TALE OF TWO HOSTS: *VIBRIO PARAHAEMOLYTICUS* TOXICITY AGAINST HUMAN AND SHRIMP

APPROVED BY SUPERVISORY COMMITTEE

Kim Orth, Ph.D.

Beatriz Fontoura, Ph.D.

Qinghua Liu, Ph.D.

Anthony Michael, Ph.D.

DEDICATION

To my dearest family

ACKNOWLEDGEMENTS

It is really difficult to believe a great journey is coming to an end and I have been extremely lucky to be surrounded by all of my wonderful and talented friends. There are many people to whom I owe my thanks.

First and foremost, my mentor Dr. Kim Orth, you are the most amazing mentor I could ever dream for. You have always been patient, supportive and inspiring. Without your passion and guidance leading the way, I could not be where I am today. It has been an honor to be your student and I hope I have made you proud.

I would like to thank all my committee members, Dr. Beatriz Fontoura, Dr. Qinghua Liu and Dr. Anthony Michael, for all their support and brilliant ideas. I would also like to thank Dr. Andrew Camilli from Tufts University for his great advice during our collaboration. Special thanks to Lisa from Dr. Nick Grishin's lab for her great help with my projects.

To my former and current lab members, these years would not be the same without all of you guys. We have shared so many good memories both in and outside of the lab. Giomar, I am truly grateful to have the opportunity working with you and really proud of what we have achieved together. Dor, you are a great mentor and always full of wonderful ideas. Your high standards for scientific work have affected me in a positive way that I can not thank you enough for. I am so lucky to know you. Anne Marie, you are my first mentor in the Orth lab. During my rotation, you were so patient with all my questions. You truly opened my eyes about what I can achieve and helped me make the decision to join the lab. Phi, you are also a great mentor and helped me a lot in my early years before you left for your new journey.

Although you were super busy yourself during that time, you were always there for me when I needed help. Marcela, we stayed as bay mates for years and you have been a great mentor and friend. You are the first person I talk to when I get some crazy ideas. Thank you for being so patient all the time. TJ and Hyeilin, you guys were the first people I talked to when I walked into the lab for the very first time thinking about doing my rotation here. Thank you for welcoming me to the big family and having made my life here great. I know you are best friends to each other and I would like to believe I have successfully set up a spot for myself and we can stay close forever. Lingling, we were bay mates for a few years and you are always nice, patient and full of great ideas. Anju, Andrew, Nicole, Jessie, Alex, Amanda, Ann, Sunneta, Karo, Hillery, Junmei, Yiheng, Herman and Chris, thank you all so much.

I would like to say huge thanks to my wonderful and great family. Mom and dad, thank you for your unconditional love. Leaving home and coming to Dallas is one of the most important decisions I have ever made in my life. I could never do this without your support and encouragement.

Last but not least, my wonderful and beautiful wife Shi, thank you for believing in me all the time even when I doubted myself. I am lucky and blessed to have you in my life. Love you so much.

TALE OF TWO HOSTS: *VIBRIO PARAHAEMOLYTICUS* TOXICITY AGAINST HUMAN AND SHRIMP

by

PENG LI

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

December, 2016

TALE OF TWO HOSTS: *VIBRIO PARAHAEMOLYTICUS* TOXICITY AGAINST HUMAN AND SHRIMP

Publication No. _____

PENG LI

The University of Texas Southwestern Medical Center at Dallas, 2016

Supervising Professor: Kim Orth, Ph.D.

Bile is an important component of the human gastrointestinal tract with an essential role in food absorption and antimicrobial activities. Enteric bacterial pathogens have developed strategies to sense bile as an environmental cue to regulate virulence genes during infection. We discovered that *Vibrio parahaemolyticus* VtrC, along with VtrA and VtrB, are required for activating the virulence type III secretion system 2 (T3SS2) in response to bile salts. The VtrA/VtrC complex activates VtrB in the presence of bile salts. The crystal structure of the periplasmic domains of the VtrA/VtrC heterodimer reveals a β-barrel with a hydrophobic inner chamber. A co-crystal structure of VtrA/VtrC with bile

salt, along with biophysical and mutational analysis, demonstrates that the hydrophobic chamber binds bile salts and activates the virulence network. As part of a family of conserved signaling receptors, VtrA/VtrC provides structural and functional insights into the evolutionarily conserved mechanism used by bacteria to sense their environment. Acute hepatopancreatic necrosis disease (AHPND) is a newly emerging shrimp disease that has generated severe damage to the global shrimp industry. AHPND is caused by toxic strains of *Vibrio paraehaemolyticus* that have acquired a "selfish plasmid" encoding the deadly binary toxins $PirA^{vp}$/$PirB^{vp}$. To better understand the genetic features of AHPND causing *Vibrio paraehaemolyticus*, we conducted a comparative analysis using the genome sequences of the clinical isolate RIMD2210633, environmental non-AHPND and toxic AHPND isolates of *Vibrio paraehaemolyticus*. Our studies discovered a distinguishing feature that a virulent Type VI Secretions System (T6SS1), like that in RIMD2210633, is present in all of the AHPND but none of the non-AHPND strains. This T6SS1 is potentially conserved in these virulent strains because of its anti-bacterial activities to compete against other commensal bacteria during host infection.

TABLE OF CONTENTS

LIST OF PUBLICATIONS

Li P*, Rivera-Cancel G*, Kinch LN, Salomon D, Tomchick DR, Grishin NV, Orth K. Bile salt receptor complex activates a pathogenic type III secretion system. eLife 2016;5:e15718

Li P, Kinch LN, Ray A, Dalia AB, Cong Q, Nunan L, Camilli A, Grishin NV, Salomon D, Orth K. Acute Hepatopancreatic Necrosis Disease (AHPND)-causing *Vibrio parahaemolyticus* strains maintain an antibacterial Type VI Secretion System with versatile effector repertoires. In preparation.

* Indicates co-authorship

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ADP	Adenosine diphosphate
AHH	Ala-His-His
AHPND	Acute hepatopancreatic necrosis disease
AMP	Adenosine 5′-monophosphate
CA	Cholate
CDC	Chenodeoxycholate
CN	China
CT	Cholera toxin
DBD	DNA binding domain
DMEM	Dulbecco's Modified Eagle Medium
DOC	Deoxycholate
EMS	Early mortality syndrome
FBS	Fetal bovine serum
GEF	Guanine nucleotide exchange factor
GDC	Glycodeoxycholate
Hcp	Haemolysin coregulated protein

IM	Inner membrane
JP	Japan
LB	Luria Bertani
LC	Loading control
LDH	Lactate dehydrogenase
MFS	Major facilitator superfamily
MIX	Marker for type six effectors
MLB	Marine Luria Bertani
MOI	Multiplicity of infection
MX	Mexico
OM	Outer membrane
PAAR	Pro-Ala-Ala-Arg
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SST	sea surface temperature
TCA	Trichloroacetic acid
TCP	Toxin-coregulated pilus

TDC	Taurodeoxycholate
TDH	Thermostable direct hemolysin
TH	Thailand
TRH	TDH-related hemolysin
T3SS	Type III secretion system
T3SS1	Type III secretion system 1
T3SS2	Type III secretion system 2
T6SS	Type VI secretion system
T6SS1	Type VI secretion system 1
T6SS2	Type VI secretion system 2
VAI	*Vibrio*-relative abundance index
VgrG	Valine-glycine repeat protein G
Vop	*Vibrio* outer protein
VT	Vietnam
WHTH	Winged-helix-turn-helix
YT	Yeast extract and Tryptone

CHAPTER ONE

Introduction and Review of Literature

Physiology of bile production

Bile is a significant component of the human gastrointestinal tract that plays an essential role in the emulsification and solubilization of lipids during food digestion (*1*). Its main components are bile salts, cholesterol, and phospholipids that together possess antimicrobial activities due to detergent-like characteristics, functioning as a potent defense system to challenge the survival and colonization of both pathogenic and commensal bacteria (*1*). Bile is synthesized in the liver, stored and concentrated in the gallbladder during interdigestive periods and then released into the duodenum upon food intake **(Figure 1)** (*1, 2*). During the process of enterohepatic circulation, bile acids are absorbed by passive diffusion along the entire gut and by active transport in the distal ileum and then returned to the liver (*1, 2*).

Bile acids represent approximately 60~70% of the compounds of bile (*1*). They are saturated, hydoxylated C-24 cyclopentanephenanthrene sterols synthesized from cholesterol in the liver (*1, 2*). Two main mechanisms regulate the process of bile acids synthesis for the production of various types of bile acids (*1*). The first one is the generation of primary bile acids in the liver and the second is the production of secondary bile acids by bacteria in the intestine through modifications of primary bile

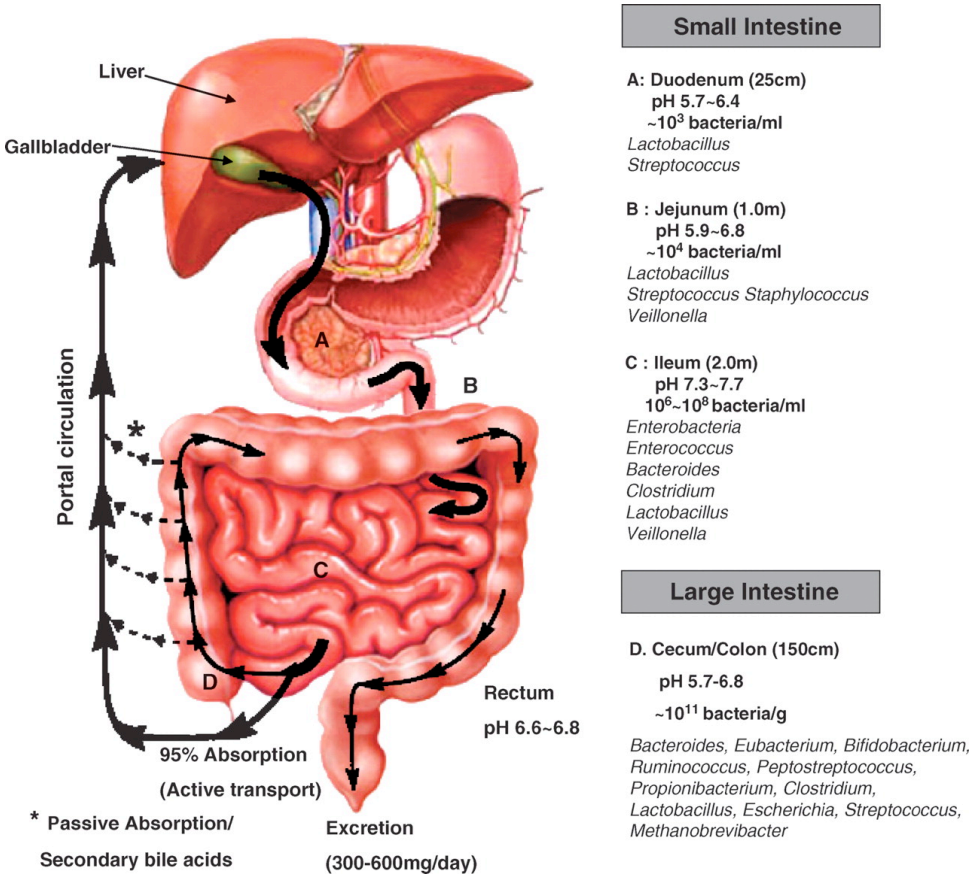


Figure 1. Enterohepatic circulation of bile

Anatomy, physiology and microbiology of the gastrointestinal tract. Large arrows denote the enterohepatic circulation of bile acids, which begins with contraction of the gallbladder, releasing bile into the duodenum. Small arrows denote the passive absorption of bile acids that escape active transport. * Secondary bile acids produced by 7α-dehydroxylation are passively absorbed in the large intestine and returned to the liver. The genera of predominant bacteria isolated from each region of the lower gastrointestinal tract are listed. From (*2*).

acids **(Figure 2)** (*1, 2*). Primary bile acids include cholic acid and chenodeoxycholic acid and they can be modified by bacterial enzymes to produce secondary bile acids deoxycholic acid and lithocholic acid **(Figure 2A)** (*1*). All bile acids are conjugated with either glycine (glycoconjugated) or taurine (tauroconjugated) as N-acyl amidates **(Figure 2B)** (*1*). Conjugated bile acids are also called bile salts.

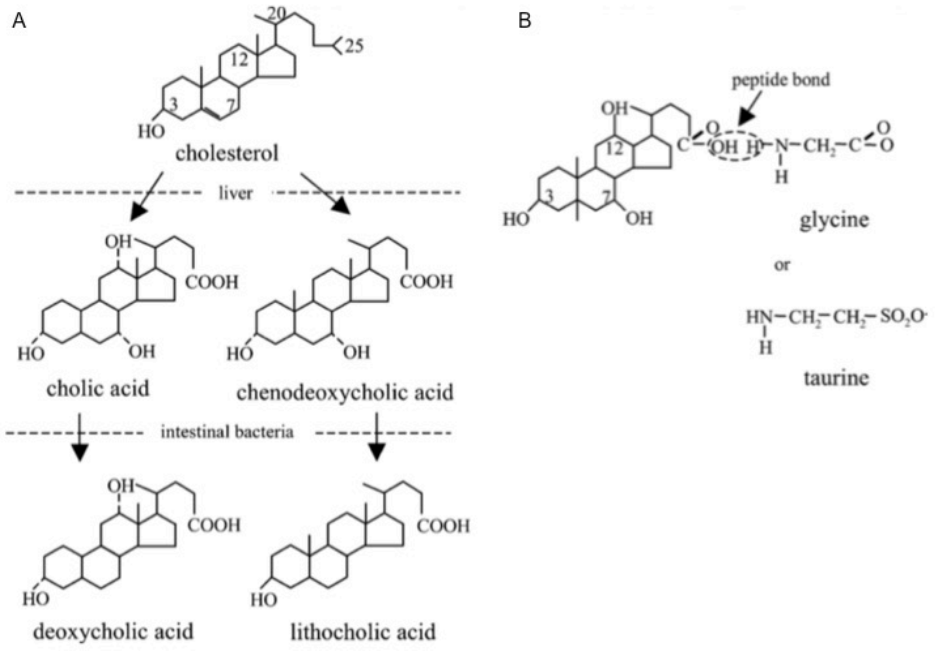


Figure 2. Biosynthesis and chemical structure of bile acids

A. Chemical structure of the major bile acids of human bile. Primary bile acids (cholic acid and chenodeoxycholic acid) are synthesized in the liver from cholesterol. They can be modified by bacterial enzymes in the intestine to form secondary bile acids (deoxycholic acid and lithocholic acid). **B.** All bile acids are conjugated with either glycine or taurine. The carboxyl group of the bile acid and the amino group of the amino acid are linked by an amide bond (peptide linkage). From (*1*).

Various enteric bacterial pathogens sense bile to regulate the activation of virulence systems

Increasing evidence has shown that enteric bacterial pathogens can utilize bile as a signaling cue to regulate their gene expression profiles during host infection. For example, *Salmonella typhimurium* senses bile to repress the genes involved in invasion when growing in the intestinal lumen and induces the expression of genes to promote invasion upon the penetration of the mucosal layer wherein the concentration of bile is decreased (*3, 4*). In *Shigella spp*., bile salts increase the pathogen's adherence and invasion of epithelial cells (*5*). Bile also has important effects on the pathogenicity of *Vibrio* species, as observed in multiple studies with pathogenic strains of both *V. cholerae* and *V. parahaemolyticus* (*6-9*). *V. cholerae* produces two major virulence factors during infection, cholera toxin (CT) and toxin-coregulated pilus (TCP), and these factors are repressed by ToxT in the presence of bile (*6-9*). However, it was also reported that bile can activate the production of CT independent of ToxT (*10*). Non-O1/non-O139 *V. cholerae* strains that do not encode CT and TCP can cause gastroenteritis utilizing a bile-activated pathogenic type III secretion system (*11-13*). Similarly, *V. parahaemolyticus* type III secretion system 2 (T3SS2) is induced specifically by bile salts during infection, leading to acute gastroenteritis (*14*). Despite all these important observations and discoveries, the mechanism underlying bile salts sensing by pathogenic bacteria still remains unknown. To investigate this mystery, we decided to use *V. parahaemolyticus* as a model to elucidate how *Vibrio spp.* senses bile salts as a signal to regulate the expression of virulence genes.

***Vibrio parahaemolyticus*, an emerging bacterial pathogen**

V. parahaemolyticus is a global, Gram-negative, halophilic bacterium that naturally lives in warm marine and estuarine environments (*15, 16*). It was first discovered by Tsunesaburo Fujino after a severe shirasu food poisoning outbreak in Japan in 1950 (*17*). *V. parahaemolyticus* is the world's leading cause of seafood-borne illness (*15, 16*). It leads to acute gastroenteritis due to the consumption of raw or undercooked seafood (*18*). It can also cause infection through the exposure to warm seawater of open wounds (*19*). In immune-competent individuals, *V. parahaemoyticus* infection is normally self-limiting and lasts for about 2 to 3 days (*18, 19*). But for individuals with underlying health conditions, the infection can lead to serious outcomes such as severe diarrhea and septicemia and in some cases subsequent death *(18, 19)*. *V. parahaemoyticus* was also shown to be the causal agent of a newly emerging shrimp disease, Acute Hepatopancreatic Necrosis Disease (AHPND), which has caused significant financial losses all worldwide (*20*). Rising ocean temperature during recent decades has contributed to fast global dissemination of *V. parahaemoyticus* and the emergence of human diseases caused by it **(Figure 3)** (*18, 19, 21-24*).

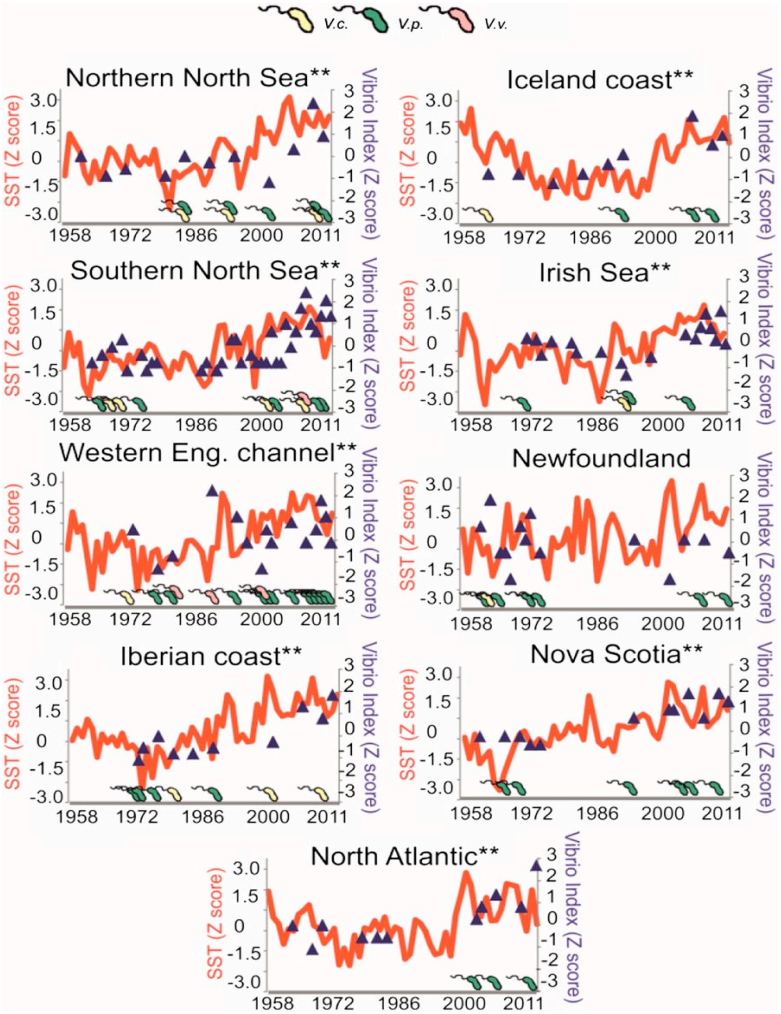


Figure 3. Relationship between *Vibrio* prokaryote abundance and sea surface temperature (SST) in the temperate North Atlantic and North Sea.

Standardized (*Z*) *Vibrio*-relative abundance index (VAI) data (blue triangles) are superimposed to standardized (*Z*) SST monthly means time series data (red line) for nine geographic areas in the temperate North Atlantic. The presence of the human pathogenic species *V. cholerae* (*V.c.*, yellow), *V. parahaemolyticus* (*V.p.*, green), and *V. vulnificus* (*V.v.*, pink) is shown. *Z*-values were obtained by subtracting the population mean and dividing the difference by the SD. $^{**}P < 0.05$, Pearson correlation analysis. From (*24*).

Key virulence factors of *Vibrio parahaemolyticus*

The whole genome sequence of a clinical *V. parahaemolyticus* strain RIMD2210633 was established in 2003 (*25*). Genome sequence analysis revealed that *V. parahaemolyticus* possesses a variety of important virulence factors including thermostable direct hemolysin (TDH) and TDH-related hemolysin (TRH), two type III secretion systems (T3SSs) and two type VI secretion systems (T6SSs) **(Figure 4)** (*26*).

TDH and TRH toxins are pore-forming exotoxins with hemolytic activities that are thought to affect ion flux in intestinal cells, leading to diarrhea (*15, 27, 28*). *V. parahaemolyticus* strains with TDH/TRH toxins are denoted as KP+ (Kanagawa Phenomenon) which is the hemolytic activity observed on blood agar (*29*). This is used as a classical method to identify virulent strains of *V. parahaemolyticus (29)*. However, there are KP- isolates of *V. parahaemolyticus* that also cause disease utilizing toxic factors other than TDH/TRH toxins. T3SS forms a secretory apparatus that directly delivers bacterial virulence factors, called effectors, into the cytoplasm of host cells to exert various functions (*30*). T6SS is another protein secretion apparatus discovered in various Gram-negative bacteria (*31*). Increasing evidence suggests that the majority of T6SSs possess anti-bacterial activities by injecting toxic effectors into neighboring cells to mediate inter-bacterial competition (*31-33*).

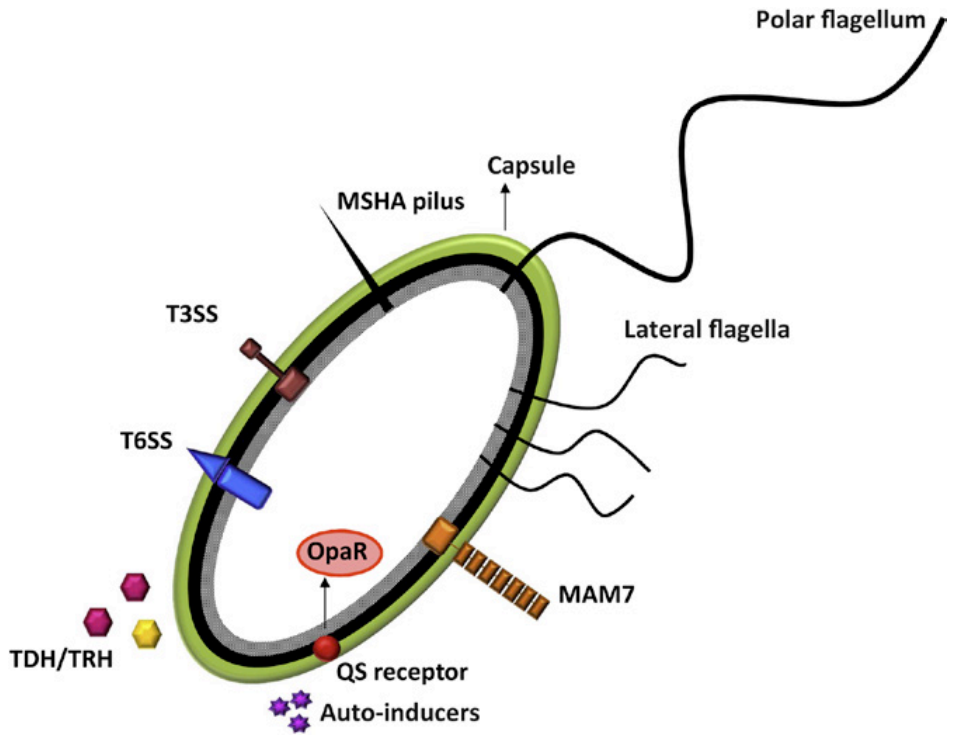


Figure 4. Virulence factors of *Vibrio parahaemolyticus*

Key virulence determinants of *Vibrio parahaemolyticus* include TDH and TRH toxins, T3SS and T6SS. TDH and TRH are toxins with hemolytic activity and likely form a pore in the host cell membrane. T3SS is a conserved protein secretion apparatus that directly inject effectors into the cytosol of host cells. T6SS is another conserved protein secretion apparatus with anti-bacterial activities that deliver toxic effectors into neighboring cells to mediate inter-bacterial competition. From (*26*).

Type III secretion system (T3SS)

T3SS is a conserved protein secretion apparatus discovered in many Gram-negative bacteria that directly delivers bacterial toxic proteins, called effectors, into the cytosol of host cells during infection **(Figure 5)** (*34*). When inside the bacteria, T3SS effectors are kept quiescent through different mechanisms including binding to chaperones. Upon activation of T3SS, the needle-like secretion apparatus is assembled and directly connects the cytosol of bacteria and host cell. T3SS effectors can then be recognized, unfolded and translocated into the host cell. The translocated effectors re-fold within the host cell to carry out their function. *V. parahaemolyticus* contains two different T3SSs (T3SS1 and T3SS2) that are differentially regulated (*15, 35*). *V. parahaemolyticus* T3SS effectors are named *Vibrio* outer proteins (Vops).

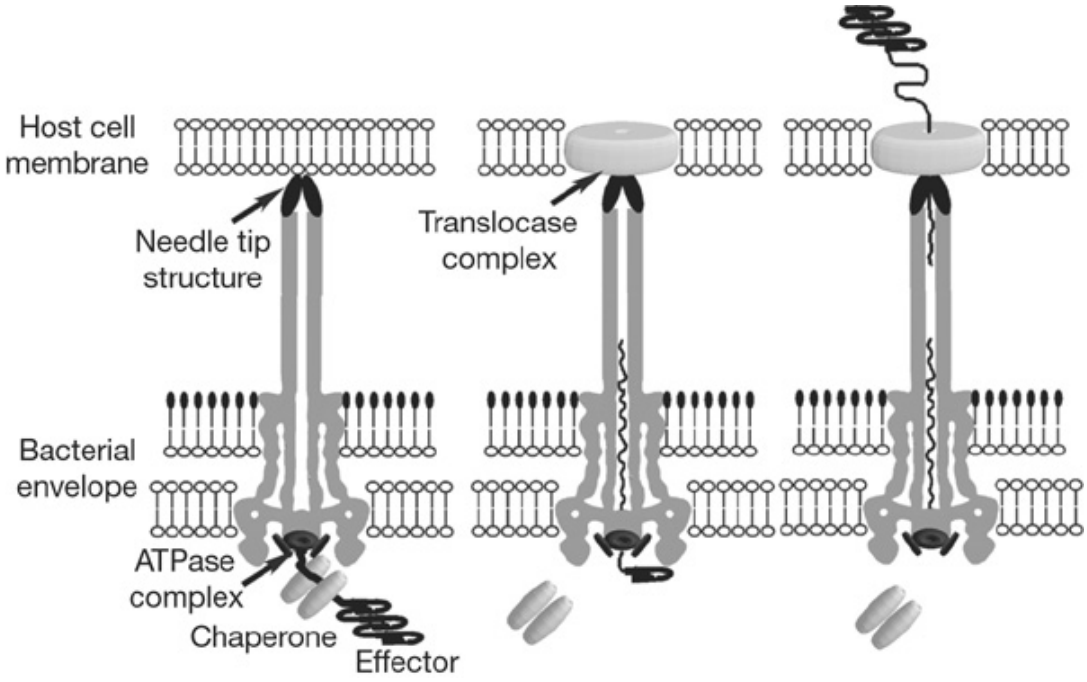


Figure 5. Type III secretion system (T3SS)

Upon activation of T3SS, the needle-like secretion apparatus is assembled directly connecting the cytosol of bacteria and host cell. T3SS effector, usually bound to chaperone when inside bacteria, is then recognized, unfolded and translocated into the host cell in an ATPase-dependent manner. The translocated effector re-folds within the host cell to carry out its function. From (*34*).

T3SS1

T3SS1 is highly conserved and present in all sequenced *V. parahaemolyticus* strains, including both environmental and clinical isolates *(36)*. It is activated by low Ca^{2+} environment as with the serum-free Dulbecco's Modified Eagle Medium and possesses cytotoxicity against various cultured eukaryotic cell lines (*37, 38*).

Three T3SS1 effectors have been functionally characterized, VopQ (VP1680), VopS (VP1686) and VPA0450 **(Table 1)** (*26*). While VopQ and VopS are located inside the T3SS1 cluster in chromosome one, VPA0450 is found in chromosome two and outside the T3SS1 cluster. Upon T3SS1 activation, these effectors orchestrate a temporally regulated cell death: First, VopQ blocks autophagy causing a rapid accumulation of autophagic vesicles and disruption of ion homeostasis. Second, VPA0450-induced plasma membrane blebbing mediated by the dissociation of the actin cytoskeleton from the plasma membrane. Third, VopS then causes cell shrinkage and rounding by transferring adenosine 5′-monophosphate (AMP) moiety to Rho GTPases, all of which culminates with cell lysis **(Figure 6)** (*38-41*). A fourth T3SS1 effector, VopR (VP1683), has been identified and shown to interact with PI(4,5)P2 via its N-terminal bacterial phosphoinositide-binding domain (BPD) and induces cell rounding **(Table 1)** (*26, 42*). However, its host cell target as well as its cellular function remains unknown. It is predicted, based on in-house proteomic data, that there may be other effectors secreted by this system as well.

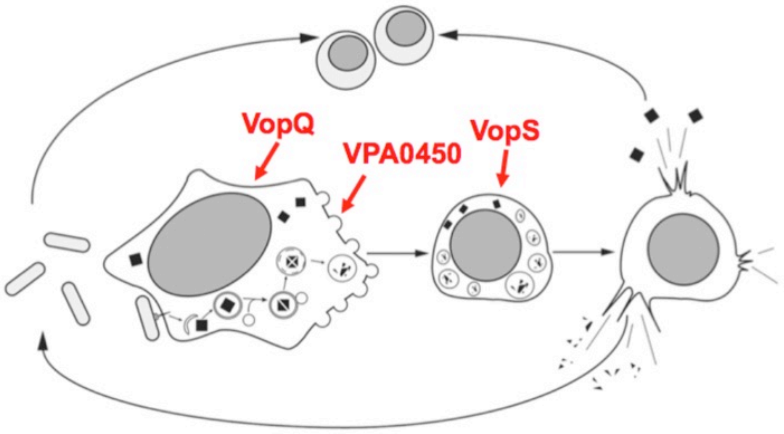


Figure 6. T3SS1 orchestrates multifaceted cell death.

Upon infection, *V. parahaemolyticus* uses its T3SS1 to inject effectors into the host cell to induce autophagy (VopQ), membrane blebbing (VPA0450), cell rounding (VopS) and cell lysis. Adapted from (*38*).

T3SS1 Effector	**Molecular activity**	**Host target**	**Biological function**
VopQ (VP1680)	Binds V-ATPases and creates gated channels on lysosomal membranes	lysosome	Autophagosome accumulation and disruption of ion homeostasis
VopS (VP1686)	AMPylator	RhoGTPases	Disruption of the actin cytoskeleton and host immunity
VPA0450	5′-phosphatase	PI(4,5)P2	Loss of plasma membrane integrity
VopR (VP1683)	unkown	unkown	Host cell rounding

Table 1. Characterized T3SS1 effectors of *V. parahaemolyticus*. From (*26*).

T3SS2

V. parahaemolyticus T3SS2 is encoded in the chromosome two and only discovered in clinical isolates of *V. parahaemolyticus (36)*. It is very similar to that present in other *Vibrio* species such as *V. mimicus* and non-O1/non-O139 *V. cholera* (*36*). T3SS2 is activated by bile salts and has been shown to be the primary virulence system causing enterotoxicity and symptoms associated with gastroenteritis during infection (*14, 43-45*). Bile salts not only induce T3SS2 effectors such as VopA and VopC, but also the needle-like secretion apparatus, including components like T3SS2 translocon VopD2 (*15, 16, 46*).

Seven T3SS2 effectors have been discovered and characterized **(Table 2)** (*26*). Four out of seven T3SS2 effectors are shown to be involved in actin polymerization: VopL and VopV mediate actin nucleation and bundling, respectively (*47, 48*); VopC and VopO function upstream of actin and target small Rho GTPases Rac and CDC42 and the RhoA guanine nucleotide exchange factor (GEF) GEF-H1, respectively (*49, 50*). Both VopA and VopZ are implicated in blocking MAPK signaling pathway. VopA is an acetyltransferase that inhibits MKKs REF (*46*). VopZ inhibits by an unknown function TAK1 (*51*). VopT is an adenosine diphosphate (ADP) – ribosyl transferase targeting the small GTPase Ras (*52*). Another T3SS2 effector, VPA1380, exhibits protein sequence similarity with cysteine protease domains utilizing inositol-6-phosphate as a co-factor and is toxic to yeast (*53*). However, its eukaryotic targets and cellular activities remain unknown.

T3SS2 Effector	Molecular activity	Host target	Biological function
VopA (VPA1346)	Acetyltransferase	MKKs	Inhibition of host cell immune responses
VopL (VPA1370)	Nucleates actin polymerization	actin	Induction stress fiber formation
VopV (VPA1357)	Binds actin and filamin	Actin, filamin	Microvilli remodeling, required for enterotoxicity
VopT (VPA1327)	ADP-ribosyl transferase	Ras	unknown
VopC (VPA1321)	Deamidase	Rac1 and CDC42	Cytoskeleton remodeling for host cell invasion
VopZ (VPA1336)	TAK1 inhibitor	TAK1	Inhibition of MAPK and NFκB pathways, required for enterotoxicity
VPA1380	Cysteine protease	unknown	Toxic to yeast
VopO (VPA1329)	Binds GEF-H1	GEF-H1	Induces actin stress fiber formation, mediate bacterial invasion and epithelial disruption

Table 2. Characterized T3SS2 effectors of *V. parahaemolyticus*. From (*26*).

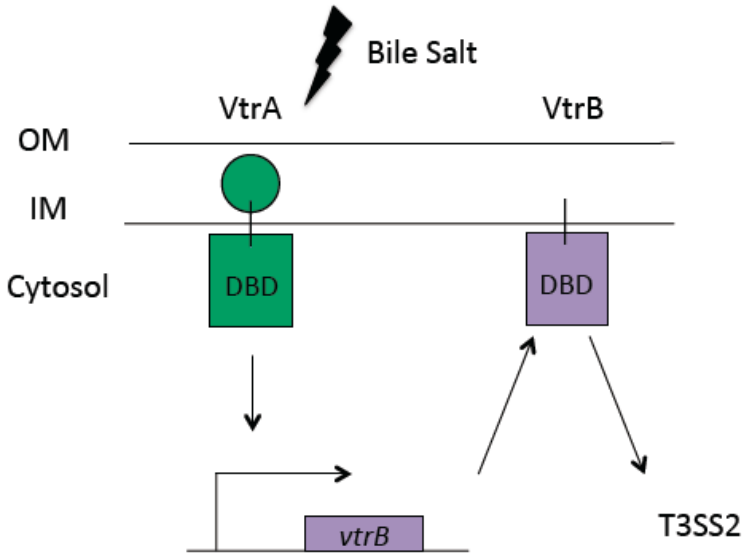


Figure 7. Bile salt actives T3SS2 via VtrA and VtrB

VtrA and VtrB are ToxR-like transmembrane transcription factors that regulate the activety of T3SS2. Upon sensing of bile salt, VtrA induces VtrB, which then activates T3SS2. DBD-DNA binding domain, OM-outer membrane, IM-inner membrane.

In *V. parahaemolyticus,* the activation of T3SS2 by bile salts is regulated by two transmembrane ToxR-like transcription factors, VtrA (VPA1332) and VtrB (VPA1348) (*14, 54*). Both VtrA and VtrB localize to the inner membrane with the N-terminal winged-helix-turn-helix (WHTH) DNA binding domain (DBD) in the cytoplasm (*54*). VtrA contains an uncharacterized C-terminal domain in the periplasm and VtrB has a short C-terminus after the transmembrane region. Upon sensing of bile salt, VtrA induces VtrB, which then activates T3SS2 **(Figure 7)**. Homologues of VtrA and VtrB, named VttRA and VttRB respectively, have been identified in T3SS-containing non-O1/non-O139 *V. cholerae* strains and shown to function in a similar way (*13*). Despite the

identification of these transcription factors, the molecular mechanism underlying bile salts sensing and T3SS2 signal propagation by *V. parahaemolyticus* and *V. cholerae* is unknown.

Type VI secretion system (T6SS)

T6SS is a conserved protein secretion apparatus discovered in various Gram-negative bacteria (*31*). A single bacterial genome may possess up to six different T6SSs (*55*). T6SS contains 13 core components that are highly conserved and additional genes which may vary between different systems (*55*). The secretion machinery of T6SS is similar to an intracellular membrane-attached contractile phage tail **(Figure 8)** (*31, 56*). It possesses an inner tail tube structure containing stacked hexameric rings of the haemolysin coregulated protein (Hcp) with a spike complex on the top made of a valine-glycine repeat protein G (VgrG) trimer and a Pro-Ala-Ala-Arg (PAAR)-repeat-containing protein (*31, 56*). T6SS effectors are positioned at the tip of the inner tail tube complex within a sheathlike structure and attached to a baseplate platform that is connected to the bacterial membrane (*31, 56*). Contraction of the sheath propels the inner tube out of the cell and delivers the effectors directly into neighboring cells (*32, 57*).

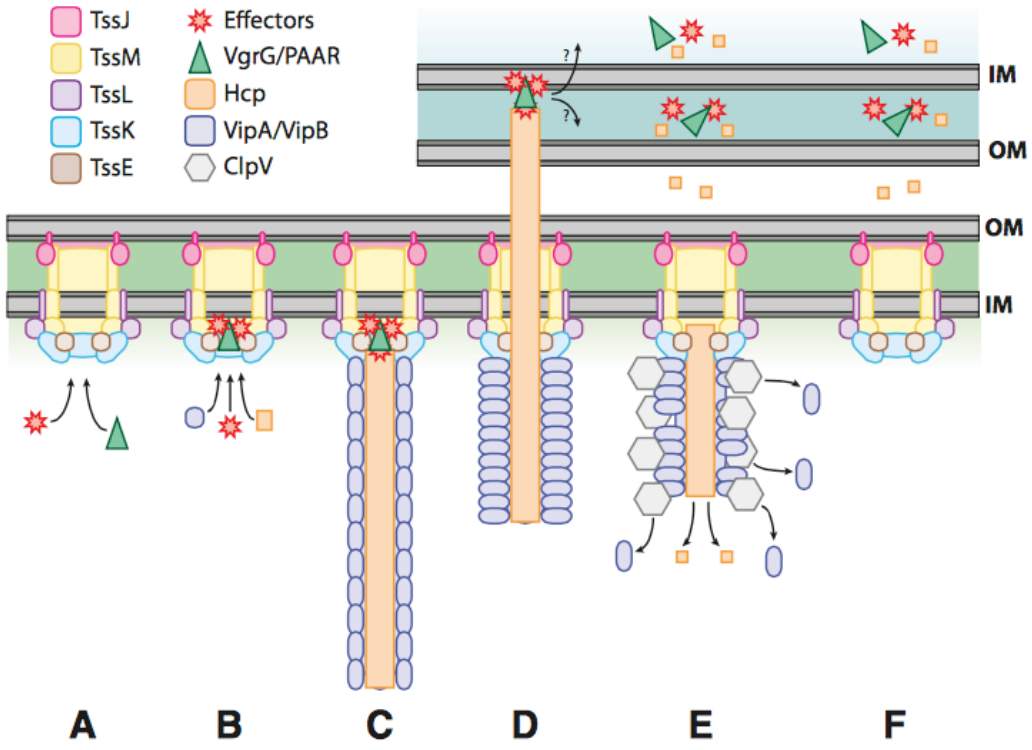


Figure 8. Mechanism of action of the T6SS

A. Assembly of the membrane and baseplate complexes at the site of secretion. **B.** VgrG, PAAR and effectors are recruited to this complex and assembled into the structure. **C.** Elongation of the cytoplasmic tubular structure built by Hcp hexamers stacking on each other coupled to the polymerization of the sheath. **D.** Once in contact with a prey cell, the T6SS sheath contracts hence propulsing the internal tube towards the target prey cell to deliver effector proteins. **E.** ClpV uses ATP to remodel the contracted sheath, restoring the pool of available sheath subunits. The unsheathed Hcp tube disassembles; parts of the tube that are not expelled from the cell are recycled within the cytosol. **F.** The naked baseplate complex is then ready to be recycled or disassembled, depending on the T6SS and its activation state. From (*31, 56*).

T6SS is strictly regulated and multiple T6SSs contained in the same bacteria are often differentially regulated (*58*). Environmental conditions that have been shown to affect the expression and secretion of T6SS components include temperature, salinity, surface sensing, quorum sensing, and membrane disruption (*31, 58-63*). Increasing evidence suggests that the majority of T6SSs possess anti-bacterial activities by injecting toxic effectors into neighboring cells to mediate inter-bacterial competition (*31-33*). Genes encoding toxic effectors and immunity proteins are often located directly next to each other and the effector–immunity protein pairs help prevent self-toxicity (*31*). *V. parahaemolyticus* RIMD2210633 strain contains two T6SSs, T6SS1 and T6SS2 (*58*).

T6SS1

T6SS1 has previously been associated predominantly with clinical isolates of *V. parahaemolyticus* and shown to possess anti-bacterial activities against various bacterial competitors (*58, 64, 65*). T6SS1 is activated by marine-like conditions, i.e., high-salt (3% NaCl) medium at 30°C (*58*). It is further induced by surface sensing indicated by increased expression and secretion of T6SS1 components and antibacterial activities (*58*). T6SS1 is likely utilized to provide growth advantage for *V. parahaemolyticus*, given its antibacterial role and the conditions that activate it (*58*).

Three T6SS1 effectors have been identified with comparative proteomics: VP1388, a MIX effector with uncharacterized activity that is encoded within the T6SS1 gene cluster; VP1415, a PAAR repeat-containing protein with a C-terminal Ala-His-His

(AHH) nuclease domain; VPA1263, a mobile MIX effector encoded outside the T6SS1 gene cluster, which contains PyocinS and colicin DNase domains (*66*).

T6SS2

T6SS2 is found in all sequenced *V. parahaemolyticus* isolates including both environmental and clinical strains (*58, 64*). Different from T6SS1, T6SS2 is activated by low-salt (1% NaCl) medium at 30°C, and repressed by surface sensing (*58*). No antibacterial activities have been shown for T6SS2 and no T6SS2 effectors have been discovered. Because T6SS1 and T6SS2 are differentially regulated, they probably function under different life styles of *V. parahaemolyticus*.

Shrimp disease acute hepatopancreatic necrosis disease (AHPND) caused by *Vibrio parahaemolyticus*

Acute hepatopancreatic necrosis disease (AHPND), previously named as early mortality syndrome (EMS), is a newly emerging shrimp disease that has caused serious production reduction and financial losses of the global shrimp aquaculture industry (*20, 67*). Since the disease outbreak firstly appeared in China in 2009, it was quickly spread to Vietnam in 2010, Malaysia in 2011, Thailand in 2012, Mexico in 2013 and Philippines in 2015 (*20, 68-70*). AHPND can cause up to 100% mortality within about 20~30 days after a pond is stocked with shrimp post-larvae (*71*). Important symptoms of affected shrimp include an empty gut and atrophied pale hepatopancreas (**Figure 9)** (*20, 67*). Histopathology analysis shows sloughing of the hepatopancreatic tubule epithelial cells

and hemocytic infiltration (*20*). In 2013, it was revealed that AHPND was caused by a specific set of *Vibrio parahaemolyticus* isolates (*20*). AHPND-causing *V. parahaemolyticus* acquired a 63~70 kbp plasmid encoding the binary toxins $PirA^{vp}$/$PirB^{vp}$ that are homologous to the *Photorhabdus luminescens* insect-related (Pir) toxins PirA/PirB (*72, 73*). $PirA^{vp}$/$PirB^{vp}$ were secreted toxins and determined to be the primary toxic factors causing AHPND (*72, 73*).

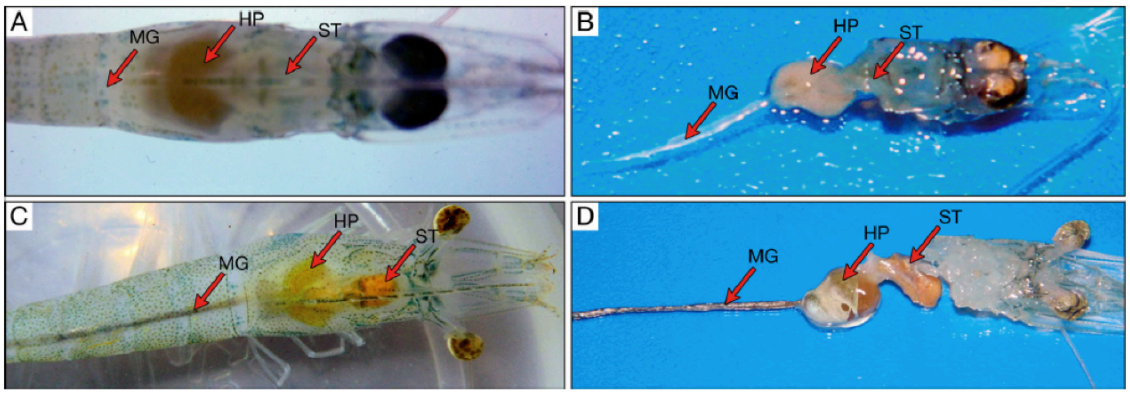


Figure 9. Gross symptoms of AHPND-infected shrimp

A and **B.** AHPND-infected shrimp shows pale and atrophied hepatopancreas (HP), and an empty stomach (ST) and midgut (MG). **C** and **D.** Normal shrimp shows a normal size HP with dark orange color and a full stomach and midgut. From (*20*).

In 2015, a strain of *Vibrio harveyi* isolated from Vietnam and a strain of *Vibrio owensii* isolated from China were shown to cause AHPND (*74, 75*). Both strains contain the plasmid that is highly similar to the one discovered in AHPND causing *V. parahaemolyticus* and encodes the binary toxins homologous to PirA/PirB (*74, 75*). This supports the discovery that $PirA^{vp}$/$PirB^{vp}$ are the main virulence factors causing AHPND.

It is also consistent with the hypothesis that this plasmid could be spread between different *Vibrio* species, probably mediated by the conjugative transfer and mobilization genes encoded by the plasmid (*72*).

Aims of this study

This thesis includes two different studies. The first study aims to reveal and characterize the mechanism how *V. parahaemolyticus* senses bile salts to regulate the activity of T3SS2 during infection. Evidence has shown that various pathogenic bacteria utilize bile as a signaling cue for the regulation of their gene expression profiles to facilitate host infection, leading to severe diseases. However, the mechanisms used by bacteria to sense bile remain largely unknown, which hinders our understanding of the infection process and design of effective therapeutic strategies. To help solve this important problem, we chose a well-studied bacterial pathogen, *V. parahaemolyticus*, utilized a combination of different techniques including bacterial genetics, biochemistry and biophysics to elucidate the process of bile salts-induced T3SS2 activation. This study aims to promote the discovery of similar and novel mechanisms used by other enteric pathogenic bacteria for sensing bile.

The second study aims to characterize the genetic features and virulence factors of AHPND causing *V. parahaemolyticus* strains. AHPND is a newly emerging shrimp disease that has devastated the global shrimp industry. Identification of *V. parahaemolyticus* secreted $PirA^{vp}$/$PirB^{vp}$ toxins as the causal factors significantly increased our understanding of this shrimp disease. However, it is far from enough for

curing the disease and managing the losses. This study aims to discover novel features of AHPND-causing *V. parahaemolyticus* strains, utilizing techniques including bacterial whole genome sequencing, bioinformatics and bacterial genetics.

CHAPTER TWO

MATERIALS AND METHODS

Bacterial strains

V. parahaemolyticus strains used in this study are listed in **Table 3** and are derived from RIMD2210633, which was a kind gift provided by Dr. Tetsuya Iida and Takeshi Honda. Strain genotypes and characteristics are detailed in **Table 3**. Non-AHPND strain A2 and AHPND strains A3, 1335, 12297B and D4 of *V. parahaemolyticus* were generously provided by Dr. Donald Lightner. *V. parahaemolyticus* strains were cultured in marine LB (MLB) with 3% NaCl at 30°C. *E.coli* Rosetta 2(DE3) used for protein purification was cultured in LB at 37°C. *E.coli* DH5α used for T6SS1 killing assay was cultured in 2xYT at 37°C.

Mammalian cell culture

HeLa cells were maintained in Dulbecco's Modified Eagle Medium (DMEM) containing 10% fetal bovine serum (FBS) and penicillin/streptomycin/glutamine at 37°C with 5% CO_2.

Antibodies

FLAG antibodies were purchased from Sigma-Aldrich (F3165 and F7425). Rabbit polyclonal antibodies were produced in house with the recombinant proteins for VopA

and VopC. Rabbit polyclonal antibodies were produced against the peptides of targets proteins by PIERCE for VtrA (VPA1332) peptide aa 235-253, VgrG1 (VP1394) peptide aa 504-522 and Hcp2 (VPA1027) peptide aa 8-20.

Deletion and knock-in strains

To generate a deletion strain of *V. parahaemolyticus*, the 1kb upstream and downstream of the nucleotide sequence to be deleted were cloned into the pDM4, a Cm^ROriR6K suicide plasmid containing the *sacB* gene, which encodes an enzyme that metabolizes sucrose into a toxic product. For C-terminal FLAG tag knock-in, the 1kb upstream of the knock-in site, followed by the nucleotide sequence encoding the FLAG tag and the 1kb downstream of the knock-in site into pDM4. The resulting pDM4 plasmids were conjugated into the POR1 strain from *E. coli* S17 (λ pir) and the transconjugants were selected on medium containing 25 μg/ml chloramphenicol. Bacteria were then counterselected by growing on medium containing 15% sucrose to select for clones were the suicide plasmid had recombined out of the chromosome. PCR analysis was performed to confirm successful deletions and knock-ins.

Quantitative RT-PCR

Total RNA was extracted from *V. parahaemolyticus* strains using RNeasy Plus Mini Kit (QIAGEN 74134). Extracted RNA was then reverse transcribed into cDNA using ProtoScript First Strand cDNA Synthesis Kit (NEB E6300S) utilizing Random Primer Mix. The resulting cDNA served as the template for quantitative RT-PCR

analysis using iTaq Universal SYBR Green Supermix (Bio-Rad) and the ViiA7 Real-Time PCR System (Applied Biosystems). The $2^{-\Delta\Delta CT}$ method was used to determine the mRNA level of *vtrA* and *vtrB* in each sample relative to POR1 grown in LB without bile salts. The expression of *vtrA* and *vtrB* was normalized against the expression of *fliA*. Primers for specific genes were as follows: *vtrA*, 5'-TTGGAACCCACGAACATCTC-3' and 5'-CAGTCACAAATTTTCCTGGCC-3'; *vtrB*, 5'-ATTATCAGCTTAGGTGGGCG-3' and 5'-ACTTTACCCCACACTTTGTCG-3'; control gene *fliA*, 5'-AAGCGATAACC TATGACCAGC-3' and 5'-TCCTCTACCTGAACACTCGG-3'. For the PCR that tests if *vtrA* and *vtrC* are cotranscribed, the cDNA served as the template and the primers used were: 5'- AATTGTTCCAGAAAGGCTCTATGTCATGCTTAATG'-3 and 5'- GTTTCA TAAAAATGAACTGGTTGAAAAAAATTG-3'.

T3SS1 expression and secretion assay

V. parahaemolyticus strains were grown overnight in MLB at 30°C. Overnight cultures were diluted to OD_{600nm}=0.3 in LB or Dulbecco's Modified Eagle Medium (DMEM) and induced for 4 hours at 37°C. For the expression fraction (cell), the same OD_{600nm} of bacterial cultures were collected and cell pellets were resuspended in 2x protein sample buffer (100 mM Tris·HCl pH 6.8, 20% glycerol, 2% sodium dodecyl sulfate (SDS), 2% β-mercaptoethanol, 150 mM sodium hydroxide, bromophenol blue). For the secretion fraction (sup), bacterial culture supernatants were filtered and precipitated with deoxycholate and trichloroacetic acid (*76*). Precipitated proteins were

pelleted and washed with acetone and then resuspended in 2x protein sample buffer. Protein expression and secretion were detected by western blot analysis.

T3SS2 expression and secretion assay

V. parahaemolyticus strains were grown overnight in MLB at 30°C. For experiments that involved vector induced expression of VtrC under its endogenous promoter, overnight cultures were diluted to OD_{600nm}=0.3 in LB supplemented with 0.05% bile salts or 0.5 mM individual bile acid taurodeoxycholate (TDC), glycodeoxycholate (GDC), chenodeoxycholate (CDC) or cholate (CA) and induced for 4 hours at 37°C. For experiments that involved vector induced expression of VtrC under the arabinose inducible promoter of pBAD, overnight cultures were diluted to OD_{600nm}=0.6 in MLB supplemented with 0.1% arabinose and induced for 3 hours at 30°C. The cultures were then diluted to OD_{600nm}=0.3 in LB supplemented with 0.1% arabinose and 0.05% bile salts and induced for 4 hour at 37°C. For the expression fraction (cell), the same OD_{600nm} of bacterial cultures were collected and cell pellets were resuspended in 2x protein sample buffer. For the secretion fraction (sup), bacterial culture supernatants were filtered and precipitated with deoxycholate and trichloroacetic acid (*76*). Precipitated proteins were pelleted and washed with acetone and then resuspended in 2x protein sample buffer. Protein expression and secretion were detected by western blot analysis.

T6SS1 expression and secretion assay

V. parahaemolyticus strains were grown in MLB at 30°C for overnight. Overnight bacterial cultures were diluted to OD_{600nm} of 0.18 with MLB -/+ 20 μM phenamil and grown at 30°C for 5h. For the expression fraction (cell), the same OD_{600nm} of bacterial cultures were collected and cell pellets were resuspended in 2x protein sample buffer. For the secretion fraction (sup), bacterial culture supernatants were filtered and precipitated with deoxycholate and trichloroacetic acid (*76*). Precipitated proteins were pelleted and washed with acetone and then resuspended in 2x protein sample buffer. VgrG1 expression and secretion were detected by western blot analysis.

T6SS2 expression and secretion assay

V. parahaemolyticus strains were grown in MLB at 30°C for overnight. Overnight bacterial cultures were diluted to OD_{600nm} of 0.9 with LB and grown at 30°C or 37°C for 5h. For the expression fraction (cell), the same OD_{600nm} of bacterial cultures were collected and cell pellets were resuspended in 2x protein sample buffer. For the secretion fraction (sup), bacterial culture supernatants were filtered and precipitated with deoxycholate and trichloroacetic acid (*76*). Precipitated proteins were pelleted and washed with acetone and then resuspended in 2x protein sample buffer. Hcp2 expression and secretion were detected by western blot analysis.

Bioinformatics

In order to identify VtrA and VtrC homologues, we performed PSI-BLAST (*77*) searches against a non-redundant (NR) database (threshold E-value cutoff 0.02) with

query sequences corresponding to the predicted periplasmic domains of VtrA (VPA1332: gi|28901187, aa 150-253) and VtrC (VPA1333: gi|28901188, aa 27-161). Because our VtrC PSI-BLAST hits were from different species than our VtrA hits, we initiated a transitive PSI-BLAST search against the NR database (threshold E-value cutoff 0.01) using the most distant identified VtrA sequence as a query (gi|494726765, aa 160-272). The transitive search found the VtrA sequences from the remaining species containing VtrC. All identified VtrA and VtrC hits were found in tandem in their respective genomes, and representative sequences from non-redundant species were collected. VtrB sequences were identified by searching VtrA/VtrC-containing genomes for the VtrB sequence (VPA1348: gi|28901203). Due to widespread nature and strong conservation of the VtrB helix-turn-helix sequence in many transcription factors, we required VtrB orthologs to: 1. be the top-scoring hit in the respective genomes, 2. retain the predicted C-terminal transmembrane helix and 3. be in close proximity to the VtrA/VtrC gene pair. Representatives from each sequence family were aligned using the MAFFT server (http://mafft.cbrc.jp/alignment/server/).

To calculate the structure-based distance tree, representative structures were chosen from the lipocalins/streptavidin group in the Evolutionary Classification Of protein Domain structures (ECOD) database. We chose structures with bound ligands, when available, from the main branches of the calycin superfamily (*78*): lipocalins (1kt7:1-175, 2aco:10-177), triabin (4n7c:3-176), fatty acid binding proteins (FABPs) (1o1v: 1-126, 3wvm:0-132), streptavidin (2f01:14-134), and metalloprotease inhibitor

(MPI) (1jiw:1-105, chain I). We also included the HRI1 N-terminal domain structure (HRIN-like; 3rby:1-151) with a unique binding site that traverses the lipocalin-like barrel core, as well as the VtrC structure. The representative lipocalin-like domains were compared pairwise all-against-all using DaliLite (*79*). Pairwise DaliLite Z-scores (Z_{AB}) were transformed to distances by comparing to self DaliLite Z-scores (Z_{AA}, Z_{BB}) using the following equation: $-\ln(Z_{AB}/(0.5*Z_{AA}+0.5*Z_{BB}))$. The structure-based tree was produced using the FITCH program (with global optimization) of the Phylip package (*80*).

Alkaline phosphatase PhoA assay

V. parahaemolyticus strains containing the empty pBAD vector, pBAD-*phoA-vtrC,* or pBAD-*vtrC-phoA* were grown overnight in MLB at 30°C. Overnight cultures were diluted to OD_{600nm}=0.6 in MLB supplemented with 0.1% arabinose and induced for 5 hours at 30°C. 1 ml of bacterial culture was collected, washed once, and resuspended with 1 ml Tris solution (1M Tris·HCl, pH 8.0). 50 µl of resuspended cells were transfered to 1 ml Tris solution and permeabilized by adding 30 µL 0.1% SDS and vortexing for 10 seconds. The mixture was then incubated at 37°C for 5 min. 100 µl of p-nitrophenyl phosphate (PNPP, Thermo Scientific Pierce 37621) was added to the samples to start the reaction and the time was recorded as T1 (min). Samples were incubated at 37°C until a pale yellow color develops. The reaction was stopped by adding 100 µl of cold 1 M KH_2PO_4 and the time was recorded as T2 (min). The supernatant was collected after centrifugation at 20000 x g for 5 min. Measure the absorbance of the supernatant with a

spectrophotometer at wavelength 420nm. Alkaline phosphatase activity was calculated as below:

Alkaline phosphatase activity (Miller unit) = 1000* OD_{420nm}/0.05* OD_{600nm} *(T2-T1)

β-galactosidase LacZ assay

V. parahaemolyticus strains containing the empty pBAD vector, pBAD-*lacZ-vtrC* or pBAD-*vtrC-lacZ* were grown overnight in MLB at 30°C. Overnight cultures were diluted to OD_{600nm}=0.6 in MLB supplemented with 0.1% arabinose and induced for 5 hours at 30°C. Bacterial cells were permeablized by mixing 20 µl of bacterial culture with 80 µl permeabilization solution (100 mM Na_2HPO_4, 20 mM KCl, 2 mM $MgSO_4$, 0.8 mg/ml hexadecyltrimethylammonium bromide, 0.4 mg/ml sodium deoxycholate, 5.4 µl/ml β-mercaptoethanol). 600 µl substrate solution (60 mM Na_2HPO_4, 40 mM NaH2PO_4, 1 mg/ml 2-Nitrophenyl β-D-galactopyranoside (ONPG, Sigma-Aldrich N1127), 2.7 µl/ml β-mercaptoethanol) was added to the samples to start the reaction and the time was recorded as T1 (min). Samples were incubated at 30°C until a pale yellow color develops. The reaction was stopped by adding 100 µl 1 M Na_2CO_3 and the time was recorded as T2 (min). The supernatant was collected after centrifugation at 20000 x g for 5 min. Measure the OD_{420nm} of the supernatant. β-galactosidase activity was calculated as below:

β-galactosidase activity (Miller unit) = 1000* OD_{420nm}/0.02* OD_{600nm} *(T2-T1)

Immunoprecipitation

The following strains were used for the Co-IP experiments on Figure 18: V. parahaemolyticus strains POR1 + pBAD, POR1ΔvtrA + pBAD-FLAG-vtrC and POR1 + pBAD-FLAG-vtrC. The following strains were used for the Co-IP experiments on **Figure 28C**: V. parahaemolyticus strains POR1 + pBAD, POR1ΔvtrA + pBAD-FLAG-vtrC and POR1 ΔvtrC + pBAD-FLAG-vtrC wild type and mutants H50R, Y90A and Q32A. Strains were grown overnight in MLB at 30°C. Overnight cultures were diluted to OD_{600nm}=0.6 in MLB supplemented with 0.1% arabinose and induced for 3 hours at 30°C. The cultures were then diluted to OD_{600nm}=0.3 in LB supplemented with 0.1% arabinose and 0.05% bile salts and induced for 4 hours at 37°C. 200 ml bacterial culture were resuspended with lysis buffer (50 mM Tris·HCl pH 7.4, 100 mM NaCl, 0.2% Triton X-100, 5 mM EDTA, 1 mg/ml lysozyme and 1 mM PMSF). Resuspended samples were shaken at 22°C for 30 min and then subjected to three freeze/thaw cycles. The supernatant was collected after centrifugation at 25000 x g, 4°C for 30 min. Immunoprecipitation was performed by incubating the supernatant with Anti-FLAG M2 beads at 4°C with gentle shaking for 4 hours. Beads were collected and washed with wash buffer (50 mM Tris·HCl pH 7.4, 100 mM NaCl, 0.2% Triton X-100) 3 times. Proteins bound to beads were eluted with 2x protein sample buffer. VtrA and VtrC were detected by western blot analysis using Anti-VtrA and Anti-FLAG antibodies.

Protein expression and purification

The periplasmic domain of VtrA (aa 161-253) was cloned as an N-terminally hexahistidine and maltose binding protein-tagged fusion protein into pET28b (Novagen)

to produce pET28b-HisMBP-VtrA, which contains a Tobacco Etch Virus (TEV) protease cleavage site between the HisMBP-tag and VtrA. The pET28b-HisMBP-VtrA construct was expressed in *E. coli* Rosetta 2(DE3) cells (Novagen). For VtrC/VtrA coexpression, the periplasmic domains of VtrC (aa 31-161) and VtrA (aa 161-253) were cloned into the first and second multiple cloning site, respectively, of pACYCDuet-1 (Novagen) to produce pACYCDuet-VtrC/VtrA, where VtrC is N-terminally hexahistidine-tagged. The pACYCDuet-VtrC/VtrA construct was expressed in *E. coli* BL21(DE3) cells. All cultures were grown in LB at 37°C until OD_{600nm} 0.5-0.6 and induced with 0.4 mM isopropyl β-D-thiogalactopyranoside (IPTG) overnight at 22°C for pET28b-HisMBP-VtrA, and 17°C for pACYCDuet-VtrC/VtrA. Selenomethionyl-derivatized VtrA/VtrC complex was expressed in *E. coli* B834(DE3) cells (Novagen) in SelenoMet medium (Molecular Dimensions). Cells were harvested by centrifugation, resuspended in buffer A (50 mM Tris pH 8.0 and 100 mM NaCl), and lysed by extrusion. Lysates were clarified by centrifugation and filtered (0.45-μm pore size). All proteins were purified by nickel-affinity purification using HisPur Ni-NTA resin (ThermoFisher) on a gravity flow column. Briefly, lysates were incubated with the resin for 30 min at 4°C with nutation. Lysate and beads were applied to the column and washed with 20 column volumes of buffer A supplemented with 15 mM imidazole. Proteins were eluted with 5 column volumes of buffer B (50 mM Tris pH 8.0, 100 mM NaCl and 250 mM imidazole). For VtrA, the HisMBP fusion domain was removed by treating with TEV protease overnight at 4°C, followed by a second round of nickel-affinity chromatography. Proteins were further purified by size exclusion chromatography (SEC) on a Superdex 75 16/600

column (Pharmacia Biotech) with buffer A. All proteins were buffer exchanged to 10 mM Tris pH 8 and 10 mM NaCl for crystallographic studies. Selenomethionyl-derivatized VtrA/VtrC complex was purified in a similar manner with the addition of 1 mM DTT to all buffers.

Gel Filtration assay

Gel filtration was performed on a Superdex S75 16/600 column (Pharmacia Biotech) with buffer containing 50 mM Tris pH 8.0 and 100 mM NaCl at a flow rate of 1 ml/min. The protein complex solution was injected into the column at a final concentration of 1 mg/ml in a total volume of 1 ml. Fractions were analyzed by SDS-PAGE and Coomassie blue staining. For protein molecular weight determination, the column was calibrated with a gel filtration LMW calibration kit (GE Healthcare) using the running buffer described above.

Crystallization and X-ray data collection

Crystals of native VtrA/VtrC periplasmic domain heterodimer were grown using the sitting-drop vapor diffusion method from drops containing 0.2 µl protein (13 mg/ml) and 0.2 µl reservoir solution (1.0 M lithium sulfate, 0.5 M ammonium sulfate, 0.1 M sodium citrate pH 5.6) and equilibrated over 50 µl reservoir solution at 20°C. Crystals of selenomethionyl-derivatized heterodimer were grown using the hanging-drop vapor diffusion method from drops containing 1 µl protein (10 mg/ml) and 1 µl of reservoir solution (1.0 M lithium sulfate, 0.5 M ammonium sulfate, 0.1 M sodium citrate pH 5.9)

and equilibrated over 500 µl of reservoir solution. Crystals appeared after 2 days at 20°C and grew to their maximal extent by 1 week. Native and selenomethionyl-derivatized crystals were cryoprotected by transferring to a final solution of 28.7% ethylene glycol, 1.1 M lithium sulfate, 0.6 M ammonium sulfate, and 0.1 M sodium citrate pH 5.6-5.8, then the crystals were flash-cooled in liquid nitrogen. Crystals of the VtrA/VtrC heterodimer bound to the bile salt TDC were grown with protein that had been loaded with TDC by performing purification in the presence of 0.5 mM TDC in all buffers and buffer exchanging into 10 mM Tris pH 8, 10 mM NaCl. Drops containing 1 µl protein (5 mg/ml) and 1 µl of reservoir solution (2.0 M ammonium sulfate, 0.1 M sodium acetate pH 4.6) were set up by the hanging-drop vapor diffusion method, and equilibrated over 500 µl of reservoir solution. Crystals appeared after 6 days at 20°C and grew to their maximal extent by 2 weeks. Crystals were cryoprotected by transferring to a final solution of 22.5% ethylene glycol, 2.1 M ammonium sulfate, 0.1 M sodium acetate pH 4.6, and 0.5 mM TDC, and flash-cooled in liquid nitrogen.

Data were collected at APS beamline 19-ID at 100 K, and were indexed, integrated and scaled using the HKL-3000 program package (*81*). Native and selenomethionyl-derivatized VtrA/VtrC heterodimer crystals exhibited the symmetry of space group F432 with unit cell parameters of a = 211.01 Å and contained one molecule each of VtrA and VtrC per asymmetric unit, with a solvent content of 65%. TDC-containing crystals belonged to space group $P2_12_12_1$ with unit cell parameters of a = 55.39 Å, b = 71.28 Å and c = 203.73 Å and contained three molecules each of VtrA/VtrC

heterodimer per asymmetric unit, with a solvent content of 50%. Native, selenomethionyl-derivatized and TDC-containing crystals diffracted isotropically to a d_{min} of 2.70 Å, 2.60 Å and 2.10 Å, respectively, when exposed to synchrotron radiation. Data collection statistics are provided in **Table 6**.

Phase determination and structure refinement

Phases for the native VtrA/VtrC heterodimer were obtained from a two-wavelength anomalous dispersion experiment using selenomethionyl-derivatized heterodimer protein with data to a d_{min} of 2.60 Å. Two selenium sites were located using the program *SHELXD* (*82*), and phases were refined with the program *Mlphare* (*83*), resulting in an overall figure-of-merit of 0.25 for data between 121.8 and 2.60 Å. Phases were further improved by density modification in the program *Parrot* (*84*) resulting in a figure-of-merit of 0.88. An initial model containing 84% of all VtrA/VtrC heterodimer residues was automatically generated in the program *Buccaneer* (*85*).

As the selenomethionyl-derivatized and native crystals were isomorphous, all further calculations for the native structure were performed versus the native data. Additional residues for the VtrA/VtrC heterodimer were manually modeled in the program *Coot* (*86*). Positional and isotropic atomic displacement parameter (ADP) as well as TLS ADP refinement was performed to a resolution of 2.70 Å using the program *Phenix* (*87*) with a random 10% of all data set aside for an R_{free} calculation. The current model contains one VtrA/VtrC heterodimer; included are residues residues 164-253 of

VtrA 31-112 and 118-161 of VtrC, and one sulfate ion. The R_{work} is 0.260, and the R_{free} is 0.298. A Ramachandran plot generated with *Molprobity* (*88*) indicated that 94.3% of all protein residues are in the most favored regions and 0.5% (one residue) in disallowed regions.

Phases for the TDC-containing heterodimer were obtained by the molecular replacement method in the program *Phaser* (*89*) using the coordinates for the native VtrA/VtrC heterodimer. Model building and refinement were performed to a resolution of 2.1 Å using a similar protocol to the native structure. Four TDC molecules were located in the asymmetric unit, one each bound to the VtrC monomer of a VtrA/VtrC heterodimer and one at a lattice contact between VtrA molecules. The current model contains three VtrA/VtrC heterodimers in the asymmetric unit, four TDC molecules, seven sulfate ions and 204 water molecules. A *Molprobity* (*88*) generated Ramachandran plot indicates that 96.5% of all protein residues are in the most favored regions and 0.3% (two residues) in disallowed regions. Phasing and model refinement statistics for all structures are provided in **Table 6**.

Isothermal titration calorimetry (ITC)

The VtrA/VtrC periplasmic domain complex was dialyzed at 4°C overnight against the assay buffer (50 mM Tris pH8, 100 mM NaCl). Taurodeoxycholic acid (400 μM) was prepared by dissolving the dry powder (Sigma) with the same dialysis buffer. ITC experiments were performed at 25°C on a MicroCal iTC200 system (Malvern), with

reference power at 5 μcal/s and stirring rate at 750 rpm. Measurements were performed as 19 injections of 400 μM taurodeoxycholic acid (1 μl for the first injection and 2 μl for injections 2-19) into approximately 200 μl of 36 μM VtrA/VtrC. ITC data were integrated and analyzed using NITPIC 1.1.5 (*90, 91*) and ITCsy version 1a (*92*). ITC data plots were prepared with GUSSI 1.1.0 (*93*).

Bacterial killing assay

Bacterial strains were grown overnight in MLB (*V. parahaemolyticus*) or 2xYT (*E. coli*). Bacterial cultures were mixed and spotted on LB or MLB plates as previously described (*58*). CFU of the prey spotted at t = 0h were determined by plating 10-fold serial dilutions on selective medium plates. Bacterial spots were harvested from plates after 4 hours incubation and the CFU of the surviving prey cells were determined. Assays were repeated at least twice with similar results, and results of a representative experiment are shown.

Bacterial Growth Assay

Triplicates of over-night cultures of *V. parahaemolyticus* strains were normalized to OD_{600}=0.1 in MLB and grown at 30°C. Growth was measured as OD_{600nm}.

Table 3. List of bacterial strains

MX: Mexico; VT: Vietnam.

Strains	Description	Source
Vibrio parahaemolyticus		
RIMD2210633	Clinical isolate of *V. parahaemolyticus; serotype O3:K6*	Dr. Takeshi Honda
POR1	RIMD2210633 *ΔtdhA*: RIMD2210633 strain with genes encoding TDH/TRH toxins deleted	Dr. Takeshi Honda
POR1+pBAD	POR1 containing the empty pBAD	This study
POR1*ΔvtrC*+pBAD	POR1*ΔvtrC* complemented with the empty pBAD	This study
POR1*ΔvtrC*+p-*vtrC*	POR1*ΔvtrC* complemented with the pBAD vector expressing VtrC under the putative promoter of its operon	This study
POR1 *vopD2-FLAG*	C-terminal endogenously FLAG-tagged *vopD2*	This study
POR1 *vopS-FLAG*	C-terminal endogenously FLAG-tagged *vopS*	This study
POR1 p-*phoA-vtrC*	POR1 expressing the N terminal PhoA-VtrC fusion protein	This study
POR1 p-*vtrC-phoA*	POR1 expressing the C terminal VtrC-PhoA fusion protein	This study
POR1 p-*lacZ-vtrC*	POR1 expressing the N terminal LacZ-VtrC fusion protein	This study
POR1 p-*vtrC-lacZ*	POR1 expressing the C terminal VtrC-LacZ fusion protein	This study
POR1*ΔvtrA*+pBAD	POR1*ΔvtrA* complemented with the empty pBAD	This study
POR1*ΔvtrA*+p-	POR1*ΔvtrA* complemented with the pBAD	This study

FLAG-*vtrC*	vector expressing FLAG-VtrC	
POR1+p-FLAG-*vtrC*	POR1 with the pBAD vector expressing FLAG-VtrC	This study
POR1*ΔvtrC*+p-FLAG-*vtrC*	POR1*ΔvtrC* complemented with the pBAD vector expressing FLAG-VtrC	This study
POR1*ΔvtrC*+p-FLAG-*vtrC* ΔN30	POR1*ΔvtrC* complemented with the pBAD vector expressing FLAG-VtrC ΔN30	This study
POR1*ΔvtrC*+p-FLAG-*vtrC* Q42A	POR1*ΔvtrC* complemented with the pBAD vector expressing FLAG-VtrC Q42A	This study
POR1*ΔvtrC*+p-FLAG-*vtrC* H50R	POR1*ΔvtrC* complemented with the pBAD vector expressing FLAG-VtrC H50R	This study
POR1*ΔvtrC*+p-FLAG-*vtrC* Y81A	POR1*ΔvtrC* complemented with the pBAD vector expressing FLAG-VtrC Y81A	This study
A2	Non-AHPND strain isolated from VT	Dr. Lightner
A3	AHPND strain isolated from VT	Dr. Lightner
1335	AHPND strain isolated from VT	Dr. Lightner
12297B	AHPND strain isolated from VT	Dr. Lightner
D4	AHPND strain isolated from MX	Dr. Lightner
POR1*ΔvgrG1*	POR1 with *vgrG1* gene encoding the T6SS1 marker VgrG1 deleted	(*58*)
POR1*Δhcp1*	POR1 with *hcp1* gene encoding the T6SS1 marker Hcp1 deleted	(*58*)
POR1*Δhcp2*	POR1 with *hcp2* gene encoding the T6SS2 marker Hcp2 deleted	(*58*)
Escherichia coli		
Rosetta 2(DE3)	Allows for high level expression of recombinant protein in E. coli	Novagen
Rosetta 2(DE3)+ pACYCDuet-VtrC/VtrA	Rosetta 2(DE3) containing the pACYCDuet that coexpresses VtrC (aa31-161) and VtrA (aa-161-253)	This study
DH5α	Used as the prey for T6SS1 killing assay	This study

Table 4. List of primers

Name	RE site	Sequence (5'-3')	Purpose
Cloning primers			
PL553	Sal1	CACAGTCGACCACATGTAAAAC GAAATCAACGGC	Clone of (-1kb) of *vtrC* into pDM4 for the deletion of *vtrC*
PL554	BamH1	CATGGGATCCGCTCAGGTGCAA TCTTTTAATATTC	Clone of (-1kb) of *vtrC* into pDM4 for the deletion of *vtrC*
PL555	BamH1	CATGGGATCCTTATAGTTTCACT TCAGAAGCC	Clone of (+1kb) of *vtrC* into pDM4 for the deletion of *vtrC*
PL556	Sac1	CATCGAGCTCATACTCCTGGAC TTAGAGTATTC	Clone of (+1kb) of *vtrC* into pDM4 for the deletion of *vtrC*
PL606	Nco1	CACACCATGGTCTGCTATTATG TTTAAATCCACC	Clone of (-1kb) of *vtrA* into pBAD to make pBAD-(-1kb) of *vtrA-vtrC*
PL561	BamH1	CTACGGATCCTTCACATCTACCT GCTGTTGTATG	Clone of (-1kb) of *vtrA* into pBAD to make pBAD-(-1kb) of *vtrA-vtrC*
PL607	BamH1	CACAGGATCCATGAAATTGAAT ATTAAAAGATTGC	Clone of *vtrC* into pBAD to make pBAD-(-1kb) of *vtrA-vtrC*
PL587	EcoR1	CATGGAATTCTTAATATCTATGT AGAGTTATTATTTG	Clone of *vtrC* into pBAD to make pBAD-(-1kb) of *vtrA-vtrC*
PL643	Sal1	CATGGTCGACTGAAGGCAAACT CAGCATTGGTGG	Clone of (-1kb) of *vopS* stop codon-FLAG into pDM4 for the endogenous tagging of *vopS*
PL644	BamH1	CATGGGATCCTTACTTGTCATC GTCGTCCTTGTAGTCTTTGATAC CGTGAAGGCTATTTTCAG	Clone of (-1kb) of *vopS* stop codon-FLAG into pDM4 for the endogenous tagging of *vopS*
PL645	BamH1	CATGGGATCCCTTAGCGCACGA AATAACTCAGTG	Clone of (+1kb) of *vopS* stop codon- into pDM4 for the endogenous tagging of *vopS*
PL646	Sac1	CATGGAGCTCAGATTTTTCAAA TTGCCGAGCGTC	Clone of (+1kb) of *vopS* stop codon- into pDM4 for the endogenous tagging of *vopS*

PL651	Sal1	GTACGTCGACATGAGTGCTGCACGCATTGGACAAG	Clone of (-1kb) of *vopD2* stop codon-FLAG into pDM4 for the endogenous tagging of *vopD2*
PL652	Xho1	GTACCTCGAGTTACTTGTCATCGTCGTCCTTGTAGTCACGGTTAGTAACCAAACTTGCCCTTTG	Clone of (-1kb) of *vopD2* stop codon-FLAG into pDM4 for the endogenous tagging of *vopD2*
PL653	Xho1	GTACCTCGAGTTAGTTTAAAAGGGAGAGCGGGTG	Clone of (+1kb) of *vopD2* stop codon- into pDM4 for the endogenous tagging of *vopD2*
PL654	Sac1	GTACGAGCTCTAAGATAGCTAACACATCACGACTC	Clone of (+1kb) of *vopD2* stop codon- into pDM4 for the endogenous tagging of *vopD2*
PL814	Nco1	CATGCCATGGCACGGACACCAGAAATGCCTGTTC	Clone of *phoA* into pBAD to make pBAD-*phoA*-*vtrC*
PL815	BamH1	CATGGGATCCTTTCAGCCCCAGAGCGGCTTTCATG	Clone of *phoA* into pBAD to make pBAD-*phoA*-*vtrC*
PL813	BamH1	CATGGGATCCATGAAATTGAATATTAAAAGATTGC	Clone of *vtrC* into pBAD to make pBAD-*phoA*-*vtrC*
PL579	Nco1	CATGCCATGGCAATGAAATTGAATATTAAAAGATTGC	Clone of *vtrC* into pBAD to make pBAD- *vtrC*-*phoA*
PL808	BamH1	CATGGGATCCATATCTATGTAGAGTTATTATTTGTCC	Clone of *vtrC* into pBAD to make pBAD- *vtrC*-*phoA*
PL816	BamH1	CATGGGATCCCGGACACCAGAAATGCCTGTTC	Clone of *phoA* into pBAD to make pBAD- *vtrC*-*phoA*
PL817	HindIII	CATGAAGCTTTTATTTCAGCCCCAGAGCGGCTTTC	Clone of *phoA* into pBAD to make pBAD- *vtrC*-*phoA*
PL826	Nco1	CATGCCATGGCAATGACCATGATTACGGATTCACTGGCC	Clone of *lacZ* into pBAD to make pBAD-*lacZ*-*vtrC*
PL827	BamH1	CATGGGATCCTTTTTGACACCAGACCAACTGGTAATG	Clone of *lacZ* into pBAD to make pBAD-*lacZ*-*vtrC*
PL828	BamH1	GTACGGATCCATGACCATGATT	Clone of *lacZ* into pBAD to make

		ACGGATTCACTGGCC	pBAD- *vtrC-lacZ*
PL829	HindIII	CATGAAGCTTTTATTTTTGACAC CAGACCAACTGGTAATG	Clone of *lacZ* into pBAD to make pBAD- *vtrC-lacZ*
PL665	Nco1	CATGCCATGGCAGACTACAAGG ACGACGATGACAAGATGAAATT GAATATTAAAAGATTGC	Clone of pBAD-FLAG-*vtrC*
GC1	BamH1	CATGGATCCAGTTCATTTTTATG AAACCAG	Clone of *vtrC* (aa 31-161) into pACYCDuet
GC2	Not1	AGCGGCCGCCTAATATCTATGT AGAGTTATTATTTG	Clone of *vtrC* (aa 31-161) into pACYCDuet
GC3	Nde1	ATACATATGACAGCCAAAGATG ACTATC	Clone of *vtrA* (aa 161-253) into pACYCDuet
GC4	Xho1	ATGCCTCGAGTTAATATTCAAT TTCATTTAC	Clone of *vtrA* (aa 161-253) into pACYCDuet
Mutagenesis primers			
GC5	/	TGTGCTGTCAGCGGCCGCGTAT TTATAGCTGGTTTCATAA	Mutate wild type VtrC to Q42A
GC6	/	TTATGAAACCAGCTATAAATAC GCGGCCGCTGACAGCACA	Mutate wild type VtrC to Q42A
GC7	/	ACGTTAATGGCTACGTCACGCA TATATGTGCTGTCAG	Mutate wild type VtrC to H50R
GC8	/	CTGACAGCACATATATGCGTGA CGTAGCCATTAACGT	Mutate wild type VtrC to H50R
GC9	/	ATCACCATGCCCTATTACATTA GCATAGTTTTTATTCTCACTTTT TACTA	Mutate wild type VtrC to Y81A
GC10	/	TAGTAAAAAGTGAGAATAAAA ACTATGCTAATGTAATAGGGCA TGGTGAT	Mutate wild type VtrC to Y81A
RT-PCR primers			
PL615	/	AATTGTTCCAGAAAGGCTCTAT	Amplify the DNA fragment

		GTCATGCTTAATG	spanning *vtrA* and *vtrC*
PL616	/	GTTTCATAAAAATGAACTGG TTGAAAAAAATTG	Amplify the DNA fragment spanning *vtrA* and *vtrC*
qRT-PCR primers			
PL752	/	TTGGAACCCACGAACATCTC	Quantify *vtrA* mRNA
PL753	/	CAGTCACAAATTTTCCTGGCC	Quantify *vtrA* mRNA
PL754	/	ATTATCAGCTTAGGTGGGCG	Quantify *vtrB* mRNA
PL755	/	ACTTTACCCCACACTTTGTCG	Quantify *vtrB* mRNA
PL756	/	AAGCGATAACCTATGACCAGC	Quantify *fliA* mRNA
PL757	/	TCCTCTACCTGAACACTCGG	Quantify *fliA* mRNA
Sequencing primers			
PL842	/	GAGCTCGGTTCTGAGTTCGATA ACCTC	Amplify *vp1396* and *vp1397*
PL843	/	GACGATCTCTTCCCATTCATACT CGAC	Amplify *vp1396* and *vp1397*

Table 5. List of constructs

Plasmid Construct	**Description**	**Source**
pDM4-(-1kb)-(+1kb) of *vtrC*	To deleted *vtrC* in POR1	This study
pBAD-(-1kb) of *vtrA-vtrC*	To complement *vtrC* deletion by expressing VtrC under its putative endogenous promoter	This study
pDM4-(-1kb) of *vopS* stop codon-FLAG-(+1kb) of *vopS* stop codon	To endogenously tag *vopS* with C-terminal FLAG tag	This study
pDM4-(-1kb) of *vopD2* stop codon-FLAG-(+1kb) of *vopD2* stop codon	To endogenously tag *vopD2* with C-terminal FLAG tag	This study
pBAD-*phoA-vtrC*	To express PhoA-VtrC fusion protein and study cellular localization of VtrC	This study
pBAD- *vtrC-phoA*	To express VtrC-PhoA fusion protein and study cellular localization of VtrC	This study
pBAD-*lacZ-vtrC*	To express LacZ-VtrC fusion protein and study cellular localization of VtrC	This study
pBAD- *lacZ-phoA*	To express VtrC-LacZ fusion protein and study cellular localization of VtrC	This study
pBAD-FLAG-*vtrC*	To express FLAG-VtrC under arabinose inducible promoter	This study
pBAD- FLAG-VtrCΔN30	To express the mutant FLAG-VtrCΔN30 under arabinose inducible promoter	This study
pBAD-FLAG-*vtrC* Q42A	To express the mutant FLAG-VtrC Q42A under arabinose inducible promoter	This study
pBAD-FLAG-*vtrC* H50R	To express the mutant FLAG-VtrC H50R under arabinose inducible promoter	This study
pBAD-FLAG-*vtrC* Y81A	To express the mutant FLAG-VtrC Y81A under arabinose inducible promoter	This study
pACYCDuet-*vtrC* (aa 31-161)-*vtrA* (aa 161-253)	To co-express the periplasmic domains of VtrA and VtrC	This study

CHAPTER THREE

Bile Salt Receptor Complex Activates a Pathogenic Type III Secretion System

Introduction

Bile is a significant component of the human gastrointestinal tract that plays an important role in the emulsification and solubilization of lipids during food digestion (*1*). Various enteric pathogenic bacteria have been shown to sense bile for the regulation of their virulence systems during host infection, including *Salmonella typhimurium, Shigella spp*., *Vibrio parahaemolyticus and Vibrio cholera* (*3-14*). However, the mechanisms underlying bile sensing by theses bacterial pathogens remain largely unknown, which hinders our understanding of the infectious diseases they induce.

V. parahaemolyticus is the world's leading cause of seafood borne illness (*15, 16*). It causes acute gastroenteritis due to the consumption of raw or undercooked seafood (*18*). *V. parahaemolyticus* contains two different T3SSs, T3SS1 and T3SS2 (*15, 35*). T3SS2 is only discovered in clinical isolates of *V. parahaemolyticus* and has been shown to be the primary virulence system during infection (*36*). Activation of T3SS2 by bile salts is regulated by two transmembrane ToxR-like transcription factors, VtrA and VtrB (*14, 54*). Upon sensing of bile salt, VtrA induces VtrB, which then activates T3SS2 (*14,*

54). Using *V. parahaemolyticus* as the model, this study aims to elucidate how *Vibrio spp.* sense bile salts as a signal to regulate the expression of virulence genes.

Results

VtrC is conserved in various bacterial species with VtrA-like genes and encodes a predicted transmembrane protein

vtrC (*vpa1333*) is a previously uncharacterized gene in *V. parahaemolyticus* that is located directly downstream of *vtrA* **(Figure 10)**. The open reading frames for *vtrA* and *vtrC* overlap by 17 nucleotides, suggesting that they are likely in the same operon and function in the same biological process. We observed that the gene organization of *vtrA*, *vtrC* and *vtrB* in the genome is highly conserved in other *Vibrio* and related species **(Figure 10)**. Both *V. cholerae* non-O1/O139 strains and *Grimontia hollisae* possess a T3SS similar to T3SS2 in *V. parahaemolyticus* and cause gastroenteritis during human infection (*11*). RT-PCR of the region spanning *vtrA* and *vtrC* showed that these genes are indeed in the same operon and co-transcribed before and after bile salt-mediated T3SS2 induction **(Figure 11, lanes 3, 5)**. PSI-BLAST analysis with the predicted VtrC protein sequence revealed that it is conserved in various *Vibrio* and related species, as well as in *Moritella*. VtrC has a highly conserved N-terminal transmembrane signal anchor that is predicted to direct and retain proteins in the bacterial membrane **(Figure 12)**. PSI-BLAST analysis using VtrA's sequence without the cytoplasmic DNA binding domain identified homologues in the same group of bacteria where VtrC is found **(Figure 13)**. Furthermore, ten other strains of bacteria contain homologous genes of *vtrA* and *vtrC* but

not *vtrB* **(Figure 14)**. These results support the hypothesis that VtrA and VtrC evolved independently of VtrB.

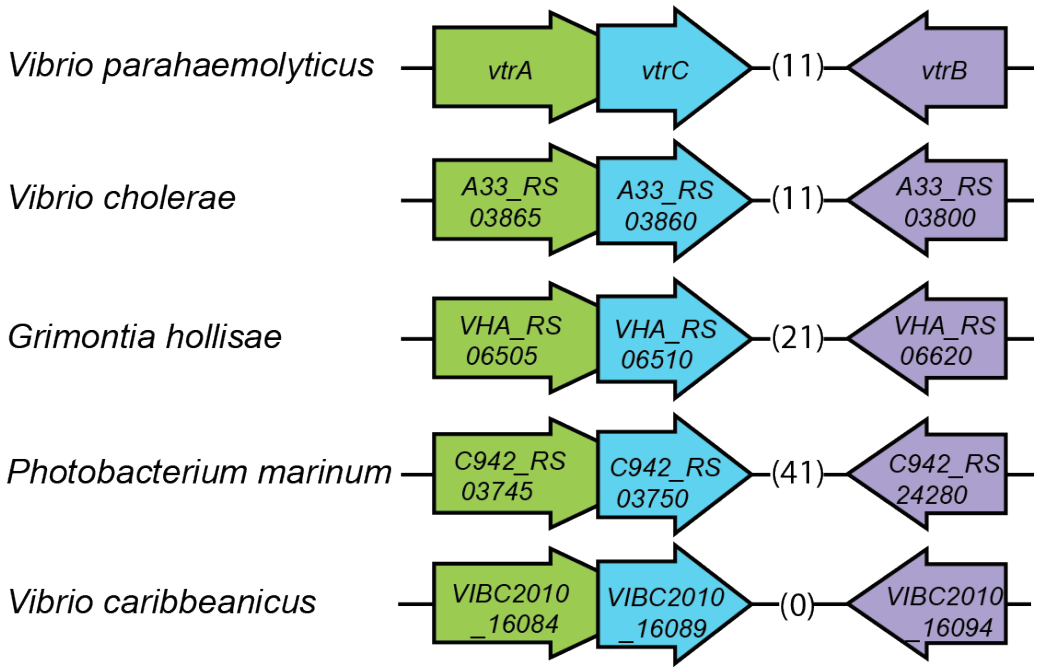


Figure 10. VtrC is conserved in various *Vibrio* and other species with VtrA-like sequences

The gene organization of *vtrA*, *vtrC* and *vtrB* is conserved in the T3SS2-like pathogenicity island of *Vibrionaceae* family species, with variable numbers of inserted genes (indicated by numbers in parentheses) between *vtrC* (or its homologous gene) and *vtrB* (or its homologous gene).

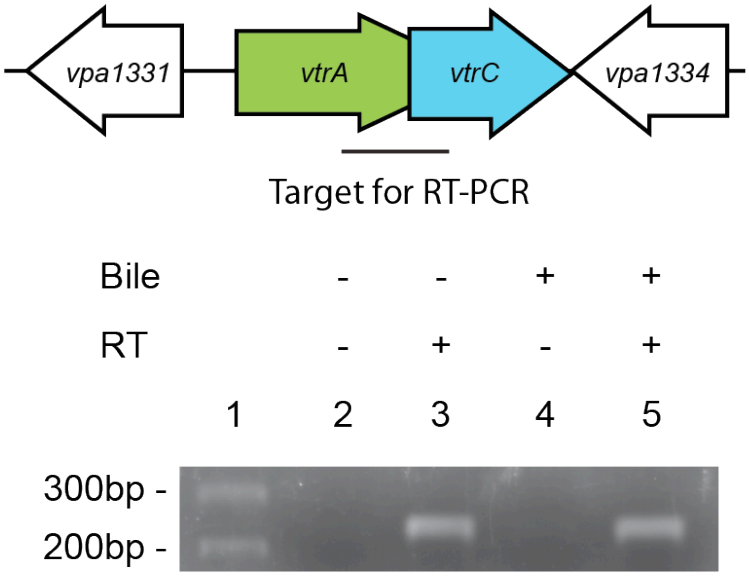


Figure 11. VtrA and VtrC are in the same operon and co-transcribed

RT-PCR to amplify the target region shown by the bar that spans *vtrA* and *vtrC*. - Bile, *Vibrio parahaemolyticus* POR1 strain growing in LB without bile salts; + Bile, POR1 growing in LB supplemented with 0.05% bile salts. - RT, without reverse transcriptase; + RT, with reverse transcriptase; lane 1: DNA marker. RT-PCR is representative of three independent experiments.

Figure 12. Multiple protein sequence alignment of VtrC and its homologues

GI number of each protein is listed before the species names. Residues are highlighted according to group-wise conservations: mainly hydrophobic (yellow) and small (gray). Signal anchor: the predicted N-terminal transmembrane domain. An empty line is inserted between the species that contain (top) or lack (bottom) *vtrB* (or *vtrB* homologous gene). * marks protein sequences that were translated from nucleotide to include the entire N-terminus.

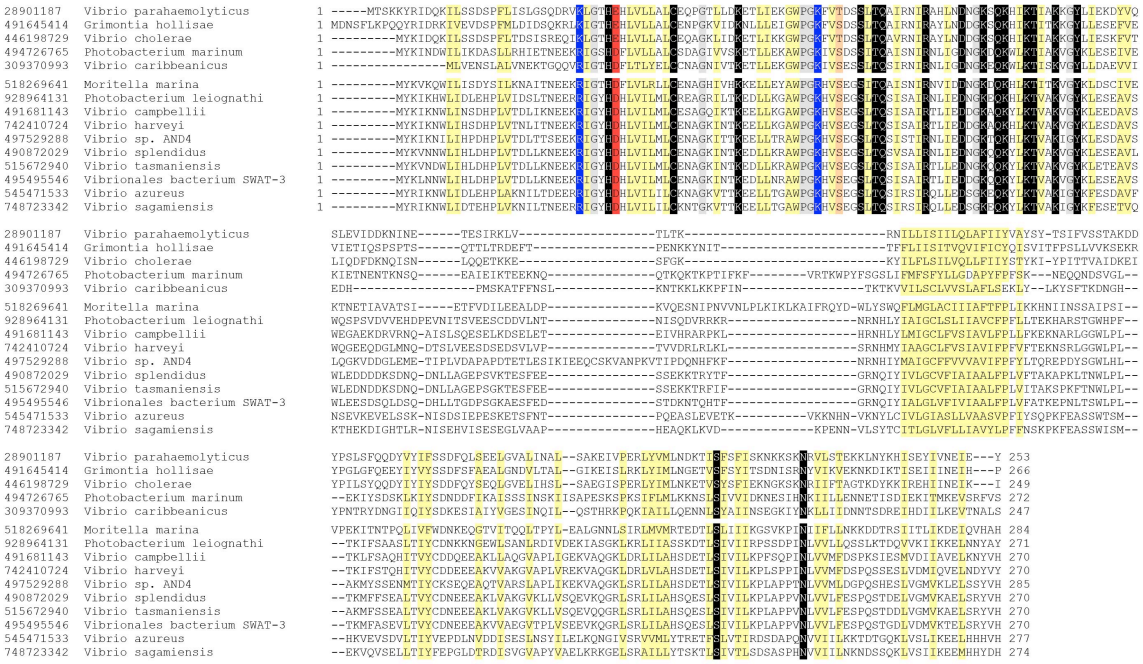


Figure 13. Multiple protein sequence alignment of VtrA and its homologues

GI number of each protein is listed before the species names. Residues are highlighted according to group-wise conservations: mainly hydrophobic (yellow), small (gray), conserved positive charge (blue), conserved negative charge (red), conserved S/T (orange), and invariant polar position (black). An empty line is inserted between the two groups of species that either contain (top) or lack (bottom) *vtrB* (or *vtrB* homologous gene).

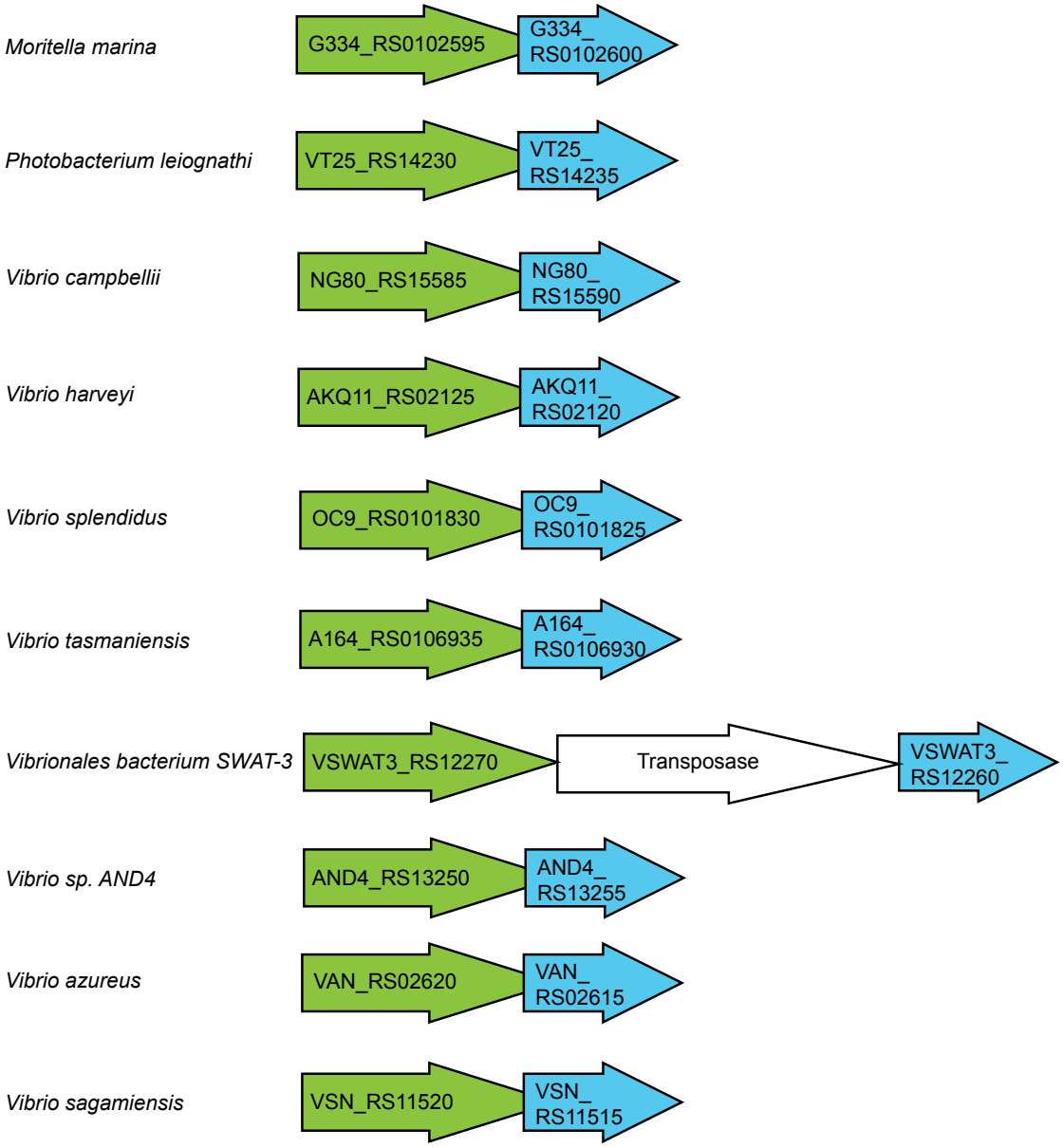


Figure 14. The gene organization of *vtrA* and *vtrC* in species that lack *vtrB*

The gene organization of *vtrA* and *vtrC* homologous genes is conserved in *Vibrionaceae* and *Moritellaceae* family species.

VtrC is essential for the activation of T3SS2 by bile salts

VtrA and VtrB are involved in the activation of the T3SS by bile salts. Therefore, we set out to determine whether VtrC also plays a role in this pathway. Because *vtrA* and *vtrC* overlap by 17 nucleotides, we generated a deletion of *vtrC* from the *V. parahaemolyticus* POR1 strain by retaining *vtrC*'s first 33 nucleotides thereby leaving the *vtrA* open reading frame intact. Deletion of *vtrC* completely abolished the activation of T3SS2 by bile salts, as shown by the loss of expression and secretion of the T3SS2 effectors VopA and VopC, and of the T3SS2 translocon VopD2 **(Figure 15A,B)**. Complementation of *vtrC* deletion by a vector expressing VtrC fully restored the activity of T3SS2 **(Figure 15A,B)**, indicating that the phenotypes observed in the *vtrC* deletion were not caused by a polar effect on neighboring genes. The effect of VtrC is specific to T3SS2 because its deletion had no impact on the expression and secretion of the T3SS1 effector VopS **(Figure 15C)**. Taken together, these results indicate that VtrC is required for T3SS2 activation by bile salts.

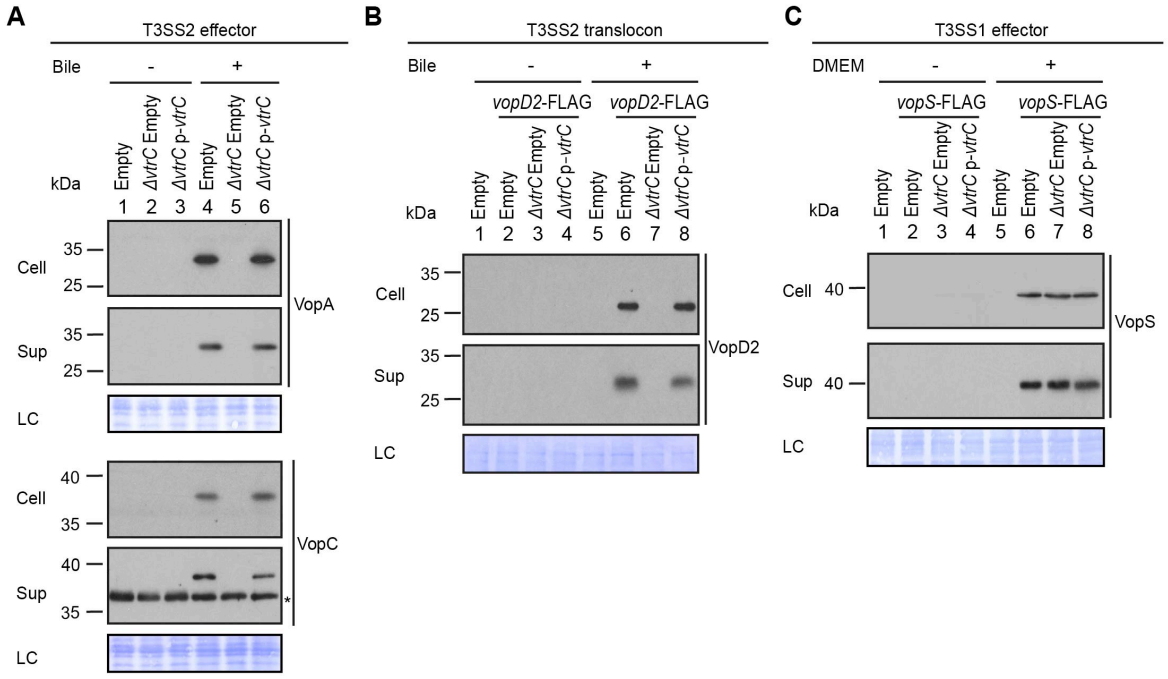


Figure 15. VtrC is essential for the activity of *V. parahaemolyticus* T3SS2 in the presence of bile salts

Expression (Cell) and secretion (Sup) of *V. parahaemolyticus* T3SS components were analyzed by Western blot. Loading control (LC) is shown for total protein lysate. **A-B.** Expression and secretion of T3SS2 effectors VopA and VopC, and translocon VopD2 by *V. parahaemolyticus* POR1 derivative strains with the empty pBAD vector (Empty), *vtrC* deletion (*ΔvtrC* Empty) or *vtrC* complementation (*ΔvtrC* p-*vtrC*) containing a pBAD vector expressing VtrC under the putative promoter of its operon (1kb upstream of *vtrA*). Protein-specific antibodies were used to detect VopA and VopC. Anti-FLAG antibody was used to detect C-terminal endogenously FLAG-tagged VopD2. -/+ Bile, *V. parahaemolyticus* grown in LB without bile salts (-) or supplemented with 0.05% bile salts (+). Non-specific band is indicated with an asterisk. **C.** Expression and secretion of

T3SS1 effector VopS by *V. parahaemolyticus* POR1 derivative strains with the empty pBAD vector (Empty), *vtrC* deletion (*ΔvtrC* Empty) or *vtrC* complementation (*ΔvtrC* p-*vtrC*). Anti-FLAG antibody was used to detect endogenously C-terminal FLAG-tagged VopS. -/+ DMEM, *V. parahaemolyticus* grown in LB (-) or DMEM (+). Data is representative of three independent experiments.

VtrC localizes to the inner membrane with the N-terminus in the cytoplasm and the C-terminus in the periplasm

Analysis of the VtrC protein sequence using membrane topology prediction programs, TMpred, TMHMM 2.0, HMMtop 2.0, and Phobius indicates VtrC is an inner membrane protein with the N-terminal 1-9 amino acids facing the cytoplasm, a single transmembrane helix (10-24 amino acids) and the C-terminal domain (25-161 amino acids) in the periplasm **(Figure 16A)**. To verify the cellular localization and orientation of VtrC, we used two reporter proteins from *Escherichia coli*, alkaline phosphatase PhoA and β-galactosidase LacZ, whose activities depend on their subcellular localization. PhoA is active only in the periplasm after disulfide bond formation and dimerization, whereas LacZ only exhibits enzymatic activity in the cytoplasm where it can fold properly (*94*). PhoA or LacZ were fused to either the N-terminus or the C-terminus of VtrC and their activities were then measured to determine the localization of VtrC (*95*). As expected, PhoA was active when fused to the C-terminus of VtrC, but not N-terminus **(Figure 16B)**. In addition, LacZ was more active when fused to the N-terminus of VtrC **(Figure 16C)**. These observations demonstrate that VtrC is localized to the inner membrane with the N-terminal 9 amino acids in the cytoplasm and the C-terminal domain in the periplasm.

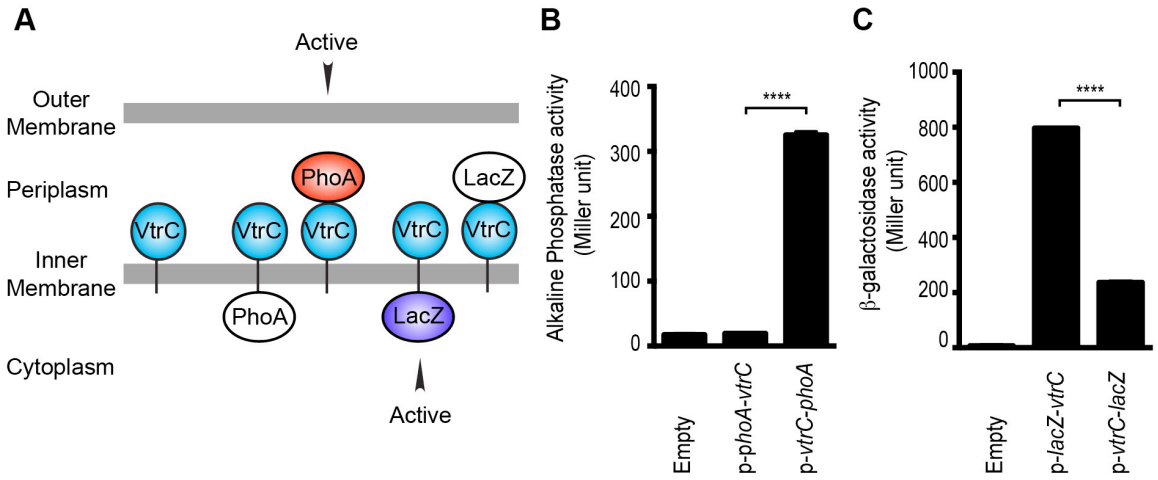


Figure 16. VtrC is an inner membrane protein with the N-terminus in the cytoplasm and the C-terminus in the periplasm

A. Predicted cellular localization and orientation of VtrC. Active form of PhoA or LacZ fused to VtrC based on prediction. **B.** Alkaline phosphatase PhoA activity of POR1 with the empty pBAD vector (Empty), expressing N terminal PhoA-VtrC (p-*phoA-vtrC*) or C terminal VtrC-PhoA (p-*vtrC-phoA*) fusion protein. **C.** β-galactosidase LacZ activity of POR1 with the empty pBAD vector (Empty), expressing N terminal LacZ-VtrC (p-*lacZ-vtrC*) or C terminal VtrC-LacZ (p-*vtrC-lacZ*) fusion protein. **** indicates $p < 0.0001$, n=3, +/- S.E.M. Data is representative of three independent experiments.

VtrC, like VtrA, functions upstream of VtrB

Previous studies have shown that VtrA directly controls the expression of VtrB and that both proteins are inner membrane proteins with their N-terminal DNA binding domain oriented in the cytoplasm and their C-terminal region in the periplasm **(Figure 17A)** (*54*). Given that VtrC is constitutively expressed and its C-terminal domain is oriented in the periplasm, we hypothesized that VtrC, like VtrA, functions upstream of VtrB. The deletion of *vtrC* did not change the mRNA level of VtrA **(Figure 17B)**. Interestingly, deletion of *vtrC* resulted in decreased protein levels of VtrA independent of bile salts **(Figure 17C, lanes 3 and 7)**, suggesting that the presence of VtrC might stabilize VtrA. Ectopic expression of VtrC from a plasmid rescued VtrA levels **(Figure 17C, lanes 4 and 8)**. As was previously observed for VtrA, the absence of VtrC prevented the induction of *vtrB* upon T3SS2 activation with bile salts **(Figure 12D)**. Thus, VtrC appears to function upstream of VtrB via a signaling cascade including VtrA. The observed degradation of VtrA in the absence of VtrC suggested that there might be a direct physical interaction between VtrA and VtrC.

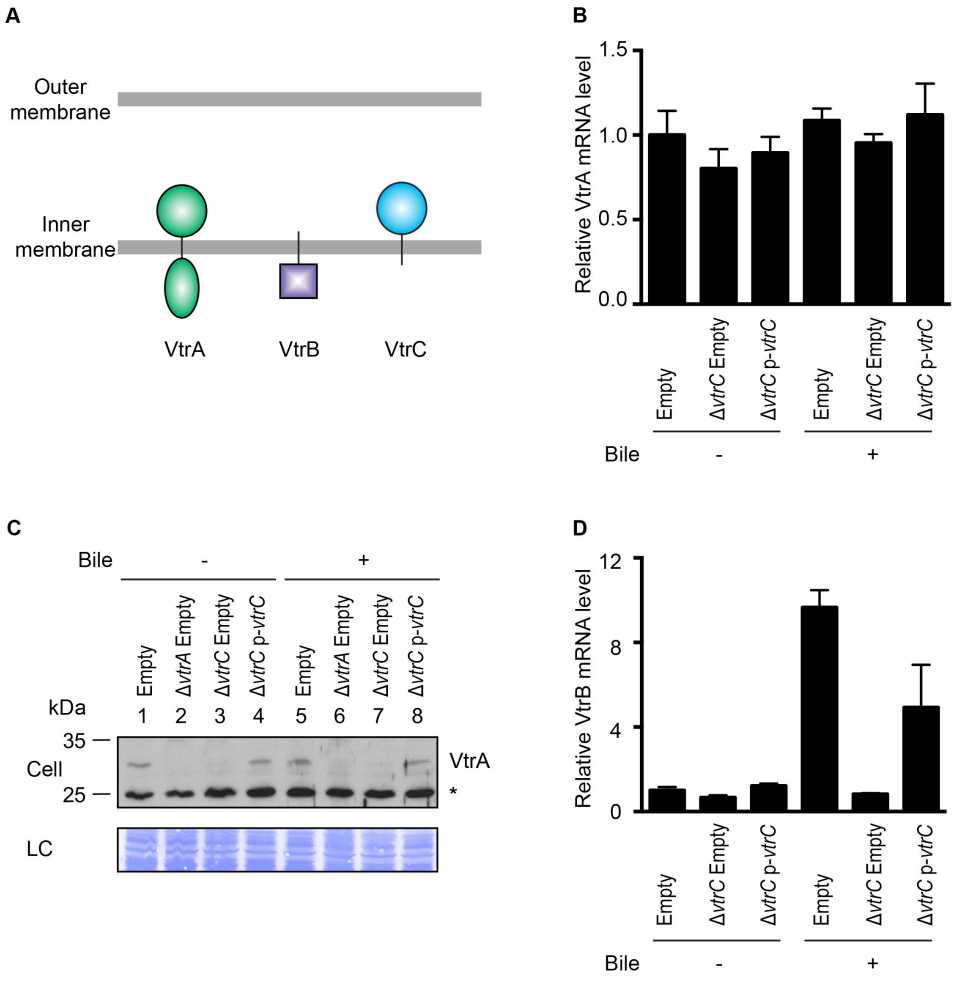


Figure 17. VtrC is necessary for maintaining the protein level of VtrA and the induction of VtrB

The effects of VtrC on VtrA and VtrB were characterized using *V. parahaemolyticus* POR1 derivative strains with the empty pBAD vector (Empty), *vtrA* deletion (*ΔvtrA* Empty), *vtrC* deletion (*ΔvtrC* Empty) or *vtrC* complementation (*ΔvtrC* p-*vtrC*) containing a pBAD vector expressing VtrC under the putative promoter of its operon (1kb upstream of *vtrA*). **A.** Cellular localization and orientation of VtrA, VtrB and VtrC. **B.** qRT-PCR

analysis of VtrA mRNA level relative to POR1 with the empty pBAD vector (Empty) grown in LB without bile salts. **C.** Western blot analysis of VtrA protein level. Protein specific antibody was used to detect VtrA. Non-specific band is indicated with an asterisk. **D.** qRT-PCR analysis of VtrB mRNA level relative to POR1 with the empty pBAD vector (Empty) grown in LB without bile salts. -/+ Bile, *V. parahaemolyticus* grown in LB without bile salts (-) or supplemented with 0.05% bile salts (+). For qRT-PCR analysis, expression of *vtrA* and *vtrB* was normalized to the expression of the control gene *fliA*. Data is representative of three independent experiments.

VtrA and VtrC form a complex

To test whether VtrA and VtrC form a complex *in vivo*, we performed co-immunoprecipitation using endogenous VtrA and ectopically expressed N-terminal FLAG-tagged VtrC. Analysis of FLAG immunoprecipitates revealed that endogenous VtrA co-precipitated with VtrC, indicating that these two proteins interact *in vivo* **(Figure 18, lanes 3 and 6)**. VtrA and VtrC interact in the presence and absence of bile salts, suggesting that these proteins interact independently of bile salts before activation of T3SS2 by bile salts, and remain in a complex after treatment with bile salts **(Figure 18)**. Notably, ectopic expression of FLAG-VtrC, but not FLAG-VtrC that is deleted for its N-terminal transmembrane domain (FLAG-VtrCΔN30), fully complemented the deletion of *vtrC* and restored the activity of T3SS2, confirming that the wild type FLAG-tagged is functional **(Figure 19)**. However, we noticed that the apparent molecular weight of full-length FLAG-VtrC (12 kDa) is smaller than the expected size of this protein (20 kDa) suggesting that VtrC may be processed in the periplasm. In any event, our results support that VtrC and VtrA interact *in vivo*.

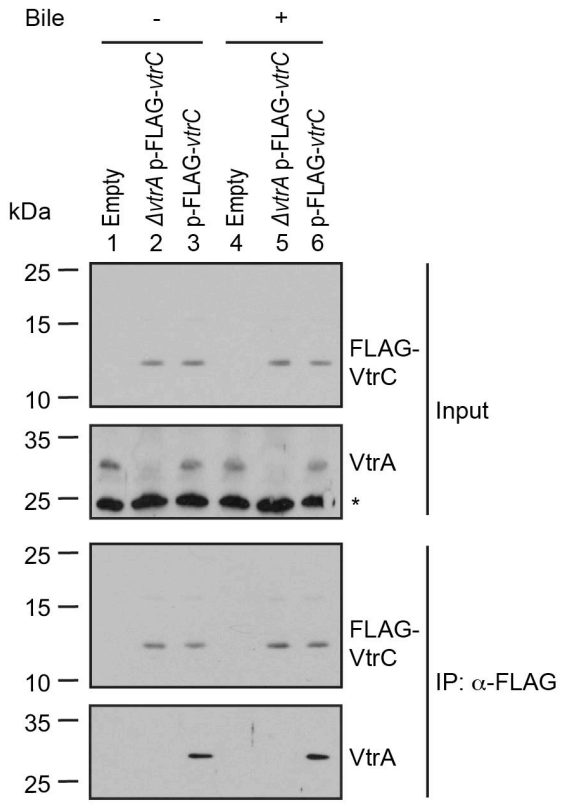


Figure 18. VtrA and VtrC form a complex *in vivo*

Co-immunoprecipitation (Co-IP) of endogenous VtrA and vector induced FLAG-VtrC. pBAD vector induced N-terminal FLAG-tagged VtrC was immunoprecipitated with anti-FLAG affinity gel from *V. parahaemolyticus* POR1 derivative strains that express only VtrA (Empty), only N-terminal FLAG-tagged VtrC (*ΔvtrA* p-FLAG-*vtrC)*, or both (p-FLAG-*vtrC*). Protein-specific antibody was used to detect VtrA. Anti-FLAG antibody was used to detect FLAG-VtrC. -/+ Bile, *V. parahaemolyticus* grown in LB without bile salts (-) or supplemented with 0.05% bile salts (+). Non-specific band is indicated with an asterisk. Data is representative of three independent experiments.

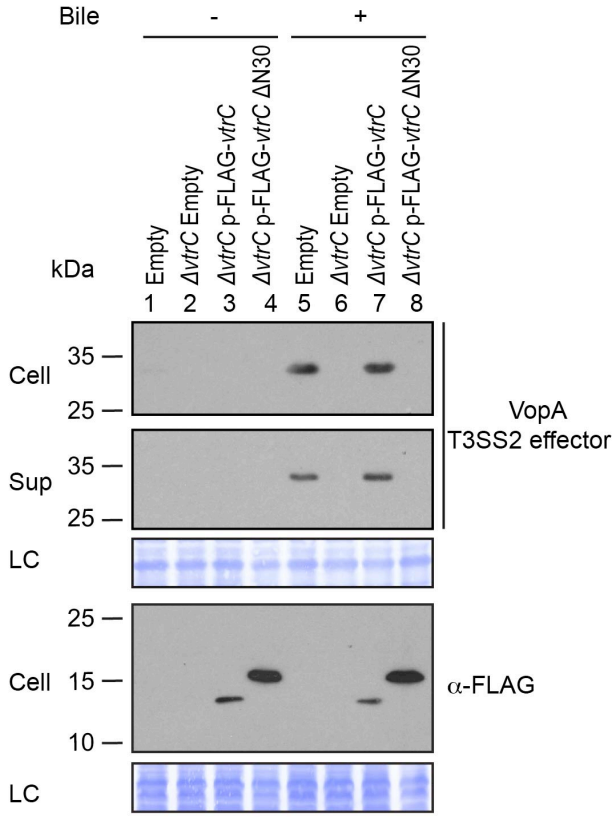


Figure 19. N-terminal FLAG-tagged VtrC is functional

Expression (Cell) and secretion (Sup) of *V. parahaemolyticus* T3SS2 effector VopA by POR1 derivative strains with the empty pBAD vector (Empty), *vtrC* deletion (*ΔvtrC* Empty), *vtrC* complementation by N-terminal FLAG-tagged wild type (*ΔvtrC* p-FLAG-*vtrC)* or mutant *vtrC* (*ΔvtrC* p-FLAG-*vtrC* ΔN30*)* that expresses pBAD vector induced protein under the arabinose inducible promoter. Protein-specific antibody was used to detect VopA. Anti-FLAG antibody was used to detect N-terminal FLAG-tagged VtrC (both wild type and ΔN30 mutant). -/+ Bile, *V. parahaemolyticus* grown in LB without bile salts (-) or supplemented with 0.05% bile salts (+). Loading control (LC) is shown for total protein lysate. Data is representative of three independent experiments.

Based on the orientation of VtrA and VtrC, we hypothesized these proteins would interact through their periplasmic domains. We found that the VtrA periplasmic domain (aa 161-253) is soluble when expressed in *E. coli*; however, attempts to express and purify the periplasmic domain of VtrC were unsuccessful. We predicted that the VtrA/VtrC interaction might stabilize the VtrC periplasmic domain. Therefore, we co-expressed the periplasmic domains of these two proteins in *E. coli*: amino acids 161-253 of VtrA and N-terminal His-tagged 31-161 of VtrC. Both proteins co-purified by Ni-NTA-affinity chromatography and remained as a stable and soluble complex during size exclusion chromatography (SEC). The VtrA/VtrC complex eluted as a 29.5 kDa species by SEC, indicating a 1:1 heterodimer of VtrA (11.0 kDa) and VtrC (16.9 kDa) periplasmic domains **(Figure 20)**. VtrA alone appeared to elute as a monomer of approximately 11 kDa **(Figure 20)**.

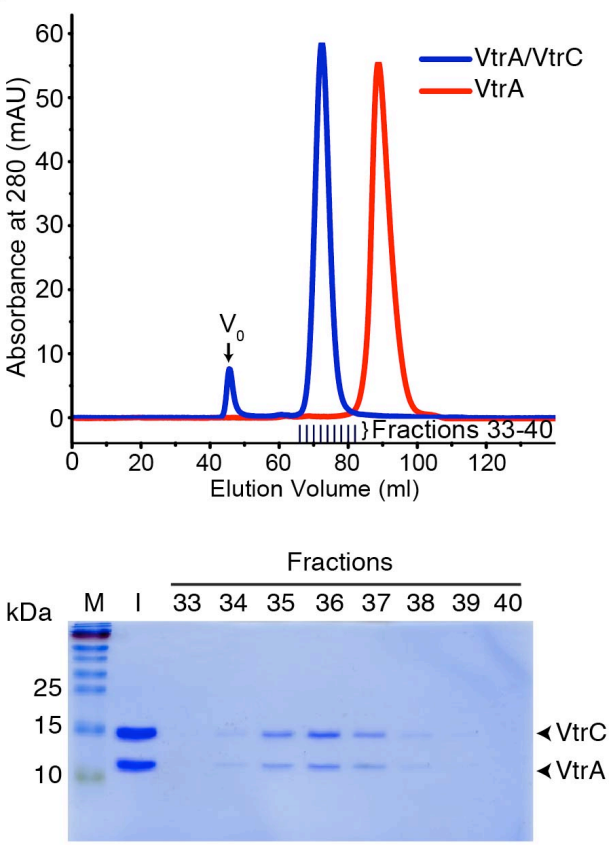


Figure 20. VtrA and VtrC periplasmic domains form a complex *in vitro*

Top, gel filtration analysis of the VtrA/VtrC periplasmic domain complex, void volume of the column is indicated as V_0. Bottom, SDS-PAGE analysis of samples from the fractions corresponding to the elution peak of the complex.

The VtrA/VtrC complex is an obligate heterodimer

To further understand the nature of the interaction between the periplasmic domains of VtrA and VtrC, we crystallized the complex and obtained its X-ray structure. The structure was solved by multiple-wavelength anomalous dispersion phasing using anomalous signals from selenomethionine and refined to a resolution of 2.70 Å using native data from an isomorphous crystal. Two of the five methionines, VtrA Se-Met 210 at the interface with VtrC and VtrC Se-Met 49, were used for refinement. Another two methionines showed no electron density and we presume they are highly disordered because they are N-terminal methionines. The final missing Se-Met is Se-Met 115 from VtrC, which is in the disordered loop that covers the beta-barrel. The asymmetric unit contains one complex with one copy each of the VtrA and VtrC periplasmic domains **(Figure 21A)**. The VtrA and VtrC subunits in the heterodimer make extensive interactions with each other **(Figure 21B)**, with an interface area of 1,149 $Å^2$ (*96*). An analysis of the macromolecular interfaces in this crystalline lattice by the web server PDBePISA (http://www.ebi.ac.uk/pdbe/pisa/pistart.html) (*96*) indicate that the VtrA/VtrC heterodimer is the only stable quaternary structure in solution.

The VtrC subunit consists of eight β-strands and a short α-helix. The eight β-strands of VtrC form a β-sheet meander that folds into itself to form a β-barrel with a small space between the first and last β-strands (Aβ and Iβ). The β-barrel is lined on the inside by hydrophobic side chains **(Figure 21C)** and is covered on one side by the short α-helix (Fα) and an adjacent disordered loop, which could not be modeled in its entirety

due to the poor electron density of this region **(Figure 21A)**. The Hβ-Iβ loop is particularly rich in aromatic side-chains, prompting us to hypothesize that this loop could be involved in binding the steroid ring structure of bile acids **(Figure 21A,B)**. Despite the lack of sequence similarity to proteins of known function, VtrC shares structural similarity with lipocalins, fatty acid binding proteins (FABPs), avidins, metalloprotease inhibitors (MPIs), and other members of the calycin superfamily of proteins **(Figure 22)** (*78*). The inside of the calycin β-barrel often acts as a binding site for fatty acids and other hydrophobic molecules, such as retinol and biotin (*78*). Comparison of the overall VtrC structure with other calycins suggests that it adopts a new fold within the superfamily. VtrC falls in between the lipocalins, which include a C-terminal helix that packs against the barrel on the corresponding VtrA interacting surface; and the FABPs, which include a β-hairpin insertion into the typical 8-stranded β-barrel in the position of the VtrC disordered loop. Interestingly, VtrA/VtrC would be the first member of this family that is an obligate dimer and not only binds a hydrophobic ligand but also transmits a signal upon binding.

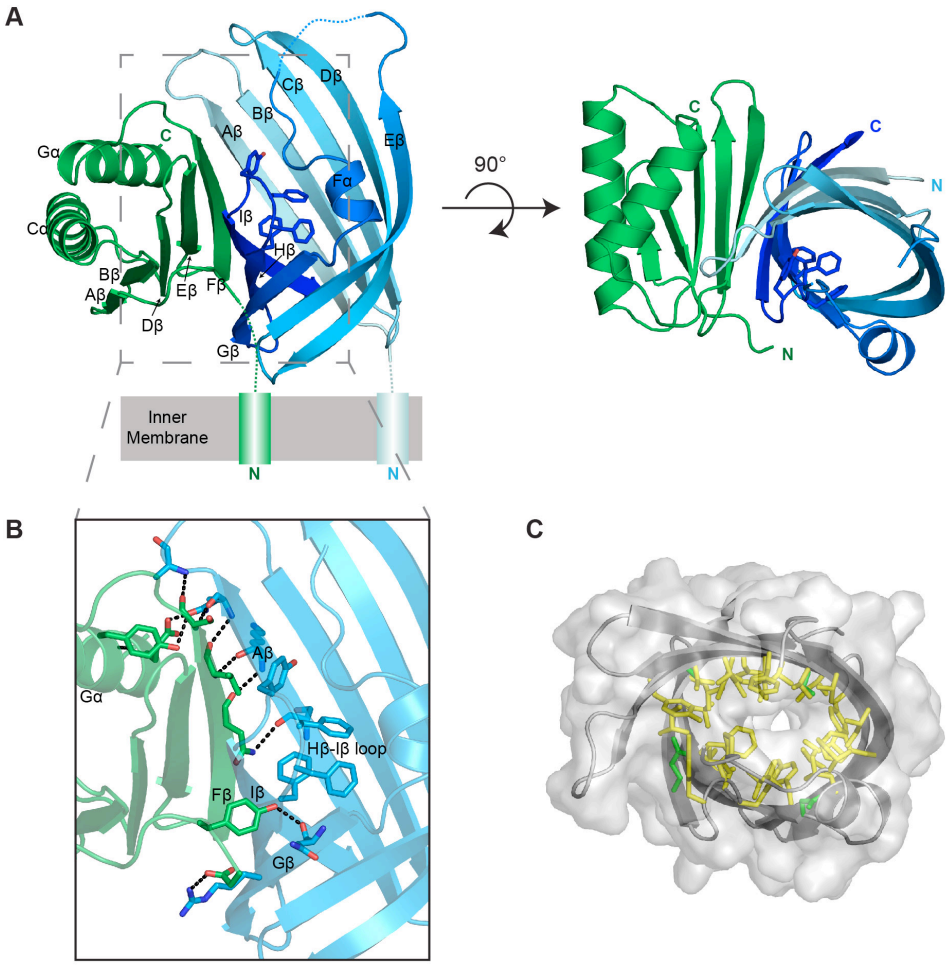


Figure 21. Structure of the VtrA/VtrC complex

A. Cartoon representation of the periplasmic domain complex formed by VtrA (green) and VtrC (blue, light to dark gradient from N-terminus to C-terminus). Side chains of Hβ-Iβ loop residues are shown as sticks. **B.** Detailed view of the VtrA/VtrC interface. Selected residues that form polar contacts (black dashed lines), as well as potential bile salt binding residues are shown as sticks. **C.** Overlay of surface and ribbon models of VtrC showing interior cavity. Side chains of residues lining the cavity are shown as sticks in yellow for hydrophobic residues (Ala, Val, Ile, Leu, Met, Phe, Tyr, Trp) and green for all other.

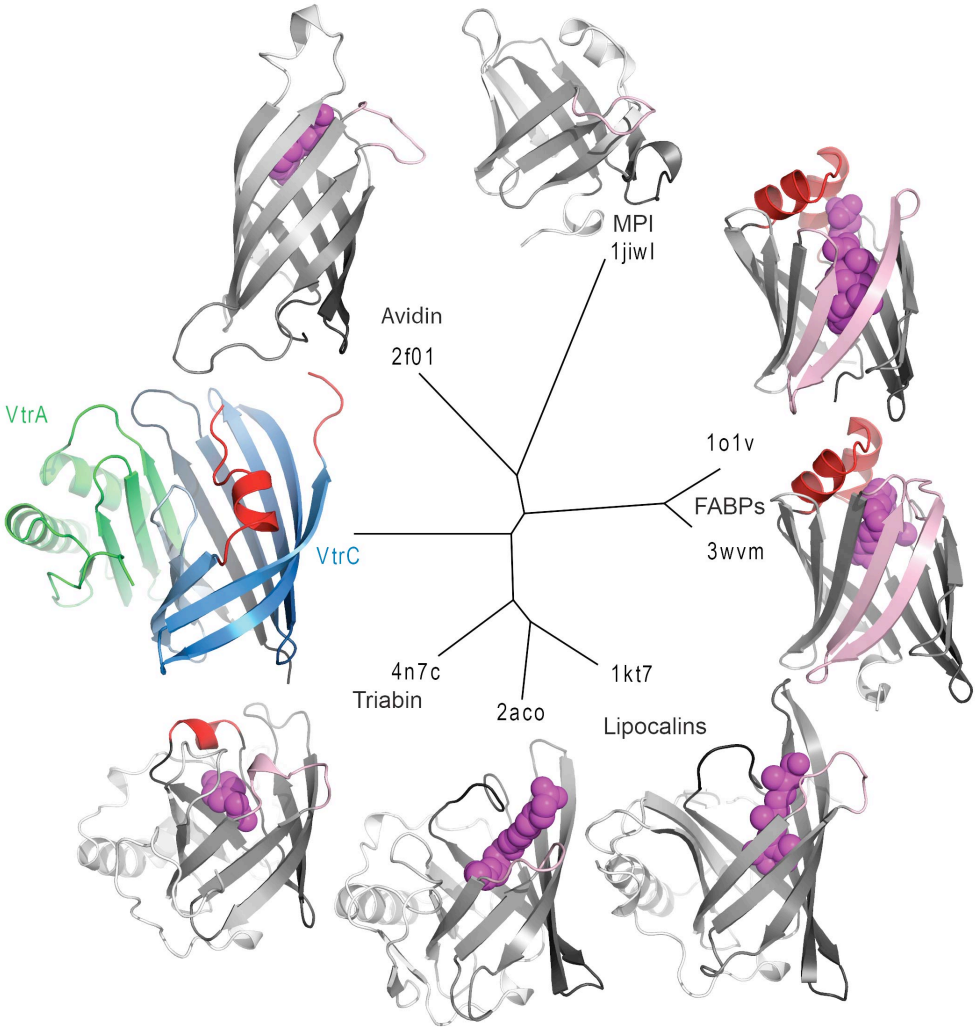


Figure 22. Structure based distance tree of VtrA/VtrC heterodimer and members of the calycin superfamily

Structures representatives of calycin superfamilies: MPI, avidin, FABPs, lipocalins, and triabin with bound ligands in the centers of the barrels (magenta spheres) were chosen. The conserved 8-stranded barrel core found in all structures is colored from dark to light grey for all representatives and dark to light blue for VtrC. The Hβ-Iβ loop containing the presumed VtrC helical lid (red) is unique to the subunit, as compared to the

corresponding loops from representative structures (light pink). This loop also corresponds to the position of the inserted FABP β-hairpin (light pink). The functionally analogous helical lid in FABPs (red) is located after the first β-strand of the barrel. The triabin structure closes the binding pocket with hydrophobic residues from both the FABP lid loop (red) and the VtrC loop (pink).

The VtrA subunit adopts an alpha/beta fold with a five-stranded β-sheet and two α-helices on one side of the sheet **(Figure 21A)**. The bulk of the contacts with VtrC is through VtrA's five-stranded β-sheet on the side opposite to the α-helices and involves both hydrophobic and polar contacts. VtrA Fβ intercalates between VtrC Aβ and Iβ, making numerous polar contacts with VtrC Aβ and closing the gap in the VtrC β-barrel **(Figure 21B)**. The incorporation of structural elements from VtrA into the β-barrel fold suggests that VtrC may not fold properly in isolation and is consistent with our inability to purify the VtrC periplasmic domain in the absence of VtrA. Interestingly, the N-terminus of the VtrA periplasmic domain, leading to the transmembrane helix, contacts VtrC Gβ and is in close proximity to the aromatic-rich Hβ-Iβ loop **(Figure 21B)**, leading us to hypothesize that these residues might be involved in signal transmission across the membrane in response to ligand binding.

The VtrA/VtrC complex binds bile salt

Based on the structural similarities between the monomeric calycins and the VtrA/VtrC heterodimer, we predicted that VtrA/VtrC complex could bind bile salts.

Previously, Gotoh et al. (*14*) established that bile salts are the component of bile that activates the T3SS2. We have recapitulated this data and confirmed which purified bile salts could activate T3SS2. We validated that taurodeoxycholate (TDC) and glycodeoxycholate (GDC), but not chenodeoxycholate (CDC) or cholate (CA), could activate the *V. parahaemolyticus* T3SS2 **(Figure 23)**.

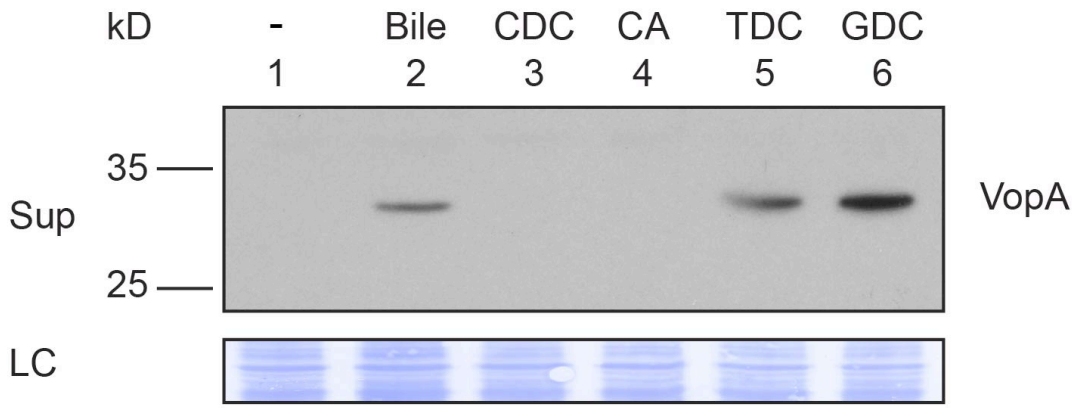


Figure 23. Activation of T3SS2 by individual bile acid

Secretion (Sup) of *V. parahaemolyticus* T3SS2 effector VopA by POR1. Protein-specific antibody was used to detect VopA. *V. parahaemolyticus* grown in LB without bile salts (-), supplemented with 0.05% crude bile (Bile) or 0.5 mM individual bile acids. CDC: chenodeoxycholate, CA: cholate, TDC: taurodeoxycholate, GDC: glycodeoxycholate. Loading control (LC) is shown for total protein lysate. Data is representative of three independent experiments.

Based on this information we moved forward with biophysical experiments to test whether the VtrA/VtrC heterodimer is the bile salt receptor using a relevant bile salt, TDC. We first tested whether the VtrA/VtrC periplasmic domain heterodimer could bind bile salts using isothermal titration calorimetry (ITC). Negative power deflections were observed throughout the titration of the bile salt TDC into the VtrA/VtrC solution. These results indicated that TDC binds to the VtrA/VtrC heterodimer in an exothermic manner, with a dissociation constant (K_D) of 315.4 nM **(Figure 24)**. The stoichiometry of TDC binding to VtrA/VtrC is approximately 1:1 (n = 0.94). Taken together our results suggest that VtrA and VtrC form a functional complex that can bind bile salts.

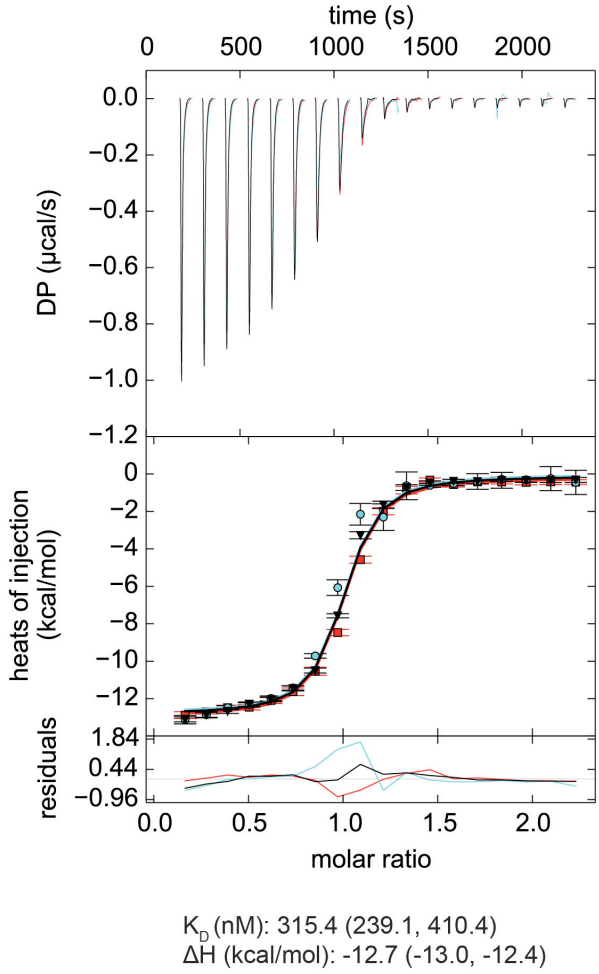


Figure 24. VtrA/VtrC periplasmic domain complex binds the bile salt taurodeoxycholate (TDC)

ITC-derived binding curves of the VtrA/VtrC complex with TDC. Thermodynamic parameters were determined by global fitting of triplicate isotherms (presented in black, red, and cyan). The dissociation constant (K_D) and enthalpy (ΔH) values are reported followed by the 1σ error intervals in parenthesis. Data is representative of two independent experiments.

The VtrA/VtrC complex binds bile salts in a primarily hydrophobic cleft in the VtrC β-barrel

We next purified the VtrA/VtrC heterodimer in the presence of the bile salt TDC and crystallized it. The crystal structure was solved via molecular replacement by using our original VtrA/VtrC heterodimer as a search model, and refined to a resolution of 2.10 Å. The model reveals three VtrA/VtrC heterodimers in the asymmetric unit, each with clear electron density for one TDC molecule bound inside the β-barrel **(Figure 25A, Figure 26A)**. A fourth TDC molecule mediated lattice contacts between VtrA chain A and its symmetry mate **(Figure 26B)**. An analysis of the macromolecular interfaces found in this crystalline lattice by the web server PDBePISA (*96*) indicated that the heterodimer is the only stable quaternary structure in solution. Although the three heterodimers in the asymmetric unit are highly similar, average root mean squared deviations (RMSD) of 0.5 Å for VtrA and 0.6-1.5 Å for VtrC indicates that there are differences in the VtrC monomer **(Figure 27A,E)**. RSMD values of 0.4-0.5 Å for VtrA and 1.2-1.9 Å were obtained by superimposing each of the TDC-bound heterodimers to the apo structure **(Figure 27B-E)**, indicating that the overall fold is maintained but there are considerable changes in the VtrC monomer. The largest differences between the VtrA/VtrC-TDC and the apo heterodimers are in the VtrC disordered loop and Fα that follows it. In the VtrA/VtrC-TDC heterodimer, the disordered loop has been displaced by the TDC molecule and is now observable in two out of the three heterodimers in the crystallographic asymmetric unit. Residues Thr122 and Ser123, which were in the disordered loop in the apo structure, now form part of Fα in two of the TDC-bound

heterodimers. Remarkably, one of these residues (Ser123) now forms a hydrogen bond with the 12α-hydroxyl group of TDC **(Figure 25B)**. The main chain amide between F150 and Tyr151 (Hβ-Iβ loop) coordinates the 3α-hydroxyl group **(Figure 25B)**, which is present in all bile salts, suggesting that this interaction is important for specificity towards bile salts. Several VtrC residues in the binding pocket maintain hydrophobic interactions with TDC, including the side chains of Phe150 with steroid ring A and Tyr81 with the hydrophobic part of the bile acid **(Figure 25B)**.

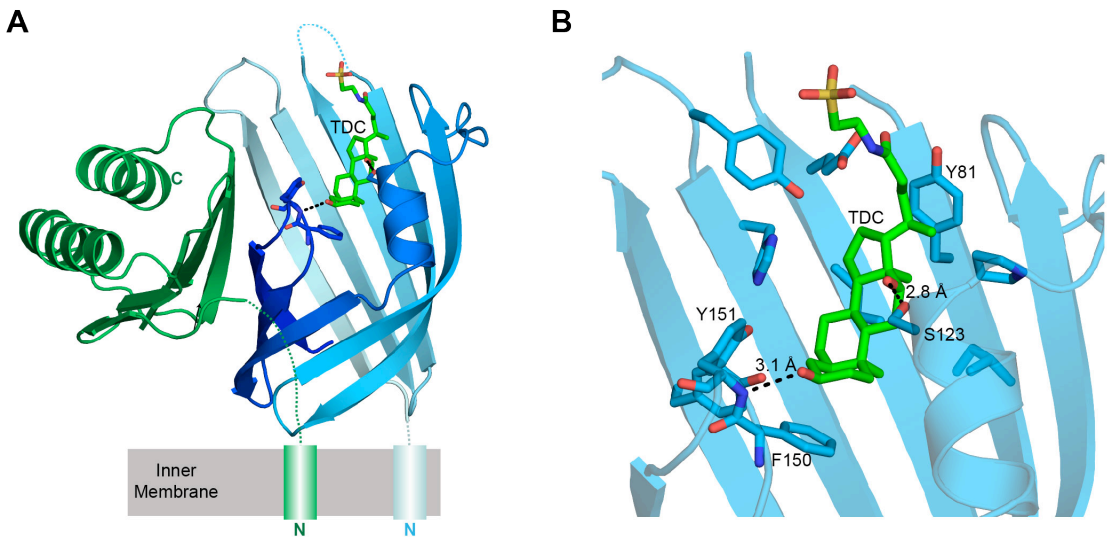


Figure 25. Structure of VtrA/VtrC periplasmic domain complex with the bile salt taurodeoxycholate (TDC)

A. Structure of the VtrA/VtrC periplasmic domain complex binding TDC (green). VtrA and VtrC follow same coloring scheme as in Figure 21. **B.** Detailed view of the TDC binding site. Hydrogen bonds between protein and TDC are represented as dashed lines.

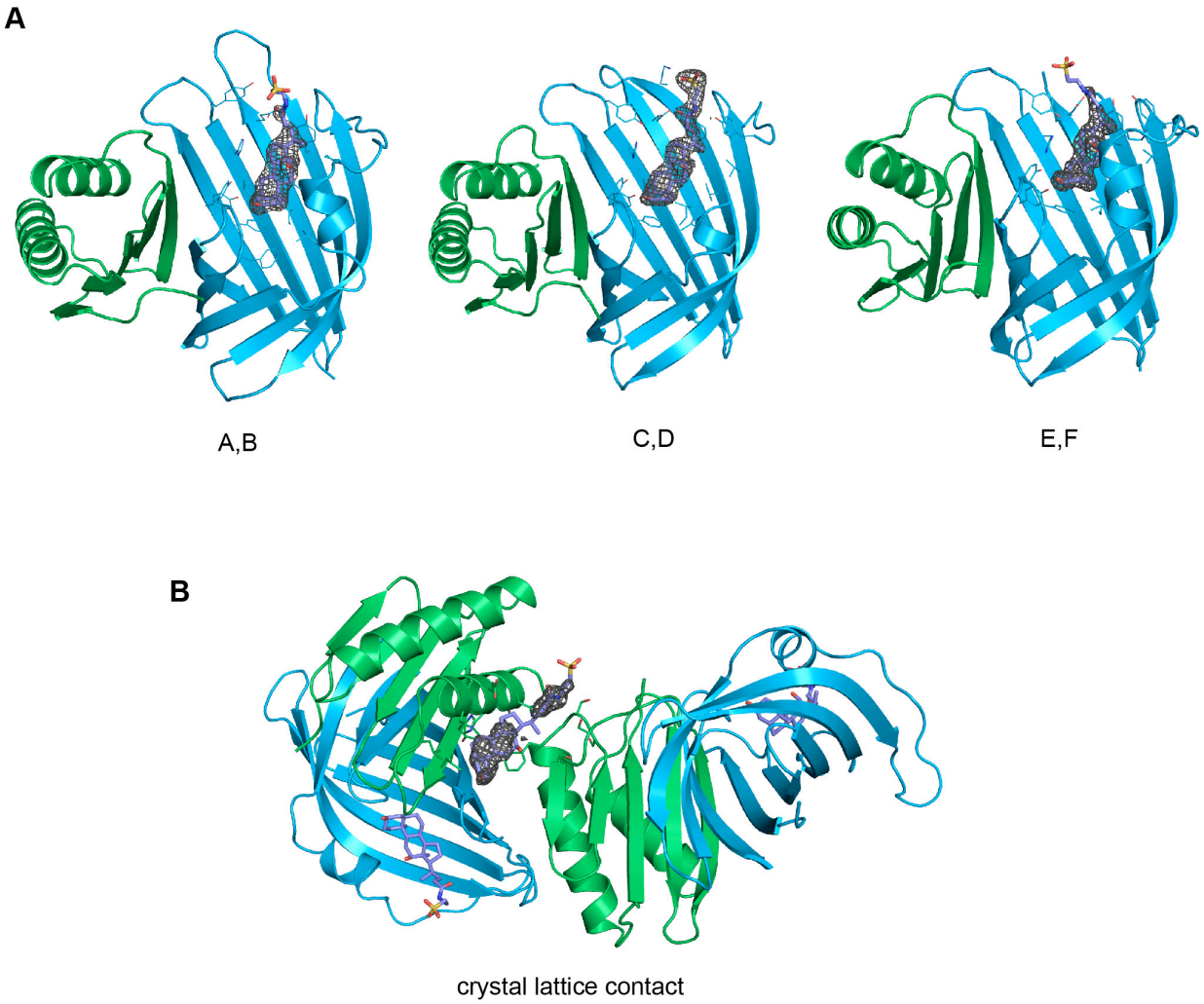
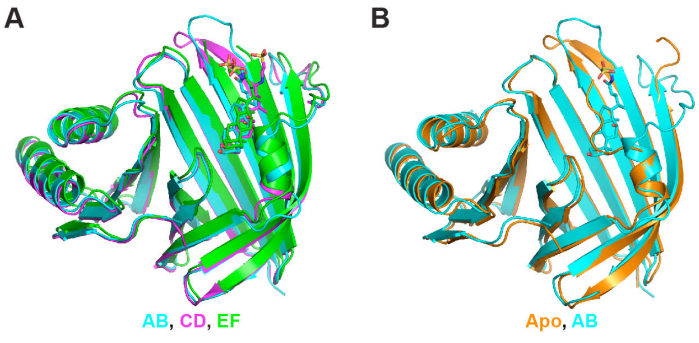


Figure 26. Electron density around taurodeoxycholate (TDC) molecules

Kicked F_o-F_c omit maps of the three TDC molecules bound to the A/B, C/D and E/F heterodimers **(A)** and the TDC mediating lattice contacts **(B)** in the TDC-containing crystal. The maps are shown as grey mesh (contoured at 3 σ in **A** and 1.5 σ in **B**) and carved around the ligand at a 1.6 Å radius.



E

Alignment statistics for VtrA/VtrC heterodimers

Heterodimers	RMSD (Å) for VtrA (# C-α atoms)	RMSD (Å) for VtrC (# C-α atoms)
A:B/C:D	0.5 (90)	1.5 (125)
A:B/E:F	0.5 (88)	0.7 (127)
C:D/E:F	0.5 (88)	0.6 (127)
Apo/A:B	0.4 (89)	1.9 (132)
Apo/C:D	0.5 (90)	1.2 (125)
Apo/E:F	0.5 (88)	1.4 (127)

Figure 27. Comparison between apo and taurodeoxycholate (TDC) bound **structure heterodimers**

A. Superposition of all the heterodimers in the asymmetric unit of the TDC-bound crystal. **B-D.** Superposition of AB **(B)**, CD **(C)** and EF **(D)** heterodimers from the TDC-bound complex crystal structure and apo heterodimer. **E.** Root-mean-square deviations (RMSD) for each superposition with the number of aligned α-carbon atoms in parethesis. Superpositions were made via the DaliLite server (http://ekhidna.biocenter.helsinki.fi/dali_lite /start) (*79*).

Mutations of the VtrA/VtrC complex with the bile salt taurodeoxycholate

Based on our co-crystal structure of VtrA/VtrC with TDC, the role of several residues that contact the bile salt were investigated for their ability to affect the binding of bile salts and activation of the T3SS2 **(Figure 28)**. The mutation of His50, located at the top of the hydrophobic cleft, to an arginine residue results in significant steric hindrance by insertion of a large side chain into the bile salt binding pocket **(Figure 28A)**. When VtrC-H50R was expressed in the *vtrC* deletion strain and bile salts were added, the T3SS2 could not be induced as indicated by the lack of expression and secretion of the T3SS2 effector VopA **(Figure 28B, lanes 5 and 11)**. We also changed Tyr81 to an alanine residue, which we predicted would disrupt bile salt binding by eliminating the interaction between the Tyr81 aromatic ring and the bile salt side chain **(Figure 28A)**. When VtrC-Y81A was expressed in the *vtrC* deletion strain and bile salts were added, the T3SS2 was not induced as indicated by the lack of expression and secretion of VopA **(Figure 28B, lanes 6 and 12)**. As a control, we chose to mutate a surface exposed residue, Gln42, to an alanine **(Figure 28A)**. As expected, this mutant rescued the *vtrC* deletion strain and upon addition of bile salts the T3SS2 was activated as indicated by the expression and secretion of VopA **(Figure 28B, lanes 4 and 10)**. To assess whether these mutations had any effect on the VtrA/VtrC complex formation, we co-immunoprecitated the complex with the N-terminal FLAG-tagged VtrC. As seen in **Figure 28C**, the wild type and mutant VtrC proteins formed a stable complex with endogenous VtrA, as indicated by their co-immunopreciptation.

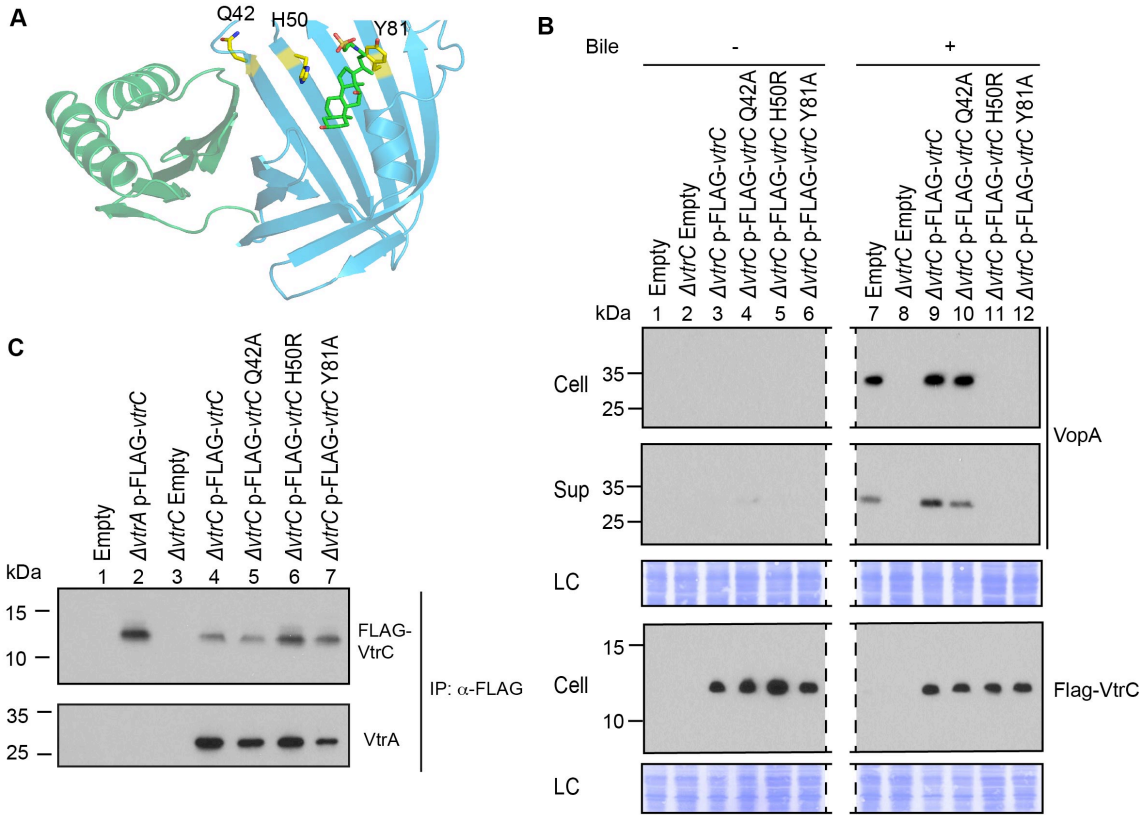


Figure 28. Mutations in the hydrophobic chamber of VtrA/VtrC heterodimer disrupt signaling mediated by bile salts

A. Structure of VtrA/VtrC heterodimer. Q42, H50 and Y81 of VtrC are highlighted as yellow. **B.** Restoration of T3SS2 activation in POR1*ΔvtrC* by VtrC mutant Q42A but not H50R or Y81A. Expression (Cell) and secretion (Sup) of *V. parahaemolyticus* T3SS2 effector VopA by POR1 derivative strains with the empty pBAD vector (Empty), *vtrC* deletion (*ΔvtrC* Empty), *vtrC* complementation by N-terminal FLAG-tagged wild type (*ΔvtrC* p-FLAG-*vtrC)* or mutant (Q42A, H50R or Y81A) VtrC that expresses pBAD vector induced protein under the arabinose inducible promoter. Protein-specific antibody

was used to detect VopA. Anti-FLAG antibody was used to detect N-terminal FLAG-tagged VtrC. -/+ Bile, *V. parahaemolyticus* grown in LB without bile salts (-) or supplemented with 0.05% bile salts (+). Loading control (LC) is shown for total protein lysate. Dashed line crops out one lane skipped in the blot. **C.** All three VtrC mutants form a stable complex with VtrA.

Data collection				
Crystal	SeMet peak[a]	SeMet inflection point[a]	Native	Native + bile salt
Space group	F432	F432	F432	$P2_12_12_1$
Cell constants (Å)	a = 211.01	a = 211.46	a = 211.39	a = 55.39, b = 71.28, c = 203.73
Wavelength (Å)	0.97927	0.97943	0.97935	0.97926
Resolution range (Å)	35.67 – 2.60 (2.64 – 2.60)	35.74 – 2.65 (2.70 – 2.65)	40.68 – 2.65 (2.70 – 2.65)	35.64 – 2.10 (2.14 – 2.10)
Unique reflections	12,839 (620)	12,267 (593)	12,322 (594)	45,762 (1,868)
Multiplicity	22.6 (20.2)	22.7 (21.8)	36.1 (36.5)	3.9 (2.8)
Data completeness (%)	99.9 (100.0)	99.9 (100.0)	99.9 (100.0)	95.6 (79.4)
R_{merge} (%)[b]	7.2 (369.2)	7.7 (292.5)	7.6 (172.4)	7.9 (36.2)
R_{pim} (%)[c]	1.5 (83.7)	1.7 (63.5)	1.3 (28.8)	4.3 (24.9)
I/σ(I)	49.9 (0.8)	46.8 (1.1)	63.6 (4.2)	15.9 (2.5)
Wilson *B*-value (Å^2)	35.8	35.4	37.5	24.2
Phase determination				
Anomalous scatterers	selenium, 2 out of 5 possible sites			
Figure of merit (121.8 – 2.60 Å)	0.25			
Refinement statistics				
Resolution range (Å)			40.68 – 2.70 (2.79 – 2.70)	35.64 – 2.10 (2.16 – 2.10)
No. of reflections R_{work}/R_{free}			11,593/1,470 (899/129)	41,304/2,000 (1,807/91)
Data completeness (%)			99.9 (100.0)	86.3 (57.0)
Atoms (non-H protein/ions/ligands/ solvent)			1,797/5/NA/NA	5,411/NA/375/204

R_{work} (%)		26.0 (34.0)	20.1 (22.3)
R_{free} (%)		29.8 (39.2)	23.3 (28.3)
R.m.s.d. bond length (Å)		0.002	0.002
R.m.s.d. bond angle (°)		0.42	0.51
Mean B-value (Å^2) (protein chain ID) (ligands/ions/solvent)		VtrA(A): 77.2 VtrC(B): 62.2. ions: 70.2	VtrA(A): 25.7 VtrC(B): 29.8 VtrA(C): 34.9 VtrC(D): 37.8 VtrA(E): 39.1 VtrC(F): 36.2 ligands: 43.3 ions: 86.2 solvent: 37.0
Ramachandran plot (%) (favored/additional/disallowed)[d]		94.3/5.2/0.5	96.5/3.2/0.3
Maximum likelihood coordinate error		0.36	0.23
Missing residues, protein (chain ID)		VtrA(A): 161 – 163. VtrC(B): -13 – 0, 113 – 117.	VtrA(A): 161 – 164. VtrC(B): -13 – 0. VtrA(C): 161 – 163. VtrC(D): -13 – 0, 119 – 124. VtrA(E): 161 – 165. VtrC(F): -13 – 0.

Table 6. Data collection and refinement statistics, VtrA/VtrC complex

Data for the outermost shell are given in parentheses.

[a]Bijvoet-pairs were kept separate for data processing.

[b] $R_{merge} = 100\ \Sigma_h\Sigma_i|I_{h,i} - \langle I_h\rangle|/\Sigma_h\Sigma_i\ \langle I_{h,i}\rangle$, where the outer sum (h) is over the unique reflections and the inner sum (i) is over the set of independent observations of each unique reflection.

[c] $R_{pim} = 100\ \Sigma_h\Sigma_i\ [1/(n_h - 1)]^{1/2}|I_{h,i} - \langle I_h\rangle|/\Sigma_h\Sigma_i\ \langle I_{h,i}\rangle$, where n_h is the number of observations of reflections **h**.

[d] As defined by the validation suite MolProbity (Chen, V.B., Arendall, W.B.A., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., Richardson, D.C. (2010) *MolProbity*: all-atom structure validation for macromolecular crystallography. *Acta Cryst*. D66, 12-21.).

Discussion

Bile sensing plays a significant role in human infection caused by various enteric bacterial pathogens (*1*). Deciphering how bacteria use bile as a signal to regulate the expression of virulence genes is essential to our understanding of disease mechanisms and also can offer insights for designing novel therapeutic strategies. Bile salt-induced T3SSs are the primary virulence factor in many *Vibrio* species (*15*). In this study, we uncovered a previously unidentified component of a widespread, conserved signaling platform that is essential for bile salt sensing and activation of virulence factors in *Vibrios*. Our biochemical data and the structural analysis demonstrated that the periplasmic domains of the inner membrane proteins VtrA and VtrC form a stable complex that functions as a receptor for bile salts.

The crystal structure of the VtrA/VtrC complex reveals that their periplasmic domains form a obligate heterodimer where VtrC folds into an open eight-stranded β-barrel and VtrA adopts a mixed α/β structure whereby one of its a β-strands contributes to the formation of the β-barrel. The VtrA/VtrC heterodimer fold is similar to members of the calycin superfamily, which contain eight to ten-stranded β-barrels that often bind hydrophobic molecules (*78*). Calycins have low sequence similarity, but relatively high structural conservation. They can be subdivided into several families based on structural details characteristic of each group **(Figure 22)** (*78*). Calycin families are markedly different in their distribution among biological kingdoms, cellular localization, and function, e.g.: (1) FABPs are cytosolic proteins involved in lipid homeostasis in animals,

(2) MPIs are bacterial proteins that are secreted into the periplasm to inhibit metalloproteases, (3) lipocalins are found in both eukaryotes and prokaryotes, and they are mostly extracellular or attached to the outer membrane and bind a variety of ligands (*78*). In terms of structure, VtrC does not seem to cluster with any particular family and it forms its own clade **(Figure 22)**. Although calycins have been found to form complexes with other proteins, our inspection of the available structures of complexes did not reveal any arrangement that resembled that of the VtrA/VtrC obligate heterodimer. This complex is unique for this family because it is formed from two proteins and it is a complex that transmits a signal upon binding its ligand.

Interestingly, members of the FABP family have been found to bind bile salts. This is exemplified by human and porcine ileal lipid binding proteins (ILBPs), and their NMR structures in complex with taurocholate (*97*) and glycocholate (*98*), respectively. Analogous to VtrC's disordered loop and short α-helix, FABPs have a helix-turn-helix motif that forms a lid over one side of the barrel and is proposed to act as a portal for ligand access (*99*). In the case of FABPs, it has been suggested that ligand binding induces subtle conformational changes that promote protein-membrane or protein-protein interactions (*99*). We observe a similar scenario for the VtrA/VtrC obligate heterodimer bound to bile salts where, upon binding bile salt, it appears an unstructured loop is translocated from the inside of the hydrophobic chamber to the outside of the chamber, where it becomes structured. Future studies will focus on how a conformational change

during bile salt binding may induce a switch in the VtrA/VtrC interaction that will result in activation of the VtrA cytoplasmic transcription activator domain **(Figure 28)**.

The closest example with regard to complex formation and function is the GrlR/GrlA complex, a regulator of the T3SS-encoding locus of enterocyte effacement (LEE) in enterohaemorrhagic and enteropathogenic *E. coli* strains (EHEC and EPEC, respectively) (*100*). GrlR is a cytosolic lipocalin that represses the LEE repressor GrlA by binding to its helix-turn-helix (HTH) domain. Similar to *vtrA* and *vtrC*, *grlR* and *grlA* are transcribed from a bicistronic operon, although the order of the lipocalin and transcriptional regulator genes is reversed in the latter pair. Analytical ultracentrifugation and the structure of GrlR in complex with the GrlA HTH domain revealed a heterotrimeric complex where a GrlR dimer binds one GrlA HTH monomer (PDB ID: 4KT5) (*100*). It has been shown that the inner cavity of GrlR binds lipids (*101*). However, the role of lipid binding in this system is unknown. Although GrlR and VtrC have analogous roles regulating T3SS through their interaction with a transcription factor, they do so from different cellular compartments at different stoichiometries, and their binding partners (GrlA HTH domain and VtrA periplasmic domain, respectively) have different folds.

Several signaling features of VtrA/VtrC appear to parallel two other pairs of regulatory proteins found in *V. cholerae*, ToxR/ToxS and TcpP/TcpH, which control the expression of numerous virulence factors including CT and TCP. First, the genes

encoding the proteins in each pair are arranged in bicistronic operons (*102, 103*). Second, VtrA/ToxR/TcpP and VtrC/ToxS/TcpH have the same cellular localization and membrane topology (*102-104*). Third, VtrC, ToxS and TcpH affect the activity of VtrA, ToxR and TcpP, respectively. On this last point, we showed that VtrA and VtrC form a complex *in vivo* both before and after T3SS2 activation and *in vitro* through their periplasmic domains.

Our data supports a model in which VtrA and VtrC form a complex both in the absence and presence of bile salts. Upon binding bile salts to the hydrophobic chamber in the VtrA/VtrC complex, the cytoplasmic DNA binding domain of VtrA is activated and in turn induces VtrB to activate the T3SS2 **(Figure 29)**. The specific changes caused by binding bile salts and their effects on the VtrA DNA-binding domain are yet to be determined. Although VtrA contains no cysteine residues in its periplasmic domain, it is still possible that bile salts increase VtrA activity by inducing its dimerization or oligomerization, similar to what has been proposed for the ToxR/ToxS and TcpP/TcpH systems (*105, 106*).

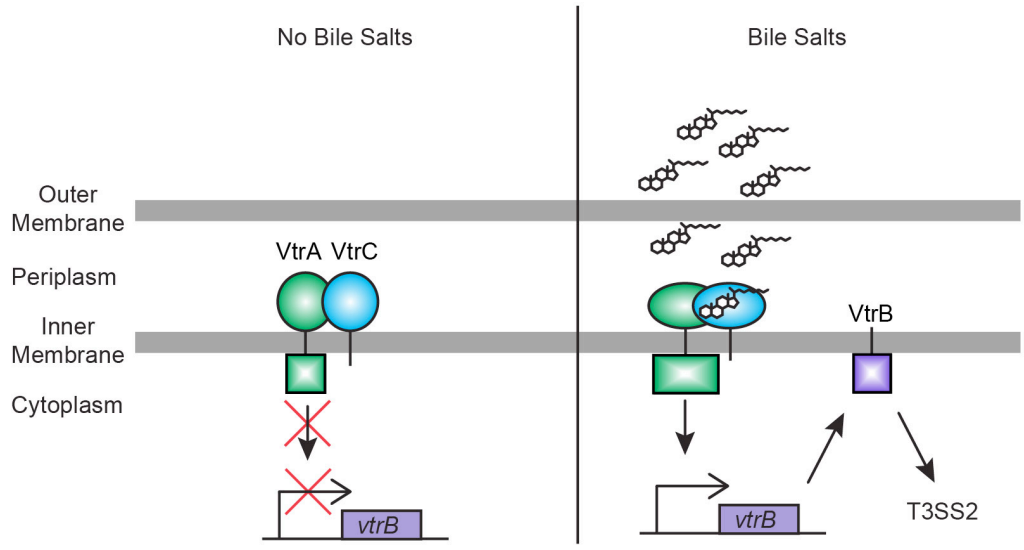


Figure 29. Model for bile salts sensing and T3SS2 regulation in *V. parahaemolyticus*

In the absence of bile salts, VtrA (green) and VtrC (blue) form a complex where the VtrA cytoplasmic DNA binding domain is kept inactive. Upon binding of bile salts, the VtrA/VtrC complex activates the VtrA DNA binding domain by a yet to be determined mechanism, which induces VtrB (purple) expression. VtrB, in turn, activates T3SS2.

The VtrA/VtrC complex is highly conserved in a group of diverse *Vibrionaceae* family species, as well as in *Moritella marina*. While the bile salt-induced T3SS in *V. cholerae* non-O1/O139 strains is necessary for bacterial colonization and the primary cause of disease during infection (*12*), the role of T3SS in *G. hollisae* infection is yet to be characterized. For other bacteria encoding *vtrA*, *vtrC* and *vtrB*, such as *Photobacterium marinum* and *V. caribbeanicus,* only a few T3SS related genes or no T3SS related genes are observed in their genomes, respectively. The differences between T3SS gene representation in these bacteria demonstrate this system's diversity and imply that the T3SS may be evolving in response to different living environments. Based on surrounding genes, this bile salt receptor complex may have adapted in other species to bind other hydrophilic ligands. For example, in the ten strains that lack a *vtrB* homolog but possess a *vtrA*/*vtrC* pair, the gene downstream of the *vtrA/vtrC* operon encodes a putative sphingomyelin phosphodiesterase. This enzyme is a hydrolase involved in the metabolism of sphingolipids (*107*) and is a putative virulence factor (*108*). It is tempting to hypothesize that the homologues of VtrA/VtrC in these bacteria function as a sensor for sphingolipids and regulate the induction of the downstream gene that produces sphingomyelin phosphodiesterase.

In conclusion, using microbial genetics and biophysics, we have identified and characterized a bile salt sensor that is widely distributed and used to induce virulence by many *Vibrio*s. Additionally, we have found that a family of monomeric lipid binding

calycin domain proteins has expanded to include an obligate heterodimer that binds bile salts and can be used to transmit a signal. To our knowledge, our study provides the first biochemical and structural analysis of a prokaryotic receptor involved in mediating response to bile salts, a significant environmental cue during infection.

CHAPTER FOUR

Acute Hepatopancreatic Necrosis Disease Causing *Vibrio parahaemolyticus* Contains A Virulent Type VI Secretion System With Antibacterial Activities

Introduction

Acute hepatopancreatic necrosis disease (AHPND) is a newly emerging shrimp disease that is causing serious reduction in shrimp production and financial losses for the global shrimp aquaculture industry (*20, 67*). Since the disease outbreak first appeared in China in 2009, it quickly spread to Vietnam, Malaysia, Thailand, Philippines and Mexico (*20, 68-70*). In 2013, it was revealed that AHPND was caused by a specific set of *Vibrio parahaemolyticus* isolates (*20*). Since then, various environmental non-AHPND and toxic AHPND strains of *V. parahaemolyticus* were isolated from affected shrimp farms coast around the world. Their whole genomes were sequenced and analyzed for the identification of toxic factors causing AHPND (*69, 109-112*).

In 2015, Dr. Lo's group from National Cheng Kung University discovered that AHPND-causing *V. parahaemolyticus* acquired a 63~70 kbp plasmid encoding the binary toxins $PirA^{vp}$/$PirB^{vp}$, which are homologous to the *Photorhabdus luminescens* insect-related (Pir) toxins PirA/PirB (*72, 73*). $PirA^{vp}$/$PirB^{vp}$ are secreted toxins and determined to be the primary toxic factors causing AHPND (*72, 73*). This plasmid encodes a set of

conjugative transfer and mobilization genes that could facilitate its spread between different *Vibrio* species (*72*). It is also interesting to note that this is a "selfish plasmid" that contains a toxin/antitoxin system ensuring the acquisition of this plasmid in bacterial progeny for survival (*72*).

Despite the discovery of this toxic plasmid and binary toxins $PirA^{vp}$/$PirB^{vp}$, no comprehensive study was done to characterize the other virulence systems commonly found in *V. parahaemolyticus* such as TDH/TRH toxins, T3SS and T6SS. To better understand the genetic features of AHPND causing *V. parahaemolyticus*, we selected a pool of 18 isolates of *V. parahaemolyticus* including one clinical, five environmental non-AHPND and twelve toxic AHPND strains. Using their genome sequences, we performed a thorough comparative analysis and found interesting correlations with key virulence determinants.

Results

Vibrio parahaemolyticus clinical, environmental non-AHPND and toxic AHPND strains contain different virulence factors

Genome sequences of one clinical, five environmental non-AHPND and twelve toxic AHPND strains of *V. parahaemolyticus* were collected for comparative analysis. Among these strains, we sequenced four, including one non-AHPND strain A2 and three AHPND strains 1335, 12297B and D4. The genome sequence data for the other strains were obtained from the NCBI database. RIMD2210633 is a clinical isolate and the others

strains are isolated from various countries with shrimp affected by AHPND including China (CN), Vietnam (VT), Thailand (TH) and Mexico (MX). We searched the genomes of all these *V. parahaemolyticus* strains for known key virulence factors such as the binary toxins $PirA^{vp}$/$PirB^{vp}$, T3SSs (T3SS1 and T3SS2), TDH/TRH toxins and T6SSs (T6SS1 and T6SS1). The results are summarized in **Table 7**.

Table 7 Summary of toxins and virulence systems in different *V. parahaemolyticus* strains analyzed in this study

	Origin	Strain	PirA^{vp}/ PirB^{vp}	T3SS1	T3SS2	TDH/ TRH	T6SS1	T6SS2
Clinical	JP	RIMD2210633	-	+	+	+	+	+
Non-AHPND	VT	A2	-	+	-	-	-	+
	TH	TUMSAT_H01_S4	-	+	-	-	-	+
		TUMSAT_H10_S6	-	+	-	-	-	+
		NCKU_TN_S02	-	+	-	-	-	+
	MX	FIM-S1392-	-	+	-	-	-	+
AHPND	VT	A3	+	+	-	-	+	+
		1335	+	+	-	-	+	+
		12297B	+	+	-	-	+	+
	TH	TUMSAT_DE1_S1	+	+	-	-	+	+
		TUMSAT_DE2_S2	+	+	-	-	+	+
		TUMSAT_D06_S3	+	+	-	-	+	+
		NCKU_TV_3HP	+	+	-	-	+	+
		NCKU_TV_5HP	+	+	-	-	+	+
	CN	NCKU_CV_CHN	+	+	-	-	+	+
	MX	D4	+	+	-	-	+	+
		M0605	+	+	-	-	+	+
		FIM-S1708+	+	+	-	-	+	+

JP: Japan; VT: Vietnam; TH: Thailand; MX: Mexico; CN: China

T3SS and TDH/TRH toxins

V. parahaemolyticus possesses two different T3SSs, called T3SS1 and T3SS2 (*15, 35*). T3SS1 is found in all sequenced strains of *V. parahaemolyticus*, including both environmental and clinical isolates (*36*). The T3SS1 pathogenicity island is on chromosome one and contains 49 genes (*vp1656-vp1702, vpa0450* and *vpa0451*) **(Figure 30)** (*15*). All of the *V. parahaemolyticus* strains analyzed in this study contain this highly conserved T3SS1 that is further categorized into two subtypes, T3SS1a and T3SS1b **(Figure 30)**. RIMD2210633, all five non-AHPND strains and seven out of the twelve AHPND strains have T3SS1a that contains *vp1676* to *vp1679*, shown as blue **(Figure 30)**. The other five AHPND strains possess T3SS1b that contains three hypothetical genes, shown as green in **Figure 30**, at the locus from *vp1676* to *vp1679* in T3SS1a. The functions of VP1676 to VP1679 remain unknown. VP1676 is a putative LysR family transcriptional regulator. VP1678 is a putative dienelactone hydrolase. VP1677 and VP1679 are both annotated as hypothetical proteins. The three newly identified proteins in T3SS1b are annotated as a putative transcriptional regulator, a putative Major Facilitator Superfamily (MFS) transporter and a putative multidrug transporter, respectively.

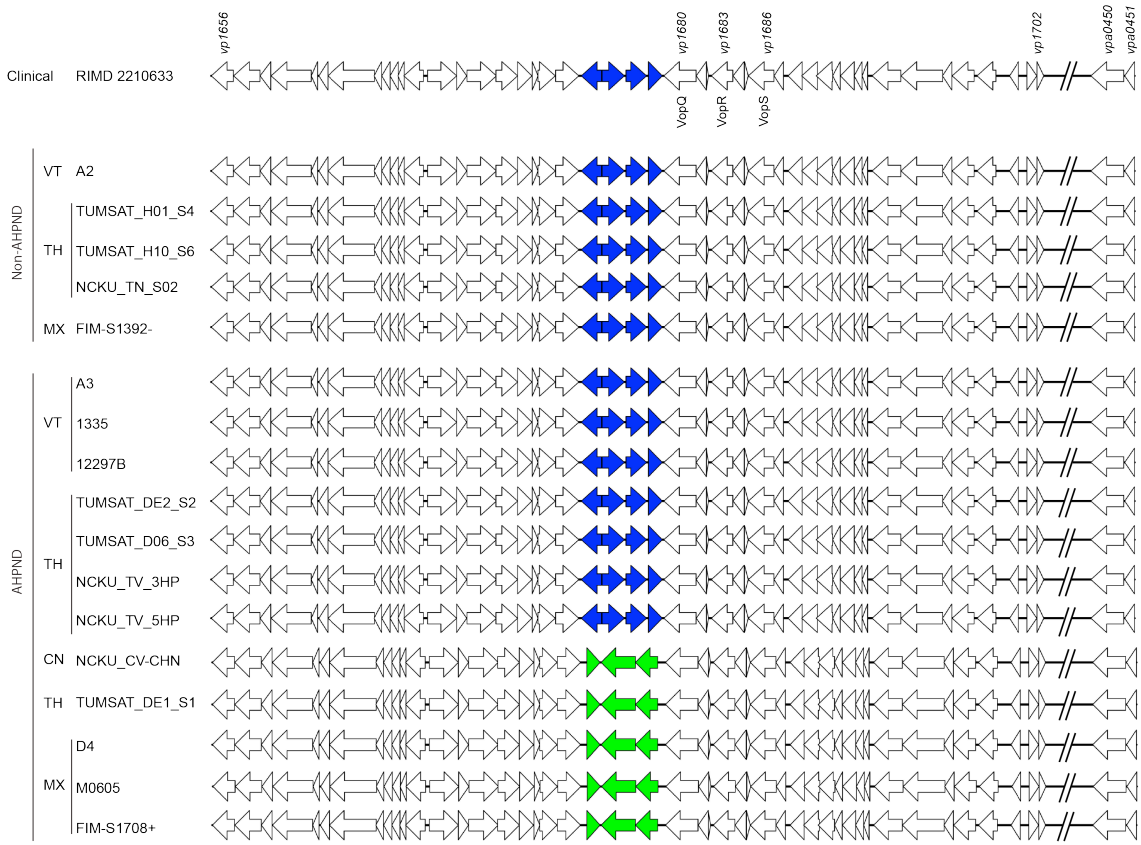


Figure 30. *V. parahaemoyticus* clinical, non-AHPND and AHPND strains contain a conserved T3SS1

Comparison of T3SS1 cluster of clinical, non-AHPND and AHPND strains of *V. parahaemoyticus*. All *V. parahaemoyticus* strains analyzed in this study contain a highly conserved T3SS1 that is further categorized into two subtypes, T3SS1a and T3SS1b. T3SS1a contains *vp1676* to *vp1679* (blue) and T3SS1b contains three newly identified genes (green). JP: Japan; VT: Vietnam; TH: Thailand; MX: Mexico; CN: China.

T3SS1 contains four characterized effectors VopQ (VP1680), VopR (VP1683), VopS (VP1683) and VPA0450 that orchestrate rapid cell death against multiple eukaryotic cell lines (*15, 16*). To test whether the T3SS1 of the *V. parahaemolyticus* strains analyzed in this study is functional, RIMD2210633-derived POR1 strain (containing deletions for the TDH/TRH toxins genes), one non-AHNPD strain and three AHPND strains were selected to infect HeLa cells. Their cytotoxicity was evaluated at 4h post infection by measuring the release of lactate dehydrogenase (LDH) into the culture medium during cell death. All of the strains tested exhibited similar cytotoxicity towards HeLa cells, which suggests their T3SS1 is active **(Figure 31)**.

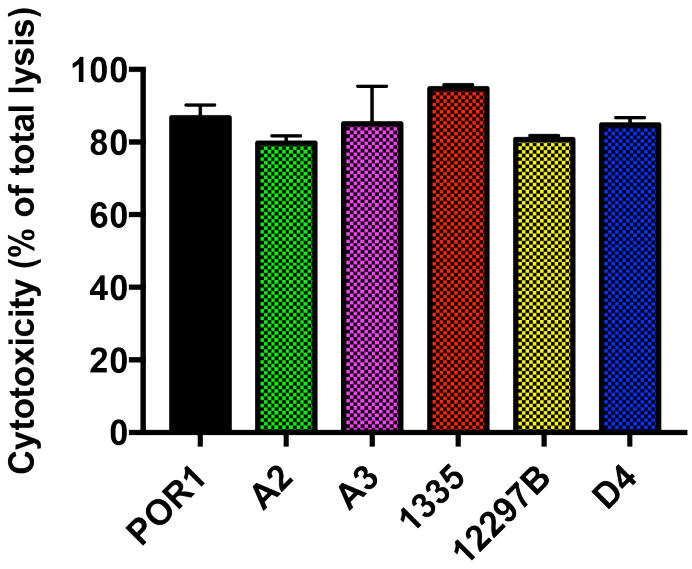


Figure 31. Cytotoxicity of T3SS1 against HeLa cells

HeLa cells were infected with POR1, A2, A3, 1335, 12297B and D4 stains of *V. parahaemolyticus* for 4h at MOI of 10. Lactate dehydrogenase (LDH) release was evaluated as the measure of cytotoxicity against host cells. Data is representative of three independent experiments.

None of the *V. parahaemolyticus* strains we analyzed except for RIMD2210633 contains T3SS2, consistent with previous reports that this system is only found in clinical isolates of *V. parahaemolyticus* (*69*). This indicates that the AHPND *V. parahaemolyticus* strains identified so far would not infect human beings and cause gastroenteritis. TDH and TRH toxins are generally associated and encoded along with T3SS2 (*36*). Consistent with the absence of T3SS2 in all of the other *V. parahaemolyticus* strains analyzed, they also lack TDH and TRH toxins.

T6SS2

T6SS2 is highly conserved, located on chromosome 2 and found in all sequenced isolates of *V. parahaemolyticus (58, 64)*. RIMD2210633 T6SS2 contains 22 genes (*vpa1025-vpa1046*). Consistent with previous discoveries, all of the *V. parahaemolyticus* strains analyzed in this study possess this conserved T6SS2 **(Figure 32A)**. T6SS2 is activated by low-salt (1% NaCl) medium and more active at 30°C compared to 37°C (*58*). To test whether the T6SS2 in the *V. parahaemolyticus* isolates analyzed in this study is active, clinical strain RIMD2210633-derived POR1 strain, non-AHPND strain A2 and three AHPND strains A3, 1335 and 12297B were selected. The expression and secretion of T6SS2 secreted protein Hcp2 was monitored as the marker for the activity of T6SS2. Similar to POR1, strains A2, 12297B and D4 expressed Hcp2 when grown at 30°C and the expression level of Hcp2 decreased at 37°C **(Figure 32B)** (*58*). Strains A3 and 1335 had no detectable level of Hcp2 expression in the cells indicating the regulatory mechanisms for the expression of T6SS2 components are different for these two strains

(Figure 32B). Only POR1, but none of the other strains, secreted Hcp2, supporting the hypothesis that secretion of T6SS2 components is also differentially regulated **(Figure 32B)**. The mechanisms underlying these differences remain unknown and need further characterization.

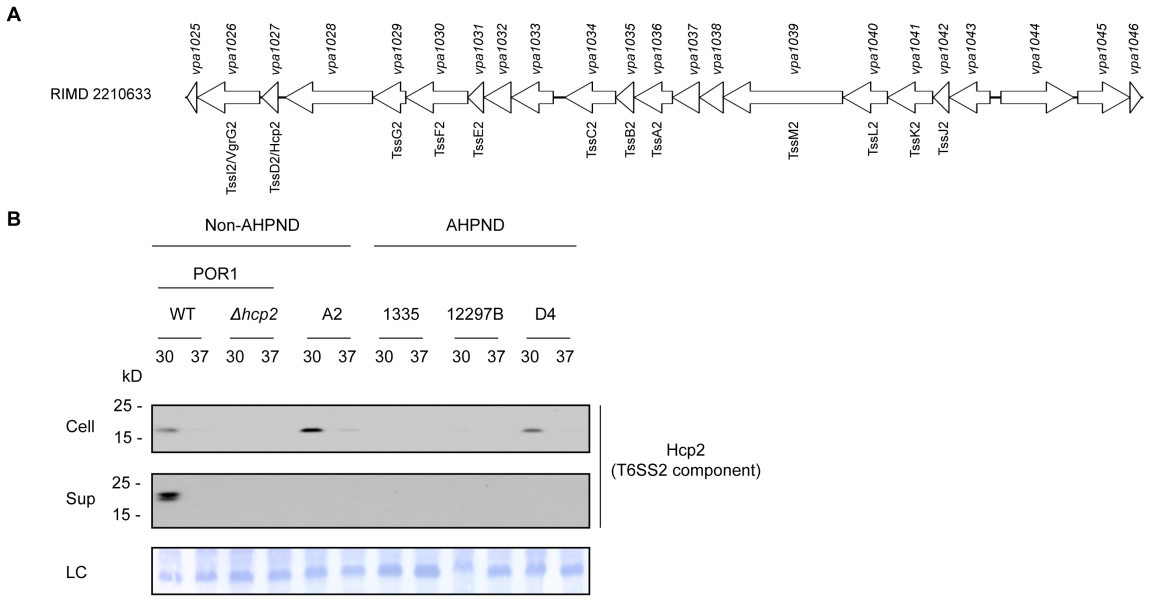


Figure 32. *V. parahaemoyticus* clinical, non-AHPND and AHPND strains contain a conserved but differentially regulated T6SS2

A. T6SS2 of *V. parahaemolyticus* RIMD2210633. **B.** Expression (Cell) and secretion (Sup) of *V. parahaemolyticus* T6SS2 component Hcp2 by *V. parahaemolyticus* strains POR1, POR1*Δhcp2*, A2, A3, 1335, 12297B and D4 grown in LB at 30°C or 37°C were analyzed by western blot using α-Hcp2 antibody. LC: loading control.

T6SS1

T6SS1 has previously been associated predominantly with clinical isolates of *V. parahaemolyticus* and shown to possess anti-bacterial activities against various bacterial competitors (*58, 64, 65*). RIMD2210633 T6SS1 contains a main cluster of 37 genes (*vp1386-vp1420*) and an orphan effector/immunity module (*vpa1263* and *vti2*), including three effectors that mediate anti-bacterial activities (*58, 66*). VP1388 is a Marker for type sIX (MIX) effector encoded at the beginning of the T6SS1 cluster with unknown function; VP1415 is a Pro-Ala-Ala-Arg (PAAR)-repeat-containing protein with a C-terminal Ala-His-His (AHH) nuclease domain and encoded at the end of the cluster; VPA1263 is another MIX effector located outside the T6SS1 cluster and contains PyocinS and colicin DNase domains **(Figure 33)** (*66*). Genes encoding immunity proteins that protect against self-toxicity are normally located next to their cognate effectors (*32*). In contrast to previous discoveries that T6SS1 is predominantly associated with clinical isolates of *V. parahaemolyticus*, all of the twelve AHPND strains analyzed in this study contain a similar T6SS1 **(Figure 33)**.

Comparative analysis of the T6SS1 between *V. parahaemolyticus* RIMD2210633 and the AHPND strains revealed variances at four different sites, which are highlighted as blue in RIMD2210633 T6SS1 **(Figure 33)**. Corresponding genes in the AHPND strains are labeled with blue when they are the same as the RIMD2210633 genes and different colors when otherwise. Three out of the four sites contain genes encoding the

T6SS1 effectors, which is consistent with previous discoveries that the T6SS1 effector repertoire varies between different isolates (*113*).

Site one contains genes *vp1388-vp1390*. VP1388 is a MIX effector and VP1389 is its putative immunity protein. VP1390 is a hypothetical protein with unknown function that contains a peptidoglycan binding domain similar to the C-terminal domain of outer-membrane protein OmpA. Six out of twelve AHPND strains contain the same three genes as RIMD2210633. Strains TUMSAT_DE2_S2, NCKU_TV_3HP and NCKU_TV_5HP contain a different set of genes labeled as purple while strains TUMSAT_D06_S3 and D4 have another set of genes labeled as red. In both sets of newly identified genes, they encode a MIX effector, a putative immunity protein and a hypothetical protein with unknown function that contains a peptidoglycan binding domain similar to the C-terminal domain of OmpA. The beginning of the T6SS1 cluster that contains site one is missing in the genome sequencing result of strain M0605 **(Figure 33)**. Site two is *vp1396* and *vp1397*. In RIMD2210633, VP1396 and VP1397 are two hypothetical proteins with unknown functions. In contrast, all twelve AHPND strains contain one gene at this site encoding a single fusion protein of VP1396 and VP1397. To confirm this difference, we amplified *vp1396* and *vp1397* by PCR for sequencing. Sequencing result shows that the current sequence of RIMD2210633 in the database wrongly inserted a 22nt fragment, which resulted in two separate genes, *vp1396* and *vp1397*. After correcting the error, RIMD2210633 contains a gene at site two that encodes the same protein as that in all AHPND strains. Site three is *vp1415-vp1420*. VP1415 is a T6SS1 effector containing a

N-terminal PAAR domain and a C-terminal AHH nuclease domain. VP1416-VP1420 are five copies of homologous immunity proteins. Site three is highly diverse between the AHPND strains. This T6SS1 effector can either be a single protein containing the PAAR and AHH nuclease domain similar to VP1415 (PAAR-AHH) or be two separate proteins (PAAR+AHH). In the later case, the first protein is the same in all cases and contains the PAAR domain while the second protein varies and has the nuclease domain. There are seven different types of site three among the twelve AHPND strains. Type one, A3 contains two T6SS1 effectors. The first one is a PAAR-AHH effector followed by one immunity protein. The second one is a PAAR+AHH effector followed by one immunity protein and *vp1420*. Type two, 1335 and 12297B contain the same PAAR+AHH effector followed by an immunity protein and *vp1420* as that in A3. Type three, found in NCKU_CV_CHN, FIM-S1708+ and TUMSAT_DE1_S1, contains two MIX effectors. The first one is *vp1415* followed by *vp1416* and *vp1417*. The second one is a PAAR+AHH effector followed by three immunity proteins and *vp1420*. The fourth type, found in TUMSAT_DE2_S2, NCKU_TV_3HP and NCKU_TV_5HP, contains a PAAR+AHH effector followed by *vp1419* and *vp1420*. The fifth type is found in TUMSAT_D06_S3 and contains two MIX effectors. The first one is the same PAAR-AHH effector followed by one immunity protein as that in A3. The second is a PAAR+AHH effector followed by one immunity protein and *vp1420*. The sixth type, found in D4, has a PAAR+AHH effector followed by one immunity protein and *vp1420*. And finally the seventh type of this site found in M0605 has the same PAAR+AHH effector followed by one immunity protein and *vp1420* as that in D4 with another

immunity protein and *vp1420* after. Site four contains *vpa1263* and *vti2*. It is located outside the T6SS1 cluster in RIMD2210633 but none of the AHPND strains analyzed contains this MIX effector VPA1263.

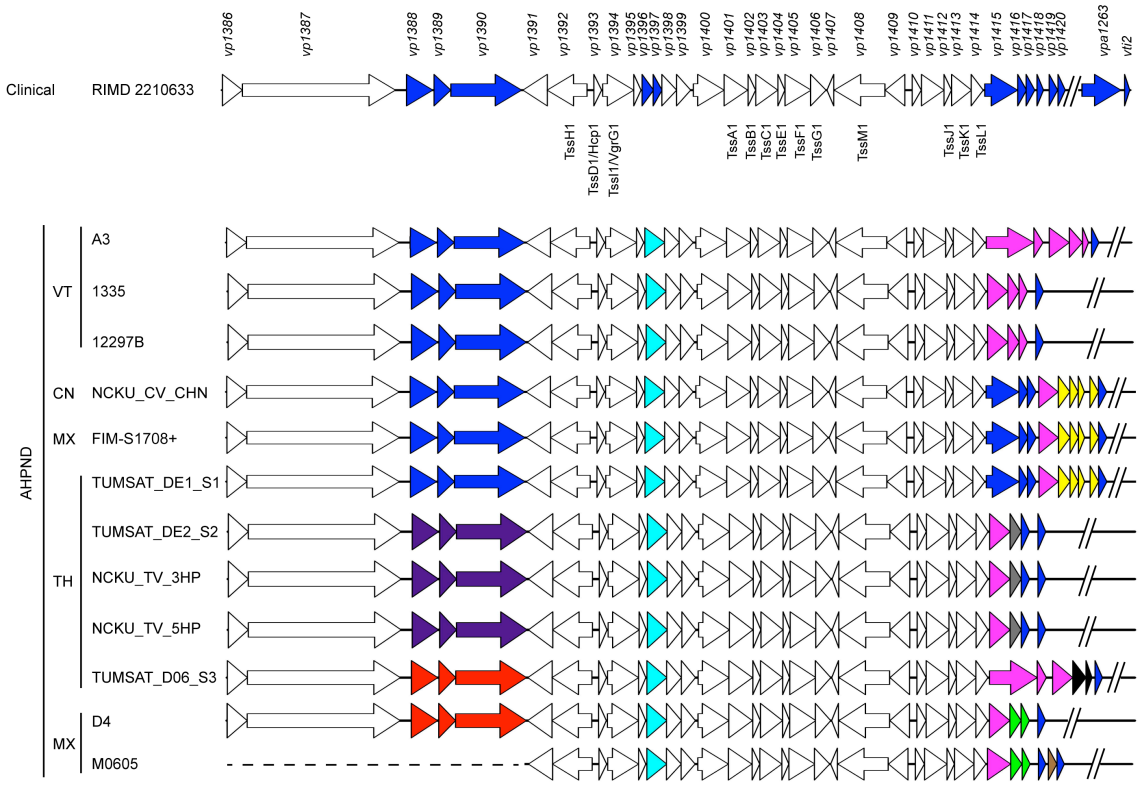


Figure 33. *V. parahaemoyticus* clinical strain RIMD2210633 and all AHPND strains contain a similar T6SS1

Comparison of the T6SS1 cluster of clinical and AHPND strains of *V. parahaemoyticus*. Genes with variations between the strains analyzed are highlighted with different colors. RIMD2210633 genes are labeled with blue and corresponding genes in AHPND strains are labeled with blue when they are the same and other colors when otherwise. The same color indicates identical or highly similar genes. JP: Japan; VT: Vietnam; TH: Thailand; MX: Mexico; CN: China.

Mobile T6SS1 MIX effectors

T6SS1 has been shown to contain mobile MIX effectors that are located outside the T6SS1 cluster such as VPA1263 in RIMD2210633 (*66, 114*). To test whether any of the AHPND strains contains mobile T6SS1 MIX effectors, we searched the protein libraries of all twelve AHPND strains using the MIX domain as the query. Analysis showed that each of four AHPND strains, TUMSAT_DE1_S1, NCKU_CV_CHN, FIM-S1708+ and M0605, possesses one mobile effector/immunity protein element (Figure 5). The mobile effectors from TUMSAT_DE1_S1, NCKU_CV_CHN and FIM-S1708+ all contain the PyocinS and colicin DNase domains, similar to VPA1263 found in the clinical isolate RIMD2210633. The effectors in TUMSAT_DE1_S1 and NCKU_CV_CHN are identical and found in the same genetic location. M0605 contains a novel mobile MIX effector with the LysM domain and is predicted to be a pore-forming toxin.

We noticed two interesting things about the genetic background of these mobile MIX effectors. First, the effector/immunity protein elements are flanked by genes encoding proteins that mediate DNA rearrangement and mobility, consistent with previous discoveries (*114*). Genes encoding transposase or resolvase, labeled as red, are found upstream of the effector/immunity protein element **(Figure 34)**. Downstream of the effector/immunity protein element, there are genes encoding integrase, invertase or helicase, shown as magenta **(Figure 34)**. For TUMSAT_DE1_S1 and NCKU_CV_CHN, no genes encoding transposase or resolvase are discovered upstream of their mobile MIX

effector. The assembled contig starts with two and one tRNA genes, respectively. Second, both the RIMD2210633 and FIM-S1708+ strains contain two genes encoding a toxin/anti-toxin module upstream of the mobile effector, which provides an effective mechanism for bacteria to maintain their mobile genetic element.

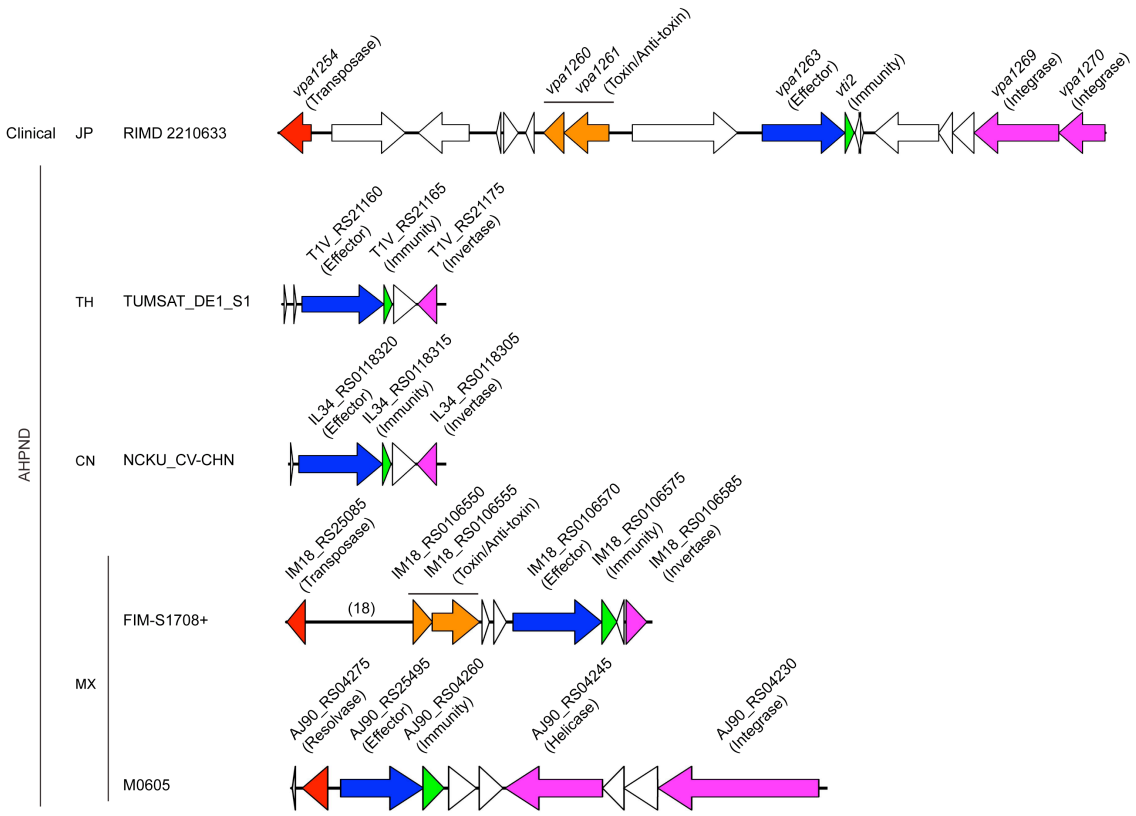


Figure 34. RIMD2210633 and four AHPND strains of *V. parahaemolyticus* contain mobile T6SS1 MIX effectors

Mobile genetic elements containing T6SS1 MIX effectors from the clinical isolate RIMD2210633 and four AHPND strains including TUMSAT_DE1_S1, NCKU_CV_CHN, FIM-S1708+ and M0605. Important components within the mobile elements are highlighted with different colors. T6SS1 mobile MIX effector (blue); immunity protein (green); toxin/anti-toxin (brown); genes encoding proteins that mediate DNA rearrangement and mobility upstream (red) and downstream (magenta) of the effector/immunity protein pair. Number shown in the parentheses indicates the number of genes in between.

AHPND *Vibrio parahaemolyticus* strains A3, 12297B and D4 contain an active T6SS1 that possesses anti-bacterial activity

RIMD2210633 T6SS1 is activated by high-salt (3% NaCl) medium at 30°C (*58*). It is further induced by surface sensing, which can be mimicked by adding the polar flagellar motor inhibitor, phenamil, to the growth medium (*58*). To test whether the T6SS1 in AHPND *V. parahaemolyticus* strains is active, four toxic AHPND isolates were selected and the expression and secretion of a T6SS1 secreted protein VgrG1 was monitored. As shown in **Figure 35A**, A3, 1335, 12297B and D4 all expressed and secreted VgrG1 when growing in high-salt phenamil containing medium at 30°C, similar to POR1, albeit, with 1335 to a much lesser extent. Next, the anti-bacterial activity of the T6SS1 in these four AHPND strains against other bacteria was tested using *Escherichia coli* as the prey. A3, D4 and 12297B killed *E.coli* similarly to POR1 but 1335 did not show any effect **(Figure 35B)**. The growth rate of POR1 and all four AHPND strains were similar **(Figure 36)**. The mechanism underlying the lack of anti-bacterial activities with 1335 needs further characterization and is likely related to lower expression and secretion level of T6SS1 components as indicated by VgrG1 **(Figure 35B)**. Given the important role of T6SS1 and the fact that all AHPND strains contain it, this virulence system is likely to kill other competitor bacteria and confer AHPND *V. parahaemolyticus* strains with a growth advantage to facilitate infection in shrimp.

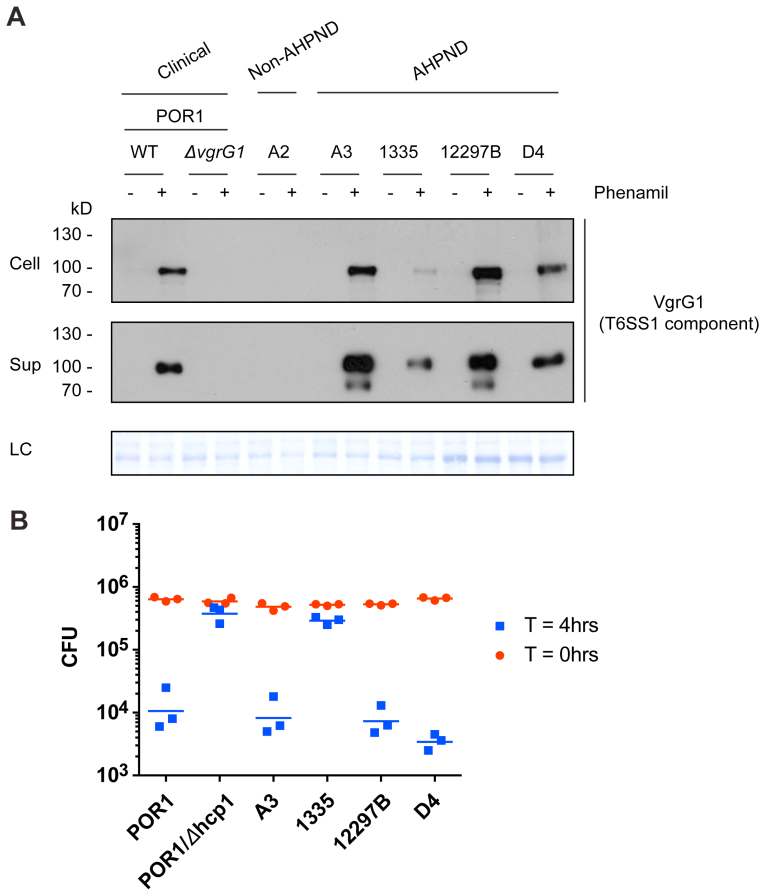


Figure 35. AHPND *V. parahaemolyticus* strains contain a virulent T6SS1 with anti-bacterial activity

A. Expression (Cell) and secretion (Sup) of *V. parahaemolyticus* T6SS1 component VgrG1 by *V. parahaemolyticus* strains POR1, POR1*ΔvgrG1*, A2, A3, 1335, 12297B and D4 grown in MLB at 30°C -/+ 20 μM phenamil were analyzed by western blot using α-VgrG1 antibody. LC: loading control. **B.** Viability counts before (0 h) and after (4 h) co-culture of *E. coli* with *V. parahaemolyticus* strains POR1, POR1Δ*hcp1* A3, 1335, 12297B and D4. Mixed cultures at an OD_{600} ratio of 1 : 4 (*E. coli*:*V. parahaemolyticus*) were spotted on MLB plates at 30°C .

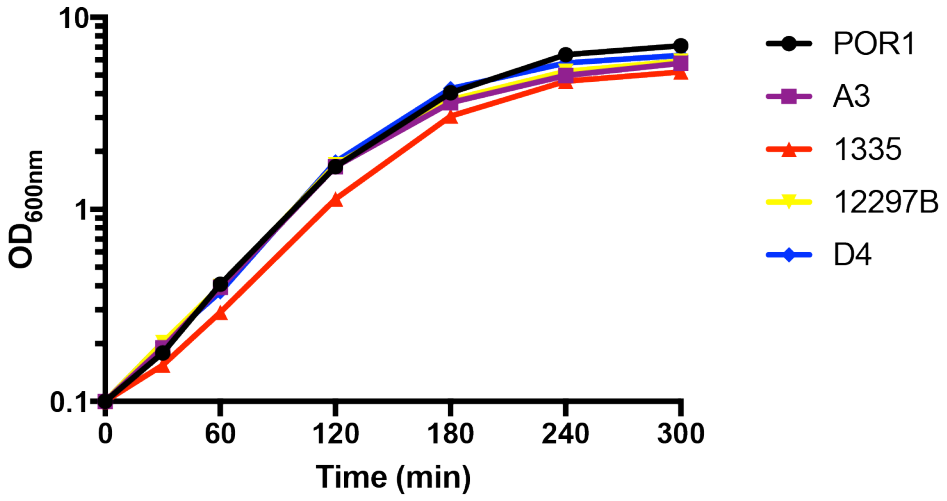


Figure 36. Different *V. parahaemolyticus* strains exhibit similar growth rate

Growth of *V. parahaemolyticus* strains POR1, A3, 1335, 12297B and D4 in MLB at 30°C based on OD_{600} measurements. Data are mean ± SD, n=3.

Discussion

AHPND is a newly emerging shrimp disease that has generated devastating impact on the global shrimp industry (*20, 67*). Identification of the plasmid encoded binary toxins $PirA^{vp}/PirB^{vp}$ as the deadly cause of AHPND provided significant insights for understanding the mechanism of this disease. However, to better prevent AHPND and design effective therapeutic strategies, it is necessary to gain better understanding of the genetic features and toxic systems used by AHPND *V. parahaemolyticus*. In this study, we conducted thorough comparative analysis of the genome sequences for clinical, environmental non-AHPND and toxic AHPND strains of *V. parahaemolyticus* and revealed important differences between their repertoire of virulence factors.

All of the *V. parahaemolyticus* strains analyzed in this study contain a highly conserved T3SS1 that has been discovered in all sequenced isolates of *V. parahaemolyticus* **(Figure 30)** (*36*). T3SS can be further categorized into T3SS1a and T3SS1b. Similar to clinical isolate RIMD2210633-derived POR1, infection of HeLa cells by non-AHPND strain A2 and four AHPND strains A3, 1335, 12297B and D4 showed that their T3SS1 was active and exhibited cytotoxicity in a similar fashion **(Figure 31)**. Except for the clinical isolate RIMD2210633, none of the strains used in this study possess T3SS2 and TDH/TRH toxins.

A highly conserved T6SS2 was found in all of the *V. parahaemolyticus* strains analyzed in this study **(Figure 32A)**. However, the expression and secretion of T6SS2

components in the *V. parahaemolyticus* strains tested are differentially regulated. Using Hcp2 as a marker for the activity of T6SS2, our studies showed that selected *V. parahaemolyticus* strains expressed and secreted Hcp2 differently than RIMD2210633-derived POR1 strain **(Figure 32B)**. To understand the mechanisms underlying these differences, it will be necessary to further characterize the regulation of T6SS2 mediated by components outside the T6SS2 cluster. It has been shown in *Vibrio cholerae* that the T6SS in clinical and environmental isolates can be differentially regulated (*115*).

T6SS1 has previously been shown to be primarily associated with clinical isolates of *V. parahaemolyticus* and contain anti-bacterial activities (*58, 64, 65*). However, among the *V. parahaemolyticus* strains used in our study, all of the twelve AHPND strains contain a T6SS1 similar to that found in RIMD2210633, whereas, none of the non-AHPND strains contains T6SS1 **(Figure 33)**. Given this distinguishing feature, it is likely that T6SS1 may play an important role during shrimp infection by killing competitor bacteria thereby providing a growth advantage for toxic AHPND *V. parahaemolyticus*. Furthermore, we demonstrated that three AHPND strains, A3, 12297B and D4, possess anti-bacterial activities and can kill *E. coli* in a fashion similar to RIMD2210633-derived POR1 strain **(Figure 35B)**. Interestingly, another AHPND strain, 1335, does not kill *E. coli* probably due to lower expression and secretion level of T6SS1 components indicated by a marker protein VgrG1 **(Figure 35A)**. Another possible explanation for this phenomenon could be that the environmental conditions that activate this T6SS1 differ from those used in our in vitro test. To further characterize the role of T6SS1 in causing

AHPND, the toxicity of wild type and mutant of AHPND strains with inactivated T6SS1 against the shrimp needs to be evaluated.

Various *V. parahaemolyticus* strains may contain different repertoires of effector and immunity proteins (*58, 113*). Consistent with this, comparison of the T6SS1 from RIMD2210633 and all of the AHPND strains showed that major variations are primarily located at genes encoding the T6SS1 effector and immunity proteins **(Figure 33)**. RIMD2210633 T6SS1 contains three effector/immunity protein elements. Genes encoding VP1388-VP1390 and VP1415-VP1420 are inside the T6SS1 cluster and genes encoding the mobile element VPA1263-Vti2 are located outside of the T6SS1 cluster. Our analysis revealed three and eight different versions for VP1388-VP1390 and VP1415-VP1420, respectively. In the search for mobile T6SS1 effectors, we discovered that each of the four AHPND strains, TUMSAT_DE1_S1, NCKU_CV_CHN, FIM-S1708+ and M0605, contains one mobile T6SS1 effector/immunity protein element **(Figure 34)**. Three of the effectors contain the PyocinS and colicin DNase domains, similar to VPA1263 and the fourth effector is predicted to be a pore-forming toxin with the LysM domain. The mobile effector/immunity protein elements are normally located within mobile genetic units that contain genes encoding proteins involved in DNA rearrangement and mobility, such as transposase, resolvase, integrase, invertase and helicase. Interestingly, both RIMD2210633 and FIM-S1708+ strains possess genes that encode a toxin/anti-toxin pair in their mobile genetic units, providing an effective mechanism for the maintenance of these mobile units.

In conclusion, our studies revealed different repertoires of virulence factors in non-AHPND and AHPND *V. parahaemolyticus* that provided important insights for persistance of this devastating shrimp disease.

CHAPTER FIVE

Conclusions and Future Directions

Conclusions

This thesis work characterized the toxicity of *V. parahaemolyticus* against two different hosts, human and shrimp, and provided important information about the infectious diseases caused by *V. parahaemolyticus* and similar bacterial pathogens. Key discoveries from this study are summarized below.

VtrC is a newly identified component of a widespread, conserved signaling pathway for bile salt sensing and activation of virulence factors in *Vibrios*

V. parahaemolyticus vtrC is a previously unidentified gene located downstream of *vtrA*, which together with *vtrB* have been shown to regulate a virulent T3SS2 during human infection (*14, 54*). Bioinformatics studies showed that *vtrA* and *vtrC* are conserved within a group of 15 different bacteria including *Vibrios* and other species. Among them, five bacteria species also contain the homologous gene of *vtrB*. Three out of these five bacteria species, *V. parahaemolyticus, V. cholerae* non-O1/O139 strains and *Grimontia hollisae*, possess a similar T3SS and cause gastroenteritis during human infection (*11*). VtrC plays an essential role in the activation of T3SS2 by bile salts because deletion of *vtrC* completely abolished the expression and secretion of different components of

T3SS2 including the T3SS2 effectors and translocon. Complementation of *vtrC* deletion fully restored the T3SS2 activity.

Periplasmic domains of VtrA and VtrC form an obligate heterodimer that functions as a receptor for bile salts

The crystal structure of the VtrA/VtrC complex reveals that their periplasmic domains form a obligate heterodimer where VtrC folds into an open eight-stranded β-barrel and VtrA adopts a mixed α/β structure whereby one of its a β-strands contributes to the formation of the β-barrel. A co-crystal structure of VtrA/VtrC with the bile salt TDC, along with biophysical and mutational analysis, demonstrates that the hydrophobic chamber within the β-barrel binds bile salts and activates the virulence network.

AHPND *V. parahaemolyticus* strains contain a virulent T6SS1 with anti-bacterial activity

T6SS1 has previously been shown to be primarily associated with clinical isolates of *V. parahaemolyticus* and contain anti-bacterial activities (*58, 64, 65*). Through a comparative analysis using the genome sequences of clinical, non-AHPND and AHPND strains of *V. parahaemolyticus*, it was discovered that all of the AHPND but none of the non-AHPND strains contain a virulent T6SS1, like that in the clinical isolate RIMD2210633. We further demonstrated that three out of four AHPND strains tested possess anti-bacterial activities and can kill *E. coli* in a fashion similar to RIMD2210633-derived POR1 strain **(Figure 35B)**. The fourth AHPND strain 1335, although able to

induce T6SS1, does not kill *E. coli* probably due to lower expression and secretion level of T6SS1 components indicated by a marker protein VgrG1 **(Figure 35A)**.

Future directions

How is the activity of VtrA regulated upon sensing of bile salts by the VtrA/VtrC complex?

Sensing of bile salts by the VtrA/VtrC complex increases the activity of VtrA, which then induces VtrB to activate T3SS2. Our study focused on the periplasmic domains of VtrA and VtrC, not the cytoplasmic DBD of VtrA. As a result, the mechanism underlying the regulation of VtrA activity remains unknown and needs further characterization in the future. Two possible mechanisms are post-translational modification and protein oligmerization, which are commonly seen for transcription factors.

How do *V. parahaemolyticus* ToxR and ToxS affect T3SS2 activity?

It was shown recently that *V. parahaemolyticus* ToxR (VP0820) is essential for the activation of T3SS2 during infection (*116*). ToxS (VP0819) is cotranscribed with ToxR and was shown to affect bacterial colonization in vivo but not the T3SS2 activity in vitro. ToxR/ToxS have low protein sequence similarities with VtrA/VtrC but share the same domain organization and cellular localization. Understanding how ToxR/ToxS regulate T3SS2 activity is important for deciphering *V. parahaemolyticus* induced infectious disease. Interesting questions such as whether ToxR/ToxS form a complex that

senses bile salts and if there is crosstalk between VtrA/VtrC and ToxR/ToxS need to be answered.

Roles of VtrA/VtrC homologues in other bacterial species

VtrA/VtrC are conserved in a large group of different bacteria including *Vibrios* and other species. The roles of VtrA/VtrC homologues in other bacteria are largely unknown. For *V. cholerae* non-O1/O139 strains and *Grimontia hollisae*, which possess a T3SS similar to T3SS2 and cause gastroenteritis during human infection, it is very likely their VtrA/VtrC homologues also function as the receptor for bile salts to regulate the virulent T3SS during infection. In the ten bacterial species that lack a *vtrB* homolog but possess a *vtrA*/*vtrC* pair, the gene downstream of the *vtrA/vtrC* operon encodes a putative sphingomyelin phosphodiesterase. This enzyme is a hydrolase involved in the metabolism of sphingolipids (*107*) and is a putative virulence factor (*108*). It is tempting to hypothesize that the homologues of VtrA/VtrC in these bacteria function as a sensor for sphingolipids and regulate the induction of the downstream gene that produces sphingomyelin phosphodiesterase.

Roles of *V. cholerae* ToxR/ToxS and TcpP/TcpH

V. cholerae contains two pairs of regulatory proteins, ToxR/ToxS and TcpP/TcpH. They control the expression of numerous virulence factors including CT and TCP and it was shown that their activities were affected by bile (*10, 106*). Several signaling features of VtrA/VtrC appear to parallel ToxR/ToxS and TcpP/TcpH, raising

the question whether ToxR/ToxS and TcpP/TcpH function as the receptor for bile in *V. cholerae*. To solve this mystery in the future, we can try to solve the crystal structure of ToxR/ToxS and TcpP/TcpH with bile.

Mechanism underlying the differential regulation of T6SS2 in clinical, non-AHPND and AHPND strains of *V. parahaemolyticus*

Clinical, non-AHPND and AHPND strains of *V. parahaemolyticus* contain a conserved T6SS2. However, the strains we tested showed differential regulation of the expression and secretion of T6SS2 components indicated by a T6SS2 marker protein Hcp2. This is consistent with previous discoveries in *V. cholerae* that the T6SS in clinical and environmental isolates can be differentially regulated (*115*). To understand the mechanism underlying this phenomenon, we need to discover the regulators of T6SS2 activity that are located outside the T6SS2 and study how they vary among different *V. parahaemolyticus* strains.

Role of T6SS1 in causing AHPND

T6SS1 has previously been shown to be primarily associated with clinical isolates of *V. parahaemolyticus* and contain anti-bacterial activities (*58, 64, 65*). Our study discovered a distinguishing feature that a virulent T6SS1, like that in RIMD2210633, is present in all of the AHPND but none of the non-AHPND strains of *V. parahaemolyticus*. We demonstrated that three AHPND strains, A3, 12297B and D4, possess anti-bacterial activities and can kill *E. coli* in a fashion similar to POR1 **(Figure 35B)**. However,

another AHPND strain, 1335, does not kill *E. coli* probably due to lower expression and secretion level of T6SS1 components indicated by a marker protein VgrG1 **(Figure 35A)**. We believe this T6SS1 is potentially conserved in these virulent strains because of its anti-bacterial activities to compete against other commensal bacteria during host infection. To test this hypothesis, we can compare the toxicity of wild type and mutant of AHPND strains with inactivated T6SS1 in a shrimp infection model.

Therapeutic strategies for curing AHPND

The ultimate goal of studying the genetic features of AHPND strains of *V. parahaemolyticus* is to design an effective therapeutic strategy and cure AHPND. Although it is already determined that $PirA^{vp}/PirB^{vp}$ are secreted toxins that cause AHPND, it is not clear how $PirA^{vp}/PirB^{vp}$ are secreted (*72, 73*). With the genome sequences of various non-AHPND and AHPND strains of *V. parahaemolyticus* available, it is now feasible to identify the secretion system that control $PirA^{vp}/PirB^{vp}$ and then design inhibitors to block the their secretion. If T6SS1 is shown to play an important role in causing AHPND in the future, we can also design strategies to inhibit T6SS1 and decrease the severity of this shrimp disease.

BIBLIOGRAPHY

1. M. Begley, C. G. Gahan, C. Hill, The interaction between bacteria and bile. *FEMS Microbiol Rev* **29**, 625-651 (2005).
2. J. M. Ridlon, D. J. Kang, P. B. Hylemon, Bile salt biotransformations by human intestinal bacteria. *J Lipid Res* **47**, 241-259 (2006).
3. A. M. Prouty, J. S. Gunn, Salmonella enterica serovar typhimurium invasion is repressed in the presence of bile. *Infect Immun* **68**, 6763-6769 (2000).
4. A. M. Prouty *et al.*, Transcriptional regulation of Salmonella enterica serovar Typhimurium genes by bile. *FEMS Immunol Med Microbiol* **41**, 177-185 (2004).
5. L. M. Pope, K. E. Reed, S. M. Payne, Increased protein secretion and adherence to HeLa cells by Shigella spp. following growth in the presence of bile salts. *Infect Immun* **63**, 3642-3648 (1995).
6. S. Gupta, R. Chowdhury, Bile affects production of virulence factors and motility of Vibrio cholerae. *Infect Immun* **65**, 1131-1134 (1997).
7. D. A. Schuhmacher, K. E. Klose, Environmental signals modulate ToxT-dependent virulence factor expression in Vibrio cholerae. *J Bacteriol* **181**, 1508-1514 (1999).
8. E. S. Krukonis, V. J. DiRita, From motility to virulence: Sensing and responding to environmental signals in Vibrio cholerae. *Curr Opin Microbiol* **6**, 186-190 (2003).

9. S. M. Faruque, M. J. Albert, J. J. Mekalanos, Epidemiology, genetics, and ecology of toxigenic Vibrio cholerae. *Microbiol Mol Biol Rev* **62**, 1301-1314 (1998).

10. D. T. Hung, J. J. Mekalanos, Bile acids induce cholera toxin expression in Vibrio cholerae in a ToxT-independent manner. *Proc Natl Acad Sci U S A* **102**, 3028-3033 (2005).

11. M. Dziejman *et al.*, Genomic characterization of non-O1, non-O139 Vibrio cholerae reveals genes for a type III secretion system. *Proc Natl Acad Sci U S A* **102**, 3465-3470 (2005).

12. M. Chaand *et al.*, Type 3 Secretion System Island Encoded Proteins Required for Colonization by Non-O1/non-O139 Serogroup V. cholerae. *Infect Immun*, (2015).

13. A. Alam, V. Tam, E. Hamilton, M. Dziejman, vttRA and vttRB Encode ToxR family proteins that mediate bile-induced expression of type three secretion system genes in a non-O1/non-O139 Vibrio cholerae strain. *Infect Immun* **78**, 2554-2570 (2010).

14. K. Gotoh *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).

15. C. A. Broberg, T. J. Calder, K. Orth, Vibrio parahaemolyticus cell biology and pathogenicity determinants. *Microbes Infect* **13**, 992-1001 (2011).

16. L. Zhang, K. Orth, Virulence determinants for Vibrio parahaemolyticus infection. *Curr Opin Microbiol* **16**, 70-77 (2013).

17. S. Shinoda, Sixty years from the discovery of Vibrio parahaemolyticus and some recollections. *Biocontrol Sci* **16**, 129-137 (2011).

18. N. O'Boyle, A. Boyd, Manipulation of intestinal epithelial cell function by the cell contact-dependent type III secretion systems of Vibrio parahaemolyticus. *Front Cell Infect Microbiol* **3**, 114 (2014).

19. N. A. Daniels *et al.*, Vibrio parahaemolyticus infections in the United States, 1973-1998. *J Infect Dis* **181**, 1661-1666 (2000).

20. L. Tran *et al.*, Determination of the infectious nature of the agent of acute hepatopancreatic necrosis syndrome affecting penaeid shrimp. *Dis Aquat Organ* **105**, 45-55 (2013).

21. G. B. Nair *et al.*, Global dissemination of Vibrio parahaemolyticus serotype O3:K6 and its serovariants. *Clin Microbiol Rev* **20**, 39-48 (2007).

22. J. Velazquez-Roman, N. Leon-Sicairos, L. de Jesus Hernandez-Diaz, A. Canizalez-Roman, Pandemic Vibrio parahaemolyticus O3:K6 on the American continent. *Front Cell Infect Microbiol* **3**, 110 (2013).

23. J. B. McLaughlin *et al.*, Outbreak of Vibrio parahaemolyticus gastroenteritis associated with Alaskan oysters. *N Engl J Med* **353**, 1463-1470 (2005).

24. L. Vezzulli *et al.*, Climate influence on Vibrio and associated human diseases during the past half-century in the coastal North Atlantic. *Proc Natl Acad Sci U S A*, (2016).

25. K. Makino *et al.*, Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism distinct from that of V cholerae. *Lancet* **361**, 743-749 (2003).

26. M. de Souza Santos, D. Salomon, P. Li, A.-M. Krachler, K. Orth, in *The Comprehensive Sourcebook of Bacterial Protein Toxins,* J. Alouf, D. Ladant, M. R. Popoff, Eds. (Elsevier, Waltham, 2015), pp. 230-260.

27. T. Honda, Y. Ni, T. Miwatani, T. Adachi, J. Kim, The thermostable direct hemolysin of Vibrio parahaemolyticus is a pore-forming toxin. *Can J Microbiol* **38**, 1175-1180 (1992).

28. F. Raimondi *et al.*, Enterotoxicity and cytotoxicity of Vibrio parahaemolyticus thermostable direct hemolysin in in vitro systems. *Infect Immun* **68**, 3180-3185 (2000).

29. T. Fukui *et al.*, Thermostable direct hemolysin of Vibrio parahaemolyticus is a bacterial reversible amyloid toxin. *Biochemistry* **44**, 9825-9832 (2005).

30. J. E. Galan, M. Lara-Tejero, T. C. Marlovits, S. Wagner, Bacterial type III secretion systems: specialized nanomachines for protein delivery into target cells. *Annu Rev Microbiol* **68**, 415-438 (2014).

31. B. T. Ho, T. G. Dong, J. J. Mekalanos, A view to a kill: the bacterial type VI secretion system. *Cell Host Microbe* **15**, 9-21 (2014).

32. A. B. Russell, S. B. Peterson, J. D. Mougous, Type VI secretion system effectors: poisons with a purpose. *Nat Rev Microbiol* **12**, 137-148 (2014).

33. A. B. Russell *et al.*, Type VI secretion delivers bacteriolytic effectors to target cells. *Nature* **475**, 343-347 (2011).

34. J. E. Galan, H. Wolf-Watz, Protein delivery into eukaryotic cells by type III secretion machines. *Nature* **444**, 567-573 (2006).

35. N. Okada *et al.*, Identification and characterization of a novel type III secretion system in trh-positive Vibrio parahaemolyticus strain TH3996 reveal genetic lineage and diversity of pathogenic machinery beyond the species level. *Infect Immun* **77**, 904-913 (2009).

36. N. Okada *et al.*, Presence of genes for type III secretion system 2 in Vibrio mimicus strains. *BMC Microbiol* **10**, 302 (2010).

37. T. Ono, K. S. Park, M. Ueta, T. Iida, T. Honda, Identification of proteins secreted via Vibrio parahaemolyticus type III secretion system 1. *Infect Immun* **74**, 1032-1042 (2006).

38. D. L. Burdette, M. L. Yarbrough, A. Orvedahl, C. J. Gilpin, K. Orth, Vibrio parahaemolyticus orchestrates a multifaceted host cell infection by induction of autophagy, cell rounding, and then cell lysis. *Proc Natl Acad Sci U S A* **105**, 12497-12502 (2008).

39. C. A. Broberg, L. Zhang, H. Gonzalez, M. A. Laskowski-Arce, K. Orth, A Vibrio effector protein is an inositol phosphatase and disrupts host cell membrane integrity. *Science* **329**, 1660-1662 (2010).

40. M. L. Yarbrough *et al.*, AMPylation of Rho GTPases by Vibrio VopS disrupts effector binding and downstream signaling. *Science* **323**, 269-272 (2009).

41. A. Sreelatha *et al.*, Vibrio effector protein, VopQ, forms a lysosomal gated channel that disrupts host ion homeostasis and autophagic flux. *Proc Natl Acad Sci U S A* **110**, 11559-11564 (2013).

42. D. Salomon *et al.*, Effectors of animal and plant pathogens use a common domain to bind host phosphoinositides. *Nat Commun* **4**, 2973 (2013).

43. J. M. Ritchie *et al.*, Inflammation and disintegration of intestinal villi in an experimental model for Vibrio parahaemolyticus-induced diarrhea. *PLoS Pathog* **8**, e1002593 (2012).

44. K. S. Park *et al.*, Functional characterization of two type III secretion systems of Vibrio parahaemolyticus. *Infect Immun* **72**, 6659-6665 (2004).

45. H. Hiyoshi, T. Kodama, T. Iida, T. Honda, Contribution of Vibrio parahaemolyticus virulence factors to cytotoxicity, enterotoxicity, and lethality in mice. *Infect Immun* **78**, 1772-1780 (2010).

46. J. E. Trosky *et al.*, Inhibition of MAPK signaling pathways by VopA from Vibrio parahaemolyticus. *J Biol Chem* **279**, 51953-51957 (2004).

47. H. Hiyoshi *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).

48. N. Higa *et al.*, Vibrio parahaemolyticus effector proteins suppress inflammasome activation by interfering with host autophagy signaling. *PLoS Pathog* **9**, e1003142 (2013).

49. H. Hiyoshi *et al.*, Interaction between the type III effector VopO and GEF-H1 activates the RhoA-ROCK pathway. *PLoS Pathog* **11**, e1004694 (2015).

50. L. Zhang *et al.*, Type III effector VopC mediates invasion for Vibrio species. *Cell Rep* **1**, 453-460 (2012).

51. X. Zhou *et al.*, A Vibrio parahaemolyticus T3SS effector mediates pathogenesis by independently enabling intestinal colonization and inhibiting TAK1 activation. *Cell Rep* **3**, 1690-1702 (2013).

52. T. Kodama *et al.*, Identification and characterization of VopT, a novel ADP-ribosyltransferase effector protein secreted via the Vibrio parahaemolyticus type III secretion system 2. *Cell Microbiol* **9**, 2598-2609 (2007).

53. T. Calder *et al.*, Vibrio type III effector VPA1380 is related to the cysteine protease domain of large bacterial toxins. *PLoS One* **9**, e104387 (2014).

54. T. Kodama *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).

55. F. Boyer, G. Fichant, J. Berthod, Y. Vandenbrouck, I. Attree, Dissecting the bacterial type VI secretion system by a genome wide in silico analysis: what can be learned from available microbial genomic resources? *BMC Genomics* **10**, 104 (2009).

56. A. Zoued *et al.*, Architecture and assembly of the Type VI secretion system. *Biochim Biophys Acta* **1843**, 1664-1673 (2014).

57. M. Basler, M. Pilhofer, G. P. Henderson, G. J. Jensen, J. J. Mekalanos, Type VI secretion requires a dynamic contractile phage tail-like structure. *Nature* **483**, 182-186 (2012).

58. D. Salomon, H. Gonzalez, B. L. Updegraff, K. Orth, 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).

59. K. L. Hockett, A. Y. Burch, S. E. Lindow, Thermo-regulation of genes mediating motility and plant interactions in Pseudomonas syringae. *PLoS One* **8**, e59850 (2013).

60. T. Ishikawa *et al.*, Pathoadaptive conditional regulation of the type VI secretion system in Vibrio cholerae O1 strains. *Infect Immun* **80**, 575-584 (2012).

61. R. Pieper *et al.*, Temperature and growth phase influence the outer-membrane proteome and the expression of a type VI secretion system in Yersinia pestis. *Microbiology* **155**, 498-512 (2009).

62. J. M. Silverman *et al.*, Separate inputs modulate phosphorylation-dependent and -independent type VI secretion activation. *Mol Microbiol* **82**, 1277-1290 (2011).

63. T. G. Sana *et al.*, The second type VI secretion system of Pseudomonas aeruginosa strain PAO1 is regulated by quorum sensing and Fur and modulates internalization in epithelial cells. *J Biol Chem* **287**, 27095-27105 (2012).

64. Y. Yu *et al.*, Putative type VI secretion systems of Vibrio parahaemolyticus contribute to adhesion to cultured cell monolayers. *Arch Microbiol* **194**, 827-835 (2012).

65. E. F. Boyd *et al.*, Molecular analysis of the emergence of pandemic Vibrio parahaemolyticus. *BMC Microbiol* **8**, 110 (2008).

66. D. Salomon *et al.*, Marker for type VI secretion system effectors. *Proc Natl Acad Sci U S A* **111**, 9271-9276 (2014).

67. H. C. Lai *et al.*, Pathogenesis of acute hepatopancreatic necrosis disease (AHPND) in shrimp. *Fish Shellfish Immunol* **47**, 1006-1014 (2015).

68. L. D. de la Pena *et al.*, Acute hepatopancreatic necrosis disease (AHPND) outbreaks in Penaeus vannamei and P. monodon cultured in the Philippines. *Dis Aquat Organ* **116**, 251-254 (2015).

69. B. Gomez-Gil, S. Soto-Rodriguez, R. Lozano, M. Betancourt-Lozano, Draft Genome Sequence of Vibrio parahaemolyticus Strain M0605, Which Causes Severe Mortalities of Shrimps in Mexico. *Genome Announc* **2**, (2014).

70. L. Nunan, D. Lightner, C. Pantoja, S. Gomez-Jimenez, Detection of acute hepatopancreatic necrosis disease (AHPND) in Mexico. *Dis Aquat Organ* **111**, 81-86 (2014).

71. P. De Schryver, T. Defoirdt, P. Sorgeloos, Early mortality syndrome outbreaks: a microbial management issue in shrimp farming? *PLoS Pathog* **10**, e1003919 (2014).

72. C. T. Lee *et al.*, The opportunistic marine pathogen Vibrio parahaemolyticus becomes virulent by acquiring a plasmid that expresses a deadly toxin. *Proc Natl Acad Sci U S A* **112**, 10798-10803 (2015).

73. R. Sirikharin *et al.*, Characterization and PCR Detection Of Binary, Pir-Like Toxins from Vibrio parahaemolyticus Isolates that Cause Acute Hepatopancreatic Necrosis Disease (AHPND) in Shrimp. *PLoS One* **10**, e0126987 (2015).

74. H. Kondo, P. T. Van, L. T. Dang, I. Hirono, Draft Genome Sequence of Non-Vibrio parahaemolyticus Acute Hepatopancreatic Necrosis Disease Strain KC13.17.5, Isolated from Diseased Shrimp in Vietnam. *Genome Announc* **3**, (2015).

75. L. Liu *et al.*, Draft Genome Sequence of Vibrio owensii Strain SH-14, Which Causes Shrimp Acute Hepatopancreatic Necrosis Disease. *Genome Announc* **3**, (2015).

76. N. Kimata, T. Nishino, S. Suzuki, K. Kogure, Pseudomonas aeruginosa isolated from marine environments in Tokyo Bay. *Microb Ecol* **47**, 41-47 (2004).

77. S. F. Altschul *et al.*, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research* **25**, 3389-3402 (1997).

78. D. R. Flower, A. C. North, C. E. Sansom, The lipocalin protein family: structural and sequence overview. *Biochim Biophys Acta* **1482**, 9-24 (2000).

79. L. Holm, J. Park, DaliLite workbench for protein structure comparison. *Bioinformatics* **16**, 566-567 (2000).

80. J. Felsenstein, An alternating least squares approach to inferring phylogenies from pairwise distances. *Systematic biology* **46**, 101-111 (1997).

81. W. Minor, M. Cymborowski, Z. Otwinowski, M. Chruszcz, HKL-3000: the integration of data reduction and structure solution--from diffraction images to an

initial model in minutes. *Acta crystallographica. Section D, Biological crystallography* **62**, 859-866 (2006).

82. T. R. Schneider, G. M. Sheldrick, Substructure solution with SHELXD. *Acta crystallographica. Section D, Biological crystallography* **58**, 1772-1779 (2002).

83. Z. Otwinowski, in *Proceedings of the CCP4 Study Weekend,* W. Wolf, P. R. Evans, A. G. W. Leslie, Eds. (Science & Engineering Research Council, Cambridge, U.K., 1991), pp. 80-86.

84. K. Cowtan, Recent developments in classical density modification. *Acta crystallographica. Section D, Biological crystallography* **66**, 470-478 (2010).

85. K. Cowtan, The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta crystallographica. Section D, Biological crystallography* **62**, 1002-1011 (2006).

86. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography* **66**, 486-501 (2010).

87. P. V. Afonine *et al.*, Joint X-ray and neutron refinement with phenix.refine. *Acta crystallographica. Section D, Biological crystallography* **66**, 1153-1163 (2010).

88. V. B. Chen *et al.*, MolProbity: all-atom structure validation for macromolecular crystallography. *Acta crystallographica. Section D, Biological crystallography* **66**, 12-21 (2010).

89. A. J. McCoy *et al.*, Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674 (2007).

90. S. Keller *et al.*, High-precision isothermal titration calorimetry with automated peak-shape analysis. *Anal Chem* **84**, 5066-5073 (2012).

91. T. H. Scheuermann, C. A. Brautigam, High-precision, automated integration of multiple isothermal titration calorimetric thermograms: new features of NITPIC. *Methods* **76**, 87-98 (2015).

92. C. A. Brautigam, H. Zhao, C. Vargas, S. Keller, P. Schuck, Integration and global analysis of isotherm titration calorimetry data for studying macromolecular interactions. *Nature Protocolos* **In Press**, (2016).

93. C. A. Brautigam, Calculations and Publication-Quality Illustrations for Analytical Ultracentrifugation Data. *Methods Enzymol* **562**, 109-133 (2015).

94. M. van Geest, J. S. Lolkema, Membrane topology and insertion of membrane proteins: search for topogenic signals. *Microbiol Mol Biol Rev* **64**, 13-33 (2000).

95. M. Liu, in *E. coli Plasmid Vectors: Methods and Applications,* N. Casali, A. Preston, Eds. (Humana Press, Totawa, New Jersey, 2003), pp. 290 - 294.

96. E. Krissinel, K. Henrick, Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774-797 (2007).

97. M. Kurz *et al.*, Insights into the bile acid transportation system: the human ileal lipid-binding protein-cholyltaurine complex and its comparison with homologous structures. *Proteins* **50**, 312-328 (2003).

98. C. Lucke, F. Zhang, J. A. Hamilton, J. C. Sacchettini, H. Ruterjans, Solution structure of ileal lipid binding protein in complex with glycocholate. *Eur J Biochem* **267**, 2929-2938 (2000).

99. J. Storch, L. McDermott, Structural and functional analysis of fatty acid-binding proteins. *J Lipid Res* **50 Suppl**, S126-131 (2009).

100. A. Padavannil *et al.*, Structure of GrlR-GrlA complex that prevents GrlA activation of virulence genes. *Nat Commun* **4**, 2546 (2013).

101. C. Jobichen *et al.*, Identification and characterization of the lipid-binding property of GrlR, a locus of enterocyte effacement regulator. *Biochem J* **420**, 191-199 (2009).

102. V. L. Miller, V. J. DiRita, J. J. Mekalanos, Identification of toxS, a regulatory gene whose product enhances toxR-mediated activation of the cholera toxin promoter. *J Bacteriol* **171**, 1288-1293 (1989).

103. C. C. Hase, J. J. Mekalanos, TcpP protein is a positive regulator of virulence gene expression in Vibrio cholerae. *Proc Natl Acad Sci U S A* **95**, 730-734 (1998).

104. V. L. Miller, R. K. Taylor, J. J. Mekalanos, Cholera toxin transcriptional activator toxR is a transmembrane DNA binding protein. *Cell* **48**, 271-279 (1987).

105. K. M. Ottemann, J. J. Mekalanos, The ToxR protein of Vibrio cholerae forms homodimers and heterodimers. *J Bacteriol* **178**, 156-162 (1996).

106. M. Yang *et al.*, Bile salt-induced intermolecular disulfide bond formation activates Vibrio cholerae virulence. *Proc Natl Acad Sci U S A* **110**, 2348-2353 (2013).

107. F. M. Goni, A. Alonso, Sphingomyelinases: enzymology and membrane activity. *FEBS Lett* **531**, 38-46 (2002).

108. M. Oda *et al.*, Role of sphingomyelinase in infectious diseases caused by Bacillus cereus. *PLoS One* **7**, e38054 (2012).

109. H. Kondo *et al.*, Draft Genome Sequences of Six Strains of Vibrio parahaemolyticus Isolated from Early Mortality Syndrome/Acute Hepatopancreatic Necrosis Disease Shrimp in Thailand. *Genome Announc* **2**, (2014).

110. S. Gomez-Jimenez *et al.*, High-Quality Draft Genomes of Two Vibrio parahaemolyticus Strains Aid in Understanding Acute Hepatopancreatic Necrosis Disease of Cultured Shrimps in Mexico. *Genome Announc* **2**, (2014).

111. Y. T. Yang *et al.*, Draft Genome Sequences of Four Strains of Vibrio parahaemolyticus, Three of Which Cause Early Mortality Syndrome/Acute Hepatopancreatic Necrosis Disease in Shrimp in China and Thailand. *Genome Announc* **2**, (2014).

112. J. E. Han, K. F. Tang, L. H. Tran, D. V. Lightner, Photorhabdus insect-related (Pir) toxin-like genes in a plasmid of Vibrio parahaemolyticus, the causative agent of acute hepatopancreatic necrosis disease (AHPND) of shrimp. *Dis Aquat Organ* **113**, 33-40 (2015).

113. R. V. Jensen *et al.*, Complete Genome Sequence of Prepandemic Vibrio parahaemolyticus BB22OP. *Genome Announc* **1**, (2013).

114. D. Salomon *et al.*, Type VI Secretion System Toxins Horizontally Shared between Marine Bacteria. *PLoS Pathog* **11**, e1005128 (2015).

115. E. E. Bernardy, M. A. Turnsek, S. K. Wilson, C. L. Tarr, B. K. Hammer, Diversity of Clinical and Environmental Isolates of Vibrio cholerae in Natural Transformation and Contact-Dependent Bacterial Killing Indicative of Type VI Secretion System Activity. *Appl Environ Microbiol* **82**, 2833-2842 (2016).

116. T. P. Hubbard *et al.*, Genetic analysis of Vibrio parahaemolyticus intestinal colonization. *Proc Natl Acad Sci U S A* **113**, 6283-6288 (2016).