the Victor X5 Multilabel plate reader and luminometry and cell density measurements were performed every 30 min for 270 min. Statistics were performed as in A. (n=3)

with most timepoints lacking statistically significant differences in CPS/OD_{600} values compared to wildtype (Figure 4.5B). Taking together, it appears that expression of *exsA* is dependent on *crp*.

We have previously shown that HlyU acts in a kick-start mechanism to push the expression of *exsA* and generate ExsA, which in turn auto-activates itself [269, 291]. Given the early impact of Δ*nagC* on *exsA* expression, it is reasonable to consider that this phenotype may have to do with successful expression of HlyU and activation of the auto-regulation feedback loop. To investigate this, the region distal to the *hlyU* open-reading frame, containing the *hlyU* promoter, was searched for the CRP consensus binding site from *E. coli* ($TGTGAN_6TCACA$). Indeed, an almost perfect consensus sequence for CRP was identified 325bp upstream of the HlyU gene in the C-terminal coding sequence of *nhaR* (Figure 4.6A). The sequence identified (TGTGATGCTGATTTCA) varies from the consensus site by 2bp (Figure 4.6A).

To further link CRP to *hlyU* expression, a his-tagged CRP construct was generated in the pET21 vector, transformed into BL21(DE3) *E. coli*, and expressed using IPTG induction. The successful induction of an apparent 24.5kDa protein is seen in BL21(DE3) *E. coli* after a 2 hour incubation in LB containing IPTG (Figure 4.6B, lane 2) when compared to cells before induction with IPTG (Figure 4.6B, lane 1). Subsequent cell collection and lysis (Figure 4.6B, lane 3) showed a protein band corresponding to the same protein with an apparent molecular weight of roughly 24.5 kDa. This is consistent with the predicted molecular weight of CRP-HIS based on the AA sequence encoded by the *crp* gene. This protein was no longer present after column binding (Figure 4.6B, lane 4). Subsequent wash steps appeared to elute non-specific protein products (Figure 4.6B, lanes 5-7), and finally elution of the apparent 24.5kDa protein, as seen in the induction in lanes 2 and 3, is shown in Figure 4.6B, lane 8.

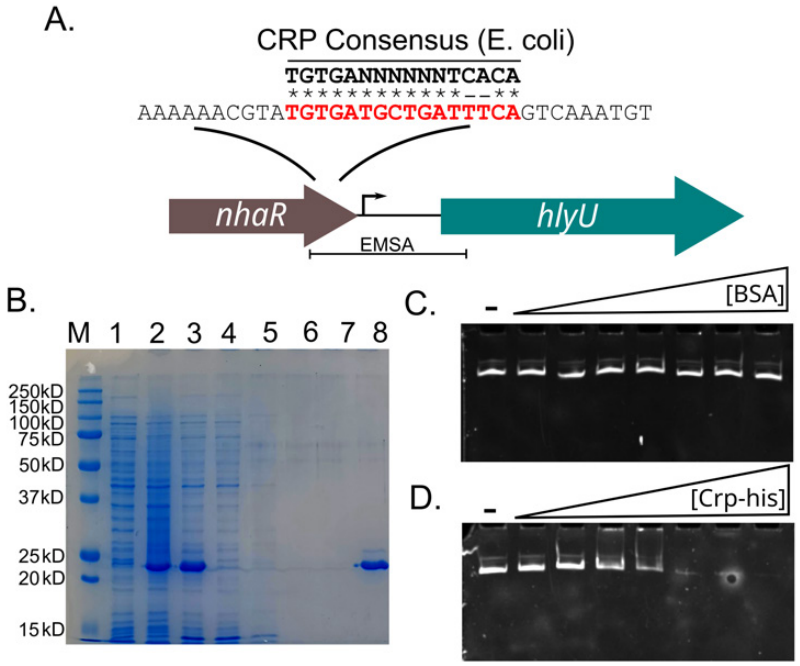

Figure 4.6: CRP-HIS EMSA demonstrates DNA-binding to the *hlyU* promoter. A. The consensus binding site for CRP from *E. coli* was identified upstream of *hlyU* in *V. parahaemolyticus*. The DNA sequence within the C-terminus coding region of *nhaR* is shown and the *E. coli* CRP consensus sequence is indicated above in bold. The sequence provided shows the consensus CRP site in the *V. parahaemolyticus* genome in red, and identical nucleotides are labelled with asterisks, while non-identical nucleotides are indicated with dashes. The EMSA fragment generated for CRP-HIS DNA-binding experiments in C. and D. is depicted schematically below. B. IPTG Induction and purification of the HIS-tagged CRP protein expressed in BL21(DE3) *E. coli* was performed and samples were subjected to SDS-PAGE, followed by Coomassie staining. The lanes contain the following samples: Marker (M), pre-IPTG induction *E. coli* lysate (1), 2 hr post-IPTG induction *E. coli* lysate (2), post-induction *E. coli* lysate before column treatment (3), lysate column-flowthrough (4), column wash steps (5-7), and CRP-HIS elution (8). The molecular weight of marker proteins is also indicated. *hlyU* promoter DNA was subjected to EMSA through an 8% TBE-polyacrylamide gel with increasing concentration of either BSA (C.) or CRP-HIS (D.). The first lane in each gel contains a DNA-only control and is indicated by a dash. Both gels were visualized using ethidium bromide staining and photos were taken on the BioRad VersaDoc platform.

The purified CRP-HIS protein product was used to perform an EMSA on a 403bp PCR product obtained from the promoter region of *hlyU*, which includes the putative CRP binding site (Figure 4.6A). EMSA performed using a non-specific protein BSA, in the presence of cAMP, generated no shift in the DNA product is shown (Figure 4.6C). However, EMSA performed using CRP-HIS, in the presence of cAMP, generated a shift in the *hlyU* promoter DNA fragment with increasing CRP-HIS concentrations (Figure 4.6D). This indicates that CRP-HIS binds to *hlyU* promoter DNA (see Discussion).

4.3 Discussion

NagC and Mlc are important regulators of sugar transport via the PTS, as well as sugar catabolism in a variety of bacterial species. Here, we show that NagC and Mlc have important functions in *V. parahaemolyticus* and share similar functions to NagC and Mlc proteins in other *Vibrio* spp., namely *V. cholerae* and *V. fischeri* [421, 423, 426, 427, 445]. We confirmed that NagC is involved in biofilm formation and chitin utilization in *Vibrio parahaemolyticus*, as expected [423]. Interestingly, *mlc* null mutants lacks biofilm defects like those seen in *Vibrio cholerae*, although the increase in biofilm formation noted upon complementation was replicated here (Figure 4.3B) [426]. Mlc regulation of *Vibrio* polysaccharide (VPS) in *V. cholerae* occurs through repression of the EI component of the PTS as EI is itself responsible for repressing the *vps* genes [445]. Two distinct possibilities arise for why *mlc* gene deletions in *Vibrio parahaemolyticus* might not disrupt biofilm formation: 1) Mlc does not regulate the same EI PTS components in *V. parahaemolyticus* as in *V. cholerae*, or 2) the EI PTS components which repress *vps* in *V. cholerae* do not repress *vps* in *V. parahaemolyticus*. This will need to be explored in future work. Further, neither NagC or Mlc appear to be involved in prey-killing in *V. parahaemolyticus* and their deletion does not effect growth in

rich media.

However, complementation with *mlc in trans* appears to cause a significant growth defect, either due to some inherent toxicity of the *mlc* protein when expressed exogenously, or through a natural growth-inhibition mechanism. Mlc is responsible for repressing a number of genes involved in glucose transport (*ptsG* and *ptsHIcrr* for example) in *E. coli* and Mlc repression is relieved by interaction with NagE (a GlcNAc-specific PTS subunit), as well as interaction with GlcNAc-6-P in *V. cholerae* [411, 425]. Given that *ptsG* and *ptsHIcrr* are likely not necessary for growth in LB, and GlcNAc-6-P is not present in our LB growth conditions, it is possible that overexpression of Mlc is causing strong repression of other genes which are necessary for growth, or activation of genes which generate a growth disadvantage. Indeed, based on the Western blot analysis of the complemented strains FLAG-tagged proteins, it appears that Mlc-FLAG may be overexpressed compared to the NagC-FLAG protein, although true loading controls for the western blot are required to confirm this in *V. parahaemolyticus*. Perhaps more pertinently, it is unclear if there are copy number effects from the complementation on a plasmid, and if expressing Mlc-FLAG would have a similar effect on cell growth in the wildtype background. This growth defect will need to be explored further in future studies.

Of particular interest to us, the *nagC* null mutant appears to have a defect in early *exsA* expression, while the *mlc* mutant appears to have enhanced *exsA* expression at both the 2 and 2.5 hour timepoints (Figure 4.3C). While the expression of *exsA* can be restored in the *nagC* null complemented mutant, *mlc* complementation completely abrogates the expression of *exsA*. The complemented *mlc* mutant does have a well-documented growth defect in our hands, and therefore the *mlc* complementation may impact *exsA* expression through a similar, unknown fashion. However, *exsA* expression changes in the Δ*nagC* and Δ*mlc* mutants do not have an impact on the function of T3SS-1, as infection assays

against both human epithelial cell lines (HeLa) and human colonic cell lines (HT29) generate cytotoxicity in both mutants (Figure 4.4). Therefore it appears that while *exsA* expression is reduced in the Δ*nagC* mutant, enough ExsA is expressed to allow for T3SS-1 expression, protein secretion, and subsequent cytotoxicity.

NagC and Mlc both regulate the expression of PTS components. Changes in the homeostasis of PTS components in the cell can, in turn, impact cAMP concentration in the cell. In *E. coli*, the production of cAMP drives catabolite repression, where cAMP will bind to cyclic-AMP receptor protein (CRP) and through this binding, CRP can interact with DNA at sites where it can either enhance expression of genes (through interaction with and recruitment of RNA polymerase) or repress genes. Virulence in a number of bacterial species is under the control of cAMP, including *Vibrio vulnificus*, *Vibrio cholerae*, *Salmonella enterica* serovars Abortusovis and Typhimurium, *Yersinia enterolitica*, and *Pseudomonas aeruginosa* [397–403, 431]. Specifically, previous work that indicts cAMP in the regulation of T3SSs is found in another pathogen, *P. aeruginosa* [403], which provided some rationale for the following: that modified cAMP homeostasis in the *nagC* and *mlc* mutants is driving the change in *exsA* expression seen in these mutants through cAMP receptor protein. Indeed in a *crp* null mutant, *exsA* expression is reduced, even when corrected for cell growth. Indeed, *crp* null mutants can be complemented, indicating that wildtype *exsA* expression is dependent on *crp*.

Given the opposing impacts on *exsA* expression in the *nagC* and *mlc* null mutants, it is possible the the differential expression of various PTS system components, by either NagC or Mlc, are driving differential changes to cAMP homeostasis. This regulatory system requires significant additional investigation. Most importantly, the regulon for both NagC and Mlc need to be characterized with experimental data in order to understand how cAMP homeostasis might be

impacted. Characterization of the NagC and Mlc regulon, perhaps by RNA-seq, ChIP-Seq, or a targeted qPCR approach, would show how expression of the various PTS systems are modified in these mutants (See Chapter 6 Discussion). As well, to provide further evidence of the role of cAMP during *exsA* expression, intracellular concentrations of cAMP in these mutants during *exsA* expression conditions should be characterized.

Finally, the delayed start of *exsA* expression in both the *nagC* mutant implicate HlyU as the target of cAMP-CRP regulation and *exsA* expression. I have previously identified that HlyU acts as a kick-start protein to drive early expression of *exsA* [269, 291]. I considered the possibility that *hlyU* expression could be under the control of cAMP-CRP, and reduced expression of *hlyU* early in *exsA* induction is driving the delay in *exsA* expression in the Δ*nagC* mutant. Indeed, a near perfect CRP consensus site was identified in the *hlyU* promoter region, and CRP appears to bind to *hlyU* promoter DNA. Our data identify that not only does a Δ*crp* mutant have reduced *exsA* expression, as well as a delayed start to *exsA* transcription, but that his-tagged CRP indeed binds to *hlyU* promoter DNA. To better identify if CRP binding to the *hlyU* promoter is indeed cAMP dependent, this experiment should be repeated in the absence of cAMP. As well, if CRP is indeed binding to the near-perfect consensus sequence at the *hlyU* promoter, mutation of the CRP consensus sequence should prevent CRP binding in the presence of cAMP. Whether cAMP-CRP binding represses or activates HlyU is also not yet clear, however, the CRP consensus site identified in Figure 4.6A is upstream of the *hlyU* open reading frame, likely implicating it in recruiting RNA polymerase and enhancing or activating expression of *hlyU*. HlyU promoter expression assays using the *luxCDABE* reporter in the *crp* null mutant could help explore this further.

While it is clear that future work will need to be undertaken to complete this study, I have thus far implicated *nagC* as a vital regulator for *Vibrio parahaemolyticus* chitin utilization, biofilm formation, and indirectly T3SS-1

expression. Further, I have shown that unlike *V. cholerae*, *mlc* in *V. parahaemolyticus* does not appear to be involved in biofilm formation, indicating an interesting difference in biofilm formation between these two related species [426]. Finally, I have provided evidence that *crp* is involved in T3SS-1 expression likely through DNA-binding, and transcriptional control of, *hlyU*.

4.4 Materials and Methods

4.4.1 Bacterial Strains and Culture Conditions

Culture conditions are as previously described (See Sections 2.5.1 and 3.5.1) [269, 291]. A list of bacterial strains and plasmids used in this study can be found in Table 4.1.

4.4.2 Recombinant DNA Techniques and Allelic Exchange

Gibson assembly was performed using previously described methods [446]. Primers for the generation of PCR products were designed using the NEBuilder tool, and were verified for the absence of hairpins which had stability greater than 45°C using the OligoAnalyzer tool (IDT). A list of oligonucleotides can be found in Table 4.2. For gibson assembly, both 5X ISO buffer and Gibson assembly master mixes were prepared. 5X ISO buffer was prepared by mixing thoroughly 3mL 1M Tris-HCl (pH 7.5; Sigma), 150µl 2M $MgCl_2$ (Sigma), 240µl 100mM dNTP mix (25mM each: dGTP, dCTP, dATP, dTTP; NEB), 300µl 1M DTT (Sigma), 1.5g PEG-8000 (Sigma), and 600µl 50 mM NAD (NEB), followed by distilled H_2O to a final volume of 6mL. 5X ISO buffer was stored at -20°C. A Gibson assembly master mix was generated by mixing 320µL of 5X ISO buffer,0.64µl 10 Unit/µl T5 exonuclease (NEB), 20µl 2 Unit/µl Phusion polymerase (NEB), 160µl 40 Unit/µl Taq DNA ligase (NEB), and distilled H_2O to a final volume of 1.2mL.

Strain or plasmid	Description	Reference or source
VpWt	Wild-type *V. parahaemolyticus*	8
DH5αλpir	Cloning strain *E. coli*	Stratagene
BL21(λDE3)	Overexpression strain of *E. coli*	Novagen
pFLAG-CTC	Cloning palsmid to create C-terminal FLAG fusions	Sigma
pVSV105	*Vibrio* shuttle vector with *cat* selection	[447]
pRE112	Allelic exchange vector	[353]
pET21	HIS-tagged purification vector for overexpression in BL21(λDE3) *E. coli*	Sigma
pEVS104	Conjugation helper plasmid	[447]
*nagC*FL-pFLAG-CTC	*vp0828* cloned into the XhoI/KpnI sites of pFLAG-CTC generating a C-terminal FLAG-tagged *nagC* allele	This Study
*mlc*FL-pFLAG-CTC	*vp2038* cloned into the HindIII/KpnI sites of pFLAG-CTC generating a C-terminal FLAG-tagged *mlc* allele	This Study
*nagC*FL-pVSV105	*vp0828* with FLAG-tag from pFLAG-CTC, cloned into Sac/KpnI sites of pVSV105	This Study
*mlc*FL-pVSV105	*vp2038* with FLAG-tag from pFLAG-CTC, cloned into Sac/KpnI sites of pVSV105	This Study
Δ*nagC*-pRE112	Δ*vp0828* allele in pRE112, for allelic exchange	This Study
Δ*mlc*-pRE112	Δ*vp2038* allele in pRE112, for allelic exchange	This Study
Δ*crp*-pRE112	Δ*crp* allele in pRE112, for allelic exchange	This Study
crp-his-pET21	*crp* ORF cloned into XhoI/EcoRI site of pET21, for his-tagged overexpression and purification of CRP-HIS	This Study
crp-pVSV105	*crp* ORF cloned into the Eco53KI site pVSV105, for complementation of Δ*crp*	This Study
Δ*nagC*	Δ*nagC* gene deletion in VpWt parent, by allelic exchange with Δ*nagC*-pRE112	This Study

Strain or plasmid	Description	Reference or source
Δ*mlc*	Δ*mlc* gene deletion in VpWt parent, by allelic exchange with Δ*mlc*-pRE112	This Study
Δ*nagC*/pVSV105	Δ*nagC* strain carrying pVSV105 by conjugation	This Study
Δ*mlc*/pVSV105	Δ*nagC* strain carrying pVSV105 by conjugation	This Study
Δ*crp*/pVSV105	Δ*crp* strain carrying pVSV105 by conjugation	This Study
Δ*nagC*/*nagC*FL-pVSV105	Δ*nagC* strain carrying *nagC*FL-pVSV105 by conjugation	This Study
Δ*mlc*/*mlc*FL-pVSV105	Δ*mlc* strain carrying *mlc*FL-pVSV105 by conjugation	This Study
Δ*crp*/*crp*-pVSV105	Δ*crp* strain carrying *crp*-pVSV105 by conjugation	This Study

Table 4.1: Bacterial strains and plasmids used in Chapter 4.

Designation	Sequence (5'→ 3')	Description or purpose
CRP1	ACCGCATGCGATATCGAGCTCTTTTGGGCTGTGAACGAGC	Gibson assembly primer
CRP2	TATTATGGTTACTCGTTAATCTCAATTTAGAGTGAC	Gibson assembly primer
CRP3	ATTAACGAGTAACCATAATATCTCACTTCCTCTGC	Gibson assembly primer
CRP4	CATGAATTCCCGGGAGAGCTAAAGGTGAAGCCGTCATCGTG	Gibson assembly primer
NT140	CTGTATCAGGCTGAAAATCTTCTCTC	pFLAG-CTC Reverse Primer
NT413	CC**AAGCTT**GGAAGTCATAGATTCTC	*mlc*FL construction
NT414	CC**GGTACC**ACCTTCCACCACCTTCATTAAC	*mlc*FL construction
NT415	CC**GGTACC**CTCAACGATTTGAGAAGT	*mlc* allelic exchange
NT416	CC**CATATG**GTGGTGGAAGGTTAATCACT	*mlc* allelic exchange
NT417	CC**CATATG**AATATGGCCCGGTTGAGC	*mlc* allelic exchange
NT418	CC**GAGCTC**ATTGCGTGGGAAAGAAATGGC	*mlc* allelic exchange
NT431	CCC**CATATG**ACCAATCTGTCCGCCATTCATGCT	*nagC* Allelic Exchange
NT432	GG**GGTACC**GCTAACTTAAGTCCTGGTTTTATC	*nagC* Allelic Exchange
NT433	CCC**GAGCTC**CGCGAAGAAAATGCC	*nagC* Allelic Exchange
NT434	CCC**CATATG**GGCGTTTTGTTACAAAAGTTAC	*nagC* Allelic Exchange
NT435	CCC**CTCGAG**CGCATGCTTGCAAACCTGACTATTTTCG	*nagC*FL construction
NT436	CC**GGTACC**GTCTTCTAGTAACTTTTGTAAC	*nagC*FL construction
NT439	CC**GCATGC**GGAAGTCATAGATTCTC	*mlc*FL PCR
NT480	CCCCC**GGTACC**CAACAGCTGATTGAAACAGG	*crp* Complementation
NT481	CCCCC**GAGCTC**GTTAGCCTCTTGCGACCTAGG	*crp* Complementation
NT482	CCCCC**CATATG**ATGGTTCTAGGTAAACCTCAAACCG	CRPhis Expression
NT483	CCCCC**CTCGAG**ACGAGTACCGTAAACAACG	CRPhis Expression
NT484	CGAATGATTCAGCATCCTGC	*hlyU* promoter EMSA
NT485	CACGACGGCTTTTGCGG	*hlyU* promoter EMSA

Table 4.2: List of oligonucleotides used in Chapter 4. Restriction enzyme recognition sites are bolded.

15µL Gibson assembly master mix aliquots were stored at -20°C. Purified PCR products, along with linearized plasmid vector were added to the Gibson assembly master mix at a 2:1 PCR product/plasmid ratio, using calculated moles of DNA product to ensure accuracy using the following formula:

$$dsDNA_{(mol)} = \frac{dsDNA_{(\mu g)}}{(dsDNA_{(bp)} \bullet 617.96_{(g/mol/bp)}) + 36.04_{g/mol}}$$

Gibson assembly reactions were made up to 20µl and incubated for 1 hr at 50°C.

Δ*crp*-pRE112 was generated using Gibson assembly. Two DNA fragments, were amplified by PCR using the RIMD2210633 genome as template, Q5 polymerase (NEB) and oligonucleotides CRP1/CRP2 for the region immediately upstream of *crp* and CRP3/CRP4 for region immediately downstream of *crp* (IDT). By design, these two PCR fragments have homology to each other at the site of the *crp* allele deletion, as well as the pRE112 vector at the opposing ends, for use in a gibson assembly reaction. The pRE112 vector was pre-treated with SacI-HF (NEB) to linearize the vector prior to gibson assembly as described above.

The *crp*-pET21 CRP-HIS expression construct and the *crp*-pVSV105 complementation construct were generated by standard cloning techniques. The *crp* ORF from the RIMD2210633 genome was first PCR amplified using Phusion polymerase (NEB) and oligonucleotides NT482/NT483 (IDT), generating a *crp* allele without its native stop codon. Following, the PCR fragment was cloned into the NdeI/XhoI sites of pET21 to generate *crp*-his-pET21. A *crp*-pVSV105 complementation construct was generated by PCR amplifying the *crp* gene and immediate upstream intergenic region from RIMD2210633 using oligos NT480/NT481 (IDT) and Phusion polymerase (NEB). Following, the PCR product was phosphorylated using T4 PNK (NEB) according to manufacturer instructions. pVSV105 was digested using Eco53KI, and dephosphorylated using recombinant shrimp alkaline phosphatase (rSAP; NEB) according to manufacturer instructions. The phosphorylated PCR product and linearized, dephosphorylated pVSV105 were subjected to blunt end ligation with T4 DNA

ligase, generating *crp*-pVSV105.

Δ*nagC*-pRE112 and Δ*mlc*-pRE112 were generated using standard cloning techniques. The immediate upstream and downstream regions of the *nagC* and *mlc* genes were amplified from the RIMD2210633 chromosome using Phusion polymerase (NEB). For the PCR products immediately upstream and downstream of *nagC*, oligos NT431/NT432 and NT433/NT434 were used, respectively. For the PCR products immediately upstream and downstream of Δ*mlc*, NT415/NT416 and NT417/NT418 were used, respectively. Following, the PCR products for each gene deletion were digested with NdeI, followed by heat-inactivation of the enzyme, and ligated with T4 DNA ligase at room temperature for 1 hr. Finally, a PCR reacton using the flanking oligonucleotides NT431/NT434 for the Δ*nagC* allele, and NT415/NT417 for Δ*mlc* allele was performed using Phusion polymerase (NEB). These null alleles were cloned into pRE112 using the SacI/KpnI restriction sites.

Complementation constructs for the *nagC* and *mlc* null mutants were generated using standard cloning techniques. The C-terminal FLAG-tag encoding *nagC* gene (*nagC*FL) was generated by first amplifying a PCR product containing the *nagC* ORF using the RIMD2210633 genome as template, Phusion polymerase (NEB), and oligonucleotides NT435/NT436 (IDT). Following PCR amplification, the fragment was cloned into the XhoI/KpnI sites of pFLAG-CTC, generating *nagC*FL-pFLAG-CTC. *nagC*FL was PCR amplified from *nagC*FL-pFLAG-CTC using NT140/NT435 and cloned into pVSV105 into the SphI/Eco53KI restriction sites, generating *nagC*FL-pVSV105. To generate a C-terminal FLAG-tag encoding *mlc* gene (*mlc*FL), *mlc* was PCR amplified using the RIMD2210633 genome as template, Phusion polymerase (NEB) and oligonucleotides NT413/NT414 (IDT). Following, the PCR fragment was cloned into the HindIII/KpnI sites of pFLAG-CTC, generating *mlc*FL-pFLAG-CTC. Following, *mlc*FL was PCR amplified from *mlc*FL-pFLAG-CTC using Phusion polymerase (NEB), and oligonucleotides

NT140/NT439 (IDT). The PCR product was then cloned into pVSV105 in the SphI/Eco53KI restriction sites, generating *mlc*FL-pVSV105.

4.4.3 Protein Sequence Comparison and Bioinformatics

Protein sequences were pulled from the NCBI gene database using the reference strains for *V. parahaemolyticus* (NCBI Accession: NC_004603.1), *V. cholerae* (NCBI Accession: NZ_LT906614.1), and *E. coli K12* (NCBI Accession: U00096.3) and compared with one another by BLAST to determine sequence identity and similarity [439, 440]. Following, the protein sequences were compared using Clustal Omega to perform an alignment. Clustal alignments were viewed and exported using JalView [448].

4.4.4 Allelic Exchange

Allelic exchange was used to generate mutant strains as previously described [269]. pRE112 suicide vectors were transferred to *V. parahaemolyticus* RIMD2210633 recipients via conjugation, and cells which crossed the plasmid into the chromosome were selected for using chloramphenicol (5µg/mL) containing LBS agar. Following, sucrose selection on 5% sucrose LB agar plates was performed, and colonies were screened for chloramphenicol sensitivity indicating that they had lost the integrated pRE112 plasmid. Strains were screened for deletion events by PCR.

4.4.5 Prey-Killing Assays

Prey-killing assays, which are T6SS dependent in *Vibrio parahaemolyticus* RIMD2210633, were performed according to previously published methods [233]. Briefly, *V. parahaemolyticus* strains were mixed with *E. coli* DH5αλpir prey cells at a 4:1 ratio after normalizing to an OD_{600} of 0.5. Cell mixtures were plated

in a puddle on LBS agar and incubated at 30°C for 4 hrs. *E. coli* prey were enumerated by serial dilution on selective media (chloramphenicol 30 μg/mL). DH5αλpir carried pVSV105 to allow for chloramphenicol selection of *E. coli* cells.

4.4.6 Immunoblotting

*nagC*FL-pVSV105 and *mlc*FL-pVSV105 carrying *nagC* null and *mlc* null strains of *V. parahaemolyticus*, as well as VSV105 empty vector controls, were subjected to whole cell lysates and immunoblotting after overnight growth in LBS at 37°C. Whole cell lysates were generated by 10-fold dilution of overnight culture in 2x ESB (4% SDS, 20% glycerol, 0.004% bromphenol blue, 0.125M Tris-Cl, (pH 6.8), 10% 2-mercaptoethanol), followed by SDS-PAGE to separate proteins by molecular weight. Following, the proteins were transferred to a nitrocellulose membrane (BioRad) using a semi-dry Transblot SD Transfer Cell (BioRad). The nitrocellulose membrane was blocked in 5% skim-milk powder in TBS-Tween (w/v) and then probed using mouse anti-FLAG antibodies followed by horse anti-igG(Ms) antibodies conjugated to horseradish peroxidase (HRP). HRP was probed using BioRad ECL Western blot substrate kit (BioRad), and visualized on the BioRad VersaDoc platform.

4.4.7 Swimming Motility Assays

Swimming motility assays were performed as previously described [449]. *V. parahaemolyticus* strains were normalized for OD_{600} to 0.05 and inoculated by stab into a 0.2% w/v agar LBS plates. The swimming ability of each strain was documented by photography after incubation at 37°C for 5 hrs.

4.4.8 Biofilm Formation Assays

Biofilm assays were performed as previously described with minor modification [165, 423]. Overnight cultures of *V. parahaemolyticus* strains, and the *E. coli* control strain, were inoculated into LB with the appropriate antibiotics, in glass tubes. Culture density was normalized to an OD_{600} of 0.025 and incubated for 16 hrs at 37°C, 200RPM. Following, the liquid culture was removed from the tubes and gently rinsed with PBS (137mM NaCl, 2.7mM KCl, 10mM Na_2HPO_4, and 1.8mM KH_2PO_4) to remove leftover planktonic cells. Following, biofilms were stained with 5mL 0.5% crystal violet (w/v, Sigma) for 30 min, followed by two rinses with 5mL PBS to remove excess crystal violet from the tubes. The biofilms were destained with 3mL of 95% ethanol. Absorbance at 565nm was then measured in a 1cm cuvette using the Biophotometer Plus (Eppendorf).

4.4.9 Chitinase Secretion/ZOC Assays

Chitinases are extracellularly secreted enzymes, and as such they degrade chitin in agar plates surrounding areas of *V. parahaemolyticus* growth. To measure chitinase activity in various *V. parahaemolyticus* strains, overnight cultures of the various strains were prepared, normalized for OD_{600}, and spot plated onto MM9 media containing colloidal chitin at 5% (w/v). Cells were allowed to incubate at 30°C for 7 days and the chitin ZOC generated by the various strains were photographed.

4.4.10 *exsA* Expression Luciferase Assays

exsA expression assays were performed in 96-well plate format following protocols previously established [269, 290]. Overnight cultures were prepared and normalized to an OD_{600} of 0.025 in LB media supplemented with 15mM Mg^{2+} and 5mM EGTA. 100µL aliquots were pipetted into triplicate wells of a clear 96-well

plate. The lid of the 96-well plate was treated with 0.05% Triton-X100 (Sigma) in 20% ethanol to prevent condensation from building up on the plate lid during incubation, and thoroughly dried before covering the 96-well plates. Cells were incubated at 30°C in the VictorX5 2030 Multilabel plate reader (Perkin-Elmer), and luminometry and cell density measurements (counts per second (CPS) and OD_{600}) were taken every 30 min for 270 min.

4.4.11 Host Cell Cytotoxicity Assays

Host cell cytoxicity assays were performed as previously described using LDH release [269, 291]. Both HeLa and HT29 cell lines (an epithelial and colonic cancer cell line respectively) were used and were cultured according to standard procedure (ATCC).

4.4.12 HIS-tagged CRP Purification

HIS-tagged purification of CRP was performed as previously described [269, 291]. Overnight culture of BL21(λDE3) strain carrying the *crp*-his-pET21 plasmid was prepared, followed by 50-fold dilution into fresh LB (ampicillin 100μg/mL). Cells were incubated at 37°C, 200RPM to an OD_{600} between 0.6 - 0.8. Following, protein expression was induced using 0.4mM IPTG for 3 hrs. Protein expression and his-tag purification was evaluated using SDS-PAGE (Figure 4.6B).

4.4.13 CRP EMSA

cAMP-CRP binding assays were performed as previously described [355]. EMSA was performed using the *hlyU*-promoter region generated by PCR amplification using oligos NT484 and NT485 (IDT), Phusion DNA polymerase (NEB), and the RIMD2210633 genome as template. Increasing amounts of purified CRP-HIS protein were combined with a constant amount of PCR product, with 10X EMSA

buffer (100 mM Tris (pH 7.5 at 20°C), 10 mM EDTA, 1 M KCl, 1 mM DTT, 5% v/v glycerol, 0.10 mg/mL BSA) and cAMP (20μM) [355]. Increasing amounts of bovine serum albumin (BSA) were also mixed with a constant concentration of PCR product, as a negative control for protein-DNA shifts, using the same EMSA conditions. DNA only controls were included for each protein. Native TBE-PAGE was performed on each sample in a 8% TBE polyacrylamide gel. cAMP (20μM) was included in the TBE running buffer. Samples were subjected to electrophoresis for 4 hrs at 100V, followed by staining with ethidium bromide. Gels were visualized using the VersaDoc platform.

Chapter 5

Interrogation of the Chitin Catabolic Pathway in *Vibrio parahaemolyticus*

5.1 Introduction

Bacterial virulence is not isolated from their ability to sense cues, as well as adapt and survive in changing environments. Properly coordinated virulence often requires sensing environmental cues to adapt to new nutrient conditions, including the ability to sense hosts and generate the virulence factors necessary for efficient and successful infection. In *Vibrionacea*, chitin is one such environmental cue. Chitin is catabolized by members of the *Vibrionacea* to be used as a source for carbon and nitrogen, but also induces cell remodelling for natural competence, biofilm formation, niche-competition, and virulence [60, 85, 91, 450–452].

Chitin, the second most abundant carbon biopolymer found on Earth, is a β-1,4-linked polymer of GlcNAc found in tremendous quantities in the world's oceans [453]. Marine arthropods (including copepods) are thought to generate in excess of billions of tons of chitin annually [75, 454]. Second to only cellulose in abundance, chitin is a rich carbon and nitrogen source for the *Vibrionacea*, and can be found either in association with other animals like crustaceans and bivalves, or in marine snow – greater than 0.5mm diameter organic aggregates which sink to the ocean floor [2]. The process through which *Vibrio* species degrade chitin into GlcNAc oligosaccharides to generate fructose-6-phosphate is well understood [27]. First, *Vibrio* spp. encode, express, and, secrete chitinases, which degrade chitin into GlcNAc oligosaccharides. These oligosaccharides are

taken up by the cell through a chitoporin and other unidentified or non-specific porins for GlcNAc [19, 455]. Exo- and endo-chitinases, chitodextrinases, and B-N-acetylhexosaminidases then degrade these polymers into mono- or disaccharides which can be moved into the cytoplasm by PTS or ABC transporters [19, 455]. Following transport and phosphorylation of the GlcNAc monomer, GlcNAc-6-P is converted to GlcN-6-P by a GlcNAc-6-P deacetylase (NagA), followed by conversion to fructose-6-P and NH_3 by a GlcN-6-P deaminanse (NagB). Fructose-6-P feeds into the parallel pathways for general metabolism, glycolysis and the pentose-phosphate pathway [19, 455].

Chitin degradation, and the many genes involved in the processing of GlcNAc into usable sugar substrate, is thought to be regulated in *Vibrio* spp. by two primary regulators: NagC and ChiS/CBP [76]. NagC and ChiS/CBP regulate the two primary GlcNAc catabolism pathways, the *nag* genes and *chb* operon, respectively. As previously discussed, NagC is a ROK family transcriptional regulator which represses the *nagE-nagAB* genes in *Vibrio* spp., as well as some chitinase genes [421, 427] (see Chapter 4). NagC itself is allosterically regulated by the presence or absence of GlcNAc-6-P where GlcNAc6-P binds to NagC and prevents NagC binding to DNA to repress or activate its targets [420, 421]. Further, a bacterial two-component hybrid sensor/kinase, ChiS, controls the *chb* operon required for GlcNAc dissacharide catabolism in *V. cholerae* and *Vibrio furnissii* [90]. CBP, a periplasmic sugar binding protein, interacts directly with ChiS in the absence of its sugar ligand and prevents ChiS from activating its regulon [90]. When CBP is bound by $(GlcNAc)_2$, and to a significantly lesser extent GlcNAc, ChiS is free to directly activate its regulon from the membrane [89].

Two additional regulatiors, HapR and TfoS, are also indicated in regulating chitin catabolism in *Vibrio* spp. TfoS regulates chitinase gene expression, while HapR is involved in repression of the *chb* operon [87, 456]. TfoS is a transmembrane receptor protein which interacts with GlcNAc disaccharides and

positively regulates the expression of a small RNA, TfoR [62, 68, 457]. TfoR is Hfq-dependent, and in turn regulates the expression of the master regulator for *V. cholerae* natural competence, TfoX, post-transcriptionally through an interaction with an upstream hairpin in the 5' untranslated region of the *tfoX* mRNA. This interaction exposes a shine-dalgarno sequence and permits translation. TfoS, through an unknown mechanism, also enhances the expression of genes in the chitin catabolism pathway, namely the 4 chitinases in *Vibrio parahaemolyticus* [87]. Finally, quorum sensing master regulator HapR represses the *chb* operon in *V. cholerae*, which can be de-repressed through proteolytic cleavage of HapR by the ClpAP protease [456]. At high cell density, the presence of abundant autoinducers activates a phosphocascade from autoinducer sensor, through LuxO and LuxP, to indirectly activate HapR expression. At low cell density, this phosphocascade is less active and therefore HapR expression is reduced. ClpAP, a bacterial protease system, degrades HapR at a steady state, and at low cell density can degrade enough HapR to indirectly activate the *chb* operon [89]. ClpAP is involved in the degradation of other quorum sensing components in other bacterial species [458].

Although much is known about chitin catabolism in *Vibrio* spp., and a genomic screen has been conducted – by microarray – to look at the induced expression of genes in *V. cholerae*, few functional genetic surveys have been performed to identify the contingent of necessary genes in *V. parahaemolyticus* [76]. Our established success in transposon mutagenesis (see Chapter 3) provides an opportunity to fill this gap [269]. Discovered as "jumping-genes" by Barbara McClintock in maize, transposons are a valuable tool for functional gene discovery [459, 460]. In traditional approaches (see Chapter 3), transposon mutants are isolated and screened for a particular outcome. This process is time expensive. Alternatively, transposon mutants can be isolated and screened as a population, and the insertional mutations and their frequency can be determined using

high-throughput DNA sequencing, in a process called transposon sequencing (TnSeq) [461]. Uniquely, this process does not require the isolation of individual mutants, but rather transposon mutants in a population, which can be screened in parallel and measured for their survival using high-throughput DNA sequencing to determine necessary genomic regions. TnSeq allows for functional genetic analysis to occur on a whole genome level, in much less time than is necessary for individual mutant isolation and screening.

Here, I set out to determine the genes that contribute to efficient chitin catabolism in *V. parahaemolyticus* growth on chitin. Here, I successfully deployed an efficient transposon mutagenesis system, and screened a population of transposon mutants for their ability to survive on glucose and colloidal chitin as their only carbon source. While these results replicate much of what is known in the literature in regards to *Vibrionacea* chitin catabolism, I identified uncharacterized genes involved in sugar transport and gene regulation as conditionally essential for chitin catabolism in *V. parahaemolyticus*. I also implicate the general secretion pathway as a necessary component of efficient chitin assimilation.

5.2 Results

5.2.1 Generation of a Streptomycin Resistant Strain of *V. parahaemolyticus*

Transposons often carry selection cassettes to allow for the selection of transposon mutants from their untransposed neighbors. Further, many transposition systems rely on *E. coli* donor strains which move transposition genes by conjugation into a recipient strain. To ensure robust selection of *V. parahaemolyticus* transposon mutants from *E. coli* donors during the simultaneous conjugation/transposition reactions, I set out to generate a

streptomycin resistant strain of *Vibrio parahaemolyticus*. Briefly, streptomycin sensitive *V. parahaemolyticus* RIMD2210633 cells were cultured overnight in LBS and subcultured into LBS media containing subinhibitory concentrations of streptomycin (25 μg/mL). These cells were incubated until cell growth was visible, around 48 hours. Subsequent subcultures were performed into LBS media containing increasing concentrations of streptomycin until the strain stably and consistently demonstrated efficient growth in overnight culture. The culture was streak purified on streptomycin containing media, and a single colony was selected and stored under the new strain name VpSm (see Table 5.10). This strain grows reliably in overnight culture containing streptomycin at 200 μg/mL.

To identify the genetic mutations generated by the conversion from a streptomycin susceptible to strepomycin resistant phenotypes, genomic DNA from both the parent strain (*V. parahaemolyticus* RIMD2210633 - VpWt) and the VpSm strain was isolated and the genome was sequenced using the Illumina MiSeq platform. Sequencing data was filtered using CutAdapt to remove reads shorter than 20bp and with a quality-score of less than 15, followed by assembly into contigs with SPAdes using default parameters [462, 463]. Contigs were assessed for quality metrics using QUAST with contigs smaller than 700bp excluded [464]. The reference genome for *V. parahaemolyticus* RIMD2210633 is 5,165,779bp in length and a GC content of 45.4%. The VpWt and VpSm assembled genomes are similar, with VpWt at 5,105,269bp and 45.3%, and VpSm at 5,107,670bp and 45.3% in length and GC content respectively (Table 5.1, Figure 5.1AB). The genome length plateued as contig numbers increased as seen in Figure 1A, and contig length was slightly longer for VpSm than for VpWt as evidenced by the steep slope of the VpSm curve compared to VpWt (Figure 5.1A). Binned contig coverage indicates that both of the genomes assembled have at least 15x coverage across most of the genome, with some contigs having a coverage of over 51x (Figure 5.1C). Overall, this data indicates that genome sequencing for the two strains was successful, and

Assembly	VpSm	VpWt
# contigs	78	98
# contigs ($\geq$ 50000 bp)	24	34
Largest contig	558, 650	318, s715
Total length	5,107,670	5,105,269
GC (%)	45.30	45.32
# N's per 100 kbp	0.00	0.00

Table 5.1: Contig analysis from genome sequencing the VpSm and VpWt strain obtained from SPAdes assembly, as analyzed by QUAST. Genomic DNA was isolated from the VpSm and VpWt strains of *V. parahaemolyticus* and subjected to sequencing by Illumina MiSeq, followed by assembly using SPAdes and quality analysis using QUAST.

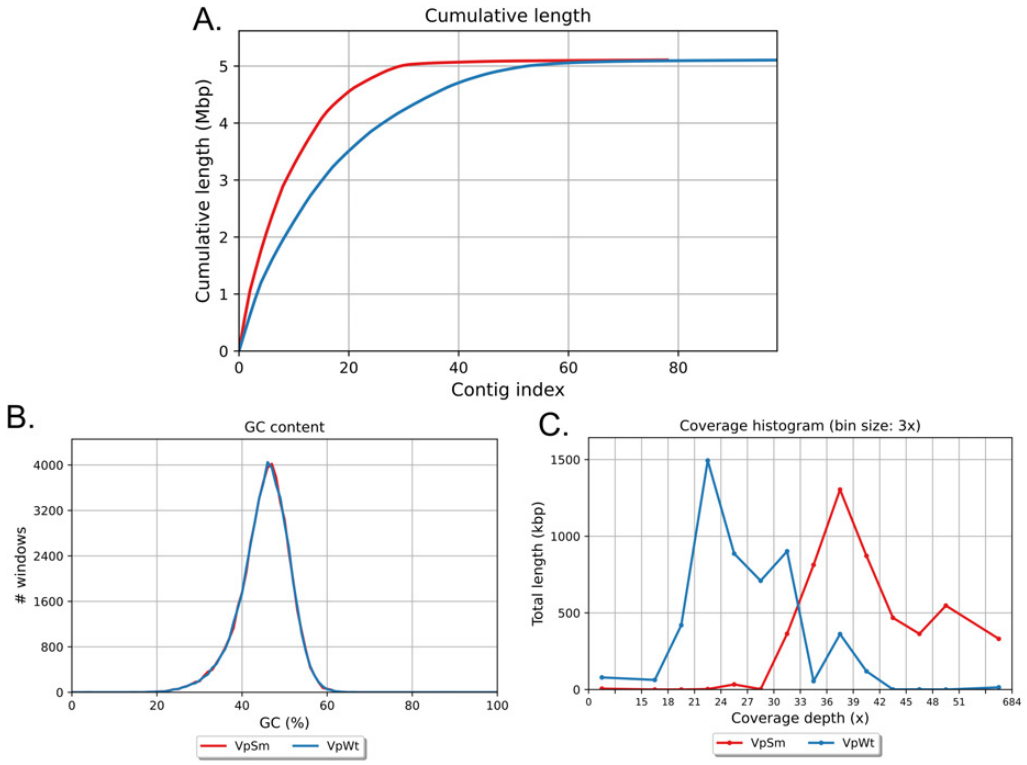

Figure 5.1: Sequencing statistics for the SPAdes assembled VpSm and VpWt parent *Vibrio parahaemolyticus* genomes by QUAST analysis. QUAST analysis was run excluding contigs smaller than 700bp on the SPAdes assembled genomes of the streptomycin resistant *V. parahaemolyticus* strain, as well as the VpWt parent. The following output from QUAST is shown: A. The cumulative length plot showing the length of the genome in Mbp with the increase in contig number, B. A GC content plot, where the %GC is shown on the y-axis and the number of non-overlapping 100bp windows on the x-axis, and C. the coverage depth (x-axis) of bins of contigs (automatically selected, in the case bin size is 3x), where the total length is shown on the y-axis in kbp.

that I have good coverage of the data to analyze the genomes for mutations.

To identify the mutations likely causing streptomycin resistance, the VpWt and VpSm strain sequences were compared to the reference genome for *V. parahaemolyticus* RIMD2210633 in the NCBI database using Snippy. The mutational differences between the sequenced RIMD2210633 strain and the reference sequences found in the NCBI database can be found in Table 5.2 and Table 5.3. Unique mutations found only in VpSm can be found in Table 5.4, and include mutations in the coding regions for cytochrome d ubiquinol oxidase subunit I, RpoS, RpsL, and GidB on chromosome 1, and two hypothetical proteins on chromosome 2.

Aminoglycoside antibiotics, including streptomycin, are known to increase the affinity of the ribosome for non-cognate tRNAs and prevent proper proofreading by the ribosome through binding to the aminoacyl-tRNA recognition site (A-site) on the 16S rRNA of the 30S subunit, preventing accurate protein synthesis and stalling cell growth [465–467]. Indeed, I identified a mutation in *rpsL*, encoding a 30S ribosomal protein (protein S12). RpsL is well known to be involved in resistance to streptomycin when mutated, most commonly at amino acids 42 and 87 in *E. coli* – and therefore is a likely culprit for the streptomycin resistant phenotype in the VpSm strain [467, 468]. The snp found in the *rpsL* coding gene generates a protein with a mutation from arginine to cysteine at position 86. Mutations in GidB, a small subunit ribosomal methyltransferase, have also been linked to streptomycin resistance through missense mutations S70R, T146M, and R187M in *Mycobacterium tuberculosis* [469]. The deletion in *gidB* causes a frameshift variation which generates a protein that is truncated by 67 amino acids and newly terminates at amino acid 144.

Gene	Locus Tag	Position	Type	Reference	Change	Effect
putative cps protein D	VP0197	207301	snp	G	C	Asp18His
two component response regulator	VP0489	505575	snp	C	G	Gly192Arg
tRNA-Met	VPt031	774113	snp	G	C	N/A
clpX	VP0918	954177	ins/snp	G	AT	frameshift
hypothetical protein	VP0926	963503	snp	G	C	Gln125Glu
hypothetical protein	VP1049	1100985	del	GA	G	frameshift Ala335fs
cytochrome d ubiquinol oxidase subunit II	VP0154	1105386	snp	C	T	Glu125Glu
hutD	VP1273	1349262	mnp	GT	TG	missense
sensor histidine kinase	VP1375	1459950	snp	G	T	Leu405Leu
hypothetical protein	VP1386	1471166	del	G A	G	Arg16fs
hypothetical protein	VP1386	1472406	ins	T	T G	Val429fs
hypothetical protein	VP1397	1499941	del	TATTTTTCATC TTCCACCCACTC	T	frameshift
hypothetical protein	VP1579	1676767	snp	C	A	Cys72Phe
amino acid ABC transporter permease	VP1622	1722561	del	TCGTGCCTTA	T	inframe deletion
SAM-dependent methyltransferase	VP1629	1727956	snp	G	T	Trp70Leu
transmembrane protein	VP1764	1878569	snp	G	C	Leu61Phe
aldehyde dehydrogenase	VP1777	1893389	del	TA	T	frameshift
short chain dehydrogenase	VP2033	2131271	snp	C	T	Arg26Arg
thiC	VP3027	3231255	snp	G	A	Ile347Ile

Table 5.2: Mutations identified in Chromosome 1 of *V. parahaemolyticus* RIMD2210633 from the Thomas Lab in 2020. Mutations identified by Snippy following whole genome sequencing of the Thomas Lab RIMD2210633 strain in 2020 on chromosome 1. Only mutations in coding genes are reported for brevity.

Gene	Locus Tag	Position	Type	Reference	Change	Effect
16S ribosomal RNA	VPAr01	133964	snp	T	A	N/A
type I secretion C-terminal target domain-containing protein	pseudogene	986999	ins	A	AG	frameshift
type I secretion C-terminal target domain-containing protein	psuedogene	988264	ins	C	CA	frameshift
FdhF/YdeP family oxidoreductase	VPA0959	1001947	snp	G	C	Gly269Gly
viaA	VPA1008	1062798	snp	G	A	Ser2Leu
DUF3763 domain-containing protein	VPA1009	1062831	mnp	CC	AA	Gly548Leu
IS4 family transposase	pseudogene	1452476	del	CG	C	frameshift
polysaccharide biosynthesis tyrosine autokinase	VPA1406	1479875	del	CG	C	frameshift
polysaccharide biosynthesis tyrosine autokinase	VPA1406	1480164	del	AC	A	frameshift
VWA domain-containing protein	VPA1455	1548543	del	ACAG	A	Gln504del

Table 5.3: Mutations identified in Chromosome 2 of *V. parahaemolyticus* RIMD2210633 from the Thomas Lab in 2020. Mutations identified by Snippy following whole genome sequencing of the Thomas Lab RIMD2210633 strain in 2020 on chromosome 2. Only mutations in coding genes are reported for brevity.

Gene	Chromosome	Locus Tag	Position	Type	Reference	Change	Effect
cytochrome d ubiquinol oxidase subunit I	1	VP1053	1104057	snp	C	A	Ala157Glu
rpoS	1	VP2553	2697884	del	CGTCCCACTT CTTCAAGCGTA	C	frameshift
rpsL	1	VP2773	2942781	snp	G	A	Arg86Cys
gidB	1	VP3079	3286127	del	CA	C	frameshift
hypothetical protein	2	VPA1369	1438813	ins	A	AT	frameshift
PH domain-containing protein	2	VPA1559	1658645	ins	G	GT	frameshift

Table 5.4: Unique mutations identified in the VpSm strain of *V. parahaemolyticus*. Mutations identified by Snippy following whole genome sequencing of the VpSm strain. Only mutations in coding genes are reported for brevity.

5.2.2 Optimization of Transposon Mutagenesis and Growth Media Conditions

While pSC189, containing a Himar1 transposon and the relevant transposase, has been used in other organisms for transposition and for transposon sequencing procedures, I was concerned about the effectiveness of the neomycin selectable marker on the mariner transposon due to common resistance to neomycin seen in *V. parahaemolyticus* [470, 471]. Alternatively, *V. parahaemolyticus* is robustly sensitive to chloramphenicol, at relatively low concentrations (2-5 μg/mL) compared to *E. coli* (30 μg/mL) [471]. To generate a chloramphenicol selectable transposon, pSC189 was PCR amplified without the neomycin cassette found in the Himar1 transposon, and was replaced by the *cat* gene from pRE112 (see Section 5.4 Materials and Methods).

Here, both a conjugation and a transposition reaction are occuring simultaneously, moving the transposition vector from *E. coli* SM10λpir, into *V. parahaemoylticus* where it is replication incompetent. Once the transposition vector is inside the *V. parahaemolyticus* cell, transposition occurs. In order to generate sufficient mutants for coverage of the VpSm genome, the transposition procedure was optimized in the VpSm strain. Transposition reactions were collected at hourly intervals after 2 hours, with the last collection at 5 hours. CFU counts for these reactions were generated on differentially selective media for the transposon mutants and for the VpSm recipient strain (Figure 5.2A). While some transposon mutants were generated at 2 and 3 hours, a significant jump in the number of transposon mutants occurs at 4 hours, and there was little change in the number of mutants between 4 and 5 hours (Figure 5.2A). A slow increase in the number of total VpSm cells was detected over the course of the entire transposition (Figure 5.2A). All future transpositions were incubated for four hours prior to collection and selection of the transposon mutants, to minimize the

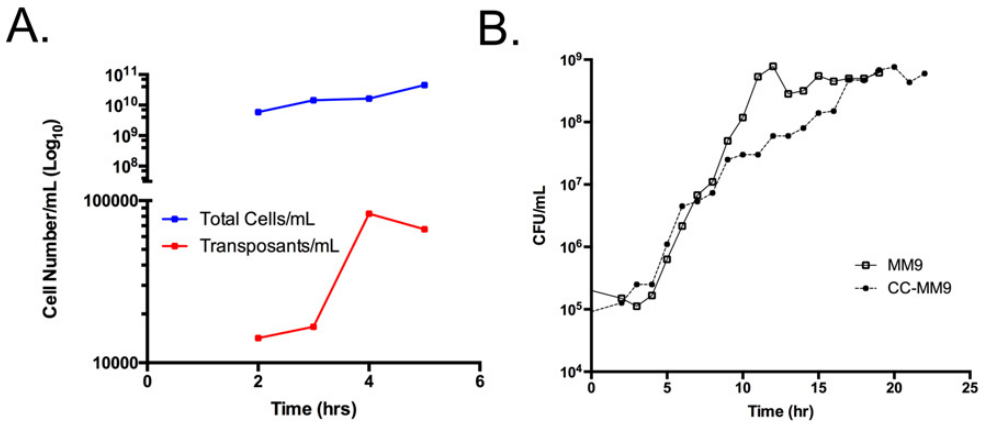

Figure 5.2: Optimization of Transposition Frequency and Growth Media Conditions. A. *V. parahaemolyticus* VpSm cells were mixed with *E. coli* SM10λpir carrying pSC189-CmR at a 1:1 ratio, washed to remove residual antibiotics, and plated on a filter placed on an LBS agar plate. The cell mixtures were incubated at 37°C for varying lengths of time, collected, and plated on selective media to determine total *V. parahaemolyticus* cell number as well as the number of transposon mutants generated. B. *V. parahaemolyticus* VpSm cells from overnight culture in LBS were inoculated into MM9 containing either glucose or colloidal chitin as the sole carbon source. Samples were collected every hr for 24 hours, vortexed to remove cells attached to chitin particles, serially diluted, and spot plated to determine the number of cells present at that timepoint.

growth of the entire population while maximizing transposon mutant generation. Using the transposon mutants generated from three replicates, a mean transposition frequency of 6.80 x 10^{-5} was determined (Table 5.5).

In preparation for selection of the transposon mutants, I performed growth curves for VpSm in minimal media containing either glucose or colloidal chitin as the exclusive carbon source. VpSm was cultured overnight in LBS, washed twice in phosphate-buffered saline (PBS), and inoculated into MM9 with either glucose of colloidal chitin as the sole carbon source. Colony forming units were calculated each hour after 2 hours time and continued until 24 hours post-inoculation (Figure 5.2B). Both colloidal chitin and glucose supported growth of VpSm as expected. Glucose allowed the cells to reach stationary phase at an earlier timepoint than the colloidal chitin condition, at twelve hours and seventeen hours respectively (Figure 5.2B). The colloidal chitin growth appears to have two phases of growth, the first period of faster exponential growth beginning around 5 hours and ending around 8 hours, followed by a period of slower exponential growth from 9 to 17 hours, when they enter stationary phase (Figure 5.2B, see Discussion).

5.2.3 Transposon Insertion Sequencing Statistics

Transposon insertion sequencing was completed using a transposon mutant library prepared according to the optimization above and subjected to growth in liquid MM9 media containing either glucose or colloidal chitin as the sole carbon source. Illumina sequencing was performed on the prepared mutant libraries for both conditions (see Materials and Methods). The total number of reads for the glucose condition was 34,712,563 and for the chitin condition was 33,384,928. In the glucose condition, a significant number of reads did not contain the transposon prefix, with valid Tn reads larger than 20bp at 14,378,650 reads. In the chitin condition, more of the reads were greater than 20bp with valid Tn prefixes, at 20,630,996 reads. Reads mapped to the *V. parahaemolyticus* genome

Replicate	Total *V. parahaemolyticus* (CFU/mL)	Transposon mutants (CFU/mL)	Transposition Frequency
1	1.76×10^{10}	1.13×10^{6}	6.42×10^{-5}
2	6.0×10^{9}	2.5×10^{5}	4.17×10^{-5}
3	2.8×10^{9}	2.75×10^{5}	9.82×10^{-5}

Table 5.5: Transposition Frequency for pSC189-CmR into *V. parahaemolyticus*. Total *V. parahaemolyticus* cells and transposon mutants were counted using spot plates on either streptomycin or streptomycin/chloramphenicol plates respectively. Colony forming units were calculated per mL and used to determine the transposition frequency.

numbered 7,743,311 reads for the glucose condition and 15,612,858 reads for the chitin condition. The *V. parahaemolyticus* reference genome contains 175,455 TA sites on chromosome 1 and 96,964 TA sites on chromosome 2. In the glucose condition, 48,415 TA sites were identified on chromosome 1, and 27,953 TA sites were identified on chromosome 2. In the chitin condition, 81,134 TA sites were identified on chromosome 1, and 50,429 were identified on chromosome 2. This translates to a 27.6% and 28.8% TA site coverage in the glucose condition for chromosome 1 and 2 respectively, and 46.2% and 52.0% in the chitin condition. This data can be found summarized in Table 5.6. Insertion counts are visualized across the entire *V. parahaemolyticus* genome for both samples in Figure 5.3A.

The TRANSIT insertion sequencing analysis software was used to determine essential genes in the glucose condition as well as the chitin condition, and compare the two conditions to find conditionally essential genes [472]. This was done by first using a hidden markov model (HMM) approach to identify essential regions of the genome in each condition, followed by a zero-inflated negative binomial (ZINB) approach to compare the two conditions. First, each condition was subjected to a HMM analysis which locally determines differences in transposon insertion density and provides essentiallity calls for genes: growth advantage, non-essential, growth disadvantage, and essential. HMM analysis of the glucose condition determined 9 genes whose mutations conferred a growth advantage, 3671 genes that were non-essential, 223 genes whose mutations generated a growth defect, and 663 genes that were essential for growth (Figure 5.3B). In the chitin condition, mutations in 82 genes conferred a growth advantage, 3725 genes were non-essential, mutations in 148 genes conferred a growth disadvantage, and 611 genes were labelled essential for growth (Figure 5.3B). 3547 genes were non-essential and 627 were essential in both conditions.

The comparative zero-inflation negative binomial (ZINB) analysis was used next to identify genes that were conditionally required for growth on chitin

Statistic	Glucose		Chitin	
	Chr 1	Chr 2	Chr 1	Chr 2
Total Reads	34,712,563		33,384,928	
Tn Reads	14,378,650		20,630,996	
Reads w/out Tn	58.6%		38.2%	
Reads Mapped	7,743,311		15,612,858	
TA Sites Hit	48,415	27,953	81,134	50,429
Tn Density	27.6%	28.8%	46.2%	52.0%

Table 5.6: Transposon Sequencing Data Statistics from TRANSIT. Data collected from the glucose and chitin condition transposon mutant libraries sequenced using Illumina sequencing and passed through the TRANSIT pipeline. Data shown appears as output of the tn_stats file from TRANSIT.

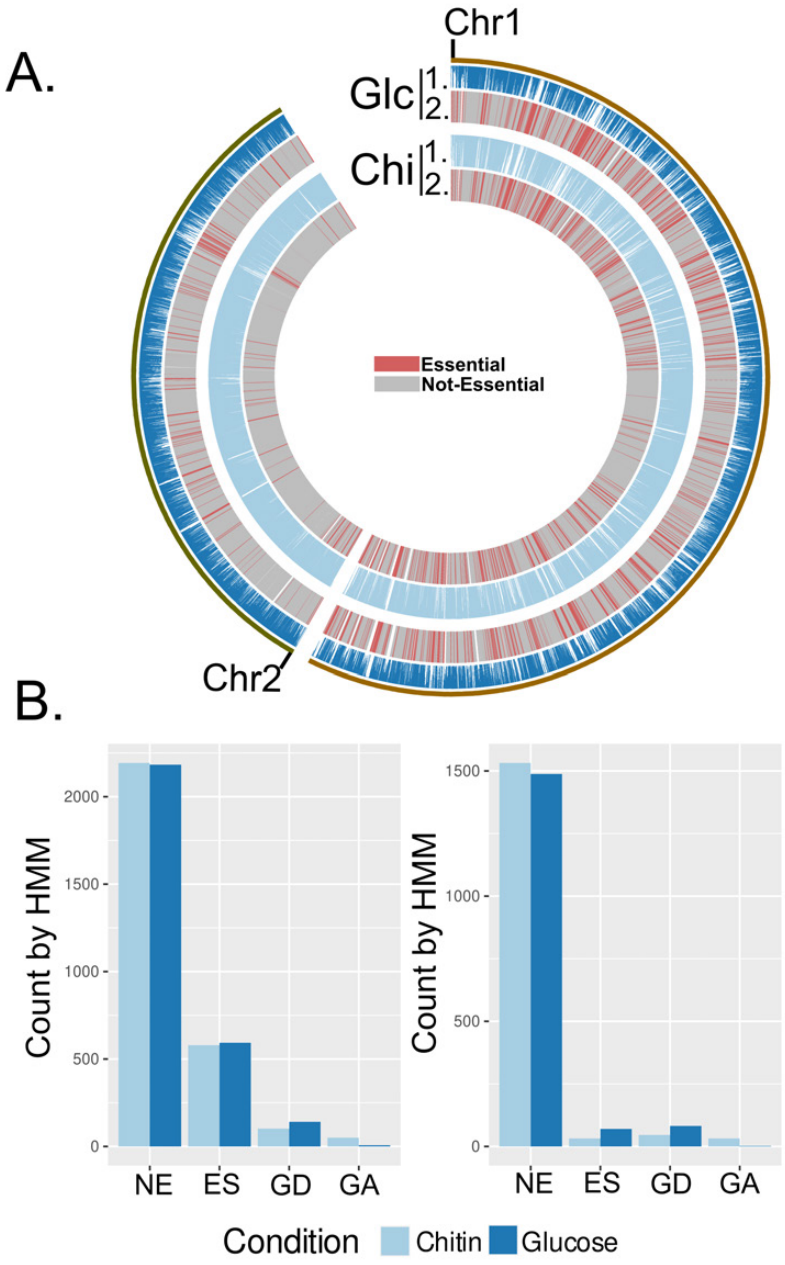

Figure 5.3: Summary of TnSeq analysis. A. A Circos plot showing transposon insertion data across both chromosome 1 and chromosome 2 (1.) and the loci indicated as essential (combined growth disadvantage and essential from HMM analysis) and non-essential (combined growth advantage and non-essential from HMM analysis) (2.). Red loci are essential, while grey loci are non-essential. Both glucose (Glc and Chitin (Chi) growth condition datasets are shown. B. Genes indicated as non-essential (NE), essential (ES), growth defect (GD), or growth advantage (GA) by HMM analysis for both chromosomes.

or glucose. ZINB analysis also provides statistical data by comparing mean insertional data across an entire gene, and therefore can be limited to those genes which have a Benjamini-Hochberg adjusted *p*-value less than 0.05. With this chosen *p*-value, ZINB analysis identified 43 genes on chromosome 1 that were statistically underrepresened in the chitin condition when compared to the glucose condition, and 14 genes that were statistically underrepresented in the glucose condition when compared to the chitin condition. On chromosome 2, ZINB analysis identified 6 genes that were underrepresented in the chitin condition, and 8 genes underrepresented in the glucose condition.

5.2.4 Transposon Sequencing Confirms Necessary Genes for Chitin Degradation in *Vibrio parahaemolyticus*

To assess if the chitin selection media was behaving as expected, the chitin catabolic pathway as modelled by Hirano et. al. was explored for essential genes in the chitin condition (Figure 5.4) [455]. Conditionally essential genes for the chitin condition include *nagA* (*vpa0038*), *nagB* (*vp0829*), *chiS* (*vp2478*), *cbp* (*vp2479*), almost the entire *chb* operon (*vp2479-vp2488*, except *vp2485*), one of three encoded β-N-acetylhexosaminidases (*vp2486*), and a chitodextrinase (*vp0832*) (Figure 5.4). Importantly, these genes the two major chitin catabolism pathways in *Vibrio* spp. and all had significant *p*-values by ZINB ($p<0.0001$, except *nagB*; $p=0.0002$, Table 5.7). On occassion, the HMM status of a gene is labelled non-essential, while ZINB analyses identifies a gene as conditionally under-represented in chitin growth media (*chiS* for example, Table 5.7). This is due to inherent differences in the analyses, and the comparative nature of the ZINB analysis (see Discussion). Interestingly, *nagE* (*vp0831*), the glucosamine kinase (*vp2485*), the cellobiose PTS (*vp2635-vp2637*), ChiP (a chitoporin, *vp0760*), two

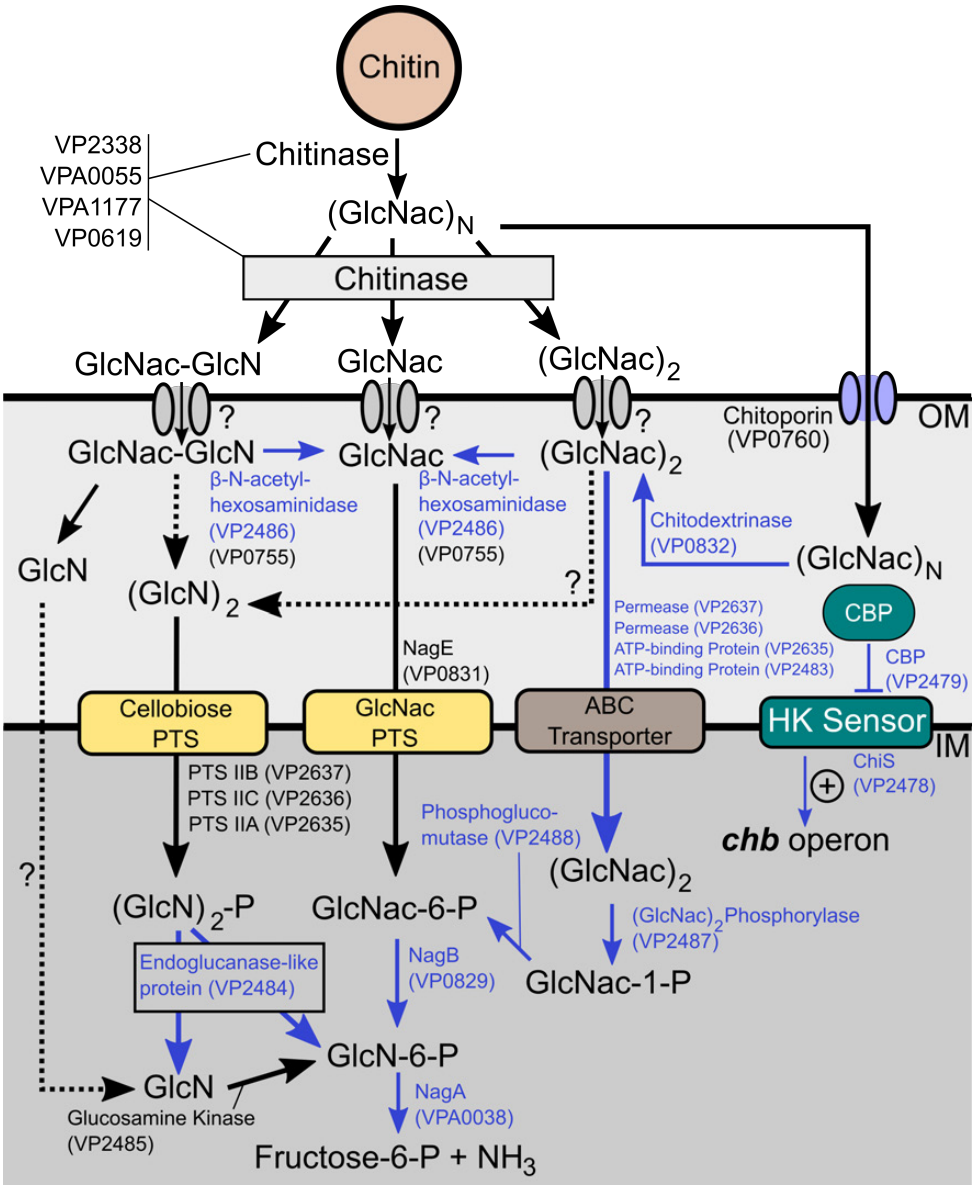

Figure 5.4: Genes essential for chitin metabolism as identified by TnSeq. A pathway diagram for catabolism of chitin to fructose-6-phosphate is shown, which component enzymes and their associated locus tags indicated, as modelled by [27, 455]. Chemical conversions which are known to occur, but do not have associated enzymes are shown with a question mark and dotted arrow. Conversions and locus tags shown in blue are indicated as conditionally essential ($p < 0.05$) for growth on chitin by ZINB analysis of the TnSeq dataset.

Locus Tag	Gene Name	Function	HMM Status	*p*-value	Sig.
VP0545	*nagZ*	β-N-acetylhexosaminidase	NE	0.407	NS
VP0619	*chiB*	Chitinase	NE	0.624	NS
VP0755	-	β-N-acetylhexosaminidase	NE	0.283	NS
VP0760	*chiP*	Chitoporin	NE	0.428	NS
VP0802	-	Porin	GD	< 0.0001	****
VP0828	*nagC*	ROK Family Transcriptional Regulator	NE	0.339	NS
VP0829	*nagA*	GlcNAc-6-P deacetylase	ES	<0.0001	****
VP0831	*nagE*	GlcNAc-specific PTS IIBC	NE	0.104	NS
VP0832	*glnS*	Chitodextrinase	ES	NA	NA
VP2338	*chiA1*	Chitinase	NE	0.875	NS
VP2478	*chiS*	hybrid His sensor/kinase	NE	<0.0001	****
VP2479	*cbp*	ABC transporter substrate-binding protein	NE	<0.0001	****
VP2480	-	ABC transporter permease	GD	<0.0001	****
VP2481	-	ABC transporter permease	GD	<0.0001	****
VP2482	-	ABC transporter ATP-binding protein	GD	<0.0001	****
VP2483	-	ABC transporter ATP-binding protein	GD	<0.0001	****
VP2484	-	Cellobiase	NE	<0.0001	****
VP2485	-	Glucosamine Kinase	NE	0.756	NS
VP2486	-	Hexoaminidase Exo I	ES	<0.0001	****
VP2487	-	$(GlcNAc)_2$ Phosphorylase	NE	<0.0001	****
VP2488	-	Phosphoglucomutase	ES	<0.0001	****
VP2635	-	$(GlcN)_2$ PTS IIB	NE	0.164	NS
VP2636	-	$(GlcN)_2$ PTS IIC	NE	0.980	NS
VP2637	-	$(GlcN)_2$ PTS IIA	GD	0.746	NS
VPA0038	*nagB*	GlcN-6-P deaminase	GD	0.0002	***
VPA0055	*chiA2*	Chitinase	NE	0.4550	NS
VPA1177	*chiA*	Chitinase	NE	0.8161	NS

Table 5.7: ZINB Analysis of the known chitin utilization program. Genes shown are underrepresented in the chitin growth condition, and their locus tag, gene name (if available), annotated function, HMM status in the chitin condition, adjusted *p*-value, and significance are shown. *p*-values were controlled for FDR using Benjamini-Hochberg. ****: $p<0.0001$). ES: Essential, GD: Growth Defect, NE: Not Essential by HMM analysis. NA: Not-Analyzed, ZINB does not analyze genes that are essential in both conditions. Sig. stands for significance.

of the three encoded β-N-acetylhexosaminidases (*vp0755*; p=0.283 and *vp0545*; p=0.407), as well as all of the encoded chitinases (*vp2338*, *vpa0055 vpa1177*, *vp0619*) were not conditionally essential in the chitin condition. The insertion data for the *nagE-nagAC* operon, *nagB*, and the *chb* operon can be found in Figure 5.5, and shows potent reduction in transposon insertion sequences in conditionally essential genes (essential genes are labelled in purple in Figure 5.5).

While the chitoporin was not identified as conditionally essential, another porin protein (VP0802) was. VP0802 has 97% sequence identity and 98% sequence similarity with an N-acetylglucoasmine regulated porin in *Vibrio* sp. JCM 18904 [473]. Phyre2 analysis of this protein revealed structural similarity to *P. aeruginosa* OprD, which transports small hydrophilic amino acids and sugars, and is involved in carbapenem resistance [474, 475]. This porin protein may represent part of the missing porin components necessary for the transport of various chitin catabolites (see Figure 5.4).

5.2.5 Type II Secretion is Essential, While Individual Chitinases are not Essential for Growth on Chitin

Strikingly, while none of the chitinase genes were essential on their own (see Table 5.7), many genes identified in the *gspCDEFGHIJKL* type II secretion system operon were essential for growth on chitin (Table 5.8 and Figure 5.6). While some of the genes in this operon did not meet the p-value cutoff of 0.05, many approached significance and were essential by HMM analysis (Table 5.8 and Figure 5.6). *gspD*, *gspE*, *gspF*, *gspG*, and *gspK* were statistically underrepresented in the chitin growth condition, while *gspC* and *gspI* approached significance (Table 5.8). *gspH*, *gspJ*, and *gspL* were not underrepresented in the chitin growth condition (Table 5.8). Importantly, *gspH*, *gspJ*, and *gspI* are all minor pseudopilin subunits, and this data would suggest that mutations in these genes do not affect the function of the T2SS with respect to supporting growth on colloidal chitin.

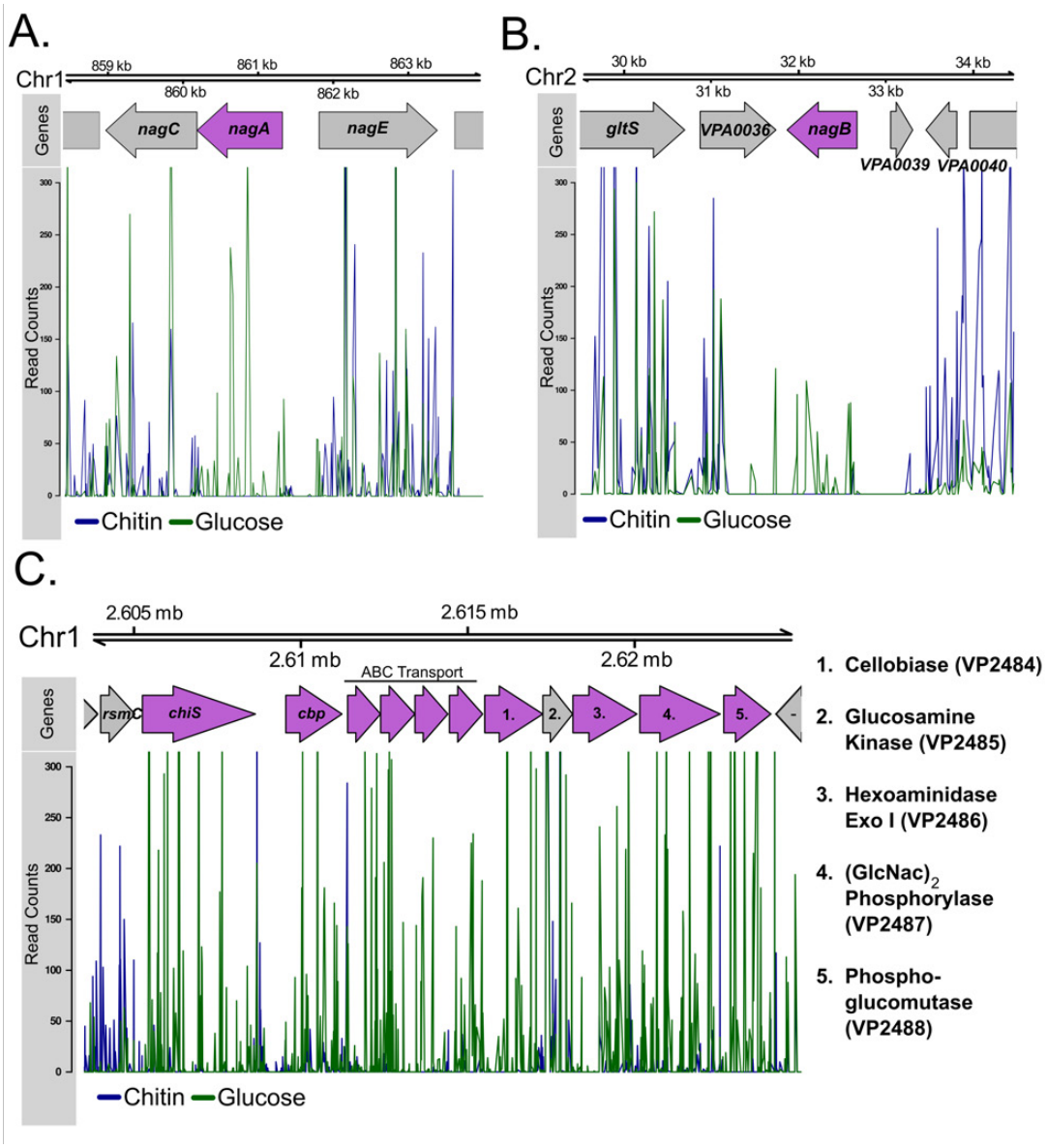

Figure 5.5: TnSeq insertion data for the GlcNAc and $(GlcNAc)_2$ transport and catabolism pathways. Chromosome numbers and locations are indicated by the parallel bars, with kbp markers for location. The gene annotation for each region is also shown, with genes with statistically underrepresented transposon insertions by ZINB analysis shown in purple. Insertion counts for each genomic region are shown for both the chitin and glucose growth condition. A. *nagE-nagAC* operon, B. *nagB* and surrounding genomic region, C. *chiS* and the *chb* operon.

Locus Tag	Gene Name	Function	HMM Status	*p*-value	Sig.
VP0122	*gspC*	T2SS Inner-membrane platform protein	ES	0.0699	NS
VP0133	*gspD*	T2SS Outer-membrane secretin	ES	0.0039	**
VP0134	*gspE*	T2SS Secretion ATPase	ES	0.0005	***
VP0135	*gspF*	T2SS Inner-membrane platform protein	ES	0.0016	**
VP0136	*gspG*	T2SS Major pseudopilin	ES	0.0075	**
VP0137	*gspH*	T2SS Minor pseudopilin	NE	1	NS
VP0138	*gspI*	T2SS Minor pseudopilin	ES	0.1074	NS
VP0139	*gspJ*	T2SS Minor pseudopilin	ES	0.2848	NS
VP0140	*gspK*	T2SS Minor pseudopilin	ES	0.0087	**
VP0141	*gspL*	T2SS Inner-membrane platform protein	ES	0.6352	NS
VP0741	*lolB*	Outer-membrane chaperone	ES	NA	NA
VP0977	*lolC*	ABC permease	ES	NA	NA
VP0978	*lolD*	ABC ATP-binding protein	ES	NA	NA
VP0979	*lolE*	ABC permease	ES	NA	NA
VP1106	*lolA*	Periplasmic chaperone	ES	NA	NA
VP2706	*gspM*	Inner-membrane platform protein	NE	0.2939	NS

Table 5.8: HMM status and ZINB analysis of the *gsp* operon and *lol* genes. Genes shown are underrepresented in the chitin growth condition, and their locus tag, gene name, annotated function, HMM status in the chitin condition, adjusted *p*-value, and significance are shown. *p*-values were controlled for FDR using Benjamini-Hochberg. **: $p<0.01$, ***: $p<0.001$. ES: Essential, NE: Not Essential by HMM analysis. NA: Not Analyzed, ZINB does not analyze genes that are essential in both conditions. Sig. stands for significance.

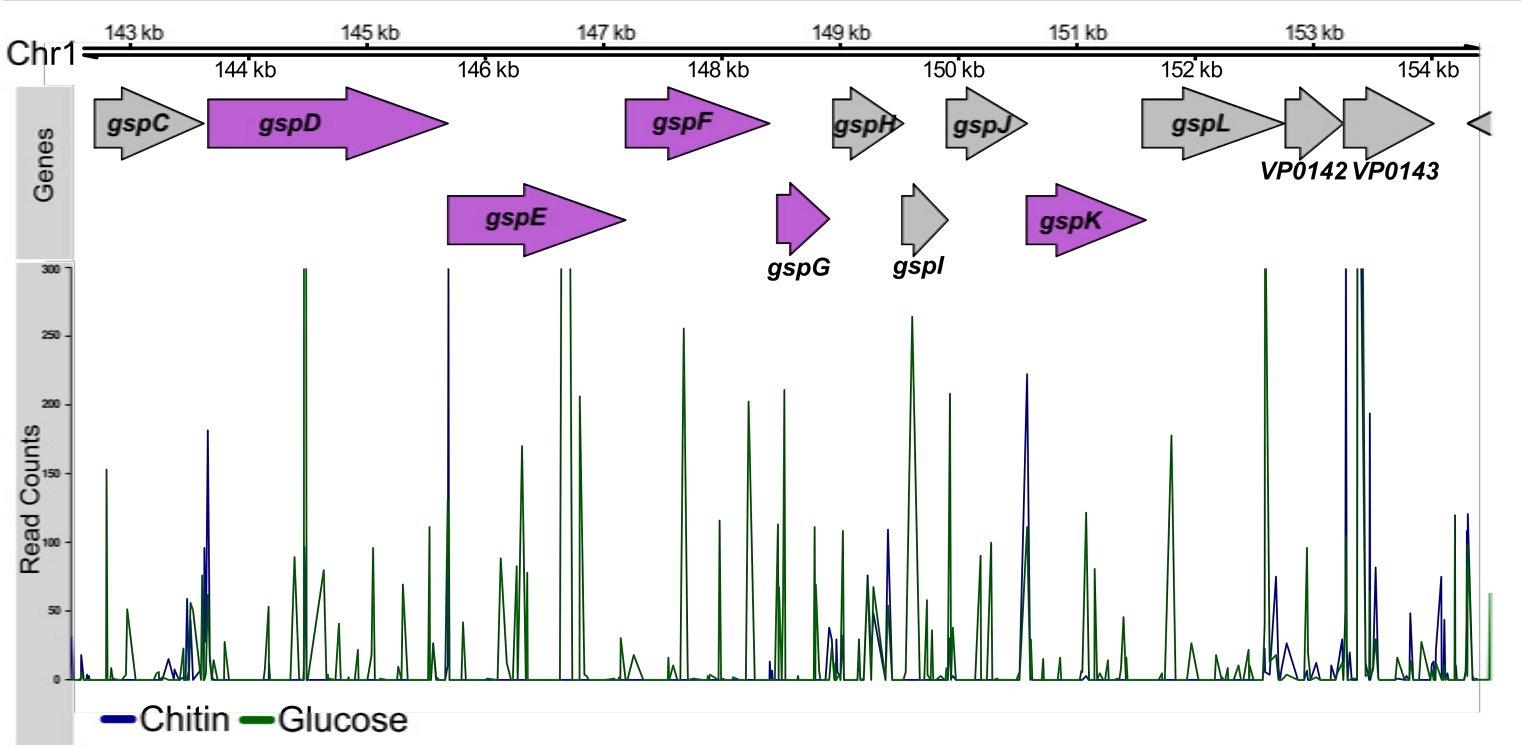

Figure 5.6: TnSeq insertion data for the *gsp* operon. Chromosome numbers and locations are indicated by the parallel bars, with kbp markers for location. The gene annotation for each region is also shown, with genes with statistically underrepresented transposon insertions by ZINB analysis shown in purple. Insertion counts for each genomic region are shown for both the chitin and glucose growth condition.

LolCDE is an ATP-binding cassette transporter that, along with LolA (a periplasmic chaperone) and LolB (an outer-membrane chaperone) aid in the process of transport and insertion of the outer-membrane secretin GspD into the outer-membrane [209]. *lolA*, *lolB*, *lolC*, *lolD*, and *lolE* are essential genes, as they coordinate the trafficking of lipoproteins other than GspD into the outer membrane (Table 5.8) [476]. ZINB analysis was therefore not performed on these genes.

General secretion pathway systems (like Type II Secretion) are involved in the secretion of chitinases, although their exact role is under-explored, particularly in *Vibrio* spp. [477]. However, identifying a T2SS as essential for growth on chitin is interesting given that the individual chitinases encoded by *V. parahaemolyticus* are not essential. *V. parahaemolyticus* encodes at least 4 chitinases, and it is possible that individual mutations in a single chitinase – where other chitinases are intact and secreted – does not slow or halt growth on chitinous surfaces. Meanwhile, a mutation in the secretion pathway for these chitinases presents a bottleneck in which no chitinase is secreted, and cells cannot breakdown or assimilate chitin locally (see Discussion).

5.2.6 Proteases and an Uncharacterized Transcriptional Regulator are Essential for Growth in Chitin

The protease genes, *clpA* and *clpP*, were identified as conditionally underrepresented by ZINB analysis, where *clpA* (p=0.065) approached significance and *clpP* (p<0.0001) was statistically significant (Table 5.9) [89]. However, related protease genes, *clpS* or *clpX*, had no impact on the cells ability to grow on chitin. In *V. cholerae*, the ClpAP protease is involved in activating the *chb* operon by degrading a quorum-sensing master regulator HapR, involved in repression of the *chb* operon [456, 478]. HapR is the master regulator of quorum-sensing in *V. cholerae*, and at high cell density it is expressed at high

levels [456, 479]. OpaR, the homologue of HapR in *V. parahaemolyticus*, is not required for growth on chitin as seen by ZINB analysis [480]. This also aligns with previous work showing that deletion of *hapR* did not induce a growth advantage on chitobiose, even with its repressive effect on the *chb* operon [456]. The other important quorum sensing proteins involved in the generation of auto-inducers (*cqsA* and *opaM*), the sensing of auto-inducers (*luxN*, *luxQ*, and *cqsS*), or the integration of the signals (*luxO* and *opaR*) were not required for growth on chitin [480, 481]. *luxU*, another protein integrator of the AI signals, was not analyzed by ZINB due to low-saturation of insertion sites at that locus (<15% saturation across the gene), but was labelled as essential for that same reason by HMM. For this reason, *luxU* is difficult to assign as essential or not for growth in chitin.

Two additional transcriptional regulators were identified as essential for growth in chitin, *tfoS* and a MurR/RpiR family transcription regulator *vp1236* (Table 5.9). MurR/RpiR family transcriptional regulators are implicated in the regulation of sugar catabolism operons, including the N-acetylmuramic acid (MurNAc) catabolism and uptake operon (*murPQ*) in *E. coli* by MurR [482]. Phyre2 analysis of the predicted structure of *vp1236* identifies NanR as a predicted homologue, which represses the N-acetylneuraminic acid (Neu5Ac) catabolic operon in *Vibrio vulnificus* [483, 484]. Further, TfoS regulates the expression of chitinases and some periplasmic GlcNAc oligosaccharide processing enzymes in *V. parahaemolyticus* and is essential for growth on $(GlcNAc)_6$, although is not required for growth on $(GlcNAc)_2$ [87], due to the dramatic downregulation of chitodextrinase (*vp0832*) and the two β-N-hexosaminidases (*vp2486* and *vp0545*) in a *tfoS* null mutant. Our data identified *tfoS* as essential for growth in the chitin condition, but two downstream elements associated with *tfoS* regulation (the small RNA *tfoR* and *tfoX*) were not essential for growth on chitin (Table 5.9).

Locus Tag	Gene Name	Function	HMM Status	*p*-value	Sig.
VP0854	*tfoS*	Membrane-bound transcriptional regulator	NE	0.000457	***
VP0917	*clpP*	Protease	GD	0.06533	NS
VP0918	*clpX*	Unfoldase	NE	0.54234	NS
VP1013	*clpS*	Protease Adapter Protein	GA	0.060657	NS
VP1014	*clpA*	Unfoldase	GD	0.000059	****
VP1236	-	MurR/RpiR Family Transcriptional Regulator	ES	0.000016	****
VP1697	*opaM*	AI synthase	NE	0.095409	NS
VP1698	*luxN*	AI Sensor	NE	0.416756	NS
VP2009	*luxO*	QS Response Regulator	NE	0.730857	NS
VP2098	*luxU*	Quorum Sensing HPt Protein	ES	NA	NA
VP2516	*opaR*	QS Master Regulator	NE	0.694376	NS
VPA0710	*cqsS*	CAI-1 HK Sensor	NE	1	NS
VPA0711	*cqsA*	CAI-1 Synthase	NE	1	NS
VPA1220	*luxQ*	HK Sensor	NE	0.137747	NS

Table 5.9: HMM status and ZINB analysis of the quorum-sensing regulators and other transcriptional regulators. Genes shown are underrepresented in the chitin growth condition, and their locus tag, gene name, annotated function, HMM status in the chitin condition, adjusted *p*-value, and significance are shown. *p*-values were controlled for FDR using Benjamini-Hochberg. **: $p<0.01$, ***: $p<0.001$. ES: Essential, NE: Not Essential by HMM analysis. NA: Not Analyzed, ZINB does not analyze genes that are low-saturation. Sig. stands for significance.

5.3 Discussion

Here, I developed a new transposon mutagenesis tool for use in *V. parahaemolyticus*, as well as generated and characterized a streptomycin-resistant *V. parahaemolyticus* strain (VpSm). I provided evidence that the transposon mutagenesis vector generates Himar1 transposon mutants at an acceptable rate, and that the VpSm strain is resistant to streptomycin through predictable mutations in *rpsL* and *gidB*. Using this transposon mutagenesis tool, combined with next generation sequencing technology, I identified an uncharacterized porin and transcriptional regulator, as well as implicated some, but not all, of the *gsp* genes of the T2SS, as essential for chitin catabolism in *V. parahaemolyticus*. Also, I successfully validated our approach, as I identified a number of previously known essential genes for chitin catabolism (namely the *chb* operon, *chiS*, *nagB*, and *nagA*). As well, I identified the known regulator *tfoS* and the protease complex *clpAP* as conditionally necessary for growth on chitin.

The identification of an uncharacterized porin is particularly interesting, and serves to expand our understanding of sugar transport for chitin catabolism in *V. parahaemolyticus*. Previous studies, including a recent chitin catabolism pathway analysis by Hirano *et al.* in *V. parahaemolyticus* showed that GlcNAc monosaccharides and $(GlcNAc)_2$ disaccharides may use a porin that has yet to be identified, and is distinct from the previously identified chitoporin (Figure 5.7) [76, 455, 485]. Our data did not identify the chitoporin (*chiP*) gene as essential for growth in our conditions, however, our data did identify another porin, *vp0802* as conditionally underrepresented by ZINB and as conferring a growth defect by HMM. The previously characterized chitoporin (ChiP) is a well characterized outer membrane porin used to import GlcNAc polymers into the periplasm to be degraded by other enzymes including the chitodextrinase (*Vp0832*) and β-N-acetylhexosaminidases (*vp0545*, *vp0755*, and *vp2486*) to

$(GlcNAc)_2$ [88, 454, 485]. Roseman and colleagues identified the chitoporin in *V. furnissi* as highly upregulated by $(GlcNAc)_{2\text{-}6}$ oligosaccharides, and later work showed that this protein is found in many different members of the *Vibrionacea* [27, 76, 88, 454]. Importantly, this channel has strong sugar specificity, and transports a number of different GlcNAc polymer molecules from $(GlcNAc)_2$ to $(GlcNAc)_6$ [485]. Given the requirement for $(GlcNAc)_2$ transport at the inner membrane, it seems likely that this porin transports at least $(GlcNAc)_2$, and perhaps GlcNAc monomers, but further molecular studies will need to be done to confirm this (FIgure 5.7). Further, VP0802 was identified as similar in predicted amino acid sequence to an N-acetylglucosamine regulated porin in *Vibrio* sp. JCM 18904 by BLAST and its predicted structure as similar to OprD – a porin for small amino acids and sugars in *P. aeroginosa* – by Phyre2 analysis [439, 440, 473–475]. These analyses provide some initial evidence that this porin may transport GlcNAc-containing sugars (Figure 5.7). OprD is involved in carbapenem resistance in *P. aeruginosa*, and the loss of VP0802 may also enhance resistance to carbapenem in *V. parahaemolyticus* and should be explored in future studies.

Further, the discovery of an uncharacterized transcription factor provides avenues to explore new pathways involved in chitin catabolism regulation. Here, I discovered a new transcriptional regulator of the MurR/RpiR family (*vp1236*) as conditionally required for growth in chitin (Figure 5.7). The predicted structure of this regulator is related to NanR in *V. vulnificus* by Phyre2 analysis [475, 483, 484]. NanR represses Neu5Ac catabolism and is de-repressed by binding to N-Acetylmannosamine-6-Phosphate, which prevents NanR binding to DNA [483]. While the nucleotide sequence for this gene does not

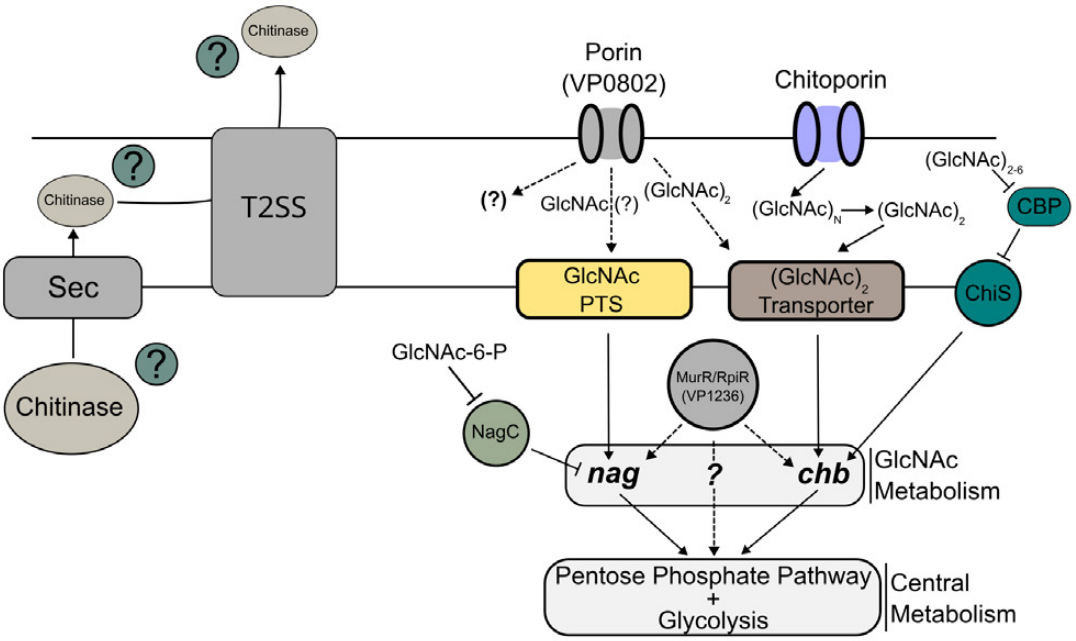

Figure 5.7: Schematic diagram of newly identified proteins and their potential roles in chitin catabolism. Chitinase secretion into the extracellular space is thought to be performed by the T2SS, following secretion into the periplasm through the Sec translocon. Other proteins involved in chitin catabolism may also be secreted via the T2SS. Important of chitin oligosaccharides occurs via outer membrane porins, which provide substrate-specific entry for GlcNAc-contianing sugars into the periplasm. The chitoporin imports $(GlcNAc)_{2\text{-}6}$ oligosaccharides, followed by degradation to GlcNAc disaccharides which are imported via a ABC transporter. A newly identified porin (VP0802) appears to be involved as well, but it is unknown what sugar specificity it has. Following important of GlcNAc mono- or disaccharides, the *nag* or *chb* encoded genes catabolize these sugars to fructose-6-P. *chb* is regulated by ChiS/CBP, while *nag* is regulated by NagC. A new regulator of the MurR/RpiR family (VP1236), may also be involved in regulating these catabolism pathways, or perhaps another currently uncharacterized pathway. All of these catabolism pathways converge on the pentose phosphate pathway and glycolysis.

indicate homology with NanR (44% similarity, 29% identified by BLAST global align), if its function remains similar, it may still be involved in regulating sugar catabolism and will be the focus of future studies [439, 440]. Importantly, it is unclear whether the newly identified MurR/RpiR family protein might regulate currently known GlcNAc catabolism pathways (*nag* or *chb*) or another as-yet identified pathway (Figure 5.7).

The other interesting set of genes identified is the T2SS *gsp* operon. Almost all of the genes in the *gsp* operon were labelled as conditionally essential for growth on chitin by ZINB analysis. Why exactly this is the case will require further exploration, although it seems likely that type II secretion would be involved in the secretion of chitinases to the extracellular environment (Figure 5.7) [486, 487]. Previous studies have indicated that ChiA2 is the major chitinase in *V. parahaemolyticus* [87]. However, our data does not identify any one chitinase as more important than other for growth in our conditions. I expected that chitinases may not be essential for growth on chitin in our conditions, as the transposon mutant population is not isogenic, and many of the bacteria in culture will contain functional chitinase genes. As such, the chitinases may act as a ”public-good”, allowing for other insertion mutants without functional chitinase genes to survive – essentially acting as ”cheaters” by exploiting the public good – as previously described [488–490]. However, chitinases obtained in this way by a T2SS deficient mutant would require diffusion, and based on our data clearly do not provide enough of a benefit quickly enough to prevent a statistically significant drop in some T2SS mutants in the transposon mutant population. Although chitinases seem to be the obvious candidates that make the T2SS essential for growth on chitin, I cannot rule out the secretion of other essential factors as well (Figure 5.7).

Further, three T2SS minor pseudopilin genes are not required for growth on chitin, *gspH*, *gspJ*, *gspK*. The role of minor pseudopilins in T2SS function is

not entirely clear, although they have been implicated in the initial formation of the pseudopilus structure [491]. Further. the minor psuedopilins have been linked to recruitment and binding to specific T2SS substrates, although this role is underexplored [492, 493]. This is of particular note, as perhaps these T2SS pseudopilin genes are essential for the secretion of some T2SS substrates, but are not essential for the secretion of chitinases. Further exploration will be required here.

Genome sequencing of the lab strain VpWt identified a number of mutations which conflict with the reference sequence, including a mutation in *clpX*. While bacteria will undoubtedly collect mutations through passaging in the lab, this *clpX* mutant is interesting considering ClpAP are indicted in regulation of the *chb* operon, although indirectly. While I did not find *clpX* or *clpS* to be essential for growth in chitin – consistent with the previous literature – it is difficult to know if this is because these genes are not essential, or if the mutation it has sustained changes its function [456]. Unique mutations in the VpSm strain allowed us to narrow the streptomycin resistance genotype to known mutations in *rpsL* and *gidB*.

Our transposon mutant libraries have acceptable density for both conditions, although the glucose condition did have less overall saturation of TA sites than the chitin condition. Further, many of the reads lacked the transposon prefix and were therefore discarded from the sequencing data. Many of the valid Tn reads did not map onto the *V. parahaemolyticus* genome, and this was more apparent in the glucose condition. One explanation for this is that *E. coli* donor cells or DNA is surviving through the selection and conditional growth process, and contaminating the sequencing run. Further optimization of the selection conditions and transposition mutant library will be required to reduce sequencing waste by donor DNA contamination, and to increase the density of insertion sequences in the libraries.

These data are from a single biological replicate, where sequencing was done on a single experiment in each condition. While statistical differences can be generated between conditions due to the number of independent TA sites within single genes (using ZINB analysis), this analysis is not in the form of biological replicates. To better support the conclusions, additional biological replicates should be pursued. However, the chitin utilization program in *Vibrio* is well defined, and I have shown that many of these genes are indeed necessary for growth on chitin (*chb* operon for example) and that our approach is working as intended. Indeed, I have also verified the involvement of other regulators such as *opaR* (indirectly through the protease *clpAP*), *chiS*/*cbp*, and *tfoS*.

Both the HMM and ZINB analyses of the transposon insertion sequencing data were valuable to assess the quality of our work and confirm that I have targeted the chitin catabolism pathway. However, there is some discrepancy between these two methods. Importantly, ZINB and HMM analysis did not always agree, and in some cases HMM analysis labelled a gene as non-essential, while ZINB analysis identified a conditionally significant reduction in transposon reads in chitin-containing media. While HMM is searching for regions of the genome that are locally different from adjacent regions in a single condition, ZINB analysis is looking more closely between the conditions for differences and involves normalization – by trimming total read-count (TTR), where the top and bottom 1% of reads are removed to reduce outliers and the entire dataset is divided by the new total and scaled up so the mean insertion count across the genome is 100 – between the two samples [472]. Therefore, although insertions may be found in genes like *cbp*, for example, and therefore are labelled as non-essential by HMM, ZINB analysis can more directly identified underrepresentation in insertion reads between conditions due to its comparative nature [472].

Much of the current understanding surrounding *V. parahaemolyticus* chitin utilization is based on homology searches and studies in other *Vibrio* spp. [27,

455]. Here, I confirmed a number of the genes involved in the process as essential (see Figure 5.4), in particular that the *GlcNAc*$_2$ assimilation pathway appears to be required for growth in our conditions (see Figure 5.4). However, the VpSm strain did not require the GlcNAc monomer PTS or the cellobiose PTS in our conditions. These mutants do not disrupt the cells ability to transport other sugar polymers like $(GlcNAc)_2$, and presumably allows for their survival using the $(GlcNAc)_2$ catabolism pathway. This finding suggests that using colloidal chitin as the GlcNAc source in our conditions may bias the cell towards use of $GlcNAc_2$ disaccharides, rather than the monomeric form. Indeed, given the rapid initial growth of VpSm in the colloidal chitin growth condition, followed by slower growth to stationary phase, it is possible that GlcNAc oligosaccharides are present in the colloidal chitin preparation and are allowing for rapid early growth (Figure 5.2B). However, this phenotype may also be caused by the initial growth conditions, where cells may survive on leftover nutrients from the overnight culture. These observations highlight that our dataset is limited to the genes absolutely necessary for the survival using our colloidal chitin preparation, and therefore may only constitute a portion of the essential genes for chitin catabolism. Future experiments should include other GlcNAc oligosaccharides, such as $(GlcNAc)_2$, as sole carbon sources to identify differences in the genes necessary for chitin versus smaller chitin oligosaccharides.

Further, three of the genes involved in regulation, *chiS*, *cbp*, and *tfoS* were identified in our data as essential for growth in colloidal chitin. This is well established in the literature, as *chiS* and *cbp* control the *chb* operon, and *tfoS* controls the β-N-acetyl-hexosaminidases, chitinase genes, and the chitin-induced DNA transformation systems [13, 62, 87, 89, 90]. Also, the *tfoS* finding aligns with previously published work in *V. parahaemolyticus* [87]. However, our data indicates that *tfoR* and *tfoX* are not essential for growth in chitin, which implicates *tfoS* regulation of the chitin catabolism operon as unique from that of natural

transformation regulation by TfoS, through both *tfoR* and *tfoX*.

Species specific quorum-sensing in *V. cholerae* represses *chb* through the quorum-sensing master regulator HapR. HapR is degraded by the protease system ClpAP after it is made, allowing for expression of the *chb* operon [456]. In the absence of ClpAP, HapR is not degraded and repression cannot be relieved [456]. Our data suggest that ClpAP may function similarly in *V. parahaemolyticus*, as mutants in these proteins are statistically underrepresented. At high cell density, HapR (or OpaR in *V. parahaemolyticus*) is highly expressed and represses the *chb* operon [456]. When *luxO* is deleted (or mutated through insertion), the cells are stuck in a ”high cell density” state and HapR (OpaR) should be highly expressed [494–496]. In opposition, specific mutations in LuxO which mimic the phosphorylated and active protein, lock the cells in a ”low cell density” state and HapR (OpaR) is not expressed at high levels [456]. In our data, *luxO* was not conditionally essential for growth on chitin, in fact none of the known quorum sensing pathway proteins are essential for growth on chitin. This is an expected outcome, as ClpAP will degrade OpaR in all cases, regardless of quorum-sensing regulation of OpaR, and relieve repression OpaR repression of the *chb* operon. More directed studies will be needed to determine how, if at all, the QS pathway impacts expression of the *chb* operon in *V. parahaemolyticus*.

Chitin catabolism is a conserved trait of the *Vibrionacea*, and is linked to a variety of the phenotypes seen in *Vibrio* spp., including natural competence, environmental survival, biofilm formation, niche-protection, and virulence [13, 27, 82, 85, 451, 452, 455, 497]. It is therefore important that I understand the necessary components of this process, in an effort to better link the lifestyle of *Vibrionacea* to their other important phenotypes, in some cases virulence. Here, I used transposon sequencing to better characterize this pathway in *Vibrio parahaemolyticus*. While transposon sequencing data is undoubtedly powerful and allows for the whole-genome interrogation of genes in various growth

conditions, the data generated is only as valuable as the annotation of the genes in question. In our case, independent verification of the MurR/RpiR family transcriptional regulator (*vp1236*) and the porin protein *vp0802* will be necessary to ensure they are annotated correctly in the genome, and to further deduce their function and relevance for chitin catabolism (see Chapter 6). Further exploration of the T2SS and its pseudopilin subunits will also need to be done to identify why many of the minor pseudopilins appear to be dispensable for T2SS essentiality in the chitin condition (see Chapter 6). Taken together, this study has provided experimental evidence for the necessary chitin catabolism components in *Vibrio parahaemolyticus* RIMD2210633. Here, I have identified a number of new avenues of exploration in this pathway, including a putative regulator and porin gene, as well as indicted Type II secretion as essential for successful growth on chitin.

5.4 Materials and Methods

5.4.1 Culturing and Bacterial Growth Conditions

Cells were cultured as previously described (Chapter 2 and 3) [269, 291]. For transposon sequencing experiments, after transposition *V. parahaemolyticus* cells were grown on Minimal M9 media (MM9) containing either glucose or colloidal chitin as the sole carbon source (420mM Na_2HPO_4, 220mM KH_2PO_4, 86mM NaCl, 187mM NH_4Cl, $MgSO_4$, $CaCl_2$, 0.2% w/v carbon). Selection for transposon mutants after transposition was performed on MM9 agar plates containing glucose, using streptomycin (100 μg/mL) and chloramphenicol (7.5 μg/mL).

Where antibiotics were required, *V. parahaemolyticus* were cultured using chloramphenicol (5μg/mL), kanamycin (100μg/mL) or streptomycin (100μg/mL). *E. coli* were cultured using chloramphenicol (30μg/mL). Strains and plasmids used in this study can be found in Table 5.10.

5.4.2 Generation of a Streptomycin Resistant Strain of *V. parahaemolyticus*

VpWt was cultured overnight in LBS media without antibiotics overnight at 37°C/200RPM. Following, VpWt cultures were subcultured into LBS containing 25 μg/mL of streptomycin and incubated at 37°C/200RPM until cell growth was visible, around 48 hrs. Once cell density was visible in the culture tubes, the cells were subcultured into LBS containing increasing concentrations of streptomycin until they reliably reached stationary phase after overnight growth at 200 μg/mL streptomycin. This culture was streak-purified and single colonies were screened for their ability to grow in LBS containing 200 μg/mL streptomycin. A colony which reliably reached stationary phase after overnight growth in media containing 200 μg/mL streptomycin was stored, and named VpSm.

5.4.3 Genome Sequencing, Quality Analysis, and Mutation Analysis

Genome sequencing of the VpWt and VpSm strain were completed following phenol-chloroform extraction of genomic DNA from both strains, as previously described [498]. Following, sequencing libraries were prepared using the Illumina XT DNA Library Preparation Kit (Illumina, FC-131-1024) and sequenced on the Illumina MiSeq platform. Following collection of the raw sequencing data, sequence files were analyzed for quality by FastQC, followed by assembly of the genomes of both strains using SPAdes, with default settings [463]. Following assembly of the genomes, the genome assembly quality was determined using QUAST, with default settings except the minimum contig length for inclusion, which was increased to 700bp based on the genome assemblies [464]. Finally, Snippy was used to identify mutations present in both strains compared to the

Strain or plasmid	Description	Reference or source
VpWt	Wild-type *V. parahaemolyticus*	[139]
pSC189/DH5αλpir	pSC189 Transposition Vector	[499]
pRE112/DH5αλpir	pRE112 Allelic Exchange Vector	[353]
pSC189-CmR/DH5αλpir	pSC189 Transposon vector carrying *cat* instead of neomycin cassette	This Study
pSC189-CmR/SM10λpir	pSC189-CmR plasmid in conjugation competent SM10λpir *E. coli*	This Study
VpSm	Streptomycin resistant strain generated from VpWt	This Study

Table 5.10: Bacterial strains and plasmids used in Chapter 5.

V. parahaemolyticus RIMD2210633 reference genome, using default settings [500].

5.4.4 Generation of pSC189-CmR

To generate a transposition vector that carried a *cat* gene instead of the neomycin resistance cassette on the encoded Himar1 transposon of pSC189, the swapped neomycin resistance cassette was swapped from pSC189 to the *cat* gene and promoter from pRE112 [353, 499]. First, the pSC189 vector was amplified by PCR using primers NT429/NT430 (IDT; see Table 5.11 for oligonucleotide sequences) using Phusion polymerase (NEB) to generate a linear fragment of pSC189 without the neomycin resistance cassette. The *cat* gene and its native promoter were generated from pRE112 via PCR with oligonucleotides NT427/NT428 (IDT) using Phusion polymerase (NEB). Following, the *cat* gene insert was phosphorylated using T4 PNK (NEB), and the pSC189 and *cat* fragments were ligated together by blunt ligation using T4 DNA Ligase (NEB).

5.4.5 Transposition Experiments

Transposition of the VpSm strain was performed using conjugation between VpSm recipients and SM10λpir *E. coli* pSC189-CmR donors. VpSm and SM10λpir were mixed at a 1:1 ratio normalized by OD_{600} and washed twice in PBS to remove overnight culture antibiotics. Cells were resuspended in LBS, plated on LBS media on an 0.45µm MLE filter (Millipore), and incubated at 37°C for 4 hours (or various times for the optimization experiments). Cells were collected in 1x MM9 salts by vortexing the filter in a 15 mL conical tube containing 1x MM9 salts, and the cells were then plated on MM9 glucose agar plates containing streptomycin (100 µg/mL) and chloramphenicol (5 µg/mL) to select only VpSm::Himar1 transposon

Designation	Sequence (5'→ 3')	Description or purpose
NT427	TT**CCCGGG**GTAGCAACCAGGCGTTTAAGG	pRE112 cat PCR
NT428	AA**GCATGC**CCTGGTGTCCCTGTTGATAC	pRE112 cat PCR
NT429	AA**CCCGGG**TTCTGAGCGGGACTCTG	pSC189 Inverse PCR
NT430	TT**GCATGC**CCACTGCAAGCTACCTGCTTTCTC	pSC189 Inverse PCR
LG17	TCGTCGGCAGCGTCAGATGTGTATAA GAGACAGTTCTAGAGACCGGGGACTTATCAGCC	HTML-PCR Transposon Primer
LG18	GTCTCGTGGGCTCGGAGATGTGTATA AGAGACAGGGGGGGGGGGGGGGG	HTML-PCR PolyC Primer
LG503	AATGATACGGCGACCACCGAGATCTA CACTATCCTCTTCGTCGGCAGCGTCAGATGTGT	HTML-PCR P5 (503)
LG510	AATGATACGGCGACCACCGAGATCTAC ACCGTCTAATTCGTCGGCAGCGTCAGATGTGT	HTML-PCR P5 (510)
LG710	CAAGCAGAAGACGGCATACGAGATTCG CCTTAGTCTCGTGGGCTCGGAGATGTGTATAA	HTML-PCR P7 (710)
LG706	CAAGCAGAAGACGGCATACGAGATTCGC CTTAGTCTCGTGGGCTCGGAGATGTGTATAA	HTML-PCR P7 (706)

Table 5.11: List of oligonucleotides used in Chapter 5. Restriction enzyme recognition sites are bolded.

mutants. Cells to be counted for colony-forming units were serially diluted, while transpositions to be used in a selective growth experiment for transposon sequencing were diluted to a total volume of 15 mL in 1x MM9 salts and plated over 150 MM9 glucose plates (FisherBrand, 100mm x 15mm), and incubated at 30°C for 36 hrs prior to collection.

5.4.6 Colloidal Chitin Preparation

Colloidal chitin was prepared using shrimp chitin flakes (Sigma) and a modified protocol from Joe *et al.* [501]. Briefly, 20g of shrimp chitin flakes were measured and blended into a fine powder using a blender. In a glass beaker, 100mL of concentrated HCl was added slowly to the powdered chitin and stirred gently with a glass rod every 15 min until the mixture was homogenous. The chitin-HCl mixture was added to 2L of distilled H_2O, and allowed to precipitate overnight at 4°C. Following, filtration through a buchner funnel and Whatman filter paper was performed using vacuum suction. The moist chitin cake was rinsed with distilled water in the Buchner funnel until the pH was neutral. The chitin cake was aliquoted into 50 mL conical tubes and sterilized by autoclave.

5.4.7 Generation of Insertion Sequencing Libraries using HTML-PCR

Following transposon mutagenesis and selection in either chitin or glucose containing minimal media, DNA was extracted from the selected population using phenol-chloroform extraction (see *Genome Sequencing, Quality Analysis, and Mutation Analysis*). HTML-PCR was performed as previously described with a few modifications [502]. Mainly, HTML-PCR does not require blunting of fragmented DNA ends, as previously published, prior to poly-C tailing (A. Camilli, personal communication, Nov 2019). 10µg of DNA was fragmented using

Fragmentase (NEB) for 30 min followed by DNA purification using AMPure XP Beads for DNA purification (1X volume, Beckman Coulter). Poly-C tailing was performed as previously described [502], followed by DNA cleanup as before, and PCR amplification using LG17 and LG18 oligos as previously described [502]. The second PCR was performed using variable primers – depending on the indices required for illumina sequencing (LG503 and LG710 for the glucose sample, along with LG510, LG706 for the chitin sample) – and 1 µL of the product from PCR1 as template. Final PCR amplifications were cleaned up, visualized using agarose gel electrophoresis, and quantified using the Illumina Quant Kit (NEB). Sequencing of these libraries was performed on the Illumina NextSeq platform, using single end reads, 150bp in length.

5.4.8 TRANSIT Analysis of Sequencing Data

Following sequencing, sequencing output was analyzed for overall quality using FastQC, followed by trimming and genome mapping using TRANSIT's tpp tool against the RIMD2210633 reference genom (NCBI Accession: NC_004603.1 and NC_004605.1), using default settings except the transposon primer window, which was increased to 40bp from the start of the read due to differences in the HTML-PCR approach and the approaches used by DeJesus *et al.* [472, 502]. Output from tpp was used to determine success of the transposon sequencing experiments, particularly considering Tn saturation and density of the libraries.

TRANSIT was then used to perform two analyses on the data for both chromosomes: a hidden markov model analysis to identify regions of the genome that are essential for growth on either condition, and a zero-inflated negative binomial approach which compared the glucose and chitin condition data following normalization. Both of these analyses were performed using default TRANSIT settings and prot_tables generated from gff3 files from the RIMD2210633 reference genome (NCBI Accession: NC_004603.1 and

NC_004605.1) [472]. Data from these analyses were explored using LibreOffice Calc to sort and identify genes that carried a ZINB adjusted *p*-value of less than 0.05, and visualized using R and the GViz package, along with modified R scripts from McCoy *et al.* [503–506].

Circos plots were generated from TRANSIT outputs for transposon insertion counts, and HMM essential gene analysis [507]. The RScript for GViz can be found in Appendix C.

Chapter 6

Discussion

6.1 Conclusions, Limitations, and Outstanding Questions

In this thesis, I have presented 4 linked but distinct studies exploring aspects of *V. parahaemolyticus* biology involving environmental sensing, T3SS-1 biogenesis and protein secretion, and sugar catabolism. While the work here begins to answer some questions about T3SS-1 activation and chitin catabolism in *V. parahaemolyticus*, it raises additional questions that will guide future work.

In Chapter 2, we hypothesized that a protein de-repressor would activate expression of the master-transcriptional regulator gene *exsA*, through de-repression of H-NS repression at that locus. Indeed, through a transposon mutagenesis screen linked to *exsA* inducing conditions (increased Mg^{2+} and EGTA), we identified HlyU, a small SmtB/ArsR family transcriptional regulator as essential for the de-repression of *exsA*, and subsequent expression of T3SS-1 [269]. We verified that in the absence of H-NS, HlyU is not necessary for *exsA* expression or T3SS-1 dependent protein secretion and cytotoxicity, confirming its role as a de-repressor of H-NS. Using the same transposon mutant screen, we also re-identified H-NS as a necessary repressor of the *exsA* gene. How HlyU acts to de-repress H-NS given the lack of overlapping DNA binding sequences remained unclear. While this study aimed to identify how the known inducing conditions for the T3SS-1 and *exsA* were transduced from extracellular environment to *exsA* activation, we were only successful in identifying and characterizing one step immediately upstream of *exsA* expression. How environmental signals are

transduced to activate the expression of HlyU, followed by *exsA* and T3SS-1 biogenesis, is not yet understood. Two other genes of interest were identified in the transposon screen, namely *phoX* and *merR*, and should also be explored in future studies (see Section 6.2 The T3SS-1 Regulatory Paradigm and HlyU).

In Chapter 3, we explored how HlyU de-represses H-NS, by studying unique sequence elements found at the HlyU binding site. While in Chapter 2 we had identified that HlyU bound to a region of DNA 56bp in length within the *exsBA* intergenic region by DNaseI footprinting, we wanted to better resolve the sequence elements necessary for HlyU-DNA interactions (Figure 2.6). We narrowed HlyU-DNA binding to a palindromic sequence centered in the DNaseI footprint by EMSA analysis, and identified a number of inverted repeat sequences surrounding it, which have the propensity to form DNA cruciform structures (Figure 3.3AC and B.1) [291]. Using *in vitro* T7 endonuclease analysis tied to DNA sequencing, we identified that a cruciform structure does indeed exist overlapping the binding site for HlyU, and that it is dependent on DNA supercoiling in that region (Figure 3.1 and 3.2). Mutational analysis of that region failed to remove a cruciform structure entirely, as mutations often recreated a cruciform structure of similar stability (Figure B.5), however mutations did have dramatic effects on *exsA* expression and T3SS-1 protein secretion (Figure 3.4). Further, we identified an additional promoter, downstream of HlyU binding and the known auto-regulatory *exsA* promoter, which is also HlyU-dependent (Figure 3.7 and B.7). Finally, without HlyU expression in a heterologous *E. coli* system, we are unable to express *exsA* from its native promoters by over-expressing the auto-regulatory ExsA protein *in trans* (Figure 3.6). Taken together, we have identified a DNA cruciform structure that, through interactions with HlyU in an as yet undefined mechanism, is involved in the regulation of *exsA*. The model we have proposed helps explain how H-NS may be de-repressed by HlyU DNA binding to either a DNA superstructure or to linear DNA, and is supported by other DNA topology studies which have shown

H-NS de-repression though changes in DNA superstructure [329, 331, 508, 509]. However, data presenting in Chapter 3 did not show direct interaction between HlyU and the cruciform structure, and did not interrogate H-NS DNA binding at the *exsA* locus, except through its de-repression by HlyU. Therefore, while it seems clear that a DNA cruciform is involved in regulation of *exsA*, it is unclear how HlyU and the cruciform might interact to remove H-NS repression. To date and to the best of our knowledge, this remains the first study to implicate a DNA cruciform in the regulation of a virulence factor in bacteria.

In Chapter 4, I explored environmental signalling of sugar catabolism through gene deletions in two related but distinct regulators NagC and Mlc. While these proteins are well studied in *E. coli*, and how NagC and Mlc function to modify the transcriptome of *E. coli* – through allosteric regulation by GlcNac-6-P and through sequestration to the membrane by the glucose EIIBC PTS subunit – is clear, the roles of NagC and Mlc in *V. parahaemolyticus* is unclear. Here, I showed that NagC is involved in regulation of biofilm formation, chitin catabolism, and – unexpectedly – *exsA* expression (Figures 4.3). Interestingly, considering the role of NagC in the regulation of PTS sugar transporters, and the role of these systems in cAMP homeostasis, we hypothesized that perhaps *exsA* was under the control of cAMP levels in the cell, as in *P. aeruginosa* T3SS expression [403, 510–512]. Indeed, CRP-cAMP binds to *hlyU* promoter DNA in EMSA experiments, presumably at an almost perfect consensus sequence (Figure 4.5), and a CRP deletion mutant has delayed and decreased activation of *exsA*, even when corrected for growth (Figure 4.5). While it appears that NagC and CRP are involved in the regulation of *hlyU* and consequently of *exsA*, additional studies to characterize cellular levels of cAMP in these mutants, along with exploration of *hlyU* expression in the CRP mutant, are necessary to understand this regulatory paradigm. Unfortunately, the phenotypes associated with the deletion of Mlc are less clear, as genetic complementation with *mlc in trans* failed to restore

wildtype phenotypes, as seen by a significant growth defect (Figure 4.2A). Further optimization of the complementation in the *mlc* null mutant will be required to fully understand the phenotypes associated with that mutant.

Finally, in Chapter 5, I used transposon mutagenesis, along with whole genome sequencing technologies, to explore the chitin catabolism pathway in *V. parahaemolyticus*. While studies on this pathway have been completed in *V. cholerae*, and extrapolated to *V. parahaemolyticus*, whole genome approaches using transposon-sequencing could provide additional resolution and avenues for exploration within this important feature of *Vibrio* spp., with implications for their ecology and environmental survival [13, 91]. Indeed, using a streptomycin resistant strain of *V. parahaemolyticus* and an optimized transposon mutagenesis system, I performed whole-genome mutagenesis followed by transposon-junction sequencing on transposon mutants cultured in either glucose or colloidal chitin containing minimal media. Here, I confirmed a number of essential genes for chitin catabolism in *V. parahaemolyticus* annotated as necessary in previous studies (Figure 5.4) [27, 76, 87, 455]. This experiment also identified an uncharacterized porin protein, genes in the general secretion pathway (T2SS), and an uncharacterized MurR/RpiR transcriptional regulator as conditionally essential for growth on colloidal chitin. While this study provides insight into chitin catabolism in *V. parahaemolyticus*, it suffers from being only from a single replicate. This experiment will need to be repeated to confirm results, along with more direct genetic studies to confirm the necessity of individual genetic factors, such as the uncharacterized porin and MurR/RpiR transcriptional regulator, as well as some genes in the general secretion pathway.

In this chapter, and to close this thesis, I will explore the data presented and provide next steps for each study moving forward. As conclusion, I will explore the inter-connected nature of the studies presented and how these studies identify that prokaryotic biology – specifically their ecology, environmental survival, and

pathogenesis – cannot and should not be studied in isolation. While some of the questions from Chapter 2 are dealt with in Chapter 3, additional questions surrounding the activation of T3SS-1 remain, and it is unclear how the other genes identified in the transposon screen in Chapter 2 – namely *merR* and *phoX* – are involved in T3SS-1 activation. Chapter 3 provides insight into how HlyU functions to de-repress H-NS, but leaves open questions about the mechanism of H-NS de-repression, and how HlyU interacts with DNA at a DNA cruciform to achieve de-repression. In Chapter 4, we explored NagC and Mlc function in *V. parahaemolyticus* and identified biofilm formation and chitin utilization as under the regulation of NagC. However, understanding how NagC influences biofilm formation and chitin utilization will require further study. Chapter 4 also identified NagC and CRP as involved in T3SS-1 expression, but additional studies are necessary to confirm *hlyU* regulation by CRP, and how a *nagC* null mutant disrupts cAMP-homeostasis. In Chapter 5, the design and implementation of genetic studies to confirm the necessity of identified factors in the chitin utilization pathway is required.

6.2 The T3SS-1 Regulatory Paradigm and HlyU

In Chapter 2, we identified an upstream activator of *exsA*, HlyU, and implicated it in the de-repression of H-NS at the *exsBA* intergenic region. In Chapter 3, we implicated a DNA cruciform structure adjacent to HlyU binding at *exsA* as involved in *exsA* regulation. In Chapter 4, I identified that HlyU expression itself appears to be under the regulation of CRP-cAMP, although additional work is required to confirm this. These studies enhance the understanding of *exsA* activation in *V. parahaemolyticus*, and bring us closer to understanding how environmental signals which induce T3SS-1 activation drive *exsA* expression (Figure 6.1).

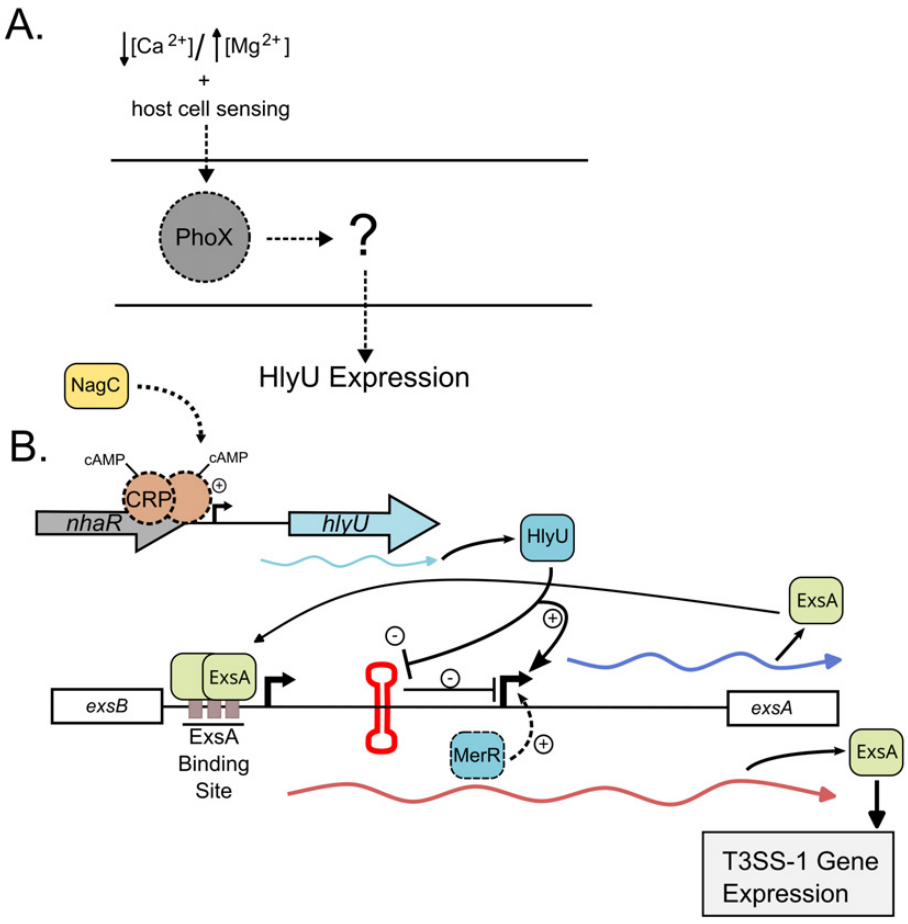

Figure 6.1: T3SS-1 Regulation in *V. parahaemolyticus*. Dashed lines around proteins indicate that they are putative/proposed roles for the protein but have yet to be proven. Dashed lines indicate an unclear connection between two steps. A. Exposure to low Ca^{2+} and high Mg^{2+} conditions, as well as host cell surface sensing, generate a signal to *hlyU* expression, through an as yet undefined mechanism. We identified *phoX* in our transposon screen (Figure A.2), and it is a known periplasmic phosphatase which appears to be under the regulation of Ca^{2+}, and therefore may be involved in the sensing of T3SS-1 inducing conditions. Many of the steps between conditions and signal transduction to *hlyU* expression are unknown, and are labelled with question marks and dashed lines. B. HlyU is expressed upon exposure to inducing conditions, possibly through regulation by CRP-cAMP, where cAMP itself appears to be regulated by NagC through an unknown mechanism (Figure 5.6C and D). Following, HlyU binds DNA at the *exsBA* intergenic region, attenuating a repressive DNA cruciform structure, through an unknown mechanism. The internal promoter is activated, potentially through the aid of MerR, and ExsA is expressed. ExsA is auto-regulatory, binding to its own upstream promoter as a dimer, and drives robust expression of its own gene. ExsA goes on to activate genes involved in T3SS-1 biogenesis.

However, while the intention of Chapter 2 was to better understand how T3SS-1 inducing conditions drive expression of *exsA* and subsequently the T3SS-1, we only explored *exsA* activation by HlyU, in part because of the robustness of the phenotypes associated with *hlyU* null mutations. Other genes were discovered in our transposon mutagenesis screen, and were reported in the publication associated with chapter 2, but were not explored in detail. A putative HTH-regulator (*vpa1473*) overlapping the open-reading frame of *merR* (*vpa1472*), as well as *phoX* (*vp0179*), were identified in the moderate luminescence group, which generated less luciferase expression than wildtype, but not as severely reduced as *hlyU* mutants (Figure A.1). MerR family transcriptional regulators have been identified in the activation of sigma70-dependent promoters, with suboptimal spacers between the -10 and -35 elements [513]. Interestingly, the locus of enterocyte effacement (LEE)-encoded regulator GrlA behaves in a similar fashion, driving expression from promoters with a sub-optimal 18-bp spacer between -10 and -35 elements in enteropathogenic *E. coli* (EPEC) [342, 514, 515]. Considering the discovery of a new, weaker, promoter for *exsA* expression downstream of *hlyU* (see Figure 3.7 and Figure B.7), it is tempting to hypothesize that perhaps polar effects from the transposition event into the upstream *vpa1473*, might disrupt *merR* expression and prevent reliable expression of the internal, cryptic promoter found in the *exsBA* intergenic region. The other gene, PhoX, is a periplasmic phosphatase which is found broadly distributed in marine prokaryotes, and appears to function in the assimilation of environmental phosphate [516, 517]. PhoX activity is stimulated by calcium sensing in other bacteria [518]. Given the activation conditions for *exsA* expression and T3SS-1 biogenesis include low Ca^{2+} conditions, PhoX may represent a functional step in the sensing of calcium conditions, and transducing that signal to expression of *exsA* or factors which activate *exsA* (Figure 6.1A). Additional genetic deletion analyses on these genes will need to be conducted to confirm their roles in *exsA* expression

and T3SS-1 biogenesis.

Conversely, a putative repeats-in-toxin gene *rtx* (*vp1633*) was identified in the high luminescence group, generating significantly more luciferase expression than wildtype. An insertion mutant in a putative *rtx* gene generated a dysregulated phenotype, that unlike H-NS, did not have a significant growth defect (Table A.1). However, since publication of this Chapter in 2018, this gene appears to have been updated as a pseudogene in the NCBI database, as it contains a stop codon in the open reading frame, and is therefore missing the C-terminus [139, 269]. Without further exploration, it is difficult to understand how this gene might be involved in *exsA* expression, if indeed the annotated gene forms a functional protein product.

While our transposon screen identified the most dramatic candidates involved in *exsA* expression, *hlyU* and *hns*, it is unlikely that we identified every gene involved in *exsA* expression. Further screening of the transposon library may identify additional candidates involved in the transduction of the environmental signal to *exsA*. Another approach would be to replicate the *exsA-luxCDABE* luciferase reporter transposon screen presented in Chapter 2, except using the promoter region for *hlyU* to identify regulators of *hlyU* in the same inducing conditions. As with *exsA*, this should identify regulators directly involved in the regulation of *hlyU*, and may identify regulators involved in the sensing of environmental conditions.

The initially identified HlyU-protected region in the *exsBA* contained a 14-bp imperfect palindrome roughly centered in the protected region, as well as a number of inverted repeat sequences, which have the propensity to form DNA superstructures, namely cruciform structures. It seems clear that a DNA cruciform exists at this site, that HlyU-binding partially overlaps the formation of the cruciform structure, and that this DNA cruciform is involved in the regulation of *exsA*. It is not clear, however, how HlyU and the cruciform interact to de-repress H-NS at the *exsBA* intergenic region. As explored in the discussion from Chapter

3, it is tempting to hypothesize that HlyU might bind to the bent DNA found at the base of the cruciform structure, however, we do not provide data supporting this hypothesis directly. On the contrary, HlyU appears to bind linear DNA *in vitro*, and it therefore seems more likely that HlyU binds to linear DNA (Figures 2.6 and 3.3). DNA cruciform binding assays, using EMSA, have been developed for HMGB-box family proteins which are known to bind bent DNA with high-affinity, and have been successfully deployed in other studies [374, 519–522]. We attempted these experiments, but had difficulty preparing a synthetic cruciform *in vitro*. Further optimization and exploration of DNA cruciform binding by HlyU using synthetic cruciforms which recapitulate the *exsBA* localized cruciforms *in vitro* may provide clarity on whether HlyU might bind to the cruciform structure itself.

Finally, HlyU itself may be under the control of catabolite repression through cAMP-receptor protein (CRP) and cAMP, as presented in Chapter 4 (Figure 4.5 and Figure 6.1B). Although the data presented here is preliminary, the regulation of T3SS-1 via the cAMP pathway is perhaps not surprising, given an analogous T3SS in *P. aeruginosa* is also regulated by cAMP through a protein called Vfr [403]. However, in *P. aeruginosa*, Vfr binds directly to the promoter region of *exsA* to direct expression of *exsA* in the presence of cAMP. Our data indicates that CRP-cAMP does not regulate *exsA* directly in *V. parahaemolyticus*, but rather indirectly through a necessary de-repressor of *exsA* (Figure 6.1B). In *E. coli*, cAMP regulates the glucose response; in the absence of glucose transport across the membrane, an adenylate cyclase generates cAMP, which binds to CRP and represses and activates genes involved in other catabolite pathways [390]. However, how cAMP production is regulated in *P. aeruginosa* and *V. parahaemolyticus*, and what the conditions are that drive cAMP increases during virulence, are unclear [301, 403]. It is clear however, that cAMP levels increase intracellularly in *P. aeruginosa* during T3SS-1 inducing conditions (low Ca^{2+} and

high salt) [402, 523]. Further study, as indicated in the discussion for Chapter 4, is required to understand the regulation of cAMP in *V. parahaemolyticus* and to confirm the role of CRP-cAMP binding in the expression of HlyU, including but not limited to cAMP level characterization in *exsA* inducing conditions, protein secretion and host-cell cytotoxicity assays against HeLa cells with the *crp* null mutant, as well as more robust EMSA assays including a CRP-binding site mutant *hlyU* promoter fragment. Given cAMP-CRP binding upstream of the *hlyU* promoter, it appears likely that cAMP enhances the expression of *hlyU*, and I expect that cAMP levels will increase inside the *V. parahaemolyticus* cell in low Ca^{2+}, high Mg^{2+} conditions, as is the case in *P. aeruginosa* [511]. Further, given the delayed *exsA* expression phenotypes, *crp* null mutants should also lack effective T3SS-1 dependent protein secretion and host cell killing. However, it is possible that *crp* null mutants will mirror *nagC* mutants in this regard, where *exsA* expression is slightly reduced, but is still expressed enough for protein secretion and cytotoxicity phenotypes similar to wildtype.

While previous studies have shown that *exsA* is auto-regulatory, and repressed by H-NS, there were considerable gaps in understanding how *exsA* was activated at the genetic level during inducing conditions to drive T3SS-1 biogenesis (Figure 6.1). Here, I have identified novel factors involved in the expression of *exsA*, including HlyU-binding and a DNA cruciform within the *exsBA* intergenic region, and cAMP-CRP regulation of *hlyU* (Figure 6.1). Genes *merR* and *phoX* may also be involved in the activation of *exsA*, although their roles remain unclear and future exploration is necessary (Figure 6.1). Taken together, this thesis has expanded the knowledge surrounding *exsA* expression, and has attempted to bridge the gap in knowledge surrounding how *exsA*, and subsequent T3SS-1 biogenesis, occurs following exposure to low Ca^{2+}, high Mg^{2+} conditions (Figure 6.1).

6.3 Chitin Metabolism in *V. parahaemolyticus*

In Chapters 4 and 5, I diverted from *V. parahaemolyticus* T3SS-1 regulation, and explored the function of two sugar catabolism genes in *V. parahaemolyticus* (NagC and Mlc), as well as the necessary genes required for *V. parahaemolyticus* growth on colloidal chitin. NagC and Mlc in *V. cholerae* have roles in GlcNAc sensing and regulation of catabolism of GlcNAc, through direct binding to the GlcNAc-6-P metabolite [425]. These two studies addressed chitin catabolism and its regulation through a directed approach (Chapter 4) and a global one (Chapter 5). While assaying NagC and Mlc, I found that NagC regulates biofilm formation and chitin utilization as noted by previous studies in other *Vibrio* spp., and that Mlc is not involved in biofilm formation in *V. parahaemolyticus*, contrary to the role of Mlc in *V. cholerae* biofilm formation [421, 423, 426–428]. In Chapter 5, I confirmed the role of a number of core chitin catabolism genes as necessary for *V. parahaemolyticus* growth on colloidal chitin, as well as implicated an uncharacterized MurR/RpiR transcriptional regulator, an uncharacterized porin gene, and some genes in the general secretion pathway (Type II Secretion), as necessary for this growth as well.

NagC is implicated in the expression of aminosugar metabolism genes in *E. coli*, primarily in the repression of the *nagE-nagBACD* divergent promoters [419]. In *Vibrio* spp., NagC has been implicated in chitin utilization and biofilm formation [421, 423, 427, 428]. In Chapter 4, I confirmed that NagC null mutants are unable to form biofilms on glass tubes, and they are de-repressed for chitinase expression, as shown by increased zone-of-clearing sizes (Figure 4.3A and B). While these phenotype appear robust and can be genetically complemented, how biofilm and the chitinase genes are regulated by NagC is not clear. Considering NagC-FLAG can be reliably expressed in *V. parahaemolyticus*, as shown with *nagC* complementation in Chapter 4, the NagC-FLAG protein could be used in

a Chromatin-immunoprecipitation Sequencing (ChIP-Seq) experiment to identify the genomic binding sites for NagC (Figure 6.2). Previous studies have identified that a conserved GR motif in the linker region of the HTH domain of NagC and Mlc are essential for recognition of their binding sites. Mutating the NagC linker sequence position G16 to R or A, and R17 to G or A leads to *nagE* expression in *E. coli*, and demonstrates that NagC cannot bind to DNA with these mutations [443]. These sites translate to the conserved GR motif in *V. parahaemolyticus* NagC at amino acids G69 and R70. Therefore, using the wildtype *nagC* gene, as well as a G69A,R70A mutant *nagC* non-specific binding control, in a ChIP-Seq experiment would identify the DNA-binding sites for NagC, and provide insight into how these processes are regulated by NagC (Figure 6.2). This experiment would involve expressing FLAG-tagged versions of the NagC proteins in *V. parahaemolyticus*, followed by cross-linking the proteins to their DNA binding sites using a crosslinking agent like formaldehyde *in vivo* [524, 525]. Following, the DNA is isolated and fragmented, NagC-bound DNA is enriched using immunoprecipitation (IP) with a FLAG-specific antibody, and the samples are protease treated to remove the protein component (Figure 6.2). Finally, DNA can be sequenced, to identify the sequences which were bound to both the NagC bound regions and non-specific control to identify binding sites for NagC across the genome (Figure 6.2). A number of tools exist for analyzing ChIP-Seq data, including MACS, CisGenome, and MOSAiCS for single samples, and DBChIP for comparing samples, as we would do here [525–529]. Here, I would expect to identify the *nagE-nagBC* divergent promoters, as well as the *nagA* promoter based on the literature, and also identify chitinase and biofilm related genes, as demonstrated by data presented in Chapter 4 [427].

The global survey of chitin catabolism genes identified a number of interesting

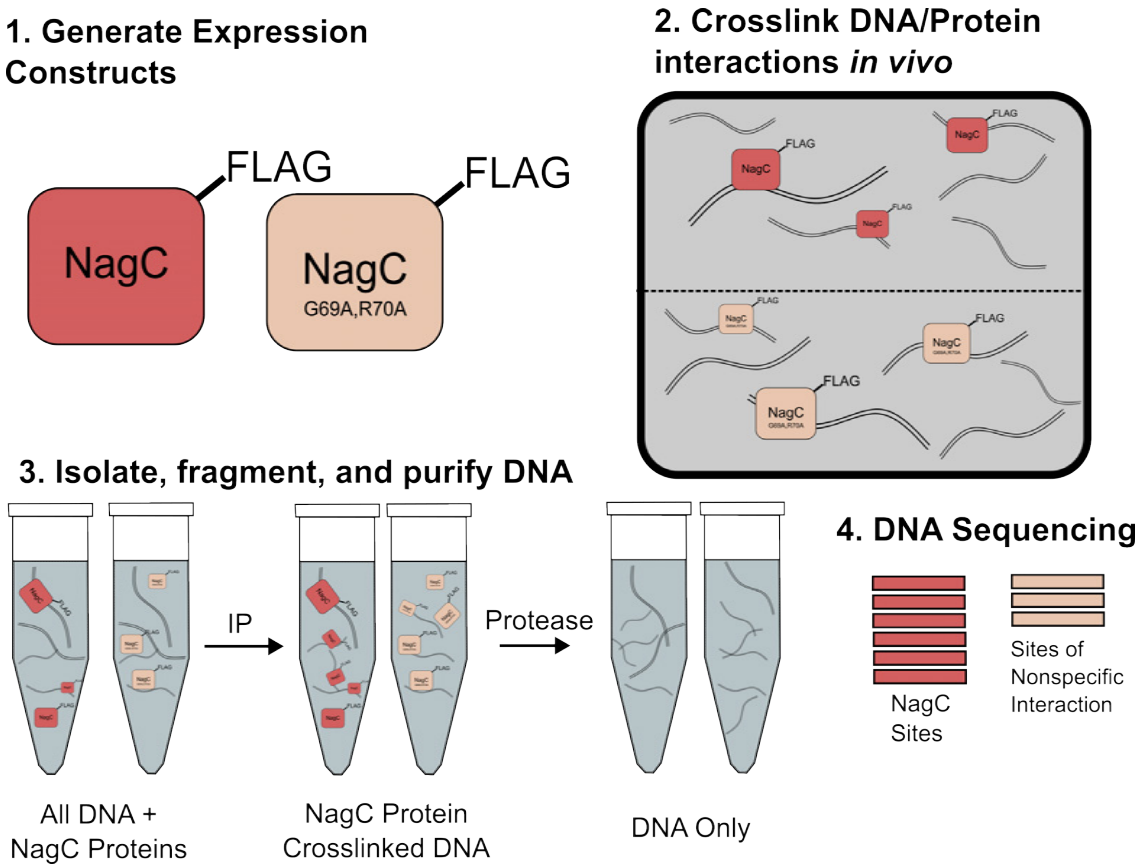

Figure 6.2: ChIP-Seq Workflow for NagC Regulon in *V. parahaemolyticus*. 1. FLAG-tagged genes are expressed in *V. parahaemolyticus*. This includes both the wildtype protein, and a binding-deficient G69A,R70A mutant. 2. Proteins are expressed in individual strains, shown here by a dotted line, and crosslinked to their DNA binding sites inside the cell. 3. DNA is isolated and fragmented, following by immunoprecipitation (IP) to purify DNA bound to FLAG-tagged protein. Following, protease treatment destroys crosslinked protein and leaves only DNA behind. 4. DNA sequencing of the remaining DNA can identify where the proteins were bound, as well as the non-specific interactions identified by the mutant protein.

and novel genes involved in the survival of *V. parahaemolyticus* on colloidal chitin. Here, we identified three novel genes/operons as involved in the chitin catabolism pathway, an uncharacterized porin gene and a MurR/RpiR family transcriptional regulator, as well as the *gsp* operon.

Beginning with the *gsp* operon, many of the genes characterized as T2SS components, including most of the core components and the major pseudopilin subunit (Figure 5.6), were essential for growth on chitin. Although this has not been confirmed for many of the chitinase genes in *V. parahaemolyticus*, chitinase genes are secreted via the T2SS in other bacteria, into the extracellular space, to degrade large chitin molecules into oligomers which can be taken up into the cell [477]. It appears likely that T2SS-dependent secretion of chitinases increases fitness during growth in colloidal chitin. Particularly interesting here is the identification of a number of minor pseudopilins (*gspH*, *gspI*, *gspJ*), as well as a few core structural genes at the inner membrane (*gspM* and *gspL*), which do not seem be conditionally essential for growth in colloidal chitin like other core T2SS components (Table 5.8). Recent reviews on T2SS biology have indicated that knowledge surrounding T2SS minor pseudopilins and their functions are still somewhat unclear, although they are involved in the initiation of pseudopilus elongation [200, 491, 530]. As well, in *P. aeruginosa*, minor pseudopilin subunits aid in substrate loading into the T2SS [492, 493]. Our finding that some minor pseudopilins are not essential may provide some clarity in the function of these minor pseudopilins, and may identify that some psuedopilins, for example GspK, are essential for chitinase secretion, while other minor pseudopilins are not.

To better resolve this, and to confirm the findings in the TnSeq study presented in Chapter 5, I propose a comprehensive mutant analysis of the *gsp* operon, taking particular care to delete genes while preserving the operon function and avoiding polar effects (Figure 6.3). This can be done by allelic exchange, as performed previously and presented in this thesis, with pRE112 allelic exchange

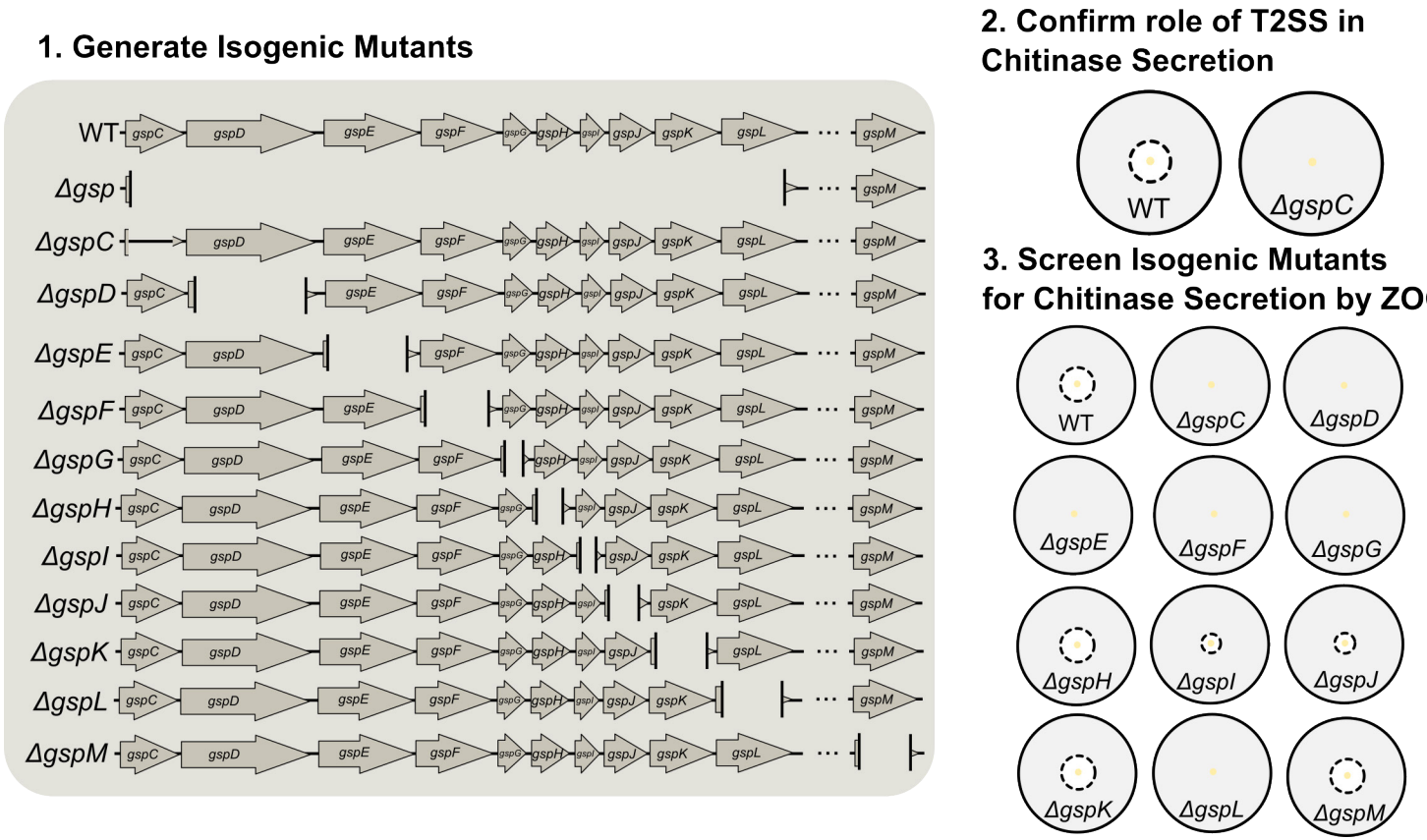

Figure 6.3: T2SS Essential Gene Screen for Chitinase Secretion in *V. parahaemolyticus*. Allelic exchange is performed to generate the mutants shown in 1. These mutants include the entire operon in Δ*gsp* and individual gene mutants. The genetic deletions leave behind N-terminal and C-terminal coding regions to avoid polar effects of the deletion. 2. Following, the wildtype (WT) and Δ*gsp* deletion strain are screened for chitinase activity using a zone-of-clearing assay. 3. Individual mutants can be assayed for their secretion of chitinases using the same zone-of-clearing assay. Various zones-of-clearing are expected based on the gene-specific results from the TnSeq experiment as shown in Table 5.8.

vectors generated quickly by gibson assembly. To avoid polar effects, the new deletion alleles should retain a few codons at both the N-terminal and C-terminal coding sequence (Figure 6.3). After generation of individual mutants, the T2SS can be assayed for its ability to secrete chitinases by a zone-of-clearing assay on agar plates containing colloidal chitin, as well as another carbon source to allow robust growth in the absence of chitinase secretion (See Section 4.4.8 and Figure 6.3). While glucose may be an acceptable carbon source, it is possible that chitinase secretion may be under the control of something that resembles catabolite repression in *E. coli*, and so the selected carbon source may need to be optimized with wildtype *V. parahaemolyticus* RIMD2210633 prior to the mutant assays. Firstly, this experiment can confirm that it is indeed chitinases that are secreted through the T2SS, which in turn generates the growth defect seen in the TnSeq analysis (Figure 6.3). If this is indeed the case, I expect that the core T2SS genes, for the outer membrane complex (*gspD*), the inner membrane complex (*gspC*, *gspF*, *gspL*), the cytoplasmic ATPase (*gspE*), and the major pseudopilin (*gspG*) will be essential for generating zones-of-clearing on agar plates supplemented with colloidal chitin, confirming the necessary genes from the TnSeq study. However, the minor pseudopilin deletions will provide more clarity on whether these pseudopilins are truly expendable for chitinase secretion, or if this is an artifact in our TnSeq data and growth conditions (Figure 6.3).

The proposed functions of the MurR/RpiR transcriptional regulator and the uncharacterized porin are a bit less clear. MurR is a MurR/RpiR transcriptional regulator found in *E. coli* are involved in MurNAc catabolism in *E. coli*, and is regulated in a similar fashion to NagC, where binding to MurNAc-6-P reduced its affinity for DNA binding [482]. In *Bacillus subtilus*, MurR was identified as a regulator for NagZ, a beta-N-acetylglucosaminidase that was essential for cell wall processing [531]. Considering the role of beta-N-acetylglucosaminidases in the catabolism of GlcNac oligomers, perhaps the MurR/RpiR transcriptional regulator

controls one or more beta-N-acetylglucosaminidase genes, such as NagZ [27, 455]. Further, the uncharacterized porin gene likely allows uptake of small aminosugar molecules, like GlcNac monomers, and perhaps contributes to the movement of GlcNAc oligomers that are not imported through the chitoporin [485]. Following a similar approach to the *gsp* operon exploration, individual gene deletions will need to be carried out to confirm the role of these genes in chitin catabolism, and will open up avenues of exploration on how exactly these genes contribute in *V. parahaemolyticus*.

6.4 Final Remarks: Virulence and Metabolism

Understanding broad prokaryotic biology is essential to predicting, controlling, and mitigating damage associated with the emergence of new bacterial pathogens, and outbreaks of existing bacterial pathogens [30, 91, 99]. While this thesis has focused specifically on *Vibrio* spp., this is not unique to *Vibrio*. A recent transposon mutagenesis study in *Mycobacterium tuberculosis* identified SNPs in metabolic genes of *M. tuberculosis* that predicted significant differences in their ability to survive antibiotic stress [532]. Indeed, links between sugar metabolism and virulence are widespread in bacterial and fungal pathogens [533–538]. This thesis also demonstrates links between metabolism and virulence, as I have identified PhoX, a periplasmic phosphatase involved in phosphate metabolism, as connected to the expression of *exsA* in inducing conditions. As well, HlyU appears to be under the control of cAMP, which may link HlyU expression to carbon sensing through PTS transport of certain sugars. While the two focus points of this thesis, chitin catabolism and T3SS-1 regulation, may appear disparate, they are both important to understanding how *Vibrio* survive and replicate in the environment and their human hosts, as well as how *Vibrio* deploy the molecular mechanisms at their disposal to cause disease.

Anthropogenic climate change is driving broad and diverse transformations to the world's oceans. *Vibrio* spp. infections – as well as new outbreaks and strains of infectious *Vibrio* – are becoming more prevalent [34, 150]. Here, I clarified, in part, how *V. parahaemolyticus* catabolizes an important carbon and nitrogen source in the environment, and how this organism integrates environmental signals to activate the highly cytotoxic T3SS-1. We identified HlyU as a necessary regulator for the master-transcriptional T3SS-1 activator ExsA, and furthered our understanding of how *exsA* is activated at the genetic level by implicating a novel internal promoter and a DNA cruciform in *exsA* regulation. Further, I identified that NagC is involved in biofilm formation, chitin utilization, and *exsA* expression in *V. parahaemolyticus*, and that the impact *nagC* null mutations have on *exsA* expression might be due to CRP-cAMP regulation of HlyU. Finally, I identified novel factors involved in chitin catabolism in *V. parahaemolyticus*, namely an uncharacterized porin and a MurR/RpiR family transcriptional regulator. While this work stands to further understanding of the biology of *V. parahaemolyticus*, future studies will be required to address questions raised by this work. Ultimately, understanding the broad biology of *Vibrio* pathogens, like *V. parahaemolyticus*, will continue to be important to mitigating, predicting, and managing future outbreaks of these consequential organisms.

References

1. Bar-On, Y. M., Phillips, R. & Milo, R. The Biomass Distribution on Earth. *Proceedings of the National Academy of Sciences* **115,** 6506–6511 (2018).
2. Azam, F. & Malfatti, F. Microbial Structuring of Marine Ecosystems. *Nature Reviews Microbiology* **5,** 782–791 (2007).
3. Field, C. B., Behrenfeld, M. J., Randerson, J. T. & Falkowski, P. Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components. *Science* **281,** 237–240 (1998).
4. Behrenfeld, M. J. & Falkowski, P. G. Photosynthetic Rates Derived from Satellite-Based Chlorophyll Concentration. *Limnology and Oceanography* **42,** 1–20 (1997).
5. Cavicchioli, R. *et al.* Scientists' Warning to Humanity: Microorganisms and Climate Change. *Nature Reviews Microbiology* **17,** 569–586 (2019).
6. Bourne, D. G., Morrow, K. M. & Webster, N. S. Insights into the Coral Microbiome: Underpinning the Health and Resilience of Reef Ecosystems. *Annual Review of Microbiology* **70,** 317–340 (2016).
7. Thurber, R. V. *et al.* Metagenomic Analysis of Stressed Coral Holobionts. *Environmental Microbiology* **11,** 2148–2163 (2009).
8. Littman, R., Willis, B. L. & Bourne, D. G. Metagenomic Analysis of the Coral Holobiont during a Natural Bleaching Event on the Great Barrier Reef: Metagenomics Analysis of Bleached Coral. *Environmental Microbiology Reports* **3,** 651–660 (2011).
9. Meron, D. *et al.* The Impact of Reduced pH on the Microbial Community of the Coral *Acropora Eurystoma*. *The ISME Journal* **5,** 51–60 (2011).
10. Howard-Jones, N. Robert Koch and the Cholera *Vibrio*: A Centenary. *British Medical Journal (Clinical research ed.)* **288,** 379–381 (1984).
11. Finkelstein, R. A. in *Medical Microbiology* (ed Baron, S.) Fourth (University of Texas Medical Branch at Galveston, Galveston (TX), 1996).
12. Lippi, D. & Gotuzzo, E. The Greatest Steps towards the Discovery of *Vibrio cholerae*. *Clinical Microbiology and Infection* **20,** 191–195 (2014).
13. Le Roux, F. & Blokesch, M. Eco-Evolutionary Dynamics Linked to Horizontal Gene Transfer in *Vibrios*. *Annual Review of Microbiology* **72,** 89–110 (2018).
14. Destoumieux-Garzón, D. *et al. Vibrio*–Bivalve Interactions in Health and Disease. *Environmental Microbiology* **22,** 4323–4341 (2020).

15. Krediet, C. J., Ritchie, K. B., Paul, V. J. & Teplitski, M. Coral-Associated Micro-Organisms and Their Roles in Promoting Coral Health and Thwarting Diseases. *Proceedings of the Royal Society B: Biological Sciences* **280,** 20122328 (2013).

16. Schryver, P. D., Defoirdt, T. & Sorgeloos, P. Early Mortality Syndrome Outbreaks: A Microbial Management Issue in Shrimp Farming? *PLOS Pathogens* **10,** e1003919 (2014).

17. Santos, H. M. *et al.* Diagnosis and Potential Treatments for Acute Hepatopancreatic Necrosis Disease (AHPND): A Review. *Aquaculture International* **28,** 169–185 (2020).

18. Colwell, R. R. & Grimes, D. J. *Vibrio* Diseases of Marine Fish Populations. *Helgoländer Meeresuntersuchungen* **37,** 265–287 (1984).

19. Erken, M., Lutz, C. & McDougald, D. Interactions of *Vibrio* Spp. with Zooplankton. *Microbiology Spectrum* **3,** 1–15 (2015).

20. Visick, K. L., Stabb, E. V. & Ruby, E. G. A Lasting Symbiosis: How *Vibrio fischeri* Finds a Squid Partner and Persists within Its Natural Host. *Nature Reviews Microbiology* **19,** 654–665 (2021).

21. Hoffmann, M. *et al.* Population Dynamics of *Vibrio* spp. Associated with Marine Sponge Microcosms. *The ISME Journal* **4,** 1608–1612 (2010).

22. Reusch, T. B. H. *et al.* Lower *Vibrio* spp. Abundances in *Zostera marina* Leaf Canopies Suggest a Novel Ecosystem Function for Temperate Seagrass Beds. *Marine Biology* **168,** 149 (2021).

23. Mishra, A. K. & Mohanraju, R. Epiphytic Bacterial Communities in Seagrass Meadows of Oligotrophic Waters of Andaman Sea. *Open Access Library Journal* **5,** 1–12 (2018).

24. Criminger, J. D., Hazen, T. H., Sobecky, P. A. & Lovell, C. R. Nitrogen Fixation by *Vibrio parahaemolyticus* and Its Implications for a New Ecological Niche. *Applied and Environmental Microbiology* **73,** 5959–5961 (2007).

25. Thompson, F. L., Iida, T. & Swings, J. Biodiversity of Vibrios. *Microbiology and Molecular Biology Reviews* **68,** 403–431 (2004).

26. Zhang, X., Lin, H., Wang, X. & Austin, B. Significance of *Vibrio* Species in the Marine Organic Carbon Cycle—A Review. *Science China Earth Sciences* **61,** 1357–1368 (2018).

27. Hunt, D. E., Gevers, D., Vahora, N. M. & Polz, M. F. Conservation of the Chitin Utilization Pathway in the *Vibrionaceae*. *Applied and Environmental Microbiology* **74,** 44–51 (2008).

28. Zampini, M. *et al. Vibrio cholerae* Persistence in Aquatic Environments and Colonization of Intestinal Cells: Involvement of a Common Adhesion Mechanism. *FEMS Microbiology Letters* **244,** 267–273 (2005).

29. Ceccarelli, D. & Colwell, R. R. *Vibrio* Ecology, Pathogenesis, and Evolution. *Frontiers in Microbiology* **5,** 256 (2014).

30. Baker-Austin, C. *et al. Vibrio* Spp. Infections. *Nature Reviews Disease Primers* **4,** 1–19 (2018).

31. Ina-Salwany, M. Y. *et al. Vibrio*sis in Fish: A Review on Disease Development and Prevention. *Journal of Aquatic Animal Health* **31,** 3–22 (2019).

32. Dash, P., Avunje, S., Tandel, R. S., K. P., S. & Panigrahi, A. Biocontrol of Luminous Vibriosis in Shrimp Aquaculture: A Review of Current Approaches and Future Perspectives. *Reviews in Fisheries Science & Aquaculture* **25,** 245–255 (2017).

33. de Souza Valente, C. & Wan, A. H. L. *Vibrio* and Major Commercially Important Vibriosis Diseases in Decapod Crustaceans. *Journal of Invertebrate Pathology* **181,** 107527 (2021).

34. Baker-Austin, C. *et al.* Emerging *Vibrio* Risk at High Latitudes in Response to Ocean Warming. *Nature Climate Change* **3,** 73–77 (2013).

35. Froelich, B. A. & Daines, D. A. In Hot Water: Effects of Climate Change on *Vibrio*–Human Interactions. *Environmental Microbiology* **22,** 4101–4111 (2020).

36. Yamaichi, Y., Iida, T., Park, K.-S., Yamamoto, K. & Honda, T. Physical and Genetic Map of the Genome of *Vibrio parahaemolyticus*: Presence of Two Chromosomes in *Vibrio* Species. *Molecular Microbiology* **31,** 1513–1521 (1999).

37. Egan, E. S., Fogel, M. A. & Waldor, M. K. MicroReview: Divided Genomes: Negotiating the Cell Cycle in Prokaryotes with Multiple Chromosomes. *Molecular Microbiology* **56,** 1129–1138 (2005).

38. Yamaichi, Y., Fogel, M. A., McLeod, S. M., Hui, M. P. & Waldor, M. K. Distinct Centromere-Like parS Sites on the Two Chromosomes of *Vibrio* Spp. *Journal of Bacteriology* **189,** 5314–5324 (2007).

39. Kaneko, T. & Colwell, R. R. Ecology of *Vibrio parahaemolyticus* in Chesapeake Bay. *Journal of Bacteriology* **113,** 24–32 (1973).

40. Vezzulli, L., Pezzati, E., Brettar, I., Höfle, M. & Pruzzo, C. Effects of Global Warming on *Vibrio* Ecology. *Microbiology Spectrum* **3,** 1–9 (2015).

41. Griffitt, K. J. & Grimes, D. J. Abundance and Distribution of *Vibrio cholerae*, *V. parahaemolyticus*, and *V. vulnificus* Following a Major Freshwater Intrusion into the Mississippi Sound. *Microbial Ecology* **65,** 578–583 (2013).

42. Ulitzur, S. *Vibrio parahaemolyticus* and *Vibrio alginolyticus*: Short Generation-Time Marine Bacteria. *Microbial Ecology* **1,** 127–135 (1974).

43. Hoff, J. *et al. Vibrio Natriegens* : An Ultrafast-growing Marine Bacterium as Emerging Synthetic Biology Chassis. *Environmental Microbiology* **22,** 4394–4408 (2020).

44. Dalia, T. N. *et al.* Multiplex Genome Editing by Natural Transformation (MuGENT) for Synthetic Biology in *Vibrio natriegens*. *ACS synthetic biology* **6,** 1650–1655 (2017).

45. Lee, H. H. *et al.* Functional Genomics of the Rapidly Replicating Bacterium *Vibrio natriegens* by CRISPRi. *Nature Microbiology* **4,** 1105–1113 (2019).

46. Westrich, J. R. *et al.* Saharan Dust Nutrients Promote *Vibrio* Bloom Formation in Marine Surface Waters. *Proceedings of the National Academy of Sciences* **113,** 5964–5969 (2016).

47. Westrich, J. R., Griffin, D. W., Westphal, D. L. & Lipp, E. K. *Vibrio* Population Dynamics in Mid-Atlantic Surface Waters during Saharan Dust Events. *Frontiers in Marine Science* **5** (2018).

48. Lukjancenko, O. & Ussery, D. *Vibrio* Chromosome-Specific Families. *Frontiers in Microbiology* **5,** 1–8 (2014).

49. Chimetto, L. A. *et al. Vibrios* Dominate as Culturable Nitrogen-Fixing Bacteria of the Brazilian Coral *Mussismilia hispida*. *Systematic and Applied Microbiology* **31,** 312–319 (2008).

50. Jones, B. W. & Nishiguchi, M. K. Counterillumination in the Hawaiian Bobtail Squid, *Euprymna scolopes* Berry (Mollusca: Cephalopoda). *Marine Biology* **144,** 1151–1155 (2004).

51. Kaper, J. B., Morris, J. G. & Levine, M. M. Cholera. *Clinical Microbiology Reviews* **8,** 48–86 (1995).

52. González-Escalona, N. *et al. Vibrio parahaemolyticus* Diarrhea, Chile, 1998 and 2004. *Emerging Infectious Diseases* **11,** 129–131 (2005).

53. Newton, A., Kendall, M., Vugia, D. J., Henao, O. L. & Mahon, B. E. Increasing Rates of Vibriosis in the United States, 1996–2010: Review of Surveillance Data From 2 Systems. *Clinical Infectious Diseases* **54,** S391–S395 (2012).

54. *Outbreak of Vibrio parahaemolyticus Infections Associated with Eating Raw Oysters – Pacific Northwest, 1997* https://www.cdc.gov/mmwr/preview/mmwrhtml/00053377.htm.

55. Baker-Austin, C., Stockley, L., Rangdale, R. & Martinez-Urtaza, J. Environmental Occurrence and Clinical Impact of *Vibrio vulnificus* and *Vibrio parahaemolyticus*: A European Perspective. *Environmental Microbiology Reports* **2,** 7–18 (2010).

56. Martinez-Urtaza, J. *et al.* Emergence of Asiatic *Vibrio* Diseases in South America in Phase with El Niño. *Epidemiology* **19,** 829–837 (2008).

57. Papke, R. T. & Gogarten, J. P. How Bacterial Lineages Emerge. *Science* **336,** 45–46 (2012).

58. Shapiro, B. J. *et al.* Population Genomics of Early Events in the Ecological Differentiation of Bacteria. *Science* **336,** 48–51 (2012).

59. Metzger, L. C. & Blokesch, M. Regulation of Competence-Mediated Horizontal Gene Transfer in the Natural Habitat of *Vibrio cholerae*. *Current Opinion in Microbiology* **30,** 1–7 (2016).

60. Meibom, K. L., Blokesch, M., Dolganov, N. A., Wu, C.-Y. & Schoolnik, G. K. Chitin Induces Natural Competence in *Vibrio cholerae*. *Science* **310,** 1824–1827 (2005).

61. Lo Scrudato, M. & Blokesch, M. A Transcriptional Regulator Linking Quorum Sensing and Chitin Induction to Render *Vibrio cholerae* Naturally Transformable. *Nucleic Acids Research* **41,** 3644–3658 (2013).

62. Dalia, A. B., Lazinski, D. W. & Camilli, A. Identification of a Membrane-Bound Transcriptional Regulator That Links Chitin and Natural Competence in *Vibrio cholerae*. *mBio* **5,** e01028–01013 (2014).

63. Borgeaud, S., Metzger, L. C., Scrignari, T. & Blokesch, M. The Type VI Secretion System of *Vibrio cholerae* Fosters Horizontal Gene Transfer. *Science* **347,** 63–67 (2015).

64. Thomas, J., Watve, S. S., Ratcliff, W. C. & Hammer, B. K. Horizontal Gene Transfer of Functional Type VI Killing Genes by Natural Transformation. *mBio* **8,** e00654–17 (2017).

65. Matthey, N. *et al.* Neighbor Predation Linked to Natural Competence Fosters the Transfer of Large Genomic Regions in *Vibrio cholerae*. *eLife* **8** (eds Mignot, T. & Weigel, D.) e48212 (2019).

66. Bisharat, N. *et al.* Clinical, Epidemiological, and Microbiological Features of *Vibrio vulnificus* Biogroup 3 Causing Outbreaks of Wound Infection and Bacteraemia in Israel. Israel Vibrio Study Group. *Lancet (London, England)* **354,** 1421–1424 (1999).

67. Efimov, V. *et al.* Insight into the Evolution of *Vibrio vulnificus* Biotype 3's Genome. *Frontiers in Microbiology* **4** (2013).

68. Yamamoto, S. *et al.* Regulation of Natural Competence by the Orphan Two-Component System Sensor Kinase ChiS Involves a Non-Canonical Transmembrane Regulator in *Vibrio cholerae*. *Molecular Microbiology* **91,** 326–347 (2014).

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

70. Ellison, C. K. *et al.* Retraction of DNA-bound Type IV Competence Pili Initiates DNA Uptake during Natural Transformation in *Vibrio cholerae*. *Nature Microbiology* **3,** 773–780 (2018).

71. Seitz, P. *et al.* ComEA Is Essential for the Transfer of External DNA into the Periplasm in Naturally Transformable *Vibrio cholerae* Cells. *PLOS Genetics* **10,** e1004066 (2014).

72. Seitz, P. & Blokesch, M. DNA Transport across the Outer and Inner Membranes of Naturally Transformable *Vibrio cholerae* Is Spatially but Not Temporally Coupled. *mBio* **5,** e01409–14 (2014).

73. Seitz, P. & Blokesch, M. DNA-uptake Machinery of Naturally Competent *Vibrio cholerae*. *Proceedings of the National Academy of Sciences* **110,** 17987–17992 (2013).

74. Blokesch, M. & Schoolnik, G. K. The Extracellular Nuclease Dns and Its Role in Natural Transformation of *Vibrio cholerae*. *Journal of Bacteriology* **190,** 7232–7240 (2008).

75. Cauchie, H.-M. Chitin Production by Arthropods in the Hydrosphere. *Hydrobiologia* **470,** 63–95 (2002).

76. Meibom, K. L. *et al.* The *Vibrio cholerae* Chitin Utilization Program. *Proceedings of the National Academy of Sciences* **101,** 2524–2529 (2004).

77. Boin, M. A., Austin, M. J. & Häse, C. C. Chemotaxis in *Vibrio cholerae*. *FEMS Microbiology Letters* **239,** 1–8 (2004).

78. Bassler, B. L., Gibbons, P. J., Yu, C. & Roseman, S. Chitin Utilization by Marine Bacteria. Chemotaxis to Chitin Oligosaccharides by *Vibrio furnissii*. *Journal of Biological Chemistry* **266,** 24268–24275 (1991).

79. Hirano, T. *et al.* Heterodisaccharide 4-O-(N-acetyl-β-d-glucosaminyl)-d-glucosamine Is an Effective Chemotactic Attractant for *Vibrio* Bacteria That Produce Chitin Oligosaccharide Deacetylase. *Letters in Applied Microbiology* **53,** 161–166 (2011).

80. Mandel, M. J. *et al.* Squid-Derived Chitin Oligosaccharides Are a Chemotactic Signal during Colonization by *Vibrio* Fischeri. *Applied and Environmental Microbiology* **78,** 4620–4626 (2012).

81. Cao, X. *et al.* The Novel Sigma Factor-Like Regulator RpoQ Controls Luminescence, Chitinase Activity, and Motility in *Vibrio fischeri*. *mBio* **3,** e00285–11 (2012).

82. Watnick, P. I., Fullner, K. J. & Kolter, R. A Role for the Mannose-Sensitive Hemagglutinin in Biofilm Formation by *Vibrio Cholerae* El Tor. *Journal of Bacteriology* **181,** 3606–3609 (1999).

83. Moorthy, S. & Watnick, P. I. Genetic Evidence That the *Vibrio cholerae* Monolayer Is a Distinct Stage in Biofilm Development. *Molecular microbiology* **52,** 573–587 (2004).

84. Shime-Hattori, A. *et al.* Two Type IV Pili of *Vibrio parahaemolyticus* Play Different Roles in Biofilm Formation. *FEMS microbiology letters* **264,** 89–97 (2006).

85. Stauder, M. *et al.* Role of GbpA Protein, an Important Virulence-Related Colonization Factor, for *Vibrio cholerae*'s Survival in the Aquatic Environment. *Environmental Microbiology Reports* **4,** 439–445 (2012).

86. Wong, E. *et al.* The *Vibrio cholerae* Colonization Factor GbpA Possesses a Modular Structure That Governs Binding to Different Host Surfaces. *PLoS Pathogens* **8,** e1002373 (2012).

87. Debnath, A., Mizuno, T. & Miyoshi, S.-I. Regulation of Chitin-Dependent Growth and Natural Competence in *Vibrio parahaemolyticus*. *Microorganisms* **8,** 1303 (2020).

88. Suginta, W. *et al.* Molecular Uptake of Chitooligosaccharides through Chitoporin from the Marine Bacterium *Vibrio harveyi*. *PLOS ONE* **8,** e55126 (2013).

89. Klancher, C. A., Yamamoto, S., Dalia, T. N. & Dalia, A. B. ChiS Is a Noncanonical DNA-binding Hybrid Sensor Kinase That Directly Regulates the Chitin Utilization Program in *Vibrio cholerae*. *Proceedings of the National Academy of Sciences* **117,** 20180–20189 (2020).

90. Li, X. & Roseman, S. The Chitinolytic Cascade in *Vibrios* Is Regulated by Chitin Oligosaccharides and a Two-Component Chitin Catabolic Sensor/Kinase. *Proceedings of the National Academy of Sciences* **101,** 627–631 (2004).

91. Pruzzo, C., Vezzulli, L. & Colwell, R. R. Global Impact of *Vibrio cholerae* Interactions with Chitin. *Environmental Microbiology* **10,** 1400–1410 (2008).

92. Matz, C. *et al.* Biofilm Formation and Phenotypic Variation Enhance Predation-Driven Persistence of *Vibrio cholerae*. *Proceedings of the National Academy of Sciences* **102,** 16819–16824 (2005).

93. Chiavelli, D. A., Marsh, J. W. & Taylor, R. K. The Mannose-Sensitive Hemagglutinin of *Vibrio cholerae* Promotes Adherence to Zooplankton. *Applied and Environmental Microbiology* **67,** 3220–3225 (2001).

94. Tarsi, R. & Pruzzo, C. Role of Surface Proteins in *Vibrio cholerae* Attachment to Chitin. *Applied and Environmental Microbiology* **65,** 1348–1351 (1999).

95. Nalin, D. R., Daya, V., Reid, A., Levine, M. M. & Cisneros, L. Adsorption and Growth of *Vibrio cholerae* on Chitin. *Infection and Immunity* **25,** 768–770 (1979).

96. Huq, A. *et al.* A Simple Filtration Method to Remove Plankton-Associated *Vibrio cholerae* in Raw Water Supplies in Developing Countries. *Applied and Environmental Microbiology* **62,** 2508–2512 (1996).

97. Colwell, R. R. *et al.* Reduction of Cholera in Bangladeshi Villages by Simple Filtration. *Proceedings of the National Academy of Sciences* **100,** 1051–1055 (2003).

98. Huq, A. *et al.* Simple Sari Cloth Filtration of Water Is Sustainable and Continues To Protect Villagers from Cholera in Matlab, Bangladesh. *mBio* **1,** e00034–10 (2010).

99. Casadevall, A., Fang, F. C. & Pirofski, L.-a. Microbial Virulence as an Emergent Property: Consequences and Opportunities. *PLoS Pathogens* **7,** e1002136 (2011).

100. FAO. *The State of World Fisheries and Aquaculture 2020: Sustainability in action The State of World Fisheries and Aquaculture (SOFIA)* **2020** (FAO, Rome, Italy, 2020).

101. FAO. *Report of the FAO/MSU/WB First Multi-Stakeholder Consultation on a Progressive Management Pathway to Improve Aquaculture Biosecurity (PMP/AB). FAO Fisheries and Aquaculture Report* **1254** (FAO, Rome, Italy, 2019).

102. Novriadi, R. Vibriosis in Aquaculture. *Omni-Akuatika* **12** (2016).

103. Chatterjee, S. & Halder, Soumya. *Vibrio* Related Diseases in Aquaculture and Development of Rapid and Accurate Identification Methods. *Journal of Marine Science: Research & Development* **s1** (2012).

104. Jayasree, L., Janakiram, P. & Madhavi, R. Characterization of *Vibrio* Spp. Associated with Diseased Shrimp from Culture Ponds of Andhra Pradesh (India). *Journal of the World Aquaculture Society* **37,** 523–532 (2006).

105. Soto-Rodríguez, S. A., Simoes, N., Roque, A. & Gómez Gil, B. Pathogenicity and Colonization of *Litopenaeus vannamei* Larvae by Luminescent *Vibrios*. *Aquaculture* **258,** 109–115 (2006).

106. Mancuso, M. *et al.* Characterization of Chitinolytic Bacteria and Histological Aspects of Shell Disease Syndrome in European Spiny Lobsters (*Palinurus elephas*) (Fabricius 1787). *Journal of Invertebrate Pathology* **104,** 242–244 (2010).

107. Cawthorn, R. J. Diseases of American Lobsters (*Homarus americanus*): A Review. *Journal of Invertebrate Pathology. Diseases of Edible Crustaceans* **106,** 71–78 (2011).

108. Sullivan, T. J. & Neigel, J. E. Effects of Temperature and Salinity on Prevalence and Intensity of Infection of Blue Crabs, *Callinectes sapidus*, by *Vibrio cholerae*, *V. parahaemolyticus*, and *V. vulnificus* in Louisiana. *Journal of Invertebrate Pathology* **151,** 82–90 (2018).

109. Radhakrishnan, E. V. & Kizhakudan, J. K. in *Lobsters: Biology, Fisheries and Aquaculture* (eds Radhakrishnan, E. V., Phillips, B. F. & Achamveetil, G.) 571–601 (Springer, Singapore, 2019).

110. Petton, B. *et al.* The Pacific Oyster Mortality Syndrome, a Polymicrobial and Multifactorial Disease: State of Knowledge and Future Directions. *Frontiers in Immunology* **12,** 1–10 (2021).

111. Cheng, A.-C., Cheng, S.-A., Chen, Y.-Y. & Chen, J.-C. Effects of Temperature Change on the Innate Cellular and Humoral Immune Responses of Orange-Spotted Grouper *Epinephelus coioides* and Its Susceptibility to *Vibrio alginolyticus*. *Fish & Shellfish Immunology* **26,** 768–772 (2009).

112. Kumar, V., Roy, S., Behera, B. K., Bossier, P. & Das, B. K. Acute Hepatopancreatic Necrosis Disease (AHPND): Virulence, Pathogenesis and Mitigation Strategies in Shrimp Aquaculture. *Toxins* **13,** 524 (2021).

113. Lee, C.-T. *et al.* The Opportunistic Marine Pathogen *Vibrio parahaemolyticus* Becomes Virulent by Acquiring a Plasmid That Expresses a Deadly Toxin. *Proceedings of the National Academy of Sciences* **112,** 10798–10803 (2015).

114. Dhar, A. K. *et al.* First Report of Acute Hepatopancreatic Necrosis Disease (AHPND) Occurring in the USA. *Diseases of Aquatic Organisms* **132,** 241–247 (2019).

115. Fao, R. & Who, G. *Risk Assessment of Vibrio vulnificus in Raw Oysters. Interpretative Summary and Technical Report* tech. rep. (Rome (Italy) FAO/WHO, 2005).

116. FAO/WHO. *Risk Assessment of Vibrio parahaemolyticus in Seafood. Interpretative Summary and Technical Report. Microbiological Risk Assessment (MRA) 16 Microbiological Risk Assessment Series (FAO/WHO)* **16** (FAO / WHO, Rome, Italy, 2011).

117. Ralston, E. P., Kite-Powell, H. & Beet, A. An Estimate of the Cost of Acute Food and Water Borne Health Effects from Marine Pathogens and Toxins in the United States. *Journal of water and health* **9,** 680–694 (2011).

118. Weis, K. E., Hammond, R. M., Hutchinson, R. & Blackmore, C. G. M. *Vibrio* Illness in Florida, 1998–2007. *Epidemiology & Infection* **139,** 591–598 (2011).

119. Rippey, S. R. Infectious Diseases Associated with Molluscan Shellfish Consumption. *Clinical Microbiology Reviews* **7,** 419–425 (1994).

120. Jones, M. K. & Oliver, J. D. *Vibrio vulnificus*: Disease and Pathogenesis. *Infection and Immunity* **77,** 1723–1733 (2009).

121. WHO. *Cholera* https://www.who.int/news-room/fact-sheets/detail/cholera.

122. Sack, D. A. *et al.* Contrasting Epidemiology of Cholera in Bangladesh and Africa. *The Journal of Infectious Diseases* **224,** S701–S709 (2021).

123. Saha, G. K. & Ganguly, N. K. Spread and Endemicity of Cholera in India: Factors Beyond the Numbers. *The Journal of Infectious Diseases* **224,** S710–S716 (2021).

124. Van Heyningen, W. E., Van Heyningen, S. & King, C. A. The Nature and Action of Cholera Toxin. *Ciba Foundation Symposium,* 73–88 (1976).

125. Cassel, D. & Pfeuffer, T. Mechanism of Cholera Toxin Action: Covalent Modification of the Guanyl Nucleotide-Binding Protein of the Adenylate Cyclase System. *Proceedings of the National Academy of Sciences* **75,** 2669–2673 (1978).

126. Rivera-Chávez, F. & Mekalanos, J. J. Cholera Toxin Promotes Pathogen Acquisition of Host-Derived Nutrients. *Nature* **572,** 244–248 (2019).

127. Gavin, H. E., Beubier, N. T. & Satchell, K. J. F. The Effector Domain Region of the *Vibrio vulnificus* MARTX Toxin Confers Biphasic Epithelial Barrier Disruption and Is Essential for Systemic Spread from the Intestine. *PLOS Pathogens* **13,** e1006119 (2017).

128. Wright, A. C., Simpson, L. M. & Oliver, J. D. Role of Iron in the Pathogenesis of *Vibrio vulnificus* Infections. *Infection and Immunity* **34,** 503–507 (1981).

129. Centres for Disease Control. *Vibrio vulnificus* Infections Associated with Raw Oyster Consumption – Florida, 1981-1992. *Morb. Mortal. Wkly. Rep.* **42,** 405–407 (1993).

130. Reilly, G. D., Reilly, C. A., Smith, E. G. & Baker-Austin, C. *Vibrio alginolyticus*-Associated Wound Infection Acquired in British Waters, Guernsey, July 2011. *Eurosurveillance* **16,** 19994 (2011).

131. Ramamurthy, T., Chowdhury, G., Pazhani, G. & Shinoda, S. *Vibrio fluvialis*: An Emerging Human Pathogen. *Frontiers in Microbiology* **5** (2014).

132. Vezzulli, L. *et al.* Long-Term Effects of Ocean Warming on the Prokaryotic Community: Evidence from the *Vibrios*. *The ISME Journal* **6,** 21–30 (2012).

133. T, F. On the Bacteriological Examination of Shirasu-Food Poisoning. *Med. J. Osaka Univ.* **4,** 299–304 (1953).

134. Daniels, N. A. *et al. Vibrio parahaemolyticus* Infections in the United States, 1973–1998. *The Journal of Infectious Diseases* **181,** 1661–1666 (2000).

135. Okuda, J. *et al.* Emergence of a Unique O3:K6 Clone of *Vibrio parahaemolyticus* in Calcutta, India, and Isolation of Strains from the Same Clonal Group from Southeast Asian Travelers Arriving in Japan. *Journal of Clinical Microbiology* **35,** 3150–3155 (1997).

136. Daniels, N. A. *et al.* Emergence of a New *Vibrio parahaemolyticus* Serotype in Raw Oysters: A Prevention Quandary. *JAMA* **284,** 1541–1545 (2000).

137. Martinez-Urtaza, J. *et al.* Pandemic *Vibrio parahaemolyticus* O3:K6, Europe. *Emerging Infectious Diseases* **11,** 1319–1320 (2005).

138. Matsumoto, C. *et al.* Pandemic Spread of an O3:K6 Clone of *Vibrio parahaemolyticus* and Emergence of Related Strains Evidenced by Arbitrarily Primed PCR and *toxRS* Sequence Analyses. *Journal of Clinical Microbiology* **38,** 578–585 (2000).

139. Makino, K. *et al.* Genome Sequence of *Vibrio parahaemolyticus*: A Pathogenic Mechanism Distinct from That of V Cholerae. *The Lancet* **361,** 743–749 (2003).

140. Myers, M. L., Panicker, G. & Bej, A. K. PCR Detection of a Newly Emerged Pandemic *Vibrio parahaemolyticus* O3:K6 Pathogen in Pure Cultures and Seeded Waters from the Gulf of Mexico. *Applied and Environmental Microbiology* **69,** 2194–2200 (2003).

141. Tuyet, D. T. *et al.* Clinical, Epidemiological, and Socioeconomic Analysis of an Outbreak of *Vibrio parahaemolyticus* in Khanh Hoa Province, Vietnam. *The Journal of Infectious Diseases* **186,** 1615–1620 (2002).

142. Velazquez-Roman, J., León-Sicairos, N., de Jesus Hernández-Díaz, L. & Canizalez-Roman, A. Pandemic *Vibrio parahaemolyticus* O3:K6 on the American Continent. *Frontiers in Cellular and Infection Microbiology* **3,** 110 (2014).

143. Li, Y. *et al. Vibrio parahaemolyticus*, Southern Coastal Region of China, 2007–2012. *Emerging Infectious Diseases* **20,** 685–688 (2014).

144. Park, K. *et al.* Food-Borne Outbreaks, Distributions, Virulence, and Antibiotic Resistance Profiles of *Vibrio parahaemolyticus* in Korea from 2003 to 2016: A Review. *Fisheries and Aquatic Sciences* **21,** 3 (2018).

145. Khaira, G. & Galanis, E. Descriptive epidemiology of *Vibrio parahaemolyticus* and other *Vibrio* species infections in British Columbia: 2001-2006. *Canada Communicable Disease Report = Releve Des Maladies Transmissibles Au Canada* **33,** 12–22 (2007).

146. Taylor, M. *et al.* Outbreak of *Vibrio parahaemolyticus* Associated with Consumption of Raw Oysters in Canada, 2015. *Foodborne Pathogens and Disease* **15,** 554–559 (2018).

147. Scallan, E. *et al.* Foodborne Illness Acquired in the United States—Major Pathogens. *Emerging Infectious Diseases* **17,** 7–15 (2011).

148. Martinez-Urtaza, J. *et al.* Epidemiological Investigation of a Foodborne Outbreak in Spain Associated with U.S. West Coast Genotypes of *Vibrio parahaemolyticus*. *SpringerPlus* **5,** 87 (2016).

149. Martinez-Urtaza, J. *et al.* Spread of Pacific Northwest *Vibrio parahaemolyticus* Strain. *The New England Journal of Medicine* **369,** 1573–1574 (2013).

150. Martinez-Urtaza, J., Bowers, J. C., Trinanes, J. & DePaola, A. Climate Anomalies and the Increasing Risk of *Vibrio parahaemolyticus* and *Vibrio vulnificus* Illnesses. *Food Research International. Climate Change and Food Science* **43,** 1780–1790 (2010).

151. Abanto, M., Gavilan, R. G., Baker-Austin, C., Gonzalez-Escalona, N. & Martinez-Urtaza, J. Global Expansion of Pacific Northwest *Vibrio parahaemolyticus* Sequence Type 36. *Emerging Infectious Diseases* **26,** 323–326 (2020).

152. O'Boyle, N., Houeix, B., Kilcoyne, M., Joshi, L. & Boyd, A. The MSHA Pilus of *Vibrio parahaemolyticus* Has Lectin Functionality and Enables TTSS-mediated Pathogenicity. *International journal of medical microbiology: IJMM* **303,** 563–573 (2013).

153. Krachler, A. M. & Orth, K. Functional Characterization of the Interaction between Bacterial Adhesin Multivalent Adhesion Molecule 7 (MAM7) Protein and Its Host Cell Ligands. *The Journal of Biological Chemistry* **286,** 38939–38947 (2011).

154. Yu, Y. *et al.* Putative Type VI Secretion Systems of *Vibrio parahaemolyticus* Contribute to Adhesion to Cultured Cell Monolayers. *Archives of Microbiology* **194,** 827–835 (2012).

155. Leighton, T. L., Buensuceso, R. N. C., Howell, P. L. & Burrows, L. L. Biogenesis of *Pseudomonas aeruginosa* Type IV Pili and Regulation of Their Function. *Environmental Microbiology* **17,** 4148–4163 (2015).

156. Chang, Y.-W. *et al.* Architecture of the *Vibrio cholerae* Toxin-Coregulated Pilus Machine Revealed by Electron Cryotomography. *Nature Microbiology* **2,** 16269 (2017).

157. Craig, L., Forest, K. T. & Maier, B. Type IV Pili: Dynamics, Biophysics and Functional Consequences. *Nature Reviews Microbiology* **17,** 429–440 (2019).

158. Aagesen, A. M. & Häse, C. C. Sequence Analyses of Type IV Pili from *Vibrio cholerae*, *Vibrio parahaemolyticus*, and *Vibrio vulnificus*. *Microbial Ecology* **64,** 509–524 (2012).

159. Collins, R. F., Davidsen, L., Derrick, J. P., Ford, R. C. & Tønjum, T. Analysis of the PilQ Secretin from *Neisseria meningitidis* by Transmission Electron Microscopy Reveals a Dodecameric Quaternary Structure. *Journal of Bacteriology* **183,** 3825–3832 (2001).

160. Koo, J. *et al.* PilF Is an Outer Membrane Lipoprotein Required for Multimerization and Localization of the *Pseudomonas aeruginosa* Type IV Pilus Secretin. *Journal of Bacteriology* **190,** 6961–6969 (2008).

161. Chiang, P., Habash, M. & Burrows, L. L. Disparate Subcellular Localization Patterns of *Pseudomonas aeruginosa* Type IV Pilus ATPases Involved in Twitching Motility. *Journal of Bacteriology* **187,** 829–839 (2005).

162. Tammam, S. *et al.* Characterization of the PilN, PilO and PilP Type IVa Pilus Subcomplex. *Molecular Microbiology* **82,** 1496–1514 (2011).

163. Strom, M. S. & Lory, S. Structure-Function and Biogenesis of the Type Iv Pili. *Annual Review of Microbiology* **47,** 565–596 (1993).

164. Adams, D. W., Pereira, J. M., Stoudmann, C., Stutzmann, S. & Blokesch, M. The Type IV Pilus Protein PilU Functions as a PilT-dependent Retraction ATPase. *PLOS Genetics* **15,** e1008393 (2019).

165. O'Toole, G. A. & Kolter, R. Flagellar and Twitching Motility Are Necessary for *Pseudomonas aeruginosa* Biofilm Development. *Molecular Microbiology* **30,** 295–304 (1998).

166. Wall, D. & Kaiser, D. Type IV Pili and Cell Motility. *Molecular Microbiology* **32,** 01–10 (1999).

167. Berne, C., Ducret, A., Hardy, G. G. & Brun, Y. V. Adhesins Involved in Attachment to Abiotic Surfaces by Gram-negative Bacteria. *Microbiology spectrum* **3,** 10.1128/microbiolspec.MB-0018–2015 (2015).

168. Jonson, G., Holmgren, J. & Svennerholm, A. M. Identification of a Mannose-Binding Pilus on *Vibrio cholerae* El Tor. *Microbial Pathogenesis* **11,** 433–441 (1991).

169. Kirn, T. J., Jude, B. A. & Taylor, R. K. A Colonization Factor Links *Vibrio cholerae* Environmental Survival and Human Infection. *Nature* **438,** 863–866 (2005).

170. O'Boyle, N. & Boyd, A. Manipulation of Intestinal Epithelial Cell Function by the Cell Contact-Dependent Type III Secretion Systems of *Vibrio parahaemolyticus*. *Frontiers in Cellular and Infection Microbiology* **3** (2014).

171. Gode-Potratz, C. J., Kustusch, R. J., Breheny, P. J., Weiss, D. S. & McCarter, L. L. Surface Sensing in *Vibrio parahaemolyticus* Triggers a Programme of Gene Expression That Promotes Colonization and Virulence. *Molecular Microbiology* **79,** 240–263 (2011).

172. Krachler, A. M., Ham, H. & Orth, K. Outer Membrane Adhesion Factor Multivalent Adhesion Molecule 7 Initiates Host Cell Binding during Infection by Gram-negative Pathogens. *Proceedings of the National Academy of Sciences* **108,** 11614–11619 (2011).

173. Chen, C., Yang, X. & Shen, X. Confirmed and Potential Roles of Bacterial T6SSs in the Intestinal Ecosystem. *Frontiers in Microbiology* **10** (2019).

174. Yu, Y., Fang, L., Zhang, Y., Sheng, H. & Fang, W. VgrG2 of Type VI Secretion System 2 of *Vibrio parahaemolyticus* Induces Autophagy in Macrophages. *Frontiers in Microbiology* **6,** 168 (2015).

175. Kalburge, S. S., Carpenter, M. R., Rozovsky, S. & Boyd, E. F. Quorum Sensing Regulators Are Required for Metabolic Fitness in *Vibrio parahaemolyticus*. *Infection and Immunity* **85,** e00930–16 (2017).

176. Sun, J. *et al.* Quorum Sensing Regulates Transcription of the Pilin Gene mshA1 of MSHA Pilus in *Vibrio parahaemolyticus*. *Gene* **807,** 145961 (2022).

177. Jude, B. A., Martinez, R. M., Skorupski, K. & Taylor, R. K. Levels of the Secreted *Vibrio cholerae* Attachment Factor GbpA Are Modulated by Quorum-Sensing-Induced Proteolysis. *Journal of Bacteriology* (2009).

178. McCarter, L. L. OpaR, a Homolog of *Vibrio harveyi* LuxR, Controls Opacity of *Vibrio parahaemolyticus*. *Journal of Bacteriology* **180,** 3166–3173 (1998).

179. Wu, K. *et al.* CqsA-introduced Quorum Sensing Inhibits Type VI Secretion System 2 through an OpaR-dependent Pathway in *Vibrio parahaemolyticus*. *Microbial Pathogenesis* **162,** 105334 (2022).

180. Zhang, Y. *et al.* Transcriptional Regulation of the Type VI Secretion System 1 Genes by Quorum Sensing and ToxR in *Vibrio parahaemolyticus*. *Frontiers in Microbiology* **8** (2017).

181. Hammer, B. K. & Bassler, B. L. Quorum Sensing Controls Biofilm Formation in *Vibrio cholerae*. *Molecular Microbiology* **50,** 101–104 (2003).

182. Salamone, M., Nicosia, A., Ghersi, G. & Tagliavia, M. *Vibrio* Proteases for Biomedical Applications: Modulating the Proteolytic Secretome of *V. alginolyticus* and *V. parahaemolyticus* for Improved Enzymes Production. *Microorganisms* **7,** 387 (2019).

183. Raghunath, P. Roles of Thermostable Direct Hemolysin (TDH) and TDH-related Hemolysin (TRH) in *Vibrio parahaemolyticus*. *Frontiers in Microbiology* **5,** 1–4 (2015).

184. Honda, T. & Iida, T. The Pathogenicity of *Vibrio parahaemolyticus* and the Role of the Thermostable Direct Haemolysin and Related Haemolysins: *Reviews in Medical Microbiology* **4,** 106–113 (1993).

185. Chimalapati, S. *et al. Vibrio* Deploys Type 2 Secreted Lipase to Esterify Cholesterol with Host Fatty Acids and Mediate Cell Egress. *eLife* **9** (eds Helaine, S. & Garrett, W. S.) e58057 (2020).

186. Luangtrakul, W. *et al.* Cytotoxicity of *Vibrio parahaemolyticus* AHPND Toxin on Shrimp Hemocytes, a Newly Identified Target Tissue, Involves Binding of Toxin to Aminopeptidase N1 Receptor. *PLOS Pathogens* **17,** e1009463 (2021).

187. Kumar, V. *et al.* $PirAB^{VP}$ Toxin Binds to Epithelial Cells of the Digestive Tract and Produce Pathognomonic AHPND Lesions in Germ-Free Brine Shrimp. *Toxins* **11,** 717 (2019).

188. Nishibuchi, M., Fasano, A., Russell, R. G. & Kaper, J. B. Enterotoxigenicity of *Vibrio parahaemolyticus* with and without Genes Encoding Thermostable Direct Hemolysin. *Infection and Immunity* **60,** 3539–3545 (1992).

189. Honda, T., Ni, Y., Miwatani, T., Adachi, T. & Kim, J. The Thermostable Direct Hemolysin of *Vibrio parahaemolyticus* Is a Pore-Forming Toxin. *Canadian Journal of Microbiology* **38,** 1175–1180 (1992).

190. Matsuda, S. *et al.* Association of *Vibrio parahaemolyticus* Thermostable Direct Hemolysin with Lipid Rafts Is Essential for Cytotoxicity but Not Hemolytic Activity. *Infection and Immunity* **78,** 603–610 (2010).

191. Ohnishi, K. *et al.* Relationship between Heat-Induced Fibrillogenicity and Hemolytic Activity of Thermostable Direct Hemolysin and a Related Hemolysin of *Vibrio parahaemolyticus*. *FEMS Microbiology Letters* **318,** 10–17 (2011).

192. Gotoh, K. *et al.* Bile Acid-Induced Virulence Gene Expression of *Vibrio parahaemolyticus* Reveals a Novel Therapeutic Potential for Bile Acid Sequestrants. *PLOS ONE* **5,** e13365 (2010).

193. Broberg, C. A., Calder, T. J. & Orth, K. *Vibrio parahaemolyticus* Cell Biology and Pathogenicity Determinants. *Microbes and Infection* **13,** 992–1001 (2011).

194. Park, K.-S. *et al.* Cytotoxicity and Enterotoxicity of the Thermostable Direct Hemolysin-Deletion Mutants of *Vibrio parahaemolyticus*. *Microbiology and Immunology* **48,** 313–318 (2004).

195. Xu, M., Yamamoto, K., Honda, T. & Xu, M. Construction and Characterization of an Isogenic Mutant of *Vibrio parahaemolyticus* Having a Deletion in the Thermostable Direct Hemolysin-Related Hemolysin Gene (*trh*). *Journal of Bacteriology* **176,** 4757–4760 (1994).

196. Johnson, T. L., Abendroth, J., Hol, W. G. & Sandkvist, M. Type II Secretion: From Structure to Function. *FEMS Microbiology Letters* **255,** 175–186 (2006).

197. Korotkov, K. V. & Sandkvist, M. Architecture, Function, and Substrates of the Type II Secretion System. *EcoSal Plus* **8** (2019).

198. Sikora, A. E. Proteins Secreted via the Type II Secretion System: Smart Strategies of *Vibrio cholerae* to Maintain Fitness in Different Ecological Niches. *PLOS Pathogens* **9,** e1003126 (2013).

199. Py, B., Loiseau, L. & Barras, F. An Inner Membrane Platform in the Type II Secretion Machinery of Gram-negative Bacteria. *EMBO Reports* **2,** 244–248 (2001).

200. Naskar, S., Hohl, M., Tassinari, M. & Low, H. H. The Structure and Mechanism of the Bacterial Type II Secretion System. *Molecular Microbiology* **115,** 412–424 (2021).

201. Possot, O. M. & Pugsley, A. P. The Conserved Tetracysteine Motif in the General Secretory Pathway Component PulE Is Required for Efficient Pullulanase secretion. *Gene* **192,** 45–50 (1997).

202. Francetić, O., Lory, S. & Pugsley, A. P. A Second Prepilin Peptidase Gene in *Escherichia Coli* K-12. *Molecular Microbiology* **27,** 763–775 (1998).

203. Chernyatina, A. A. & Low, H. H. Core Architecture of a Bacterial Type II Secretion System. *Nature Communications* **10,** 1–10 (2019).

204. Zückert, W. R. Secretion of Bacterial Lipoproteins: Through the Cytoplasmic Membrane, the Periplasm and Beyond. *Biochimica Et Biophysica Acta* **1843,** 1509–1516 (2014).

205. Okuda, S. & Tokuda, H. Lipoprotein Sorting in Bacteria. *Annual Review of Microbiology* **65,** 239–259 (2011).

206. Lybarger, S. R., Johnson, T. L., Gray, M. D., Sikora, A. E. & Sandkvist, M. Docking and Assembly of the Type II Secretion Complex of *Vibrio cholerae*. *Journal of Bacteriology* **191,** 3149–3161 (2009).

207. Pugsley, A. P., Kornacker, M. G. & Poquet, I. The General Protein-Export Pathway Is Directly Required for Extracellular Pullulanase Secretion in *Escherichia coli* K12. *Molecular Microbiology* **5,** 343–352 (1991).

208. Voulhoux, R. & Ball. Involvement of the Twin-Arginine Translocation System in Protein Secretion via the Type II Pathway. *The EMBO Journal* **20,** 6735–6741 (2001).

209. Korotkov, K. V., Sandkvist, M. & Hol, W. G. J. The Type II Secretion System: Biogenesis, Molecular Architecture and Mechanism. *Nature Reviews Microbiology* **10,** 336–351 (2012).

210. Pineau, C. *et al.* Substrate Recognition by the Bacterial Type II Secretion System: More than a Simple Interaction. *Molecular Microbiology* **94,** 126–140 (2014).

211. Shevchick E., V., Robert-Baudouy, J. & Condemine, G. Specific Interaction between OutD, an *Erwinia chrysanthemi* Outer Membrane Protein of the General Secretory Pathway, and Secreted Proteins. *The EMBO Journal* **16,** 3007–3016 (1997).

212. Hobbs, M. & Mattick, J. S. Common Components in the Assembly of Type 4 Fimbriae, DNA Transfer Systems, Filamentous Phage and Protein-Secretion Apparatus: A General System for the Formation of Surface-Associated Protein Complexes. *Molecular Microbiology* **10,** 233–243 (1993).

213. Matsuda, S. *et al.* Export of a *Vibrio parahaemolyticus* Toxin by the Sec and Type III Secretion Machineries in Tandem. *Nature Microbiology* **4,** 781–788 (2019).

214. Silverman, J. M., Brunet, Y. R., Cascales, E. & Mougous, J. D. Structure and Regulation of the Type VI Secretion System. *Annual Review of Microbiology* **66,** 453–472 (2012).

215. Jani, A. J. & Cotter, P. A. Type VI Secretion: Not Just for Pathogenesis Anymore. *Cell Host & Microbe* **8,** 2–6 (2010).

216. Logan, S. L. *et al.* The *Vibrio cholerae* Type VI Secretion System Can Modulate Host Intestinal Mechanics to Displace Gut Bacterial Symbionts. *Proceedings of the National Academy of Sciences* **115,** E3779–E3787 (2018).

217. Basler, M., Pilhofer, M., Henderson, G. P., Jensen, G. J. & Mekalanos, J. J. Type VI Secretion Requires a Dynamic Contractile Phage Tail-like Structure. *Nature* **483,** 182–186 (2012).

218. Felisberto-Rodrigues, C. *et al.* Towards a Structural Comprehension of Bacterial Type VI Secretion Systems: Characterization of the TssJ-TssM Complex of an *Escherichia coli* Pathovar. *PLoS pathogens* **7,** e1002386 (2011).

219. Logger, L., Aschtgen, M.-S., Guérin, M., Cascales, E. & Durand, E. Molecular Dissection of the Interface between the Type VI Secretion TssM Cytoplasmic Domain and the TssG Baseplate Component. *Journal of Molecular Biology* **428,** 4424–4437 (2016).

220. Aschtgen, M.-S., Bernard, C. S., De Bentzmann, S., Lloubès, R. & Cascales, E. SciN Is an Outer Membrane Lipoprotein Required for Type VI Secretion in Enteroaggregative *Escherichia coli*. *Journal of Bacteriology* **190,** 7523–7531 (2008).

221. Brunet, Y. R., Zoued, A., Boyer, F., Douzi, B. & Cascales, E. The Type VI Secretion TssEFGK-VgrG Phage-Like Baseplate Is Recruited to the TssJLM Membrane Complex via Multiple Contacts and Serves As Assembly Platform for Tail Tube/Sheath Polymerization. *PLOS Genetics* **11,** e1005545 (2015).

222. Gallique, M., Bouteiller, M. & Merieau, A. The Type VI Secretion System: A Dynamic System for Bacterial Communication? *Frontiers in Microbiology* **8,** 1–10 (2017).

223. Brunet, Y. R., Hénin, J., Celia, H. & Cascales, E. Type VI Secretion and Bacteriophage Tail Tubes Share a Common Assembly Pathway. *EMBO reports* **15,** 315–321 (2014).

224. Brackmann, M., Wang, J. & Basler, M. Type VI Secretion System Sheath Inter-subunit Interactions Modulate Its Contraction. *EMBO reports* **19,** 225–233 (2018).

225. Pietrosiuk, A. *et al.* Molecular Basis for the Unique Role of the AAA+ Chaperone ClpV in Type VI Protein Secretion. *Journal of Biological Chemistry* **286,** 30010–30021 (2011).

226. Bönemann, G., Pietrosuik, A., Diemand, A., Zentgraf, H. & Mogk, A. Remodelling of VipA/VipB Tubules by ClpV-mediated Threading Is Crucial for Type VI Protein Secretion. *The EMBO Journal* **28,** 315–325 (2009).

227. Shneider, M. M. *et al.* PAAR-repeat Proteins Sharpen and Diversify the Type VI Secretion System Spike. *Nature* **500,** 350–353 (2013).

228. Russell, A. B. *et al.* Diverse Type VI Secretion Phospholipases Are Functionally Plastic Antibacterial Effectors. *Nature* **496,** 508–512 (2013).

229. Whitney, J. C. *et al.* Identification, Structure, and Function of a Novel Type VI Secretion Peptidoglycan Glycoside Hydrolase Effector-Immunity Pair. *The Journal of Biological Chemistry* **288,** 26616–26624 (2013).

230. Ma, L.-S., Hachani, A., Lin, J.-S., Filloux, A. & Lai, E.-M. *Agrobacterium tumefaciens* Deploys a Superfamily of Type VI Secretion DNase Effectors as Weapons for Interbacterial Competition In Planta. *Cell Host & Microbe* **16,** 94–104 (2014).

231. Flaugnatti, N. *et al.* A Phospholipase A1 Antibacterial Type VI Secretion Effector Interacts Directly with the C-terminal Domain of the VgrG Spike Protein for Delivery. *Molecular Microbiology* **99,** 1099–1118 (2016).

232. Pukatzki, S., Ma, A. T., Revel, A. T., Sturtevant, D. & Mekalanos, J. J. Type VI Secretion System Translocates a Phage Tail Spike-like Protein into Target Cells Where It Cross-Links Actin. *Proceedings of the National Academy of Sciences* **104,** 15508–15513 (2007).

233. Salomon, D., Gonzalez, H., Updegraff, B. L. & Orth, K. *Vibrio parahaemolyticus* Type VI Secretion System 1 Is Activated in Marine Conditions to Target Bacteria, and Is Differentially Regulated from System 2. *PLOS ONE* **8,** e61086 (2013).

234. Li, P. *et al.* Acute Hepatopancreatic Necrosis Disease-Causing *Vibrio parahaemolyticus* Strains Maintain an Antibacterial Type VI Secretion System with Versatile Effector Repertoires. *Applied and Environmental Microbiology* **83,** e00737–17 (2017).

235. Galán, J. E., Lara-Tejero, M., Marlovits, T. C. & Wagner, S. Bacterial Type III Secretion Systems: Specialized Nanomachines for Protein Delivery into Target Cells. *Annual Review of Microbiology* **68,** 415–438 (2014).

236. Notti, R. Q. & Stebbins, C. E. The Structure and Function of Type III Secretion Systems. *Microbiology Spectrum* **4** (2016).

237. Park, D. *et al.* Visualization of the Type III Secretion Mediated *Salmonella*–Host Cell Interface Using Cryo-Electron Tomography. *eLife* **7,** e39514 (2018).

238. Deng, W. *et al.* Assembly, Structure, Function and Regulation of Type III Secretion Systems. *Nature Reviews. Microbiology* **15,** 323–337 (2017).

239. De Souza Santos, M. & Orth, K. The Role of the Type III Secretion System in the Intracellular Lifestyle of Enteric Pathogens. *Microbiology Spectrum* **7** (2019).

240. Coburn, B., Sekirov, I. & Finlay, B. B. Type III Secretion Systems and Disease. *Clinical Microbiology Reviews* **20,** 535–549 (2007).

241. Costa, T. R. D. *et al.* Secretion Systems in Gram-negative Bacteria: Structural and Mechanistic Insights. *Nature Reviews Microbiology* **13,** 343–359 (2015).

242. Portaliou, A. G., Tsolis, K. C., Loos, M. S., Zorzini, V. & Economou, A. Type III Secretion: Building and Operating a Remarkable Nanomachine. *Trends in Biochemical Sciences* **41,** 175–189 (2016).

243. Diepold, A. & Wagner, S. Assembly of the Bacterial Type III Secretion Machinery. *FEMS Microbiology Reviews* **38,** 802–822 (2014).

244. Matsuda, S., Hiyoshi, H., Tandhavanant, S. & Kodama, T. Advances on *Vibrio parahaemolyticus* Research in the Postgenomic Era. *Microbiology and Immunology* **64,** 167–181 (2020).

245. Burkinshaw, B. J. *et al.* Structural Analysis of a Specialized Type III Secretion System Peptidoglycan-cleaving Enzyme. *The Journal of Biological Chemistry* **290,** 10406–10417 (2015).

246. Journet, L., Agrain, C., Broz, P. & Cornelis, G. R. The Needle Length of Bacterial Injectisomes Is Determined by a Molecular Ruler. *Science* **302,** 1757–1760 (2003).

247. Büttner, D. Protein Export According to Schedule: Architecture, Assembly, and Regulation of Type III Secretion Systems from Plant- and Animal-Pathogenic Bacteria. *Microbiology and molecular biology reviews: MMBR* **76,** 262–310 (2012).

248. Radics, J., Königsmaier, L. & Marlovits, T. C. Structure of a Pathogenic Type 3 Secretion System in Action. *Nature Structural & Molecular Biology* **21,** 82–87 (2014).

249. Hu, J. *et al.* Cryo-EM Analysis of the T3S Injectisome Reveals the Structure of the Needle and Open Secretin. *Nature Communications* **9,** 1–11 (2018).

250. Jenkins, J., Worrall, L. J. & Strynadka, N. C. J. Recent Structural Advances towards Understanding of the Bacterial Type III Secretion Injectisome. *Trends in Biochemical Sciences* (2022).

251. Park, K.-S. *et al.* Functional Characterization of Two Type III Secretion Systems of *Vibrio parahaemolyticus*. *Infection and Immunity* **72,** 6659–6665 (2004).

252. Ham, H. & Orth, K. The Role of Type III Secretion System 2 in *Vibrio parahaemolyticus* Pathogenicity. *Journal of Microbiology* **50,** 719–725 (2012).

253. Hiyoshi, H., Kodama, T., Iida, T. & Honda, T. Contribution of *Vibrio parahaemolyticus* Virulence Factors to Cytotoxicity, Enterotoxicity, and Lethality in Mice. *Infection and Immunity* **78,** 1772–1780 (2010).

254. Zhang, L. *et al.* Type III Effector VopC Mediates Invasion for *Vibrio* Species. *Cell Reports* **1,** 453–460 (2012).

255. Wang, R. *et al.* The Pathogenesis, Detection, and Prevention of *Vibrio parahaemolyticus*. *Frontiers in Microbiology* **6,** 144 (2015).

256. Trosky, J. E. *et al.* Inhibition of MAPK Signaling Pathways by VopA from *Vibrio parahaemolyticus*. *The Journal of Biological Chemistry* **279,** 51953–51957 (2004).

257. Kodama, T. *et al.* Identification and Characterization of VopT, a Novel ADP-ribosyltransferase Effector Protein Secreted via the *Vibrio parahaemolyticus* Type III Secretion System 2. *Cellular Microbiology* **9,** 2598–2609 (2007).

258. Liverman, A. D. B. *et al.* Arp2/3-Independent Assembly of Actin by *Vibrio* Type III Effector VopL. *Proceedings of the National Academy of Sciences* **104,** 17117–17122 (2007).

259. Yu, B., Cheng, H.-C., Brautigam, C. A., Tomchick, D. R. & Rosen, M. K. Mechanism of Actin Filament Nucleation by the Bacterial Effector VopL. *Nature Structural & Molecular Biology* **18,** 1068–1074 (2011).

260. Calder, T. *et al. Vibrio* Type III Effector VPA1380 Is Related to the Cysteine Protease Domain of Large Bacterial Toxins. *PloS One* **9,** e104387 (2014).

261. Hiyoshi, H. *et al.* Interaction between the Type III Effector VopO and GEF-H1 Activates the RhoA-ROCK Pathway. *PLOS Pathogens* **11,** e1004694 (2015).

262. Zhou, X. *et al.* A *Vibrio parahaemolyticus* T3SS Effector Mediates Pathogenesis by Independently Enabling Intestinal Colonization and Inhibiting TAK1 Activation. *Cell Reports* **3,** 1690–1702 (2013).

263. Hiyoshi, H. *et al.* VopV, an F-actin-binding Type III Secretion Effector, Is Required for *Vibrio parahaemolyticus*-Induced Enterotoxicity. *Cell Host & Microbe* **10,** 401–409 (2011).

264. Ridley, A. J., Paterson, H. F., Johnston, C. L., Diekmann, D. & Hall, A. The Small GTP-binding Protein Rac Regulates Growth Factor-Induced Membrane Ruffling. *Cell* **70,** 401–410 (1992).

265. Castellano, F. *et al.* Inducible Recruitment of Cdc42 or WASP to a Cell-Surface Receptor Triggers Actin Polymerization and Filopodium Formation. *Current Biology* **9,** 351–361 (1999).

266. Santos, M. d. S., Salomon, D. & Orth, K. T3SS Effector VopL Inhibits the Host ROS Response, Promoting the Intracellular Survival of *Vibrio parahaemolyticus*. *PLOS Pathogens* **13,** e1006438 (2017).

267. Prochazkova, K. & Satchell, K. J. F. Structure-Function Analysis of Inositol Hexakisphosphate-Induced Autoprocessing of the *Vibrio cholerae* Multifunctional Autoprocessing RTX Toxin. *The Journal of Biological Chemistry* **283,** 23656–23664 (2008).

268. Sarty, D. *et al.* Characterization of the Type III Secretion Associated Low Calcium Response Genes of *Vibrio parahaemolyticus* RIMD2210633. *Canadian Journal of Microbiology* **58,** 1306–1315 (2012).

269. Getz, L. J. & Thomas, N. A. The Transcriptional Regulator HlyU Positively Regulates Expression of *exsA*, Leading to Type III Secretion System 1 Activation in *Vibrio parahaemolyticus*. *Journal of Bacteriology* **200,** e00653–17 (2018).

270. Ono, T., Park, K.-S., Ueta, M., Iida, T. & Honda, T. Identification of Proteins Secreted via *Vibrio parahaemolyticus* Type III Secretion System 1. *Infection and Immunity* **74,** 1032–1042 (2006).

271. Zhou, X. *et al.* Identification of Potential Type III Secretion Proteins via Heterologous Expression of *Vibrio parahaemolyticus* DNA. *Applied and Environmental Microbiology* **78,** 3492–3494 (2012).

272. Zhou, X., Konkel, M. E. & Call, D. R. 2. Type III Secretion System 1 of *Vibrio parahaemolyticus* Induces Oncosis in Both Epithelial and Monocytic Cell Lines. *Microbiology* **155,** 837–851 (2009).

273. Matsuda, S., Okada, N., Kodama, T., Honda, T. & Iida, T. A Cytotoxic Type III Secretion Effector of *Vibrio parahaemolyticus* Targets Vacuolar H^+-ATPase Subunit c and Ruptures Host Cell Lysosomes. *PLoS pathogens* **8,** e1002803 (2012).

274. Matlawska-Wasowska, K. *et al.* The *Vibrio parahaemolyticus* Type III Secretion Systems Manipulate Host Cell MAPK for Critical Steps in Pathogenesis. *BMC microbiology* **10,** 329 (2010).

275. Shimohata, T. *et al. Vibrio parahaemolyticus* Infection Induces Modulation of IL-8 Secretion Through Dual Pathway via VP1680 in Caco-2 Cells. *The Journal of Infectious Diseases* **203,** 537–544 (2011).

276. Broberg, C. A., Zhang, L., Gonzalez, H., Laskowski-Arce, M. A. & Orth, K. A *Vibrio* Effector Protein Is an Inositol Phosphatase and Disrupts Host Cell Membrane Integrity. *Science* **329,** 1660–1662 (2010).

277. Miller, K. A., Tomberlin, K. F. & Dziejman, M. *Vibrio* Variations on a Type Three Theme. *Current Opinion in Microbiology* **47,** 66–73 (2019).

278. Burdette, D. L., Seemann, J. & Orth, K. *Vibrio* VopQ Induces PI3-kinase-independent Autophagy and Antagonizes Phagocytosis. *Molecular Microbiology* **73,** 639–649 (2009).

279. Yarbrough, M. L. *et al.* AMPylation of Rho GTPases by *Vibrio* VopS Disrupts Effector Binding and Downstream Signaling. *Science* **323,** 269–272 (2009).

280. Higa, N. *et al.* *Vibrio parahaemolyticus* Effector Proteins Suppress Inflammasome Activation by Interfering with Host Autophagy Signaling. *PLoS pathogens* **9,** e1003142 (2013).

281. Boyd, E. F. *et al.* Molecular Analysis of the Emergence of Pandemic *Vibrio parahaemolyticus*. *BMC Microbiology* **8,** 110 (2008).

282. Ritchie, J. M. *et al.* Inflammation and Disintegration of Intestinal Villi in an Experimental Model for *Vibrio parahaemolyticus*-Induced Diarrhea. *PLOS Pathogens* **8,** e1002593 (2012).

283. Hubbard, T. P. *et al.* Genetic Analysis of *Vibrio parahaemolyticus* Intestinal Colonization. *Proceedings of the National Academy of Sciences* **113,** 6283–6288 (2016).

284. Matz, C., Nouri, B., McCarter, L. & Martinez-Urtaza, J. Acquired Type III Secretion System Determines Environmental Fitness of Epidemic *Vibrio parahaemolyticus* in the Interaction with Bacterivorous Protists. *PLOS ONE* **6,** e20275 (2011).

285. Kodama, T. *et al.* Two Regulators of *Vibrio parahaemolyticus* Play Important Roles in Enterotoxicity by Controlling the Expression of Genes in the Vp-PAI Region. *PLOS ONE* **5,** e8678 (2010).

286. Li, P. *et al.* Bile Salt Receptor Complex Activates a Pathogenic Type III Secretion System. *eLife* **5** (ed Groisman, E. A.) e15718 (2016).

287. Ottemann, K. M. & Mekalanos, J. J. The ToxR Protein of *Vibrio cholerae* Forms Homodimers and Heterodimers. *Journal of Bacteriology* **178,** 156–162 (1996).

288. Yang, M. *et al.* Bile Salt–Induced Intermolecular Disulfide Bond Formation Activates *Vibrio cholerae* Virulence. *Proceedings of the National Academy of Sciences* **110,** 2348–2353 (2013).

289. Gode-Potratz, C. J., Chodur, D. M. & McCarter, L. L. Calcium and Iron Regulate Swarming and Type III Secretion in *Vibrio parahaemolyticus*. *Journal of Bacteriology* **192,** 6025–6038 (2010).

290. Liu, A. C. & Thomas, N. A. Transcriptional Profiling of *Vibrio parahaemolyticus exsA* Reveals a Complex Activation Network for Type III Secretion. *Frontiers in Microbiology* **6,** 1089 (2015).

291. Getz, L. J. *et al.* Attenuation of a DNA Cruciform by a Conserved Regulator Directs T3SS-1 Mediated Virulence in *Vibrio parahaemolyticus*, 2022.03.07.483294 (2022).

292. Iglewski, B. H., Sadoff, J., Bjorn, M. J. & Maxwell, E. S. *Pseudomonas aeruginosa* Exoenzyme S: An Adenosine Diphosphate Ribosyltransferase Distinct from Toxin A. *Proceedings of the National Academy of Sciences* **75,** 3211–3215 (1978).

293. Frank, D. W. The Exoenzyme S Regulon of *Pseudomonas aeruginosa*. *Molecular Microbiology* **26,** 621–629 (1997).

294. Vallis, A. J., Yahr, T. L., Barbieri, J. T. & Frank, D. W. Regulation of ExoS Production and Secretion by *Pseudomonas aeruginosa* in Response to Tissue Culture Conditions. *Infection and Immunity* **67,** 914–920 (1999).

295. Brutinel, E. D., Vakulskas, C. A. & Yahr, T. L. Functional Domains of ExsA, the Transcriptional Activator of the *Pseudomonas aeruginosa* Type III Secretion System. *Journal of Bacteriology* **191,** 3811–3821 (2009).

296. Zhou, X., Shah, D. H., Konkel, M. E. & Call, D. R. Type III Secretion System 1 Genes in *Vibrio parahaemolyticus* Are Positively Regulated by ExsA and Negatively Regulated by ExsD. *Molecular Microbiology* **69,** 747–764 (2008).

297. Nydam, S. D., Shah, D. H. & Call, D. R. Transcriptome Analysis of *Vibrio parahaemolyticus* in Type III Secretion System 1 Inducing Conditions. *Frontiers in Cellular and Infection Microbiology* **4,** 1 (2014).

298. Frank, D. W., Nair, G. & Schweizer, H. P. Construction and Characterization of Chromosomal Insertional Mutations of the *Pseudomonas aeruginosa* Exoenzyme S Trans-Regulatory Locus. *Infection and Immunity* **62,** 554–563 (1994).

299. Yahr, T. L. & Frank, D. W. Transcriptional Organization of the Trans-Regulatory Locus Which Controls Exoenzyme S Synthesis in *Pseudomonas aeruginosa*. *Journal of Bacteriology* **176,** 3832–3838 (1994).

300. Hovey, A. K. & Frank, D. W. Analyses of the DNA-binding and Transcriptional Activation Properties of ExsA, the Transcriptional Activator of the *Pseudomonas aeruginosa* Exoenzyme S Regulon. *Journal of Bacteriology* **177,** 4427–4436 (1995).

301. Yahr, T. L. & Wolfgang, M. C. Transcriptional Regulation of the *Pseudomonas aeruginosa* Type III Secretion System. *Molecular Microbiology* **62,** 631–640 (2006).

302. King, J. M., Schesser Bartra, S., Plano, G. & Yahr, T. L. ExsA and LcrF Recognize Similar Consensus Binding Sites, but Differences in Their Oligomeric State Influence Interactions with Promoter DNA. *Journal of Bacteriology* **195,** 5639–5650 (2013).

303. King, J. M., Brutinel, E. D., Marsden, A. E., Schubot, F. D. & Yahr, T. L. Orientation of *Pseudomonas aeruginosa* ExsA Monomers Bound to Promoter DNA and Base-Specific Contacts with the PexoT Promoter. *Journal of Bacteriology* **194,** 2573–2585 (2012).

304. Kernell Burke, A. *et al.* OpaR Controls a Network of Downstream Transcription Factors in *Vibrio parahaemolyticus* BB22OP. *PLoS ONE* **10,** e0121863 (2015).

305. Kodama, T. *et al.* Transcription of *Vibrio parahaemolyticus* T3SS1 Genes Is Regulated by a Dual Regulation System Consisting of the ExsACDE Regulatory Cascade and H-NS. *FEMS microbiology letters* **311,** 10–17 (2010).

306. Rietsch, A., Vallet-Gely, I., Dove, S. L. & Mekalanos, J. J. ExsE, a Secreted Regulator of Type III Secretion Genes in *Pseudomonas aeruginosa*. *Proceedings of the National Academy of Sciences* **102,** 8006–8011 (2005).

307. Sun, F. *et al.* H-NS Is a Repressor of Major Virulence Gene Loci in *Vibrio parahaemolyticus*. *Frontiers in Microbiology* **5,** 675 (2014).

308. Salomon, D., Klimko, J. A. & Orth, K. H-NS Regulates the *Vibrio parahaemolyticus* Type VI Secretion System 1. *Microbiology* **160,** 1867–1873 (2014).

309. Dorman, C. J. H-NS, the Genome Sentinel. *Nature Reviews. Microbiology* **5,** 157–161 (2007).

310. Navarre, W. W., McClelland, M., Libby, S. J. & Fang, F. C. Silencing of Xenogeneic DNA by H-NS-facilitation of Lateral Gene Transfer in Bacteria by a Defense System That Recognizes Foreign DNA. *Genes & Development* **21,** 1456–1471 (2007).

311. Han, H. *et al.* Genome Plasticity of *Vibrio parahaemolyticus*: Microevolution of the 'Pandemic Group'. *BMC genomics* **9,** 570 (2008).

312. Xiao, X. *et al.* A Novel Genotyping Scheme for *Vibrio parahaemolyticus* with Combined Use of Large Variably-Presented Gene Clusters (LVPCs) and Variable-Number Tandem Repeats (VNTRs). *International Journal of Food Microbiology* **149,** 143–151 (2011).

313. Lang, B. *et al.* High-Affinity DNA Binding Sites for H-NS Provide a Molecular Basis for Selective Silencing within Proteobacterial Genomes. *Nucleic Acids Research* **35,** 6330–6337 (2007).

314. Arold, S. T., Leonard, P. G., Parkinson, G. N. & Ladbury, J. E. H-NS Forms a Superhelical Protein Scaffold for DNA Condensation. *Proceedings of the National Academy of Sciences* **107,** 15728–15732 (2010).

315. Wade, J. T. & Grainger, D. C. Waking the Neighbours: Disruption of H-NS Repression by Overlapping Transcription. *Molecular Microbiology* **108,** 221–225 (2018).

316. Kotlajich, M. V. *et al.* Bridged Filaments of Histone-like Nucleoid Structuring Protein Pause RNA Polymerase and Aid Termination in Bacteria. *eLife* **4** (ed Proudfoot, N. J.) e04970 (2015).

317. Liu, Y., Chen, H., Kenney, L. J. & Yan, J. A Divalent Switch Drives H-NS/DNA-binding Conformations between Stiffening and Bridging Modes. *Genes & Development* **24,** 339–344 (2010).

318. Stoebel, D. M., Free, A. & Dorman, C. J. Anti-Silencing: Overcoming H-NS-mediated Repression of Transcription in Gram-negative Enteric Bacteria. *Microbiology (Reading, England)* **154,** 2533–2545 (2008).

319. Chaparian, R. R., Tran, M. L. N., Miller Conrad, L. C., Rusch, D. B. & van Kessel, J. C. Global H-NS Counter-Silencing by LuxR Activates Quorum Sensing Gene Expression. *Nucleic Acids Research* **48,** 171–183 (2020).

320. Nye, M. B., Pfau, J. D., Skorupski, K. & Taylor, R. K. *Vibrio cholerae* H-NS Silences Virulence Gene Expression at Multiple Steps in the ToxR Regulatory Cascade. *Journal of Bacteriology* **182,** 4295–4303 (2000).

321. Kazi, M. I., Conrado, A. R., Mey, A. R., Payne, S. M. & Davies, B. W. ToxR Antagonizes H-NS Regulation of Horizontally Acquired Genes to Drive Host Colonization. *PLOS Pathogens* **12,** e1005570 (2016).

322. Kim, B. S. Spatiotemporal Regulation of Vibrio Exotoxins by HlyU and Other Transcriptional Regulators. *Toxins* **12,** E544 (2020).

323. Stone, J. B. & Withey, J. H. H-NS and ToxT Inversely Control Cholera Toxin Production by Binding to Overlapping DNA Sequences. *Journal of Bacteriology* **203,** e0018721 (2021).

324. Liu, M., Naka, H. & Crosa, J. H. HlyU Acts as an H-NS Antirepressor in the Regulation of the RTX Toxin Gene Essential for the Virulence of the Human Pathogen *Vibrio vulnificus* CMCP6. *Molecular Microbiology* **72,** 491–505 (2009).

325. Choi, G. *et al*. The Transcriptional Regulator IscR Integrates Host-Derived Nitrosative Stress and Iron Starvation in Activation of the *vvhBA* Operon in *Vibrio vulnificus*. *The Journal of Biological Chemistry* **295,** 5350–5361 (2020).

326. Shao, C.-P., Lo, H.-R., Lin, J.-H. & Hor, L.-I. Regulation of Cytotoxicity by Quorum-Sensing Signaling in *Vibrio vulnificus* Is Mediated by SmcR, a Repressor of *hlyU*. *Journal of Bacteriology* **193,** 2557–2565 (2011).

327. Mou, X., Spinard, E. J., Driscoll, M. V., Zhao, W. & Nelson, D. R. H-NS Is a Negative Regulator of the Two Hemolysin/Cytotoxin Gene Clusters in *Vibrio anguillarum*. *Infection and Immunity* **81,** 3566–3576 (2013).

328. Li, L., Mou, X. & Nelson, D. R. HlyU Is a Positive Regulator of Hemolysin Expression in *Vibrio anguillarum*. *Journal of Bacteriology* **193,** 4779–4789 (2011).

329. Prosseda, G. *et al*. The *virF* Promoter in *Shigella*: More than Just a Curved DNA Stretch. *Molecular Microbiology* **51,** 523–537 (2004).

330. Manna, D. & Gowrishankar, J. Evidence for Involvement of Proteins HU and RpoS in Transcription of the Osmoresponsive *proU* Operon in *Escherichia coli*. *Journal of Bacteriology* **176,** 5378–5384 (1994).

331. Bouffartigues, E., Buckle, M., Badaut, C., Travers, A. & Rimsky, S. H-NS Cooperative Binding to High-Affinity Sites in a Regulatory Element Results in Transcriptional Silencing. *Nature Structural & Molecular Biology* **14,** 441–448 (2007).

332. Nishi, K. *et al.* Crystal Structure of the Transcriptional Activator HlyU from *Vibrio vulnificus* CMCP6. *FEBS Letters* **584,** 1097–1102 (2010).

333. Mukherjee, D., Datta, A. B. & Chakrabarti, P. Crystal Structure of HlyU, the Hemolysin Gene Transcription Activator, from *Vibrio cholerae* N16961 and Functional Implications. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1844,** 2346–2354 (2014).

334. Ceccarelli, D., Hasan, N. A., Huq, A. & Colwell, R. R. Distribution and Dynamics of Epidemic and Pandemic *Vibrio parahaemolyticus* Virulence Factors. *Frontiers in Cellular and Infection Microbiology* **3,** 97 (2013).

335. Newton, A. E. *et al.* Increase in *Vibrio parahaemolyticus* Infections Associated with Consumption of Atlantic Coast Shellfish–2013. *MMWR. Morbidity and mortality weekly report* **63,** 335–336 (2014).

336. Yeung, P. S. M. & Boor, K. J. Epidemiology, Pathogenesis, and Prevention of Foodborne *Vibrio parahaemolyticus* Infections. *Foodborne Pathogens and Disease* **1,** 74–88 (2004).

337. Chen, Y. *et al.* Comparative Genomic Analysis of *Vibrio parahaemolyticus*: Serotype Conversion and Virulence. *BMC genomics* **12,** 294 (2011).

338. Li, L. *et al.* Comparative Genomic Analysis of Clinical and Environmental Strains Provides Insight into the Pathogenicity and Evolution of *Vibrio parahaemolyticus*. *BMC genomics* **15,** 1135 (2014).

339. Ritchie, J. M. *et al.* Inflammation and Disintegration of Intestinal Villi in an Experimental Model for *Vibrio parahaemolyticus*-Induced Diarrhea. *PLoS pathogens* **8,** e1002593 (2012).

340. Yang, J., Tauschek, M. & Robins-Browne, R. M. Control of Bacterial Virulence by AraC-like Regulators That Respond to Chemical Signals. *Trends in Microbiology* **19,** 128–135 (2011).

341. Vakulskas, C. A., Brady, K. M. & Yahr, T. L. Mechanism of Transcriptional Activation by *Pseudomonas aeruginosa* ExsA. *Journal of Bacteriology* **191,** 6654–6664 (2009).

342. Jiménez, R., Cruz-Migoni, S. B., Huerta-Saquero, A., Bustamante, V. H. & Puente, J. L. Molecular Characterization of GrlA, a Specific Positive Regulator of Ler Expression in Enteropathogenic *Escherichia coli*. *Journal of Bacteriology* **192,** 4627–4642 (2010).

343. Wattiau, P. & Cornelis, G. R. Identification of DNA Sequences Recognized by VirF, the Transcriptional Activator of the *Yersinia* Yop Regulon. *Journal of Bacteriology* **176,** 3878–3884 (1994).

344. Lyell, N. L., Dunn, A. K., Bose, J. L., Vescovi, S. L. & Stabb, E. V. Effective Mutagenesis of *Vibrio fischeri* by Using Hyperactive Mini-Tn5 Derivatives. *Applied and Environmental Microbiology* **74,** 7059–7063 (2008).

345. Mukherjee, D., Pal, A., Chakravarty, D. & Chakrabarti, P. Identification of the Target DNA Sequence and Characterization of DNA Binding Features of HlyU, and Suggestion of a Redox Switch for *hlyA* Expression in the Human Pathogen *Vibrio cholerae* from in Silico Studies. *Nucleic Acids Research* **43,** 1407–1417 (2015).

346. Liu, M., Rose, M. & Crosa, J. H. Homodimerization and Binding of Specific Domains to the Target DNA Are Essential Requirements for HlyU To Regulate Expression of the Virulence Gene *rtxA1*, Encoding the Repeat-in-Toxin Protein in the Human Pathogen *Vibrio vulnificus*. *Journal of Bacteriology* **193,** 6895–6901 (2011).

347. Williams, S. G., Attridge, S. R. & Manning, P. A. The Transcriptional Activator HIyU of *Vibrio cholerae*: Nucleotide Sequence and Role in Virulence Gene Expression. *Molecular Microbiology* **9,** 751–760 (1993).

348. Turner, E. C. & Dorman, C. J. H-NS Antagonism in *Shigella flexneri* by VirB, a Virulence Gene Transcription Regulator That Is Closely Related to Plasmid Partition Factors. *Journal of Bacteriology* **189,** 3403–3413 (2007).

349. Martínez, L. C., Banda, M. M., Fernández-Mora, M., Santana, F. J. & Bustamante, V. H. HilD Induces Expression of *Salmonella* Pathogenicity Island 2 Genes by Displacing the Global Negative Regulator H-NS from ssrAB. *Journal of Bacteriology* **196,** 3746–3755 (2014).

350. Winardhi, R. S., Gulvady, R., Mellies, J. L. & Yan, J. Locus of Enterocyte Effacement-Encoded Regulator (Ler) of Pathogenic *Escherichia coli* Competes off Histone-like Nucleoid-Structuring Protein (H-NS) through Noncooperative DNA Binding. *The Journal of Biological Chemistry* **289,** 13739–13750 (2014).

351. Dorman, C. J. H-NS: A Universal Regulator for a Dynamic Genome. *Nature Reviews. Microbiology* **2,** 391–400 (2004).

352. Saha, R. P. & Chakrabarti, P. Molecular Modeling and Characterization of *Vibrio cholerae* Transcription Regulator HlyU. *BMC Structural Biology* **6,** 24 (2006).

353. Edwards, R. A., Keller, L. H. & Schifferli, D. M. Improved Allelic Exchange Vectors and Their Use to Analyze 987P Fimbria Gene Expression. *Gene* **207,** 149–157 (1998).

354. Thomas, N. A., Deng, W., Baker, N., Puente, J. & Finlay, B. B. Hierarchical Delivery of an Essential Host Colonization Factor in Enteropathogenic *Escherichia coli*. *The Journal of Biological Chemistry* **282,** 29634–29645 (2007).

355. Hellman, L. M. & Fried, M. G. Electrophoretic Mobility Shift Assay (EMSA) for Detecting Protein-Nucleic Acid Interactions. *Nature Protocols* **2,** 1849–1861 (2007).

356. Thomassin, J.-L., He, X. & Thomas, N. A. Role of EscU Auto-Cleavage in Promoting Type III Effector Translocation into Host Cells by Enteropathogenic *Escherichia coli*. *BMC microbiology* **11,** 205 (2011).

357. DiRita, V. J., Parsot, C., Jander, G. & Mekalanos, J. J. Regulatory Cascade Controls Virulence in *Vibrio cholerae*. *Proceedings of the National Academy of Sciences* **88,** 5403–5407 (1991).

358. Livny, J. *et al*. Comparative RNA-Seq Based Dissection of the Regulatory Networks and Environmental Stimuli Underlying *Vibrio parahaemolyticus* Gene Expression during Infection. *Nucleic Acids Research* **42,** 12212–12223 (2014).

359. Williams, S. G., Attridge, S. R. & Manning, P. A. The Transcriptional Activator HlyU of *Vibrio cholerae*: Nucleotide Sequence and Role in Virulence Gene Expression. *Molecular Microbiology* **9,** 751–760 (1993).

360. Liu, M., Alice, A. F., Naka, H. & Crosa, J. H. The HlyU Protein Is a Positive Regulator of *rtxA1*, a Gene Responsible for Cytotoxicity and Virulence in the Human Pathogen *Vibrio vulnificus*. *Infection and Immunity* **75,** 3282–3289 (2007).

361. Imdad, S., Batool, N., Pradhan, S., Chaurasia, A. K. & Kim, K. K. Identification of 2',4'-Dihydroxychalcone as an Antivirulence Agent Targeting HlyU, a Master Virulence Regulator in *Vibrio vulnificus*. *Molecules* **23,** E1492 (2018).

362. Lee, Z.-W. *et al*. Small-Molecule Inhibitor of HlyU Attenuates Virulence of *Vibrio* Species. *Scientific Reports* **9,** 4346 (2019).

363. Stonehouse, E. A., Hulbert, R. R., Nye, M. B., Skorupski, K. & Taylor, R. K. H-NS Binding and Repression of the *ctx* Promoter in *Vibrio cholerae*. *Journal of Bacteriology* **193,** 979–988 (2011).

364. Liu, M. & Crosa, J. H. The Regulator HlyU, the Repeat-in-Toxin Gene *rtxA1*, and Their Roles in the Pathogenesis of *Vibrio vulnificus* Infections. *MicrobiologyOpen* **1,** 502–513 (2012).

365. Grainger, D. C. Structure and Function of Bacterial H-NS Protein. *Biochemical Society Transactions* **44,** 1561–1569 (2016).

366. Poggi, L. & Richard, G.-F. Alternative DNA Structures *in vivo*: Molecular Evidence and Remaining Questions. *Microbiology and molecular biology reviews: MMBR* **85,** e00110–20 (2021).

367. Hulton, C. S. *et al*. Histone-like Protein H1 (H-NS), DNA Supercoiling, and Gene Expression in Bacteria. *Cell* **63,** 631–642 (1990).

368. Dorman, C. J. Horizontally Acquired Homologues of the Nucleoid-Associated Protein H-NS: Implications for Gene Regulation. *Molecular Microbiology* **75,** 264–267 (2010).

369. Perez, J. C., Latifi, T. & Groisman, E. A. Overcoming H-NS-mediated Transcriptional Silencing of Horizontally Acquired Genes by the PhoP and SlyA Proteins in *Salmonella enterica*. *The Journal of Biological Chemistry* **283,** 10773–10783 (2008).

370. Pearson, C. E., Zorbas, H., Price, G. B. & Zannis-Hadjopoulos, M. Inverted Repeats, Stem-Loops, and Cruciforms: Significance for Initiation of DNA Replication. *Journal of Cellular Biochemistry* **63,** 1–22 (1996).

371. Matek, C., Ouldridge, T. E., Levy, A., Doye, J. P. K. & Louis, A. A. DNA Cruciform Arms Nucleate through a Correlated but Asynchronous Cooperative Mechanism. *The Journal of Physical Chemistry B* **116,** 11616–11625 (2012).

372. Brázda, V., Laister, R. C., Jagelská, E. B. & Arrowsmith, C. Cruciform Structures Are a Common DNA Feature Important for Regulating Biological Processes. *BMC molecular biology* **12,** 33 (2011).

373. Lilley, D. M. The Inverted Repeat as a Recognizable Structural Feature in Supercoiled DNA Molecules. *Proceedings of the National Academy of Sciences* **77,** 6468–6472 (1980).

374. Bianchi, M. E., Beltrame, M. & Paonessa, G. Specific Recognition of Cruciform DNA by Nuclear Protein HMG1. *Science* **243,** 1056–1059 (1989).

375. Murchie, A. I. & Lilley, D. M. The Mechanism of Cruciform Formation in Supercoiled DNA: Initial Opening of Central Basepairs in Salt-Dependent Extrusion. *Nucleic Acids Research* **15,** 9641–9654 (1987).

376. Brázda, V. *et al.* Palindrome Analyser – A New Web-Based Server for Predicting and Evaluating Inverted Repeats in Nucleotide Sequences. *Biochemical and Biophysical Research Communications* **478,** 1739–1745 (2016).

377. Guan, C. & Kumar, S. A Single Catalytic Domain of the Junction-Resolving Enzyme T7 Endonuclease I Is a Non-Specific Nicking Endonuclease. *Nucleic Acids Research* **33,** 6225–6234 (2005).

378. Dages, S., Dages, K., Zhi, X. & Leng, F. Inhibition of the *gyrA* Promoter by Transcription-Coupled DNA Supercoiling in *Escherichia coli*. *Scientific Reports* **8,** 14759 (2018).

379. Kowalczyk, P., Cieśla, J. M., Saparbaev, M., Laval, J. & Tudek, B. Sequence-Specific P53 Gene Damage by Chloroacetaldehyde and Its Repair Kinetics in *Escherichia coli*. *Acta Biochimica Polonica* **53,** 337–347 (2006).

380. Will, W. R., Whitham, P. J., Reid, P. J. & Fang, F. C. Modulation of H-NS Transcriptional Silencing by Magnesium. *Nucleic Acids Research* **46,** 5717–5725 (2018).

381. Walthers, D. *et al.* *Salmonella enterica* Response Regulator SsrB Relieves H-NS Silencing by Displacing H-NS Bound in Polymerization Mode and Directly Activates Transcription. *The Journal of Biological Chemistry* **286,** 1895–1902 (2011).

382. Lilley, D. M. J. & Hallam, L. R. Thermodynamics of the ColE1 Cruciform: Comparisons between Probing and Topological Experiments Using Single Topoisomers. *Journal of Molecular Biology* **180,** 179–200 (1984).

383. Declais, A.-C. *et al.* The Complex between a Four-Way DNA Junction and T7 Endonuclease I. *The EMBO Journal* **22,** 1398–1409 (2003).

384. Du, X. *et al.* The Genome-Wide Distribution of Non-B DNA Motifs Is Shaped by Operon Structure and Suggests the Transcriptional Importance of Non-B DNA Structures in *Escherichia coli*. *Nucleic Acids Research* **41,** 5965–5977 (2013).

385. Miura, O., Ogake, T. & Ohyama, T. Requirement or Exclusion of Inverted Repeat Sequences with Cruciform-Forming Potential in *Escherichia coli* Revealed by Genome-Wide Analyses. *Current Genetics* **64,** 945–958 (2018).

386. Scotto-Lavino, E., Du, G. & Frohman, M. A. 5' End cDNA Amplification Using Classic RACE. *Nature Protocols* **1,** 2555–2562 (2006).

387. Postma, P. W., Lengeler, J. W. & Jacobson, G. R. Phosphoenolpyruvate:Carbohydrate Phosphotransferase Systems of Bacteria. *Microbiological Reviews* **57,** 543–594 (1993).

388. Saler Jr, M. H. & Reizer, J. The Bacterial Phosphotransferase System: New Frontiers 30 Years Later. *Molecular Microbiology* **13,** 755–764 (1994).

389. Deutscher, J. *et al.* The Bacterial Phosphoenolpyruvate:Carbohydrate Phosphotransferase System: Regulation by Protein Phosphorylation and Phosphorylation-Dependent Protein-Protein Interactions. *Microbiology and Molecular Biology Reviews* **78,** 231–256 (2014).

390. Görke, B. & Stülke, J. Carbon Catabolite Repression in Bacteria: Many Ways to Make the Most out of Nutrients. *Nature Reviews Microbiology* **6,** 613–624 (2008).

391. Colton, D. M. & Stabb, E. V. Rethinking the Roles of CRP, cAMP, and Sugar-Mediated Global Regulation in the *Vibrionaceae*. *Current Genetics* **62,** 39–45 (2016).

392. Yang, J. K. & Epstein, W. Purification and Characterization of Adenylate Cyclase from *Escherichia coli* K12. *Journal of Biological Chemistry* **258,** 3750–3758 (1983).

393. Lévy, S., Zeng, G.-Q. & Danchin, A. Cyclic AMP Synthesis in *Escherichia coli* Strains Bearing Known Deletions in the Pts Phosphotransferase Operon. *Gene* **86,** 27–33 (1990).

394. Harwood, J. P. *et al.* Involvement of the Glucose Enzymes II of the Sugar Phosphotransferase System in the Regulation of Adenylate Cyclase by Glucose in *Escherichia coli*. *The Journal of Biological Chemistry* **251,** 2462–2468 (1976).

395. Feucht, B. U. & Saier, M. H. Fine Control of Adenylate Cyclase by the Phosphoenolpyruvate:Sugar Phosphotransferase Systems in *Escherichia coli* and *Salmonella typhimurium*. *Journal of Bacteriology* **141,** 603–610 (1980).

396. Bettenbrock, K. *et al.* Correlation between Growth Rates, EIIACrr Phosphorylation, and Intracellular Cyclic AMP Levels in *Escherichia coli* K-12. *Journal of Bacteriology* **189,** 6891–6900 (2007).

397. Kim, Y. R., Kim, S. Y., Kim, C. M., Lee, S. E. & Rhee, J. H. Essential Role of an Adenylate Cyclase in Regulating *Vibrio vulnificus* Virulence. *FEMS Microbiology Letters* **243,** 497–503 (2005).

398. Skorupski, K. & Taylor, R. K. Cyclic AMP and Its Receptor Protein Negatively Regulate the Coordinate Expression of Cholera Toxin and Toxin-Coregulated Pilus in *Vibrio cholerae*. *Proceedings of the National Academy of Sciences* **94,** 265–270 (1997).

399. Kariisa, A. T., Grube, A. & Tamayo, R. Two Nucleotide Second Messengers Regulate the Production of the *Vibrio cholerae* Colonization Factor GbpA. *BMC microbiology* **15,** 166 (2015).

400. Uzzau, S. *et al.* Virulence Attenuation and Live Vaccine Potential of *aroA*, *crp cdt cya*, and Plasmid-Cured Mutants of *Salmonella enterica* serovar Abortusovis in Mice and Sheep. *Infection and Immunity* **73,** 4302–4308 (2005).

401. Schmiel, D. H., Young, G. M. & Miller, V. L. The *Yersinia enterocolitica* Phospholipase Gene *yplA* Is Part of the Flagellar Regulon. *Journal of Bacteriology* **182,** 2314–2320 (2000).

402. Petersen, S. & Young, G. M. Essential Role for Cyclic AMP and Its Receptor Protein in *Yersinia enterocolitica* Virulence. *Infection and Immunity* **70,** 3665–3672 (2002).

403. Fuchs, E. L. *et al.* The *Pseudomonas aeruginosa* Vfr Regulator Controls Global Virulence Factor Expression through Cyclic AMP-Dependent and -Independent Mechanisms. *Journal of Bacteriology* **192,** 3553–3564 (2010).

404. Titgemeyer, F., Reizer, J., Reizer, A. & Saier, M. H. Evolutionary Relationships between Sugar Kinases and Transcriptional Repressors in Bacteria. *Microbiology* **140,** 2349–2354.

405. Schiefner, A. *et al.* The Crystal Structure of Mlc, a Global Regulator of Sugar Metabolism in *Escherichia coli*. *Journal of Biological Chemistry* **280,** 29073–29079 (2005).

406. Pennetier, C., Domínguez-Ramírez, L. & Plumbridge, J. Different Regions of Mlc and NagC, Homologous Transcriptional Repressors Controlling Expression of the Glucose and N-acetylglucosamine Phosphotransferase Systems in *Escherichia coli*, Are Required for Inducer Signal Recognition. *Molecular Microbiology* **67,** 364–377 (2008).

407. Tanaka, Y., Itoh, F., Kimata, K. & Aiba, H. Membrane Localization Itself but Not Binding to IICBGlc Is Directly Responsible for the Inactivation of the Global Repressor Mlc in *Escherichia coli*. *Molecular Microbiology* **53,** 941–951 (2004).

408. Plumbridge, J. Expression of *ptsG*, the Gene for the Major Glucose PTS Transporter in *Escherichia coli*, Is Repressed by Mlc and Induced by Growth on Glucose. *Molecular Microbiology* **29,** 1053–1063 (1998).

409. Plumbridge, J. Expression of the Phosphotransferase System Both Mediates and Is Mediated by Mlc Regulation in *Escherichia coli*. *Molecular Microbiology* **33,** 260–273 (1999).

410. Tanaka, Y., Kimata, K., Inada, T., Tagami, H. & Aiba, H. Negative Regulation of the Pts Operon by Mlc: Mechanism Underlying Glucose Induction in *Escherichia coli*. *Genes to Cells: Devoted to Molecular & Cellular Mechanisms* **4,** 391–399 (1999).

411. Plumbridge, J. Regulation of Gene Expression in the PTS in *Escherichia coli*: The Role and Interactions of Mlc. *Current Opinion in Microbiology* **5,** 187–193 (2002).

412. Kimata, K., Inada, T., Tagami, H. & Aiba, H. A Global Repressor (Mlc) Is Involved in Glucose Induction of the *ptsG* Gene Encoding Major Glucose Transporter in *Escherichia coli*. *Molecular Microbiology* **29,** 1509–1519 (1998).

413. Plumbridge, J. Control of the Expression of the *manXYZ* Operon in *Escherichia coli*: Mlc Is a Negative Regulator of the Mannose PTS. *Molecular Microbiology* **27,** 369–380 (1998).

414. Decker, K., Plumbridge, J. & Boos, W. Negative Transcriptional Regulation of a Positive Regulator: The Expression of *malT*, Encoding the Transcriptional Activator of the Maltose Regulon of *Escherichia coli*, Is Negatively Controlled by Mlc. *Molecular Microbiology* **27,** 381–390 (1998).

415. Kim, S.-Y. *et al.* Purification of Mlc and Analysis of Its Effects on the Pts Expression in *Escherichia coli*. *Journal of Biological Chemistry* **274,** 25398–25402 (1999).

416. Hosono, K., Kakuda, H. & Ichihara, S. Decreasing Accumulation of Acetate in a Rich Medium by *Escherichia coli* on Introduction of Genes on a Multicopy Plasmid. *Bioscience, Biotechnology, and Biochemistry* **59,** 256–261 (1995).

417. Tanaka, Y., Kimata, K. & Aiba, H. A Novel Regulatory Role of Glucose Transporter of *Escherichia coli*: Membrane Sequestration of a Global Repressor Mlc. *The EMBO journal* **19,** 5344–5352 (2000).

418. Rogers, M. J., Ohgi, T., Plumbridge, J. & Söil, D. Nucleotide Sequences of the *Escherichia coli nagE* and *nagB* Genes: The Structural Genes for the N-acetylglucosamine Transport Protein of the Bacterial Phosphoenolpyruvate: Sugar Phosphotrans-Ferase System and for Glucosamine-6-Phosphate Deaminase. *Gene* **62,** 197–207 (1988).

419. Plumbridge, J. A. Sequence of the *nagBACD* Operon in *Escherichia coli* K12 and Pattern of Transcription within the Nag Regulon. *Molecular Microbiology* **3,** 505–515 (1989).

420. Plumbridge, J. A. Repression and Induction of the Nag Regulon of *Escherichia coli* K-12: The Roles of *nagC* and *nagA* in Maintenance of the Uninduced State. *Molecular Microbiology* **5,** 2053–2062 (1991).

421. Miyashiro, T. *et al.* The N-acetyl-D-glucosamine Repressor NagC of *Vibrio fischeri* Facilitates Colonization of *Euprymna scolopes*. *Molecular microbiology* **82,** 894–903 (2011).

422. Plumbridge, J. Co-Ordinated Regulation of Amino Sugar Biosynthesis and Degradation: The NagC Repressor Acts as Both an Activator and a Repressor for the Transcription of the *glmUS* Operon and Requires Two Separated NagC Binding Sites. *The EMBO Journal* **14,** 3958–3965 (1995).

423. Enos-Berlage, J. L., Guvener, Z. T., Keenan, C. E. & McCarter, L. L. Genetic Determinants of Biofilm Development of Opaque and Translucent *Vibrio parahaemolyticus*. *Molecular Microbiology* **55,** 1160–1182 (2005).

424. Plumbridge, J. & Vimr, E. Convergent Pathways for Utilization of the Amino Sugars *N*-Acetylglucosamine, *N*-Acetylmannosamine, and *N*-Acetylneuraminic Acid by *Escherichia coli*. *Journal of Bacteriology* **181,** 47–54 (1999).

425. Bréchemier-Baey, D., Pennetier, C. & Plumbridge, J. Dual Inducer Signal Recognition by an Mlc Homologue. *Microbiology* **161,** 1694–1706 (2015).

426. Pickering, B. S., Lopilato, J. E., Smith, D. R. & Watnick, P. I. The Transcription Factor Mlc Promotes *Vibrio cholerae* Biofilm Formation through Repression of Phosphotransferase System Components. *Journal of Bacteriology* **196,** 2423–2430 (2014).

427. Sun, Y., Verma, S. C., Bogale, H. & Miyashiro, T. NagC Represses N-acetyl-glucosamine Utilization Genes in *Vibrio fischeri* within the Light Organ of *Euprymna scolopes*. *Frontiers in Microbiology* **6,** 741 (2015).

428. Ghosh, S., Rao, K. H., Sengupta, M., Bhattacharya, S. K. & Datta, A. Two Gene Clusters Co-Ordinate for a Functional N-acetylglucosamine Catabolic Pathway in *Vibrio cholerae*: New Insights into GlcNAc Catabolism of *Vibrio cholerae*. *Molecular Microbiology* **80,** 1549–1560 (2011).

429. Huang, Y. *et al.* Functional Characterization and Conditional Regulation of the Type VI Secretion System in *Vibrio fluvialis*. *Frontiers in Microbiology* **8,** 528 (2017).

430. Barbosa, V. A. A. & Lery, L. M. S. Insights into *Klebsiella pneumoniae* Type VI Secretion System Transcriptional Regulation. *BMC Genomics* **20,** 506 (2019).

431. Lim, S. *et al.* Mlc Regulation of *Salmonella* Pathogenicity Island I Gene Expression via *hilE* Repression. *Nucleic Acids Research* **35,** 1822–1832 (2007).

432. Schechter, L. M., Damrauer, S. M. & Lee, C. A. Two AraC/XylS Family Members Can Independently Counteract the Effect of Repressing Sequences Upstream of the *hilA* Promoter. *Molecular Microbiology* **32,** 629–642 (1999).

433. Lostroh, C. P. & Lee, C. A. The HilA Box and Sequences Outside It Determine the Magnitude of HilA-Dependent Activation of PprgH from *Salmonella* Pathogenicity Island 1. *Journal of Bacteriology* **183,** 4876–4885 (2001).

434. Lucas, R. L. & Lee, C. A. Roles of *hilC* and *hilD* in Regulation of hilA Expression in *Salmonella enterica* serovar Typhimurium. *Journal of Bacteriology* **183,** 2733–2745 (2001).

435. Lou, L., Zhang, P., Piao, R. & Wang, Y. *Salmonella* Pathogenicity Island 1 (SPI-1) and Its Complex Regulatory Network. *Frontiers in Cellular and Infection Microbiology* **9** (2019).

436. Ellermeier, C. D., Ellermeier, J. R. & Slauch, J. M. HilD, HilC and RtsA Constitute a Feed Forward Loop That Controls Expression of the SPI1 Type Three Secretion System Regulator *hilA* in *Salmonella enterica* serovar Typhimurium. *Molecular Microbiology* **57,** 691–705 (2005).

437. West, S. E., Sample, A. K. & Runyen-Janecky, L. J. The Vfr Gene Product, Required for *Pseudomonas aeruginosa* Exotoxin A and Protease Production, Belongs to the Cyclic AMP Receptor Protein Family. *Journal of Bacteriology* **176,** 7532–7542 (1994).

438. Novichkov, P. S. *et al.* RegPrecise 3.0 – A Resource for Genome-Scale Exploration of Transcriptional Regulation in Bacteria. *BMC Genomics* **14,** 745 (2013).

439. Camacho, C. *et al.* BLAST+: Architecture and Applications. *BMC Bioinformatics* **10,** 421 (2009).

440. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic Local Alignment Search Tool. *Journal of Molecular Biology* **215,** 403–410 (1990).

441. Madeira, F. *et al.* Search and Sequence Analysis Tools Services from EMBL-EBI in 2022. *Nucleic acids research* (2022).

442. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview Version 2–a Multiple Sequence Alignment Editor and Analysis Workbench. *Bioinformatics* **25,** 1189–1191 (2009).

443. Bréchemier-Baey, D., Domínguez-Ramírez, L., Oberto, J. & Plumbridge, J. Operator Recognition by the ROK Transcription Factor Family Members, NagC and Mlc. *Nucleic Acids Research* **43,** 361–372 (2015).

444. Choi, M.-H. *et al.* Effect of the *crp* Mutation on the Utilization of Transferrin-Bound Iron by *Vibrio Vulnificus*. *FEMS Microbiology Letters* **257,** 285–292 (2006).

445. Houot, L., Chang, S., Absalon, C. & Watnick, P. I. *Vibrio cholerae* Phosphoenolpyruvate Phosphotransferase System Control of Carbohydrate Transport, Biofilm Formation, and Colonization of the Germfree Mouse Intestine. *Infection and Immunity* **78,** 1482–1494 (2010).

446. Gibson, D. G. *et al.* Enzymatic Assembly of DNA Molecules up to Several Hundred Kilobases. *Nature Methods* **6,** 343–345 (2009).

447. Dunn, A. K., Millikan, D. S., Adin, D. M., Bose, J. L. & Stabb, E. V. New Rfp- and pES213-derived Tools for Analyzing Symbiotic *Vibrio fischeri* Reveal Patterns of Infection and Lux Expression *in situ*. *Applied and Environmental Microbiology* **72,** 802–810 (2006).

448. *Jalview Version 2—a Multiple Sequence Alignment Editor and Analysis Workbench | Bioinformatics | Oxford Academic* https://academic.oup.com/bioinformatics/article/25/9/1189/203460.

449. Zhang, H., Li, L., Zhao, Z., Peng, D. & Zhou, X. Polar Flagella Rotation in *Vibrio parahaemolyticus* Confers Resistance to Bacteriophage Infection. *Scientific Reports* **6,** 26147 (2016).

450. Adams, D. W., Stutzmann, S., Stoudmann, C. & Blokesch, M. DNA-uptake Pili of *Vibrio cholerae* Are Required for Chitin Colonization and Capable of Kin Recognition via Sequence-Specific Self-Interaction. *Nature Microbiology* **4,** 1545–1557 (2019).

451. Chourashi, R. *et al.* Chitin-Induced T6SS in *Vibrio cholerae* Is Dependent on ChiS Activation. *Microbiology* **164,** 751–763 (2018).

452. Chourashi, R. *et al.* Role of a Sensor Histidine Kinase ChiS of *Vibrio cholerae* in Pathogenesis. *International journal of medical microbiology: IJMM* **306,** 657–665 (2016).

453. Hudson, S. M. & Smith, C. in *Biopolymers from Renewable Resources* (ed Kaplan, D. L.) 96–118 (Springer, Berlin, Heidelberg, 1998).

454. Keyhani, N. O., Li, X.-B. & Roseman, S. Chitin Catabolism in the Marine Bacterium *Vibrio furnissii*: IDENTIFICATION AND MOLECULAR CLONING OF A CHITOPORIN. *Journal of Biological Chemistry* **275,** 33068–33076 (2000).

455. Hirano, T. *et al.* Chitin Heterodisaccharide, Released from Chitin by Chitinase and Chitin Oligosaccharide Deacetylase, Enhances the Chitin-Metabolizing Ability of *Vibrio parahaemolyticus*. *Journal of Bacteriology* **201,** e00270–19 (2019).

456. Klancher, C. A., Newman, J. D., Ball, A. S., van Kessel, J. C. & Dalia, A. B. Species-Specific Quorum Sensing Represses the Chitobiose Utilization Locus in *Vibrio cholerae*. *Applied and Environmental Microbiology* **86,** e00915–20 (2020).

457. Yamamoto, S. *et al.* Identification of a Chitin-Induced Small RNA That Regulates Translation of the *tfoX* Gene, Encoding a Positive Regulator of Natural Competence in *Vibrio cholerae*. *Journal of Bacteriology* **193,** 1953–1965 (2011).

458. Zhao, J. *et al.* Structural and Molecular Mechanism of CdpR Involved in Quorum-Sensing and Bacterial Virulence in *Pseudomonas aeruginosa*. *PLOS Biology* **14,** e1002449 (2016).

459. Ravindran, S. Barbara McClintock and the Discovery of Jumping Genes. *Proceedings of the National Academy of Sciences* **109,** 20198–20199 (2012).

460. Hayes, F. Transposon-Based Strategies for Microbial Functional Genomics and Proteomics. *Annual Review of Genetics* **37,** 3–29 (2003).

461. van Opijnen, T., Bodi, K. L. & Camilli, A. Tn-Seq: High-Throughput Parallel Sequencing for Fitness and Genetic Interaction Studies in Microorganisms. *Nature Methods* **6,** 767–772 (2009).

462. Martin, M. Cutadapt Removes Adapter Sequences from High-Throughput Sequencing Reads. *EMBnet.journal* **17,** 10–12 (2011).

463. Bankevich, A. *et al.* SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology* **19,** 455–477 (2012).

464. Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. QUAST: Quality Assessment Tool for Genome Assemblies. *Bioinformatics* **29,** 1072–1075 (2013).

465. Doi, Y., Wachino, J.-i. & Arakawa, Y. Aminoglycoside Resistance. *Infectious disease clinics of North America* **30,** 523–537 (2016).

466. Mingeot-Leclercq, M.-P., Glupczynski, Y. & Tulkens, P. M. Aminoglycosides: Activity and Resistance. *Antimicrobial Agents and Chemotherapy* **43,** 727–737 (1999).

467. Pelchovich, G., Schreiber, R., Zhuravlev, A. & Gophna, U. The Contribution of Common *rpsL* Mutations in *Escherichia coli* to Sensitivity to Ribosome Targeting Antibiotics. *International Journal of Medical Microbiology* **303,** 558–562 (2013).

468. Timms, A. R., Steingrimsdottir, H., Lehmann, A. R. & Bridges, B. A. Mutant Sequences in the *rpsL* Gene of *Escherichia coli* B/r: Mechanistic Implications for Spontaneous and Ultraviolet Light Mutagenesis. *Molecular & general genetics: MGG* **232,** 89–96 (1992).

469. Pandey, B. *et al.* Novel Missense Mutations in *gidB* Gene Associated with Streptomycin Resistance in *Mycobacterium tuberculosis*: Insights from Molecular Dynamics. *Journal of Biomolecular Structure & Dynamics* **37,** 20–35 (2019).

470. Tan, C. W. *et al.* Prevalence and Antibiotic Resistance Patterns of *Vibrio parahaemolyticus* Isolated from Different Types of Seafood in Selangor, Malaysia. *Saudi Journal of Biological Sciences* **27,** 1602–1608 (2020).

471. Elmahdi, S., DaSilva, L. V. & Parveen, S. Antibiotic Resistance of *Vibrio parahaemolyticus* and *Vibrio vulnificus* in Various Countries: A Review. *Food Microbiology* **57,** 128–134 (2016).

472. DeJesus, M. A., Ambadipudi, C., Baker, R., Sassetti, C. & Ioerger, T. R. TRANSIT–A Software Tool for Himar1 TnSeq Analysis. *PLoS computational biology* **11,** e1004401 (2015).

473. Kudo, T. *et al.* Draft Genome Sequences of *Vibrio* Sp. Strains Isolated from Tetrodotoxin-Bearing Scavenging Gastropod. *Genome Announcements* **2,** e00623–14 (2014).

474. Trias, J. & Nikaido, H. Outer Membrane Protein D2 Catalyzes Facilitated Diffusion of Carbapenems and Penems through the Outer Membrane of *Pseudomonas aeruginosa*. *Antimicrobial Agents and Chemotherapy* **34,** 52–57 (1990).

475. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. The Phyre2 Web Portal for Protein Modelling, Prediction and Analysis. *Nature protocols* **10,** 845–858 (2015).

476. Kaplan, E., Greene, N. P., Crow, A. & Koronakis, V. Insights into Bacterial Lipoprotein Trafficking from a Structure of LolA Bound to the LolC Periplasmic Domain. *Proceedings of the National Academy of Sciences* **115,** E7389–E7397 (2018).

477. Francetic, O., Belin, D., Badaut, C. & Pugsley, A. P. Expression of the Endogenous Type II Secretion Pathway in *Escherichia coli* Leads to Chitinase Secretion. *The EMBO journal* **19,** 6697–6703 (2000).

478. Lee, K.-J., Jung, Y.-C., Park, S.-J. & Lee, K.-H. Role of Heat Shock Proteases in Quorum-Sensing-Mediated Regulation of Biofilm Formation by *Vibrio* Species. *mBio* **9,** e02086–17 (2018).

479. Jobling, M. G. & Holmes, R. K. Characterization of *hapR*, a Positive Regulator of the *Vibrio cholerae* HA/Protease Gene Hap, and Its Identification as a Functional Homologue of the *Vibrio harveyi luxR* Gene. *Molecular Microbiology* **26,** 1023–1034 (1997).

480. Ball, A. S., Chaparian, R. R. & van Kessel, J. C. Quorum Sensing Gene Regulation by LuxR/HapR Master Regulators in *Vibrios*. *Journal of Bacteriology* **199,** e00105–17 (2017).

481. Liu, M., Zhu, X., Zhang, C. & Zhao, Z. LuxQ-LuxU-LuxO Pathway Regulates Biofilm Formation by *Vibrio parahaemolyticus*. *Microbiological Research* **250,** 126791 (2021).

482. Jaeger, T. & Mayer, C. The Transcriptional Factors MurR and Catabolite Activator Protein Regulate N-acetylmuramic Acid Catabolism in *Escherichia coli*. *Journal of Bacteriology* **190,** 6598–6608 (2008).

483. Kim, B. S., Hwang, J., Kim, M. H. & Choi, S. H. Cooperative Regulation of the *Vibrio vulnificus* Nan Gene Cluster by NanR Protein, cAMP Receptor Protein, and N-acetylmannosamine 6-Phosphate. *The Journal of Biological Chemistry* **286,** 40889–40899 (2011).

484. Hwang, J. *et al.* Structural Insights into the Regulation of Sialic Acid Catabolism by the *Vibrio vulnificus* Transcriptional Repressor NanR. *Proceedings of the National Academy of Sciences* **110,** E2829–2837 (2013).

485. Suginta, W., Chumjan, W., Mahendran, K. R., Schulte, A. & Winterhalter, M. Chitoporin from *Vibrio harveyi*, a Channel with Exceptional Sugar Specificity. *Journal of Biological Chemistry* **288,** 11038–11046 (2013).

486. Connell, T. D., Metzger, D. J., Lynch, J. & Folster, J. P. Endochitinase Is Transported to the Extracellular Milieu by the Eps-Encoded General Secretory Pathway ofVibrio Cholerae. *Journal of Bacteriology* **180,** 5591–5600 (1998).

487. Sikora, A. E., Zielke, R. A., Lawrence, D. A., Andrews, P. C. & Sandkvist, M. Proteomic Analysis of the *Vibrio cholerae* Type II Secretome Reveals New Proteins, Including Three Related Serine Proteases. *Journal of Biological Chemistry* **286,** 16555–16566 (2011).

488. Drescher, K., Nadell, C. D., Stone, H. A., Wingreen, N. S. & Bassler, B. L. Solutions to the Public Goods Dilemma in Bacterial Biofilms. *Current biology : CB* **24,** 50–55 (2014).

489. Griffin, A. S., West, S. A. & Buckling, A. Cooperation and Competition in Pathogenic Bacteria. *Nature* **430,** 1024–1027 (2004).

490. Wright, R. J., Gibson, M. I. & Christie-Oleza, J. A. Understanding Microbial Community Dynamics to Improve Optimal Microbiome Selection. *Microbiome* **7,** 85 (2019).

491. Cisneros, D. A., Bond, P. J., Pugsley, A. P., Campos, M. & Francetic, O. Minor Pseudopilin Self-Assembly Primes Type II Secretion Pseudopilus Elongation. *The EMBO Journal* **31,** 1041–1053 (2012).

492. Douzi, B., Ball, G., Cambillau, C., Tegoni, M. & Voulhoux, R. Deciphering the Xcp *Pseudomonas aeruginosa* Type II Secretion Machinery through Multiple Interactions with Substrates. *Journal of Biological Chemistry* **286,** 40792–40801 (2011).

493. Douzi, B. *et al.* The XcpV/GspI Pseudopilin Has a Central Role in the Assembly of a Quaternary Complex within the T2SS Pseudopilus. *Journal of Biological Chemistry* **284,** 34580–34589 (2009).

494. Hammer, B. K. & Bassler, B. L. Regulatory Small RNAs Circumvent the Conventional Quorum Sensing Pathway in Pandemic *Vibrio cholerae*. *Proceedings of the National Academy of Sciences* **104,** 11145–11149 (2007).

495. Lenz, D. H. *et al.* The Small RNA Chaperone Hfq and Multiple Small RNAs Control Quorum Sensing in *Vibrio harveyi* and *Vibrio cholerae*. *Cell* **118,** 69–82 (2004).

496. Jung, S. A., Chapman, C. A. & Ng, W.-L. Quadruple Quorum-Sensing Inputs Control *Vibrio cholerae* Virulence and Maintain System Robustness. *PLoS pathogens* **11,** e1004837 (2015).

497. Sultana, M. *et al.* Role of Shrimp Chitin in the Ecology of Toxigenic *Vibrio cholerae* and Cholera Transmission. *Frontiers in Microbiology* **2** (2012).

498. Minas, K., McEwan, N. R., Newbold, C. J. & Scott, K. P. Optimization of a High-Throughput CTAB-based Protocol for the Extraction of qPCR-grade DNA from Rumen Fluid, Plant and Bacterial Pure Cultures. *FEMS Microbiology Letters* **325,** 162–169 (2011).

499. Chiang, S. L. & Rubin, E. J. Construction of a Mariner-Based Transposon for Epitope-Tagging and Genomic Targeting. *Gene* **296,** 179–185 (2002).

500. Seemann, T. *Snippy* 2022.

501. Joe, S. & Sarojini, S. An Efficient Method of Production of Colloidal Chitin for Enumeration of Chitinase Producing Bacteria. *Mapana - Journal of Sciences* **16,** 37–45 (2017).

502. Lazinski, D. W. & Camilli, A. Homopolymer Tail-Mediated Ligation PCR: A Streamlined and Highly Efficient Method for DNA Cloning and Library Construction. *BioTechniques* **54,** 25–34 (2013).

503. The Document Foundation. *LibreOffice Calc* 2020.

504. R Core Team. *R: A Language and Environment for Statistical Computing* Manual. R Foundation for Statistical Computing (Vienna, Austria, 2021).

505. Hahne, F. & Ivanek, R. in *Statistical Genomics: Methods and Protocols* (eds Mathé, E. & Davis, S.) 335–351 (Springer, New York, NY, 2016).

506. McCoy, K. M., Antonio, M. L. & van Opijnen, T. MAGenTA: A Galaxy Implemented Tool for Complete Tn-Seq Analysis and Data Visualization. *Bioinformatics* **33,** 2781–2783 (2017).

507. Krzywinski, M. I. *et al.* Circos: An Information Aesthetic for Comparative Genomics. *Genome Research* (2009).

508. Tobe, T., Yoshikawa, M. & Sasakawa, C. Thermoregulation of *virB* Transcription in *Shigella flexneri* by Sensing of Changes in Local DNA Superhelicity. *Journal of Bacteriology* **177,** 1094–1097 (1995).

509. Badaut, C. *et al.* The Degree of Oligomerization of the H-NS Nucleoid Structuring Protein Is Related to Specific Binding to DNA. *Journal of Biological Chemistry* **277,** 41657–41666 (2002).

510. *in vitro* and *in vivo* Characterization.

511. Wolfgang, M. C., Lee, V. T., Gilmore, M. E. & Lory, S. Coordinate Regulation of Bacterial Virulence Genes by a Novel Adenylate Cyclase-Dependent Signaling Pathway. *Developmental Cell* **4,** 253–263 (2003).

512. Jin, Y. *et al.* NrtR Regulates the Type III Secretion System Through cAMP/Vfr Pathway in *Pseudomonas aeruginosa*. *Frontiers in Microbiology* **10** (2019).

513. Brown, N. L., Stoyanov, J. V., Kidd, S. P. & Hobman, J. L. The MerR Family of Transcriptional Regulators. *FEMS microbiology reviews* **27,** 145–163 (2003).

514. Islam, M. S., Bingle, L. E. H., Pallen, M. J. & Busby, S. J. W. Organization of the LEE1 Operon Regulatory Region of Enterohaemorrhagic *Escherichia coli* O157:H7 and Activation by GrlA. *Molecular Microbiology* **79,** 468–483 (2011).

515. Furniss, R. C. D. & Clements, A. Regulation of the Locus of Enterocyte Effacement in Attaching and Effacing Pathogens. *Journal of Bacteriology* **200,** e00336–17 (2017).

516. Sebastian, M. & Ammerman, J. W. The Alkaline Phosphatase PhoX Is More Widely Distributed in Marine Bacteria than the Classical PhoA. *The ISME Journal* **3,** 563–572 (2009).

517. McDonough, E., Kamp, H. & Camilli, A. *Vibrio cholerae* Phosphatases Required for the Utilization of Nucleotides and Extracellular DNA as Phosphate Sources. *Molecular Microbiology* **99,** 453–469 (2016).

518. Zaheer, R., Morton, R., Proudfoot, M., Yakunin, A. & Finan, T. M. Genetic and Biochemical Properties of an Alkaline Phosphatase PhoX Family Protein Found in Many Bacteria. *Environmental Microbiology* **11,** 1572–1587 (2009).

519. Ferrari, S. *et al.* SRY, like HMG1, Recognizes Sharp Angles in DNA. *The EMBO journal* **11,** 4497–4506 (1992).

520. Kuhn, A., Stefanovsky, V. & Grummt, I. The Nucleolar Transcription Activator UBF Relieves Ku Antigen-Mediated Repression of Mouse Ribosomal Gene Transcription. *Nucleic Acids Research* **21,** 2057–2063 (1993).

521. Stefanovsky, V. Y., Langlois, F., Bazett-Jones, D., Pelletier, G. & Moss, T. ERK Modulates DNA Bending and Enhancesome Structure by Phosphorylating HMG1-boxes 1 and 2 of the RNA Polymerase I Transcription Factor UBF. *Biochemistry* **45,** 3626–3634 (2006).

522. Stefanovsky, V. Y. & Moss, T. The Cruciform DNA Mobility Shift Assay: A Tool to Study Proteins That Recognize Bent DNA. *Methods in Molecular Biology (Clifton, N.J.)* **543,** 537–546 (2009).

523. Rietsch, A. & Mekalanos, J. J. Metabolic Regulation of Type III Secretion Gene Expression in *Pseudomonas aeruginosa*. *Molecular Microbiology* **59,** 807–820 (2006).

524. Lee, J. B. & Keung, A. J. in *Epigenome Editing* (eds Jeltsch, A. & Rots, M. G.) 257–269 (Springer New York, New York, NY, 2018).

525. Myers, K. S., Park, D. M., Beauchene, N. A. & Kiley, P. J. Defining Bacterial Regulons Using ChIP-seq Methods. *Methods (San Diego, Calif.)* **86,** 80–88 (2015).

526. Zhang, Y. *et al.* Model-Based Analysis of ChIP-Seq (MACS). *Genome Biology* **9,** R137 (2008).

527. Zhang, X. *et al.* PICS: Probabilistic Inference for ChIP-seq. *Biometrics* **67,** 151–163 (2011).

528. Chung, D., Kuan, P. F., Welch, R. & Keles, S. Analysis of ChIP-seq Data with 'Mosaics' Package: MOSAiCS and MOSAiCS-HMM, 38 (2016).

529. Liang, K. & Keleş, S. Detecting Differential Binding of Transcription Factors with ChIP-seq. *Bioinformatics* **28,** 121–122 (2012).

530. Nivaskumar, M. & Francetic, O. Type II Secretion System: A Magic Beanstalk or a Protein Escalator. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research. Protein Trafficking and Secretion in Bacteria* **1843,** 1568–1577 (2014).

531. Litzinger, S. *et al.* Muropeptide Rescue in Bacillus Subtilis Involves Sequential Hydrolysis by Beta-N-acetylglucosaminidase and N-acetylmuramyl-L-alanine Amidase. *Journal of Bacteriology* **192,** 3132–3143 (2010).

532. Carey, A. F. *et al.* TnSeq of *Mycobacterium tuberculosis* Clinical Isolates Reveals Strain-Specific Antibiotic Liabilities. *PLoS pathogens* **14,** e1006939 (2018).

533. Ripio, M. T., Brehm, K., Lara, M., Suárez, M. & Vázquez-Boland, J. A. Glucose-1-Phosphate Utilization by *Listeria monocytogenes* Is PrfA Dependent and Coordinately Expressed with Virulence Factors. *Journal of Bacteriology* **179,** 7174–7180 (1997).

534. Lobel, L. *et al.* The Metabolic Regulator CodY Links *Listeria monocytogenes* Metabolism to Virulence by Directly Activating the Virulence Regulatory Gene prfA. *Molecular Microbiology* **95,** 624–644 (2015).

535. Pacheco, A. R. *et al.* Fucose Sensing Regulates Bacterial Intestinal Colonization. *Nature* **492,** 113–117 (2012).

536. Van Ende, M., Wijnants, S. & Van Dijck, P. Sugar Sensing and Signaling in *Candida albicans* and *Candida glabrata*. *Frontiers in Microbiology* **10** (2019).

537. Chiang, C., Bongiorni, C. & Perego, M. Glucose-Dependent Activation of *Bacillus anthracis* Toxin Gene Expression and Virulence Requires the Carbon Catabolite Protein CcpA. *Journal of Bacteriology* **193,** 52–62 (2011).

538. Perinbam, K., Chacko, J. V., Kannan, A., Digman, M. A. & Siryaporn, A. A Shift in Central Metabolism Accompanies Virulence Activation in *Pseudomonas aeruginosa*. *mBio* **11,** e02730–18 (2020).

539. Stabb, E. V. & Ruby, E. G. in *Methods in Enzymology* 413–426 (Academic Press, 2002).

Appendix A

Chapter 2 Supplementary Data

Below is the supplementary data file for the following manuscript:

Getz, L.J., Thomas, N.A. (2017) *The Transcriptional Regulator HlyU Positively Regulates Expression of exsA, Leading to Type III Secretion System 1 Activation in Vibrio parahaemolyticus*. **Journal of Bacteriology**, Vol. 200, No. 15. DOI: 10.1128/JB.00653-17.

Low-moderate luminescence mutants (Mean:155.1; Standard Deviation 22.32)			
Isolate	CPS	OD_{600}	Gene
38-G06	128.0	0.592	*vp0529* (*hlyU*)
40-F04	132.0	0.505	*vp0529* (*hlyU*)
40-A03	214.0	0.645	*vp0529* (*hlyU*)
Moderate luminescence mutants (Mean: 542.7; Standard Deviation 126.6)			
Isolate	CPS	OD_{600}	Gene
10-C06	330.0	0.449	Vp1473 (putative HTH)
9-G03	348.0	0.434	Vp0179 (*phoX*)
1-B02	374.0	0.466	N/I
1-F08	404.0	0.556	N/I
33-A11	410.0	0.553	N/I
High luminescence mutants			
Isolate	CPS	OD_{600}	Gene
8-E11	8276.0	0.135	*vp1133* (*hns*)
9-G08	12822.0	0.206	*vp1133* (*hns*)
28-F05	8162.0	0.125	N/I
45-B11	3980.0	0.103	*vp1133* (*hns*)
31-G02	58707.0	0.564	*vp1633* (*rtx*)

Table A.1: Selected transposon mutants after statistical selection for *tdhS*::*exsAlux* library. The first number in the plate position identifier is the plate number in the library and the second letter and number is the plate position of the respective mutant. N/I (not identified); the sequence data was not of suitable quality to identify the insertion.

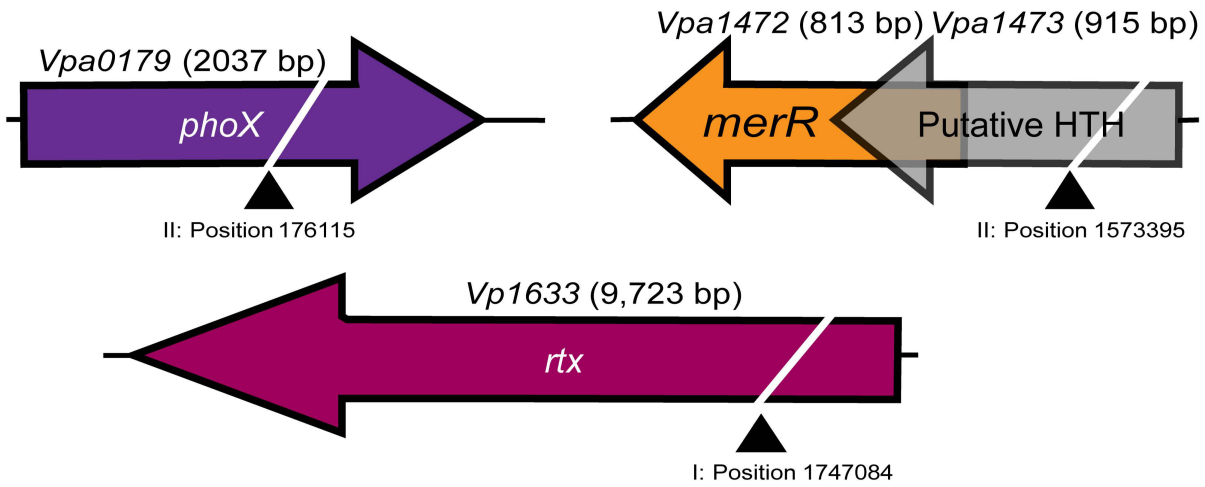

Figure A.1: Schematic diagram of transposon insertion sites within genes identified by luminescence screening of the *tdhS*::*exsAlux* Tn5 mutant library. Teh black triangles represent approximate transposon insertion sites with the approximate DNA base location in the *V. parahaemolyticus* genome (NCBI accession: NC_004603 and NC_004605; I - chromosome 1, and II - chromosome 2). *vpa0179* and *vpa1473* are two characterized mutants from the moderate luminescence group, and *vp1633* is a mutant characterized from teh high luminescence group.

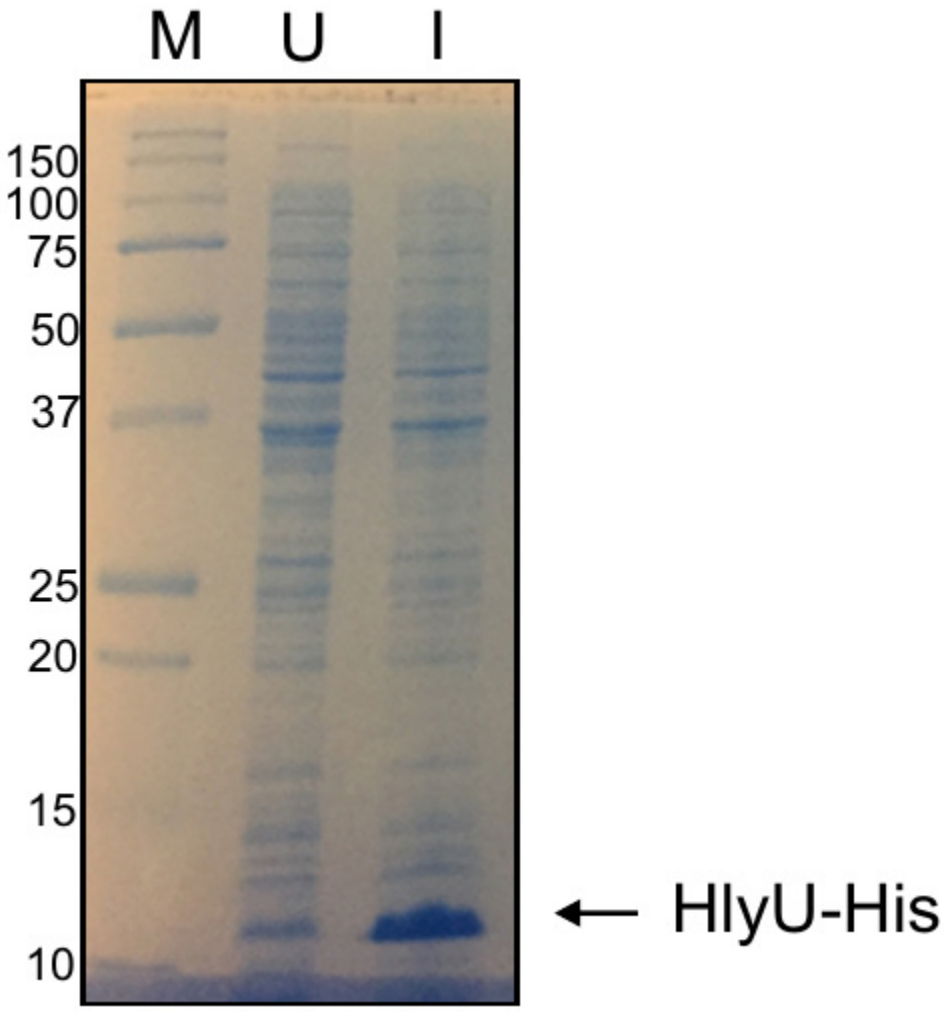

Figure A.2: Overexpression of HlyU-HIS in *E. coli* BL21(λDE3) using an IPTG inducible T7 expression system. Total cell lysates were generated and subjected to SDS-PAGE and Coomassie Blue staining. U=uninduced culture, I=2 hr post induction with 0.4 mM IPTG. M = protein standard.

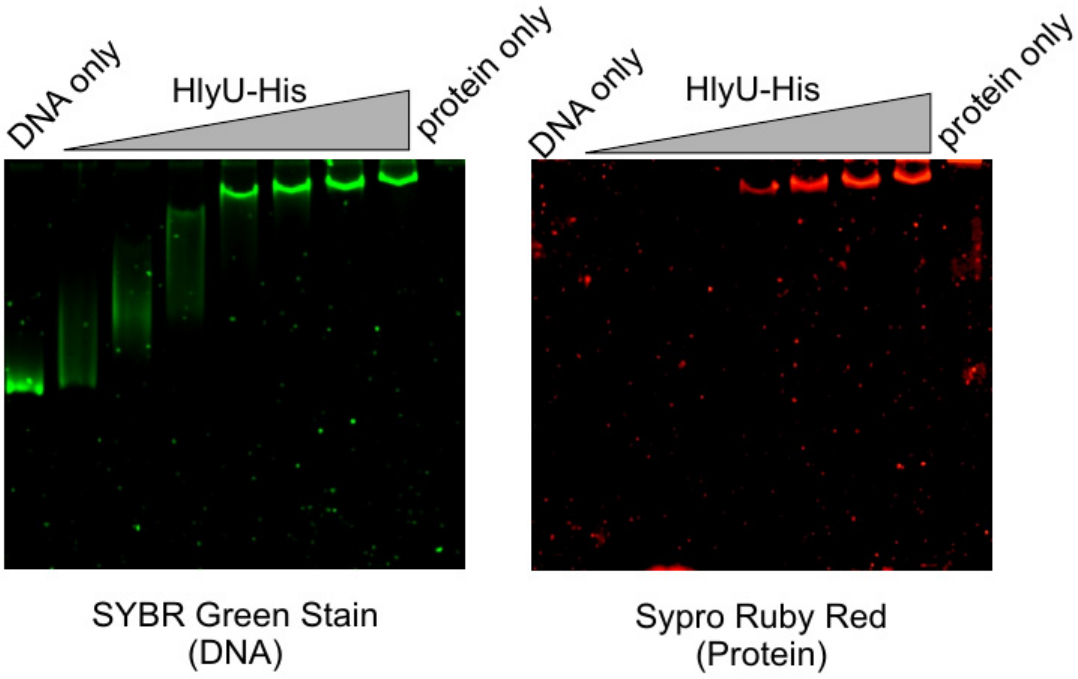

Figure A.3: Electrophoretic mobility shift assay (EMSA) with the exsA promoter region and purified HlyU-HIS. The left panel shows Sypro Green staining to specifically detect DNA within the gel. The same gel was then subjected to Sypro Ruby Red staining to specifically detect protein species within the gel. HlyU-DNA complexes are apparent as slow migrating tight bands that stained with each reagent.

Protein (ug/uL)	Protein (nM)	Unbound DNA (RFU)	Non-specifically Bound DNA (RFU))	Specifically Bound DNA (RFU)	Total (RFU)	Ratio (Bound/Total)
0	0	7380.369365	220.9168398	0	7601.286205	0
0.36	996.60	3659.839977	1255.488069	144.2732522	5059.601298	0.028514747
0.72	1993.20	584.1144885	2949.213443	108.2916266	3641.619558	0.029737216
1.44	3986.39	0	2875.867495	555.0121968	3430.879692	0.161769647
2.16	5979.59	0	756.4011737	7832.052628	8588.453801	0.911928132
2.88	7972.78	0	266.89635	11493.14347	11760.03982	0.977304809
3.6	9965.98	0	36.77605309	14500.68751	14537.46356	0.997470257
5.4	14948.97	0	2.868697518	11618.90844	11621.77714	0.999753162
B_{max} (95% CI)					11010 to 14680	
K_D (95% CI)					5046 to 6328	
H (+/- std. dev.)					8.311 (+/-) 3.547	

Table A.2: RFU units and protein concentrations used to calculate binding affinity of HlyU to DNA *in vitro*. Protein amounts and relative fluorescence units (RFU) as calculated by ImageLab2 using figure A.4 from SYBR Green stained EMSA gels containing pexsA DNA fragment and HlyU-his protein. KD here refers to the protein concentration at which 50% of the DNA is bound.Non-linear regression performed using Prism 6 according to the following equation: $y = \frac{B_{\max} \cdot X^{\mathrm{h}}}{K_{\mathrm{D}}{}^{\mathrm{H}} \cdot X^{\mathrm{h}}}$

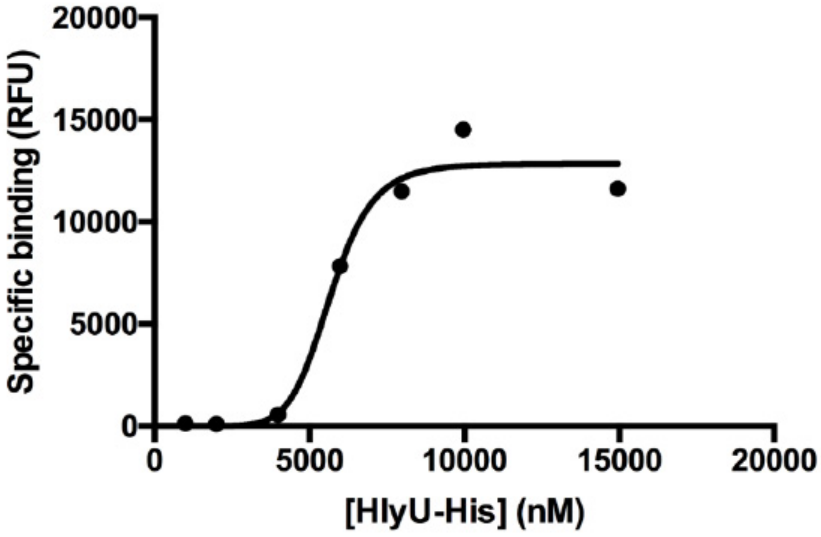

Figure A.4: Prism 6 generated graph of the non-linear regression equation above. Plotted data points are from Table A.2 specifically Bound DNA and protein concentration.

Appendix B

Chapter 3 Supplementary Data

Below is the supplementary data file for the following manuscript:

Getz, L.J., Brown, J.M., Sobot, L., Chow, A., Mahendrarajah, J., Thomas, N.A. (2021) *Attenuation of a DNA Cruciform by a Conserved Regulator Directs T3SS-1 mediated virulence in Vibrio parahaemolyticus* **BioRxiv**, DOI: 10.1101/2022.03.07.483294.

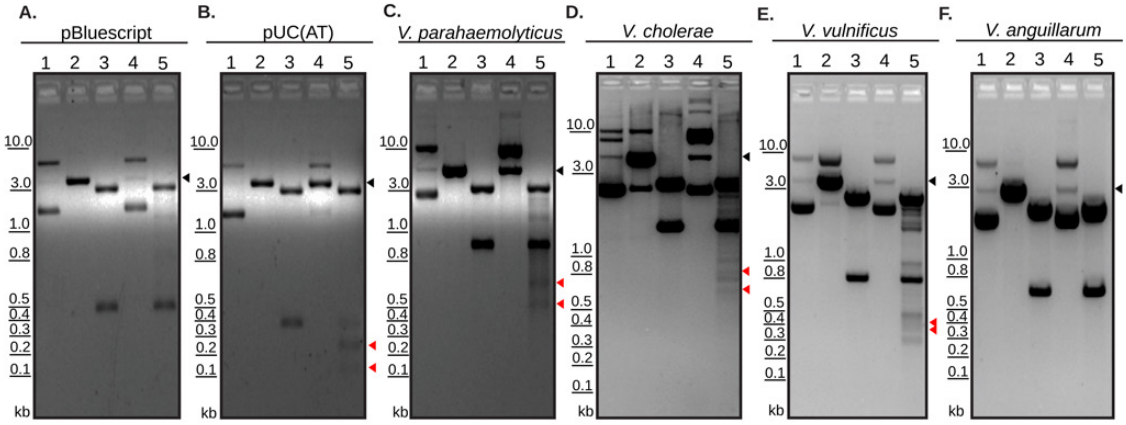

Figure B.1: Identification of cruciform structures at intergenic regions in a variety of *Vibrio* spp. Restriction digestion of cloned DNA from *Vibrio* spp. visualized by agarose gel electrophoresis. 1-undigested plasmid, 2-linearized plasmid, 3-PvuII, 4-T7 Endonuclease, 5-T7 Endonuclease followed by PvuII. T7 Endonuclease targets cruciform structures to cause double-strand break in a two-step process. PvuII was used as it flanks the cloned DNA using sequences found in the plasmid backbone and allows for restriction mapping. Red arrowheads in lanes of each panel indicate DNA fragments released by digestion by PvuII and T7 endonuclease. Black arrowheads indicate the linearized DNA fragment as shown in lane 2. pBluescript (A) and pUC(AT) (B) are negative and positive controls respectively. Other cloned DNA from *Vibrio* spp. Includes *V. parahaemolyticus exsBA* (C), *V. cholerae tlh-hlyA* (D), *V. vulnificus rtx* operon region (E), and *V. anguillarum plp-vah* (F)

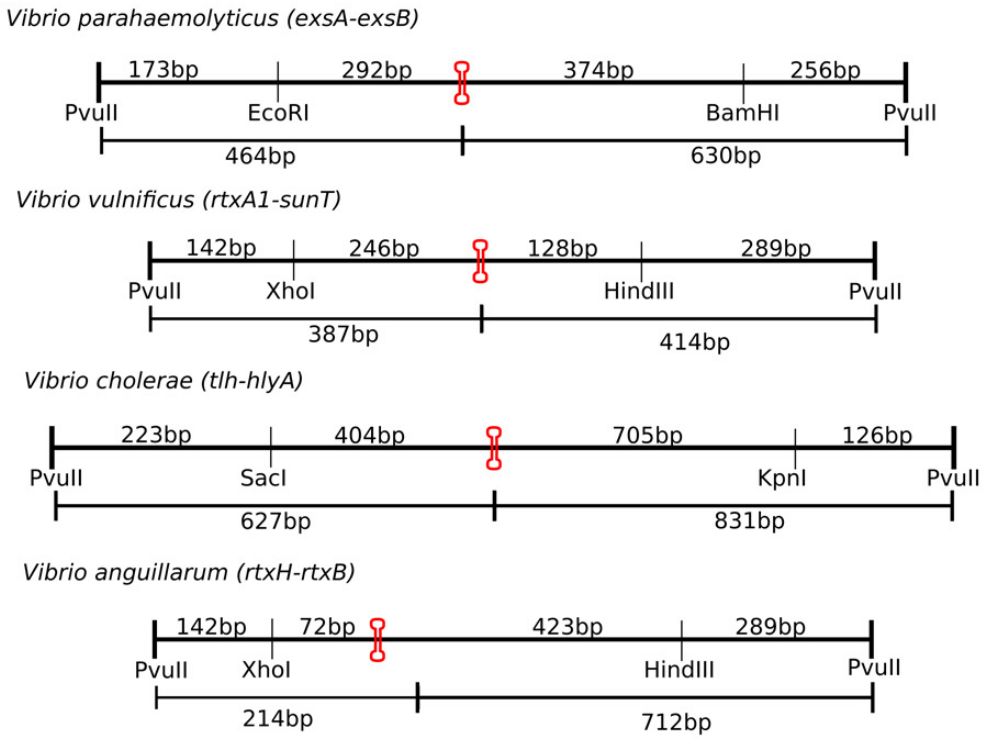

Figure B.2: Restriction maps of cloned *Vibrio* spp. DNA fragments. The terminal PvuII sites are found within pBluescript, whereas the denoted internal restriction sites were used for cloning *Vibrio* spp. DNA into the vectorś multiple cloning site. The lowest free energy cruciform for each cloned DNA fragment is shown in red. In the case of *V. anguillarum*, although cruciform structures were identified in silico, cruciform cleavage was not detected (Figure B.1).

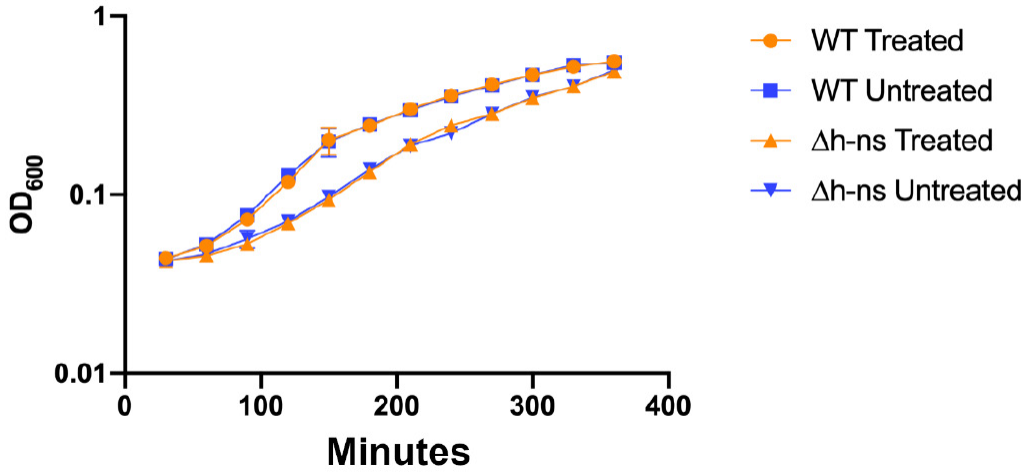

Figure B.3: Growth curve for specified *V. parahaemolyticus* strains used within the chloroacetylaldehyde pulse-chase experiment. *exsBA-lux* strains (WT) or *exsBA-lux* (Δ*hns*) were treated with chloroacetylaldehyde (CAA) or PBS (negative control, untreated). No statistical difference was apparent between treated and untreated strain growth rates over a 380-minute period.

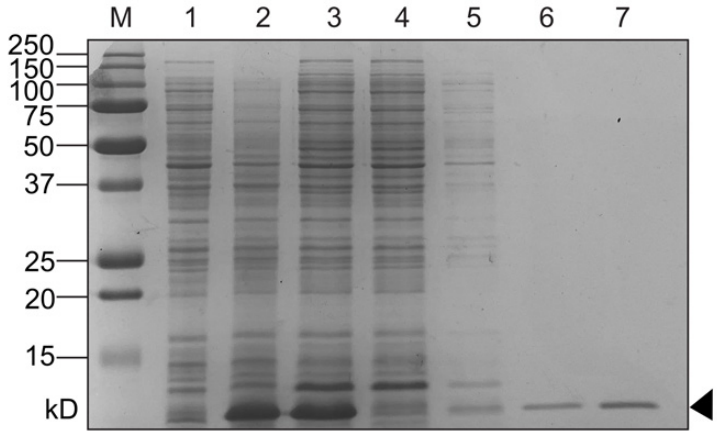

Figure B.4: HlyU-HIS Protein Purification. Nickel affinity chromatography was used to purify HIS-tagged HlyU expressed using an IPTG induction protein expression system in BL21(λDE3) E. coli. Coomassie stained SDS-PAGE lanes are as follows: M - All Blue Protein Marker (Biorad), 1 - pre-induction cell lysate, 2 - Post-IPTG induction cell lysate, 3 - pre-column soluble lysate, 4 - post-column flowthrough, 5 -Wash fraction (i), 6 - Wash fraction (ii), 7 - Elution fraction containing purified HlyU-HIS.

ΔIR1

TTTAATTTATTCTAATATAAAA (6-10-1, stem-loop-mismatch), *ΔG* 12.94

```
272:  T
273:  C
274:  T
275:  A
276:  A       TTAT
277:   TTTAAT     T
...    |||-||
298:   AAAATA     C
299:  A       TAAT
300:  C
301:  A
302:  A
303:  A
```

ΔIR2, ΔIR1ΔIR2, PAL

TTAAATAGCATAACGTTATTTAA (6-11-0, stem-loop-mismatch), *ΔG* 14.37

```
307:  A
308:  G
309:  A
310:  G
311:  T       AGCA
312:   TTAAAT     T
...    ||||||     A
334:   AATTTA     A
335:  T       TTGC
336:  A
337:  A
338:  A
339:  A
```

Figure B.5: Palindrome analyzer results for *exsBA* inverted repeats and central palindrome genetic deletions. The lowest energy cruciforms are shown. The deletion of the central palindrome alone, or inverted repeat 2 (IR2) resulted in a DNA juxtaposition that created a cruciform forming element (6-11-0) for three *exsBA* constructs (ΔIR2, ΔIR1ΔIR2, and PAL).

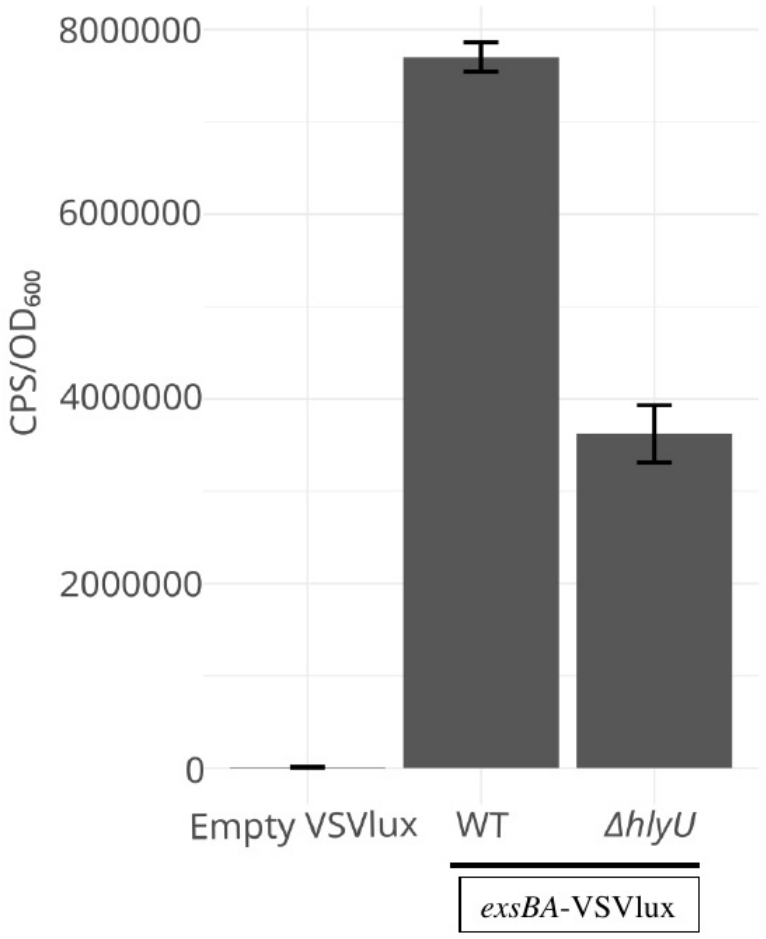

Figure B.6: A plasmid-based transcriptional fusion of the exsBA intergenic region to a *luxCDABE* cassette is dependent on HlyU for maximal activity in *V. parahaemolyticus*. Some background activity occurs in the absence of HlyU presumably due to plasmid DNA replication, and changes in supercoiling during bacterial growth. The measurements were taken 2.5 hrs post induction (magnesium +EGTA) which corresponds to the maximal activity observed for wildtype (WT) bacteria.

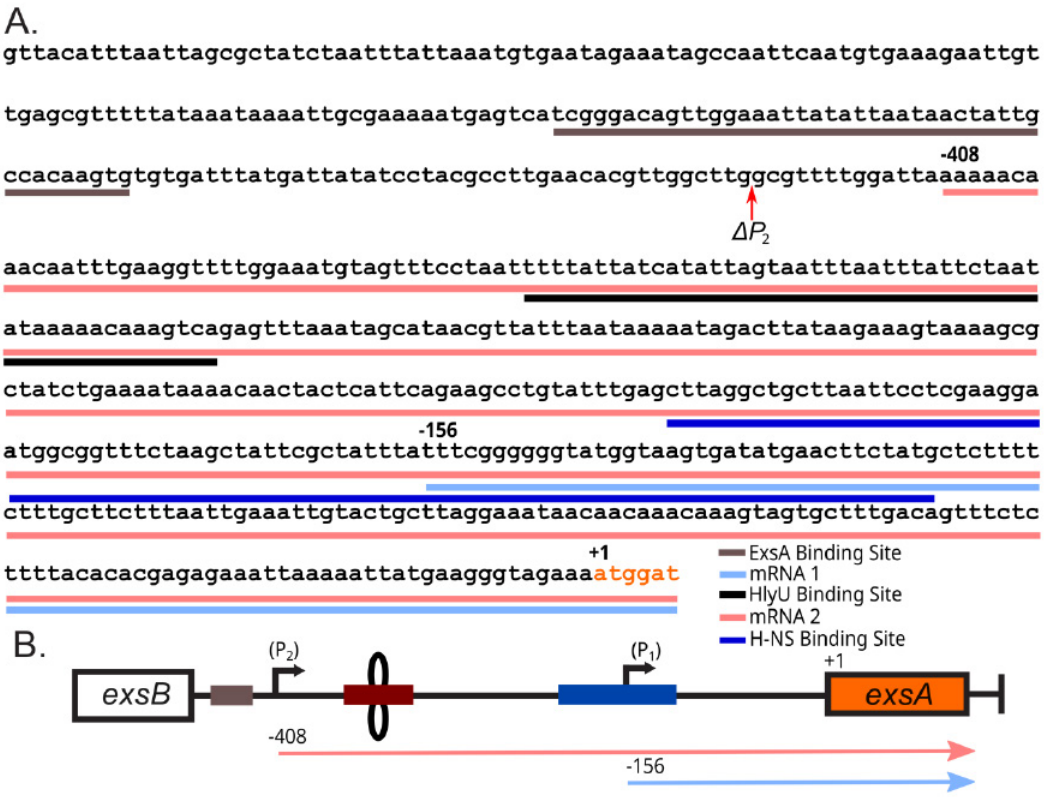

Figure B.7: Two promoters exist at the *exsBA* intergenic region that transcribe *exsA*. (A) DNA sequence of the *exsBA* intergenic region with corresponding mRNAs labelled as discovered by 5ŕRACE analysis. The light blue line indicates the shorter mRNA1 species, which begins 156bp upstream of the exsA open reading frame, while the light red line refers to the longer mRNA2 species beginning 408bp upstream of the exsA open reading frame (*exsA* ORF indicated by orange text). The ExsA protein binding site is indicated by the brown line[139, 268] , and the HlyU protected region previously reported [269] is indicated by the black line. The red arrow identifies the beginning of the ΔP_2 DNA fragment used to verify the existence of the second promoter (Figure 3.7). (B) Schematic diagram of mRNA species identified via 5'-RACE analysis in panel A. Colour scheme follows from panel A.

Sequence used in Palindrome analyser									
AAAAGTAATC	GGTTACATTT	AATTAGCGCT	ATCTAATTTA	TTAAATGTGA	ATAGAAATAG	CCAATTCAAT	GTGAAAGAAT		
TGTTGAGCGT	TTTTATAAAT	AAAATTGCGA	AAAATGAGTC	ATCGGGACAG	TTGGAAATTA	TATTAATAAC	TATTGCCACA		
AGTGTGTGAT	TTATGATTAT	ATCCTACGCC	TTGAACACGT	TGGCTTGGCG	TTTTGGATTA	AAAAACAAAC	AATTTGAAGG		
TTTTGGAAAT	GTAGTTTCCT	AAT**TTTTATT**	**ATC**ATATTAG	TAATTTAATT	TATTCTAATA	TAAAAACAAA	**GTCA**GAGTTT		
AAATAGCATA	ACGTTATTTA	ATAAAAATAG	ACTTATAAGA	AAGTAAAAGC	GCTATCTGAA	AATAAAACAA	CTACTCATTC		
AGAAGCCTGT	ATTTGAGCTT	AGGCTGCTTA	ATTCCTCGAA	GGAATGGCGG	TTTCTAAGCT	ATTCGCTATT	TATTTCGGGG		
GGTATGGTAA	GTGATATGAA	CTTCTATGCT	CTTTTCTTTG	CTTCTTTAAT	TGAAATTGTA	CTGCTTAGGA	AATAACAACA		
AACAAAGTAG	TGCTTTGACA	GTTTCTCTTT	TACACACGAG	AGAAATTAAA	AATTATGAAG	GGTAGAAAAT	GGATGTGTCA		
GGCCAACTAA	AC								

Putative Cruciform Sites Identified by Palindrome Analyser			
Sequence	Repeat-Spacer-Mismatch	ΔG	Position in Sequence
TTTAAT TTATTCTAAT ATAAAA	6-10-1	12.94	284
ATATTAG TAATTTAATTTATT CTAATAT	7-14-0	14.1	274
TTTATT CTAATATAAA AACAAA	6-10-1	14.14	289
TTAAAT AGCATAACGTT ATTTAA	6-11-0	14.37	319
ATTTAT TAAATGTGAAT AGAAAT	6-11-1	16.32	36
TTCCTAAT TTTTATTATCAT ATTAGTAA	8-12-1	16.41	256
TTAATT AGCGCTATCT AATTTA	6-10-1	16.99	19
TCTAAT TTATTAAATGTG AATAGA	6-12-1	17.34	32
AAGTAA TCGGTTACAT TTAATT	6-10-1	18.01	3

Table B.1: Palindrome Analyser output for *V. parahaemolyticus exsBA*. Intergenic sequence from *exsBA* in *Vibrio parahaemolyticus* was used as input into Palindrome Analyser - - an online bioinformatics tool which identifies putative cruciform forming sequences in provided nucleotide sequences and calculates the amount of energy required for cruciform formation. For each sequence, the 10 (or fewer) possible cruciform requiring the smallest change in free energy to form are detailed. The HlyU protected region is bolded. Cruciform structures which overlap the HlyU protected region previously identified are highlighted.

Sequence used in Palindrome analyser			
TTGGTTTAGG CACGTATCTA CTAAAGAATA AGTGCGAAAG TGAAATTAAA AAACTGGTCA GACCTATTTT GGATTTATGG TGTTAAGCGT CAATCAATAA ACGACTATGT TGAACATTAT TTGGTATTAA GATCAAACAA TGAACTGTCA GTGGAGAAAA TACGAAGGGT TTTTATAAAT CCTAATTTAG ATAAATAACA ATAAGGTTTT GCTTTATTAA TTTTGTTCTA ATGTTAAATC TTAGTAGC**TT AATAAAAATA TCAATAAAAT TATA**ATCAGT CAATATGTGT AAAAATCAAA CAGTTTTATT GTGATTTTAA GCCACATTAA AAATCATTTC TCTTTGTGTT TGTTATTTTA AGATCTATTA ATAATTCGCC ACAAAGGTGC CAAAAAACCA ATGTCCTTTT ATTTTGTGGT CTGAAAAAAT AAAAAATGGG TATTCATTTT CTTTTAAATA CTGTGATAAC AATAAAAATG CTTACGAGG			
Putative Cruciform Sites Identified by Palindrome Analyser			
Sequence	Repeat-Spacer-Mismatch	ΔG	Position in Sequence
TGATTTT AAGCCACATTA AAAATCA	7-11-0	14.17	312
TTTTATTTT GTGGTCTGAA AAAATAAAA	9-10-0	14.65	407
CCTAATT TAGATAAATAAC AATAAGG	7-12-1	15.58	181
CATTATTTG GTATTAAGAT CAAACAATG	9-10-1	15.86	115
GCACGTAT CTACTAAAGA ATAAGTGC	8-10-1	17.07	10
TAATAA AAATATCAATAAAA TTATAA	6-14-1	17.29	250
AAAAAA TGGGTATTCA TTTTCT	6-10-1	17.52	431
TGTTAA ATCTTAGTAGC TTAATA	6-11-1	17.63	232
TTGTTA TTTTAAGATCTAT TAATAA	6-13-1	17.73	350
AAAATAA AAAATGGGTAT TCATTTT	7-11-1	17.94	426

Table B.2: Palindrome Analyser output for *V. anguillarum plp-vah*. Intergenic sequence from plp-vah of V. anguillarum was used as input into Palindrome Analyser - an online bioinformatics tool which identifies putative cruciform forming sequences in provided nucleotide sequences and calculates the amount of energy required for cruciform formation. For each sequence, the 10 (or fewer) possible cruciforms requiring the smallest change in free energy to form are detailed. The DNA sequence in bold font represents the known HlyU binding site at this genetic locus. The HlyU protected region previously identified is bolded3. Cruciform structures which overlap the HlyU protected region previously identified are highlighted.

Sequence used in Palindrome analyser
CGTTTACTTC TTATCATTGA GTAATCAGAG TGAGGATCTG GTCAGATCTG TGGGTATTGT GAATGAGGTA TTCATCTGGT CAATGAAAAC ACCCAGCGGC TTCCCTAGGT TGGAAGTAAG ATCAAGTCTA GTTGATGAAA GCATCACTAT CTTCTTCTTA CTTGAAGCGG CTACTGTGTC GGTTAGGCGG GTTCATTTTG ATTGCATAGT CAATCTATGC TTATACGGGT TCACTGACTC TTAAATTATG CTTTCTTATG TGTAAGCGTA TTGAAATCTT TAGAGTTAAA ATGGAGAACT TAACTTTTTT TAAATTATTC AATTAATCTT TAAATCAACT TTATAAATTA ATTCAGACTA AATTAGTTCA AATTAAATTA GGCTCATTAA ATAATATGAA TATCAGTAAT TGTTATTTTA GTAAGAATTA TTTTACAGCA AATAAAAAGT CTTTAGAGGC TAAAATCTGT GATCCGCTGT GAATTTTCAA TTTTACCGTA TTTTACATTT AGAAACATAA GTGATATTTC AGTAAGTATG TGGTGGCAGA AAATATATAC CAAAACTCCT TGGAGTTGCA GGTAGGCGGC AAGAGAGCGA ATCCTCATGA GCATGGATAA ACTGTGTGAT TAGGATGAAC GAACGTGGCC AACACCGCTG CCGCTTCGAG TAAGAAGGGG ATATGCATTT CTGCTAAAAG GATACGCGGT AAGCCGTAGC AGTAAAGCCA CACGCAAACT CAAGGATGAC GAGGGTAACC CATGAGACAC ATGCAAAATG GGTATGTTCT AATTACTTGA AAATATAAGA ATATTACTCA ACTCAGAATT ATAGAAGAGA GTTTATTAGC AACTATTAAT TTGAGTGTTT GATATATTTC TTGTTTTTTC AGTAGTTTGA GTATAAGTCA CTTTGTTTGG AAATCTCTCT TGTAATAACA CTAAAAATAA CAGAGTCAGT GAGGTTTAT

Putative Cruciform Sites Identified by Palindrome Analyser			
Sequence	Repeat-Spacer-Mismatch	ΔG	Position in Sequence
CTAAATT AGTTCAAATT AAATTAG	7-10-1	14.41	358
TTTAAA TTATTCAATTAATC TTTAAA	6-14-0	14.41	309
TGTAAT AACACTAAAA ATAACA	6-10-1	14.96	931
TTAAAT CAACTTTATA AATTAA	6-10-1	14.98	330
TAATTGT TATTTTAGTA AGAATTA	7-10-1	15.25	407
TAAATT AATTCAGACT AAATTA	6-10-1	15.28	344
TTCTAAT TACTTGAAAAT ATAAGAA	7-11-1	15.70	787
CACTAA AAATAACAGAG TCAGTG	6-11-1	16.84	939

Sequence used in Palindrome analyser									
CGTTTACTTC	TTATCATTGA	GTAATCAGAG	TGAGGATCTG	GTCAGATCTG	TGGGTATTGT	GAATGAGGTA	TTCATCTGGT
CAATGAAAAC	ACCCAGCGGC	TTCCCTAGGT	TGGAAGTAAG	ATCAAGTCTA	GTTGATGAAA	GCATCACTAT	CTTCTTCTTA
CTTGAAGCGG	CTACTGTGTC	GGTTAGGCGG	GTTCATTTTG	ATTGCATAGT	CAATCTATGC	TTATACGGGT	TCACTGACTC
TTAAATTATG	CTTTCTTATG	TGTAAGCGTA	TTGAAATCTT	TAGAGTTAAA	ATGGAGAACT	TAACTTTTTT	TAAATTATTC
AATTAATCTT	TAAATCAACT	TTATAAAT**TA**	**ATTCAGACTA**	**AATTA**GTTCA	AATTAAATTA	GGCTCATTAA	ATAATATGAA
TATCAGTAAT	TGTTATTTTA	GTAAGAATTA	TTTTACAGCA	AATAAAAAGT	CTTTAGAGGC	TAAAATCTGT	GATCCGCTGT
GAATTTTCAA	TTTTACCGTA	TTTTACATTT	AGAAACATAA	GTGATATTTC	AGTAAGTATG	TGGTGGCAGA	AAATATATAC
CAAAACTCCT	TGGAGTTGCA	GGTAGGCGGC	AAGAGAGCGA	ATCCTCATGA	GCATGGATAA	ACTGTGTGAT	TAGGATGAAC
GAACGTGGCC	AACACCGCTG	CCGCTTCGAG	TAAGAAGGGG	ATATGCATTT	CTGCTAAAAG	GATACGCGGT	AAGCCGTAGC
AGTAAAGCCA	CACGCAAACT	CAAGGATGAC	GAGGGTAACC	CATGAGACAC	ATGCAAAATG	GGTATGTTCT	AATTACTTGA
AAATATAAGA	ATATTACTCA	ACTCAGAATT	ATAGAAGAGA	GTTTATTAGC	AACTATTAAT	TTGAGTGTTT	GATATATTTC
TTGTTTTTTC	AGTAGTTTGA	GTATAAGTCA	CTTTGTTTGG	AAATCTCTCT	TGTAATAACA	CTAAAAATAA	CAGAGTCAGT
GAGGTTTAT							

Putative Cruciform Sites Identified by Palindrome Analyser			
Sequence	Repeat-Spacer-Mismatch	ΔG	Position in Sequence
GTTTACT TCTTATCATTG AGTAATC	7-11-1	16.85	2
AATTAA ATTAGGCTCA TTAAAT	6-10-1	16.91	371

Table B.3: Palindrome Analyser output for *V. cholerae tlh-hlyA*. Intergenic sequence from tlh-hlyA of V. cholerae was used as input into Palindrome Analyser - an online bioinformatics tool which identifies putative cruciform forming sequences in provided nucleotide sequences and calculates the amount of energy required for cruciform formation. For each sequence, the 10 (or fewer) possible cruciforms requiring the smallest change in free energy to form are detailed. The lowest free energy cruciform is indicated as yellow highlighted text. The DNA sequence in bold font represents the known HlyU binding site at this genetic locus. The HlyU protected region previously identified is bolded4. Cruciform structures which overlap the HlyU protected region previously identified are highlighted.

Sequence used in Palindrome analyser			
AAATTTCCCC TACTTATTTT ATATAGATAA GACGAAAATT GTTCCTTTTA AAAGGAATCA CTCTCCGCCT GCAAGCTCAA TTAAAAGGCA AAATATAAGA ATTCAGCCAT AAATAATTAT **TGTAATGTTT ATTTTGTGTC GAAATA**TTAC ATCGTAAAAC AGTGGTCATC AATAGACTTA AATCGATTAT ATTAGAGCAA TTATTCTATT TTTATCGACC ATTATTCACT CAATTCCATA TTCATGTAAC AATCACCTTG ATTACCTATT AACGTGATAT GCATCATTCA TTTGAATCAA ATTTGTGCAA TAAACACACA ACAAAGACCA ATAAACGAGC AAAACAGTCC GCAATTGATG A			
Putative Cruciform Sites Identified by Palindrome Analyser			
Sequence	Repeat-Spacer-Mismatch	ΔG	Position in Sequence
AAATAA TTATTGTAATGT TTATTT	6-12-0	13.58	111
ATAAATA ATTATTGTAA TGTTTAT	7-10-1	15.04	109
AATAGA CTTAAATCGAT TATATT	6-11-1	17.30	171
AAGAATT CAGCCATAAAT AATTATT	7-11-1	17.66	97
TGTAATGTTT ATTTTGTGTCG AAATATTACA	10-11-1	19.07	121
ATCACCTT GATTACCTATT AACGTGAT	8-11-1	19.85	252
ATGTAA CAATCACCTTGA TTACCT	6-12-1	21.89	244
TATTCA TGTAACAATCACCT TGATTA	6-14-1	22.21	239

Table B.4: Palindrome Analyser output for *V. vulnificus rtxA1* operon. Intergenic sequence from the rtxA1 operon of V. vulnificus was used as input into Palindrome Analyser - an online bioinformatics tool which identifies putative cruciform forming sequences in provided nucleotide sequences and calculates the amount of energy required for cruciform formation. For each sequence, the 10 (or fewer) possible cruciforms requiring the smallest change in free energy to form are detailed. The HlyU binding site is indicated in bold font. Cruciform structures which overlap the HlyU protected region previously identified are highlighted.

Strain	Cytotoxicity Score
Wildtype/VSV105 (Positive)	+++
Δ*hlyU*/VSV105 (Negative)	-
WT/VSV105-*exsA*	+++
Δ*hlyU*/VSV105-*exsA*	+++
IR1/VSV105-*exsA*	+++
IR2/VSV105-*exsA*	+++
PAL2/VSV105-*exsA*	+
IV1/VSV105-*exsA*	++
IV2/VSV105-*exsA*	+++
IV3/VSV105-*exsA*	+++

Table B.5: Observed cell cytotoxicity of bacterial strains with or without complementation of *exsA*. +++: cytotoxicity equal to wildtype, ++: cytotoxicity less than wildtype, +: significantly less than wildtype, -: no cytotoxicity (relative to uninfected control). n=2

Strain Name	Organism	Genotype	Source
RIMD2210633	*V. parahaemolyticus*	Wildtype	[139]
Δ*hns*	*V. parahaemolyticus*	*hns* null	This study
Δ*vscN1*	*V. parahaemolyticus*	vscN1 null	[268]
Δ*hlyU*	*V. parahaemolyticus*	hlyU null	[269]
IR1	*V. parahaemolyticus*	IR1 mutation in *exsA* upstream region	This study
IR2	*V. parahaemolyticus*	IR2 mutation in *exsA* upstream region	This study
PAL2	*V. parahaemolyticus*	PAL2 mutation in *exsA* upstream region	This study
IV1	*V. parahaemolyticus*	IV1 mutation in *exsA* upstream region	This study
IV2	*V. parahaemolyticus*	IV2 mutation in *exsA* upstream region	This study
IV3	*V. parahaemolyticus*	IV3 mutation in *exsA* upstream region	This study
tdh::*exsBA-luxCDABE*	*V. parahaemolyticus*	*exsBA* intergenic region fused to promoter-less luxCDABE cassette and integrated into the tdh locus on chromosome	[269]
Tn5::*hns tdh*::*exsBA-luxCDABE*	*V. parahaemolyticus*	Tn5 insertion in *hns*. Strain harbors the *exsBA* intergenic region fused to promoter-less luxCDABE cassette integrated into the tdh locus on chromosome	[269]
DH5αλpir	*E. coli*		Stratagene
BL21(λDE3)	*E. coli*		Novagen

Table B.6: Bacterial Strains used in Chapter 3

Plasmid Name	Description	Source
pVSV105	Vibrio Shuttle Vector 105 contains cat, oriT, a Vibrio host ori, and oriR6Kγ. lac promoter.	[447]
pVSV105-*exsA*	VSV105 containing the *exsA* coding gene ahead of the lac promoter	This Study
pEVS104	Helper plasmid for tri-parental matings with oriT plasmids.	[539]
pVSVlux	pVSV105 containing luxCDABE cassette from pJW15 cloned blunt into the SmaI site opposite of the lac promoter.	[269]
pVSVlux-*exsBA*	pVSVlux containing the WT *exsBA* intergenic region (656 bp upstream of *exsA* start site)	[269]
pVSVlux-*exsBA*ΔIR	pVSVlux containing the *exsBA* intergenic region deleted for inverted repeat 1	This Study
pVSVlux-*exsBA*ΔIR2	pVSVlux containing the *exsBA* intergenic region deleted for inverted repeat 2	This Study
pVSVlux-*exsBA*ΔIR1ΔIR2	pVSVlux containing the *exsBA* intergenic region deleted for inverted repeats 1 and 2	This Study
pVSVlux-*exsBA*ΔPAL	pVSVlux containing the *exsBA* intergenic region deleted for central palindrome A/T rich sequence	This Study
pVSVlux-IR1	pVSVlux containing the IR1 mutation containing *exsBA* intergenic region	This Study
pVSVlux-IR2	pVSVlux containing the IR2 mutation containing *exsBA* intergenic region	This Study
pVSVlux-PAL2	pVSVlux containing the PAL2 mutation containing *exsBA* intergenic region	This Study
pVSVlux-IV1	pVSVlux containing the IV1 mutation containing *exsBA* intergenic region	This Study
pVSVlux-IV2	pVSVlux containing the IV2 mutation containing *exsBA* intergenic region	This Study
pVSVlux-IV3	pVSVlux containing the IV3 mutation containing *exsBA* intergenic region	This Study
pFLAG-CTC-hlyU	pFLAG-CTC vector containing the open reading frame coding HlyU under the control of the tac promoter	[269]
pRK415-*exsA*	pRK415 vector containing the open reading frame coding ExsA under the control of the tac promoter	This Study

Plasmid Name	Description	Source
pET21-hlyU	hlyU coding gene from *V. parahaemolyticus* cloned in frame with a C-terminal 6xHIS tag. IPTG inducible expression.	[269]
pRE112	Allelic exchange vector. Contains cat and sacB for allelic exchange procedure and oriR6Kγ origin (suicide vector)	[353]
pRE112-IR1	Allelic exchange plasmid pRE112 containing the IR1 mutation in the *exsBA* intergenic region.	This Study
pRE112-IR2	Allelic exchange plasmid pRE112 containing the IR2 mutation in the *exsBA* intergenic region.	This Study
pRE112-PAL2	Allelic exchange plasmid pRE112 containing the PAL2 mutation in the *exsBA* intergenic region.	This Study
pRE112-IV1	Allelic exchange plasmid pRE112 containing the IV1 mutation in the *exsBA* intergenic region.	This Study
pRE112-IV2	Allelic exchange plasmid pRE112 containing the IV2 mutation in the *exsBA* intergenic region.	This Study
pRE112-IV3	Allelic exchange plasmid pRE112 containing the IV3 mutation in the *exsBA* intergenic region.	This Study
pUC(AT)	Positive control cruciform construct for T7 endonuclease assays	NEB
pBluescript	General cloning vector, plasmid for cloning gBlock DNA and for use as a negative control in T7 endonuclease assays	Stratgene
p*exsBA*/pBS	PCR amplified *V. parahaemolyticus exsBA* intergenic region cloned into pBluescript for use in T7 endonuclease assays	This Study
p*hlyA*/pBS	Synthetic gene block for *V. cholerae hlyA* region cloned into pBluescript for use in T7 endonuclease assays	This Study
p*rtxA1*/pBS	Synthetic gene block for *V. vulnificus rtxA1* region cloned into pBluescript for use in T7 endonuclease assays	This Study

Plasmid Name	Description	Source
p*rtxHB*/pBS	Synthetic gene block for *V. anguillarum rtxHB* region cloned into pBluescript for use in T7 endonuclease assays	This Study
pVSVlux-ΔP2	ΔP2 deletion of the *exsBA* intergenic region cloned into pVSVlux for studying P1 expression in isolation	This Study

Table B.7: Plasmids used in Chapter 3

Name	DNA Sequence
EMSA	
IV1-1-EMSA	TGACTTTGTTTTTATTATAGAATAAATTAAATTACTATAATGATAATAAAAATTA
IV1-2-EMSA	TAATTTTTATTATCATTATAGTAATTTAATTTATTCTATAATAAAAACAAAGTCA
IV2-1-EMSA	TGACTTTGTTTTTATATTATAATAAATTAAATTAATAATATGATAATAAAAATTA
IV2-2-EMSA	TAATTTTTATTATCATATTATTAATTTAATTTATTATAATATAAAAACAAAGTCA
IV3-1-EMSA	TGACTTTGTTTTTATAGCAGAATAAATTAAATTACTGCTATGATAATAAAAATTA
IV3-2-EMSA	TAATTTTTATTATCATAGCAGTAATTTAATTTATTCTGCTATAAAAACAAAGTCA
IR1-1-EMSA	TGACTTTGTTTTTATATTAGAATAAATTAAATTATATCATAGATAATAAAAATTA
IR1-2-EMSA	TAATTTTTATTATCTATGATATAATTTAATTTATTCTAATATAAAAACAAAGTCA
IR2-1-EMSA	TGACTTTGTTTTTTATGAATAATAAATTAAATTACTAATATGATAATAAAAATTA
IR2-2-EMSA	TAATTTTTATTATCATATTAGTAATTTAATTTATTATTCATAAAAAACAAAGTCA
PAL-1-EMSA	TGACTTTGTTTTTATATTAGATTAATAATAATAACTAATATGATAATAAAAATTA
PAL-1-EMSA	TAATTTTTATTATCATATTAGTTATTATTATTAATCTAATATAAAAACAAAGTCA
PAL-2-EMSA	TGACTTTGTTTTTATATTAGATATATTAATATATCTAATATGATAATAAAAATTA
PAL-2-EMSA	TAATTTTTATTATCATATTAGATATATTAATATATCTAATATAAAAACAAAGTCA
WT-EMSA	TGACTTTGTTTTTATATTAGAATAAATTAAATTACTAATATGATAATAAAAATTA
WT-EMSA	TAATTTTTATTATCATATTAGTAATTTAATTTATTCTAATATAAAAACAAAGTCA
PCR Amplifcation and Cloning	
NT139	CATCATAACGGTTCTGGCAAATATTC
NT140	CTGTATCAGGCTGAAAATCTTCTCTC
NT337 (*exsBA* intergenic)	CC**GAATTC**AATCGGTTACATTTAATTAGCGC
NT339 (*exsBA* intergenic)	CC**GGATCC**CCGTTTCTGTGTTTAGTTGGCCTG
NT387	AAACCG**CATATG**GATGTGTCAGGCCAACTAAACAC
NT388	CC**GGTACC**ATTCGCGATGGCGACTTGCTCATCACC
NT452 (ΔP2 *exsBA* intergenic)	CCGAGCTCGCGTTTTGGATTAAAAAACAAACAATTTGAAGG
NT472	CCCC**TCTAGA**CTGTATCAGGCTGAAAATCTTCTCTC
NT473	CCCC**GGATCC**CATCATAACGGTTCTGGCAAATATTC

Name	DNA Sequence
5ŕACE	
AL400	CTGCGATAGCAAGGCATAGAGGACT
Q0	CCAGTGAGCAGAGTGACG
GSP1	ACGACATCTATGGTGCAATCGC
GSP2	CCGTTTCTGTGTTTAGTTGGC

Table B.8: Table of Oligonucleotides used in Chapter 3. Restriction enzyme recognition sites are bolded.

Below is a list of the gBlock DNA sequences ordered from IDT for the respective *Vibrio* spp using in T7 endonuclease assays and *exsBA* intergenic deletion analyses.

V. cholerae O1 biovar El Tor str. N16961; tlh-hlyA region:

CCCCGAGCTCGATAAGCTAGCTAAGCCAGCGATTAGAATAGAGAGTCTTTTTTTCATCG
TTTACTTCTTATCATTGAGTAATCAGAGTGAGGATCTGGTCAGATCTGTGGGTATTGTG
AATGAGGTATTCATCTGGTCAATGAAAACACCCAGCGGCTTCCCTAGGTTGGAAGTAAG
ATCAAGTCTAGTTGATGAAAGCATCACTATCTTCTTCTTACTTGAAGCGGCTACTGTGT
CGGTTAGGCGGGTTCATTTTGATTGCATAGTCAATCTATGCTTATACGGGTTCACTGAC
TCTTAAATTATGCTTTCTTATGTGTAAGCGTATTGAAATCTTTAGAGTTAAAATGGAGA
ACTTAACTTTTTTTAAATTATTCAATTAATCTTTAAATCAACTTTATAAATTAATTCAG
ACTAAATTAGTTCAAATTAAATTAGGCTCATTAAATAATATGAATATCAGTAATTGTTA
TTTTAGTAAGAATTATTTTACAGCAAATAAAAAGTCTTTAGAGGCTAAAATCTGTGATC
CGCTGTGAATTTTCAATTTTACCGTATTTTACATTTAGAAACATAAGTGATATTTCAGT
AAGTATGTGGTGGCAGAAAATATATACCAAAACTCCTTGGAGTTGCAGGTAGGCGGCAA
GAGAGCGAATCCTCATGAGCATGGATAAACTGTGTGATTAGGATGAACGAACGTGGCCA
ACACCGCTGCCGCTTCGAGTAAGAAGGGGATATGCATTTCTGCTAAAAGGATACGCGGT
AAGCCGTAGCAGTAAAGCCACACGCAAACTCAAGGATGACGAGGGTAACCCATGAGACA
CATGCAAAATGGGTATGTTCTAATTACTTGAAAATATAAGAATATTACTCAACTCAGAA
TTATAGAAGAGAGTTTATTAGCAACTATTAATTTGAGTGTTTGATATATTTCTTGTTTT
TTCAGTAGTTTGAGTATAAGTCACTTTGTTTGGAAATCTCTCTTGTAATAACACTAAAA
ATAACAGAGTCAGTGAGGTTTATATGCCAAAACTCAATCGTTGCGCAATCGCGATATTC
ACAATATTAAGCGCAATATCCAGTCCAACCCTGTTGGCAAATATCAATGAAGGTACCCC
CC

V.vulnificus (CMCP6, chromosome II); rtxA1 operon intergenic region:

CCCCAAGCTTAAATTTCCCCTACTTATTTTATATAGATAAGACGAAAATTGTTCCTTTT
AAAAGGAATCACTCTCCGCCTGCAAGCTCAATTAAAAGGCAAAATATAAGAATTCAGCC
ATAAATAATTATTGTAATGTTTATTTTGTGTCGAAATATTACATCGTAAAACAGTGGTC
ATCAATAGACTTAAATCGATTATATTAGAGCAATTATTCTATTTTTATCGACCATTATT
CACTCAATTCCATATTCATGTAACAATCACCTTGATTACCTATTAACGTGATATGCATC
ATTCATTTGAATCAAATTTGTGCAATAAACACACAACAAAGACCAATAAACGAGCAAAA
CAGTCCGCAATTGATGACTCGAGCCCC

V.anguillarum (J360 chromosome II); plp-vah region:

CCCCAAGCTTTTGGTTTAGGCACGTATCTACTAAAGAATAAGTGCGAAAGTGAAATTAA
AAAACTGGTCAGACCTATTTTGGATTTATGGTGTTAAGCGTCAATCAATAAACGACTAT
GTTGAACATTATTTGGTATTAAGATCAAACAATGAACTGTCAGTGGAGAAAATACGAAG
GGTTTTTATAAATCCTAATTTAGATAAATAACAATAAGGTTTTGCTTTATTAATTTTGT
TCTAATGTTAAATCTTAGTAGCTTAATAAAAATATCAATAAAATTATAATCAGTCAATA
TGTGTAAAAATCAAACAGTTTTATTGTGATTTTAAGCCACATTAAAAATCATTTCTCTT
TGTGTTTGTTATTTTAAGATCTATTAATAATTCGCCACAAAGGTGCCAAAAAACCAATG
TCCTTTTATTTTGTGGTCTGAAAAAATAAAAAATGGGTATTCATTTTCTTTTAAATACT
GTGATAACAATAAAAATGCTTACGAGGCTCGAGCCCC

exsBA deletion fragment for ΔIR1:

CCCCGAATTCGGTACCCCGTTTCTGTGTTTAGTTGGCCTGACACATCCATTTTCTACCC
TTCATAATTTTTAATTTCTCTCGTGTGTAAAAGAGAAACTGTCAAAGCACTACTTTGTT
TGTTGTTATTTCCTAAGCAGTACAATTTCAATTAAAGAAGCAAAGAAAAGAGCATAGAA
GTTCATATCACTTACCATACCCCCCGAAATAAATAGCGAATAGCTTAGAAACCGCCATT
CCTTCGAGGAATTAAGCAGCCTAAGCTCAAATACAGGCTTCTGAATGAGTAGTTGTTTT
ATTTTCAGATAGCGCTTTTACTTTCTTATAAGTCTATTTTTATTAAATAACGTTATGCT
ATTTAAACTCTGACTTTGTTTTTAATAAATTAAATTACTAATATGATAATAAAAATTAG
GAAACTACATTTCCAAAACCTTCAAATTGTTTGTTTTTTAATCCAAAACGCCAAGCCAA
CGTGTTCAAGGCGTAGGATATAATCATAAATCACACACTTGTGGCAATAGTTATTAATA
TAATTTCCAACTGTCCCGATGACTCATTTTTCGCAATTTTATTTATAAAAACGCTCAAC
AATTCTTTCACATTGAATTGGCTATTTCTATTCACATTTAATAAATTAGATAGCGCTAA
TTAAATGTAACCGATTGAGCTCCCC

exsBA deletion fragment for ΔIR2:

CCCCGAATTCGGTACCCCGTTTCTGTGTTTAGTTGGCCTGACACATCCATTTTCTACCC
TTCATAATTTTTAATTTCTCTCGTGTGTAAAAGAGAAACTGTCAAAGCACTACTTTGTT
TGTTGTTATTTCCTAAGCAGTACAATTTCAATTAAAGAAGCAAAGAAAAGAGCATAGAA
GTTCATATCACTTACCATACCCCCCGAAATAAATAGCGAATAGCTTAGAAACCGCCATT
CCTTCGAGGAATTAAGCAGCCTAAGCTCAAATACAGGCTTCTGAATGAGTAGTTGTTTT
ATTTTCAGATAGCGCTTTTACTTTCTTATAAGTCTATTTTTATTAAATAACGTTATGCT
ATTTAAACTCTGACTTTGTTTTTATATTAGAATAAATTAAATTAGATAATAAAAATTAG
GAAACTACATTTCCAAAACCTTCAAATTGTTTGTTTTTTAATCCAAAACGCCAAGCCAA
CGTGTTCAAGGCGTAGGATATAATCATAAATCACACACTTGTGGCAATAGTTATTAATA
TAATTTCCAACTGTCCCGATGACTCATTTTTCGCAATTTTATTTATAAAAACGCTCAAC
AATTCTTTCACATTGAATTGGCTATTTCTATTCACATTTAATAAATTAGATAGCGCTAA
TTAAATGTAACCGATTGAGCTCCCC

exsBA deletion fragment for ΔIR1ΔIR2:

CCCCGAATTCGGTACCCCGTTTCTGTGTTTAGTTGGCCTGACACATCCATTTTCTACCC
TTCATAATTTTTAATTTCTCTCGTGTGTAAAAGAGAAACTGTCAAAGCACTACTTTGTT
TGTTGTTATTTCCTAAGCAGTACAATTTCAATTAAAGAAGCAAAGAAAAGAGCATAGAA
GTTCATATCACTTACCATACCCCCCGAAATAAATAGCGAATAGCTTAGAAACCGCCATT
CCTTCGAGGAATTAAGCAGCCTAAGCTCAAATACAGGCTTCTGAATGAGTAGTTGTTTT
ATTTTCAGATAGCGCTTTTACTTTCTTATAAGTCTATTTTTATTAAATAACGTTATGCT
ATTTAAACTCTGACTTTGTTTTTAATAAATTAAATTAGATAATAAAAATTAGGAAACTA
CATTTCCAAAACCTTCAAATTGTTTGTTTTTTAATCCAAAACGCCAAGCCAACGTGTTC
AAGGCGTAGGATATAATCATAAATCACACACTTGTGGCAATAGTTATTAATATAATTTC
CAACTGTCCCGATGACTCATTTTTCGCAATTTTATTTATAAAAACGCTCAACAATTCTT
TCACATTGAATTGGCTATTTCTATTCACATTTAATAAATTAGATAGCGCTAATTAAATG
TAACCGATTGAGCTCCCC

exsBA deletion fragment for ΔPAL:

CCCCGAATTCGGTACCCCGTTTCTGTGTTTAGTTGGCCTGACACATCCATTTTCTACCC
TTCATAATTTTTAATTTCTCTCGTGTGTAAAAGAGAAACTGTCAAAGCACTACTTTGTT
TGTTGTTATTTCCTAAGCAGTACAATTTCAATTAAAGAAGCAAAGAAAAGAGCATAGAA
GTTCATATCACTTACCATACCCCCCGAAATAAATAGCGAATAGCTTAGAAACCGCCATT
CCTTCGAGGAATTAAGCAGCCTAAGCTCAAATACAGGCTTCTGAATGAGTAGTTGTTTT
ATTTTCAGATAGCGCTTTTACTTTCTTATAAGTCTATTTTTATTAAATAACGTTATGCT
ATTTAAACTCTGACTTTGTTTTTATATTAGCTAATATGATAATAAAAATTAGGAAACTA
CATTTCCAAAACCTTCAAATTGTTTGTTTTTTAATCCAAAACGCCAAGCCAACGTGTTC
AAGGCGTAGGATATAATCATAAATCACACACTTGTGGCAATAGTTATTAATATAATTTC
CAACTGTCCCGATGACTCATTTTTCGCAATTTTATTTATAAAAACGCTCAACAATTCTT
TCACATTGAATTGGCTATTTCTATTCACATTTAATAAATTAGATAGCGCTAATTAAATG
TAACCGATTGAGCTCCCC

Appendix C

Chapter 5 Supplementary Data

Code generated to visualize the TnSeq data, using outputs from TRANSIT's tpp are shown here [472]. Visualization was generated using R and the GViz packages, using adapted code from McCoy *et al.* [504–506].

```
# Adapted from MAGenTA Commandline tools to fit TRANSIT data outputs by Landon J. Getz (2022)
# Step 1: Prepare Environment:

# Load libraries for installed packages
library(seqinr)
library(Biostrings)
library(Gviz)

setwd('../TransitAnalysis/Viz/')
load(".RData")

# Fasta file for the reference genome, available on NCBI genomes (RIMD2210633 from NCBI)
Chr1File = "Chr1.fasta"
Chr2File = "Chr2.fasta"

# Gene Features formatted for Gviz track from a Genbank file using the
# GetGeneFeatures tool
# Format: tab-delimited text file with gene_id, gene_name, start, end , strand, # and function fields
Chr1geneFile = "gviz_Chr1.txt"
Chr2geneFile = "gviz_Chr2.txt"

# Step 2: Genome Axis and Sequence Tracks

# so Gviz doesn't look for NC_003028 in the UCSC site
options(ucscChromosomeNames = FALSE)

#Plot gTrack
gTrack <- GenomeAxisTrack(fontsize=12)
fcol <- c(A = "darkorange", C = "yellow", T = "darkred", G = "darkgreen")
```

```
Chr1sTrack <- SequenceTrack(Chr1File, name = "Chr1", fontcolor = fcol)
Chr2sTrack <-  SequenceTrack(Chr2File, name = "Chr2", fontcolor = fcol)

# Step 3: Genome Annotation Track

# A formatted file containing gene id, start, end coordinates, etc.
# Make sure coordinate columns 2 and 3 are numeric.  If not then reassign as numeric:
# column is.numeric(genes[,2])
Chr1geneDF <- as.data.frame(read.table(file = Chr1geneFile, header = TRUE))
Chr2geneDF <- as.data.frame(read.table(file = Chr2geneFile, header = TRUE))

Chr1anTrack <- AnnotationTrack(start = Chr1geneDF\$start, end = Chr1geneDF\$end, chromosome = "genome", strand = Chr1geneDF\$strand,
id = Chr1geneDF\$gene_name, showFeatureId = TRUE, name = "Genes", fill = "gray", fontcolor.item = "black", fontsize="16")

Chr2anTrack  <- AnnotationTrack(start = Chr2geneDF\$start, end = Chr2geneDF\$end, chromosome = "genome", strand = Chr2geneDF\$strand, id = Chr2geneDF\$gene_name, showFeatureId = TRUE, name = "Genes", fill = "gray", fontcolor.item = "black", fontsize="16")

# Step 4: Data Track (Reformat Chitin and Glucose Reads)

insertFit <- read.table("ChitinT2_Chr1.counts", header=TRUE)
insertFit\$seqnames <- "genome"
insertFit\$start = insertFit\$coord
insertFit\$end <- insertFit\$coord + 1
keepCols <- c("seqnames", "start", "end", "Tot_Templ_Ct")
insertFit <- insertFit[keepCols]
colnames(insertFit) <- c("seqnames","start", "end", "count")
insertFit2 <- read.table("GlucoseT2_Chr1.counts", header=TRUE)
insertFit2\$seqnames <- "genome"
insertFit2\$start = insertFit2\$coord
insertFit2\$end <- insertFit2\$coord + 1
keepCols <- c("seqnames", "start", "end", "Tot_Templ_Ct")
insertFit2 <- insertFit2[keepCols]
colnames(insertFit2) <- c("seqnames", "start", "end", "count")
merged <- merge(insertFit, insertFit2, by = c("seqnames", "start", "end"), all = TRUE)
colnames(merged) <- c("seqnames", "start", "end", "chitin", "glucose")
insertDataChr1 <- makeGRangesFromDataFrame(merged, keep.extra.columns = TRUE, ignore.strand = TRUE, seqinfo = NULL, seqnames.field = "seqnames", start.field = "start", starts.in.df.are.0based = FALSE)
```

```
insertDataChr1

insertFitChr2 <- read.table("ChitinT2_Chr2.counts", header=TRUE)
insertFitChr2\$seqnames <- "genome"
insertFitChr2\$start = insertFitChr2\$coord
insertFitChr2\$end <- insertFitChr2\$coord + 1
keepCols <- c("seqnames", "start", "end", "Tot_Templ_Ct")
insertFitChr2 <- insertFitChr2[keepCols]
colnames(insertFitChr2) <- c("seqnames","start", "end", "count")
insertFitChr2.2 <- read.table("GlucoseT2_Chr2.counts", header=TRUE)
insertFitChr2.2\$seqnames <- "genome"
insertFitChr2.2\$start = insertFitChr2.2\$coord
insertFitChr2.2\$end <- insertFitChr2.2\$coord + 1
keepCols <- c("seqnames", "start", "end", "Tot_Templ_Ct")
insertFitChr2.2 <- insertFitChr2.2[keepCols]
colnames(insertFitChr2.2) <- c("seqnames", "start", "end", "count")
merged2 <- merge(insertFitChr2, insertFitChr2.2, by = c("seqnames",
↪  "start", "end"), all = TRUE)
colnames(merged2) <- c("seqnames", "start", "end", "chitin",
↪  "glucose")
insertDataChr2 <- makeGRangesFromDataFrame(merged,
↪  keep.extra.columns = TRUE, ignore.strand = TRUE, seqinfo = NULL,
↪  seqnames.field = "seqnames", start.field = "start",
↪  starts.in.df.are.0based = FALSE)
insertDataChr2

Chr1aggTrack <- DataTrack(insertDataChr1, chromosome = "genome",
↪  groups = c("chitin", "glucose"), ylim = c(0,300), type = c("h"),
↪  col.baseline = "gray", col = c("darkblue", "darkgreen"), legend
↪  = TRUE, fontcolor.legend = "black", box.legend = TRUE, name =
↪  "Read Counts")

Chr2aggTrack <- DataTrack(insertDataChr2, chromosome = "genome",
↪  groups = c("chitin", "glucose"), ylim = c(0,300), type = c("h"),
↪  col.baseline = "gray", col = c("darkblue", "darkgreen"), legend
↪  = TRUE, fontcolor.legend = "black", box.legend = TRUE, name =
↪  "Read Counts")

# Step 5: Plot Tracks

# NagBC-NagE
# Specify viewing window start and end coordinates
fromCoord = 858900
toCoord = 863500
# Change surrounding flanking region x = 1000
y = 500
# Highlight Track
htChr1 <- HighlightTrack(trackList = list(Chr1aggTrack), fill =
↪  "white",
col = "transparent", start = fromCoord, end = toCoord, chromosome =
↪  "genome")
# Plot all Tracks
```

```
nagEnagBC <- plotTracks(list(gTrack, Chr1anTrack, htChr1), background.title = "lightgray", fontcolor = "black", fontsize = 10, from = fromCoord - y, to = toCoord + y, type = c("a"), fontcol = "black",
col.title = "black", axis.col = "black", fontface.main = 2, fontfamily = "Arial")

# NagB
# Specify viewing window start and end coordinates
fromCoordChr2 = 30000
toCoordChr2 = 34000
# Change surrounding flanking region x = 1000
y = 500
# Highlight Track
htChr2 <- HighlightTrack(trackList = list(Chr2aggTrack), fill = "white", col = "transparent", start = fromCoordChr2, end = toCoordChr2, chromosome = "genome")
# Plot all Tracks
nagB <- plotTracks(list(gTrack, Chr2anTrack, htChr2), background.title = "lightgray", fontcolor = "black", fontsize = 10, from = fromCoordChr2 - y, to = toCoordChr2 + y, type = c("a"), fontcol = "black", col.title = "black", axis.col = "black", fontface.main = 2, fontfamily = "Arial")

# GspC-L
# Specify viewing window start and end coordinates
fromCoord = 142500
toCoord = 154500
# Change surrounding flanking region x = 1000
y = 0
# Highlight Track
htChr1 <- HighlightTrack(trackList = list(Chr1aggTrack), fill = "white", col = "transparent", start = fromCoord, end = toCoord, chromosome = "genome")
# Plot all Tracks
GspCL <- plotTracks(list(gTrack, Chr1anTrack, htChr1), background.title = "lightgray", fontcolor = "black", fontsize = 10, from = fromCoord - y, to = toCoord + y, type = c("a"), fontcol = "black", col.title = "black", axis.col = "black", fontface.main = 2, fontfamily = "Arial")

# Chis-CBP-ABC transporter operon
# Specify viewing window start and end coordinates
fromCoord = 2604500
toCoord = 2624000
# Change surrounding flanking region x = 1000
y = 1000
# Highlight Track
htChr1 <- HighlightTrack(trackList = list(Chr1aggTrack), fill = "white", col = "transparent", start = fromCoord, end = toCoord, chromosome = "genome")
# Plot all Tracks
```

```
CBPOperon <- plotTracks(list(gTrack, Chr1anTrack, htChr1),
    background.title = "lightgray", fontcolor = "black", fontsize =
    10, from = fromCoord - y, to = toCoord + y, type = c("a"),
    fontcol = "black", col.title = "black", axis.col = "black",
    fontface.main = 2, fontfamily = "Arial")

save.image()
```