Sorting out the details: the roles of a novel GlyGly-CTERM domain, rhombosortase, and the type II secretion system in surface localization of a *Vibrio cholerae* serine protease, VesB

by

Shilpa Gadwal

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Microbiology and Immunology)
in the University of Michigan
2014

Doctoral Committee:

Associate Professor Maria B. Sandkvist, Chair
Professor Vernon B. Carruthers
Professor Robert S. Fuller
Associate Professor Eric S. Krukonis, University of Detroit Mercy
Assistant Professor Eric C. Martens

This dissertation is dedicated to my parents for always supporting and teaching me to do everything in life, including this, with integrity and hard work.

Acknowledgments

The work described in this dissertation would not have been possible without the help from numerous individuals.

First and foremost, I would like to acknowledge my mentor, Maria Sandkvist. Thank you Maria for taking me into the lab and mentoring me. I am grateful for all the time spent talking about data and ideas that moved my project forward. Thank you for all of your assistance with manuscripts, abstracts and presentations. The technical and translational skills I have learned while in this lab will help me with my future endeavors. Most importantly, I learned how to be a good scientist from you and how to stay positive and persevere when research got tough.

I would like to thank my committee members, Vern Carruthers, Bob Fuller, Eric Krukonis, and Eric Martens. The discussions we had helped shape my dissertation work and after every committee meeting I felt more motivated to get back into the lab and do experiments. I would like to acknowledge our collaborators, Konstantin Korotkov and Wim Hol for contributing the VesB structure for the Structure and Function paper. At the University of Michigan, Dr. Henriette Remmer at the Proteomics and Peptide Synthesis Core has given advice on mass spectrometry, proteomics, and N-terminal sequencing.

I would also like to acknowledge the funding sources that supported this work. These include the R01AI049294 and R01AI081705 awarded to Maria Sandkvist, the Molecular Mechanism of Microbial Pathogenesis Training Grant (AI007528), and the Rackham Graduate Student Research Grant. I am grateful to many people in the Microbiology and Immunology Department. I would like to mention Victor DiRita for his support and advice on my long-term career plans. The office staff has taken care of many tasks to make graduate school easier. I would especially like to thank Heidi Thompson for her encouragement and support. Tiffany Porties from the PIBS office has helped me with the transition from undergraduate to graduate school and for personal support during my time here. Paula Wishart from the Graduate Student Success office has helped with my career development.

I had the pleasure of working with a great group of people from the Sandkvist lab and for that I am grateful. First, I would like to thank Aleksandra Sikora for mentoring and teaching me so many technical skills during my rotation in the lab. More importantly, this project would not be here if it wasn't for your scientific inquiry on identifying other type II secreted proteins. Thank you for making the *ves* knockouts. I would like to thank Brian Janes for making the VesB truncations. Tanya Johnson, thank you for all of your invaluable discussions on rhombosortase and the technical help. It was always nice talking to you to get your insight and bounce ideas. Thank you for also making things run smoothly in the lab. Next, I would like to thank Ursula Waack and Racquel Domingo-Gonzalez for assisting me during their rotations. The preliminary experiments you

did, helped pave the way for more experiments. To the rest of the Sandkvist lab, both past and present members, Chelsea Rule, Aurelia Syngkon, Derek Samarian, Alana Pinksy, Marcy Patrick, Miranda Gray, and Nimesh Patel, all of the support, laughs, conversations, eating lunch and desserts lifted my spirits when research was challenging especially during the “1981 era”.

While here I had the great opportunity to meet some wonderful people that not only became good friends but my “Michigan family”. Thank you Natalie, Racquel, Randi, and Sarah for your support during my time here, it was invaluable. My friends from undergrad, Amber, Shayla, Lauren, Steve, and Brandon, are always there for laughs, good times, and encouragement. Three undergraduate programs, Meyerhoff, MARC U*STAR, and McNair have been instrumental in helping me reach graduate school and supporting me through undergraduate.

Finally, I am grateful to my family, my dad, mom, brother, and sister. My parents are the true embodiment of what great parents are. The sacrifice they made by leaving India and moving to the United States so their kids could attend school has opened up so many opportunities for us. They have instilled their values of being independent, hard working and honesty in me. With these qualities I have been able to accomplish many things, including my dissertation work. I am forever grateful for everything you have done for me. My brother has been instrumental in mentoring me when I was young. He paved the way for my sister and I in attending college. Even now, he keeps me on my toes and checks to make sure that things are going well. More importantly, his sacrifice and caring

nature for our family cannot be measured. My sister has been the person who has kept me sane while in graduate school. She is always keeping me up to date on the latest television shows and books and her visits to Ann Arbor were so much fun. My extended family has been great in supporting my family and me. For anyone that I may have forgotten, thank you.

Table of Contents

Dedication ii

Acknowledgments iii

List of Figures ix

List of Tables xi

Abstract xii

Chapter

1. Introduction
    - Secreted Proteins and the Life Cycle of *Vibrio cholerae* 1
    - Type II Secretion System 4
    - Serine Proteases 9
    - *Vibrio* extracellular serine protease B 12
    - Rhomboid Proteases 14
    - Rhombosortase and the GlyGly-CTERM domain 17
    - Protein Sorting Mechanism 19
    - Significance and Scope of this Study 22

2. Functional and Structural Characterization of *Vibrio cholerae* extracellular serine protease B, VesB
    - Abstract 25
    - Introduction 26
    - Results 30
    - Discussion 49
    - Materials and Methods 54

3. A novel protein sorting mechanism for cell surface localization in Gram-negative bacteria
    - Abstract 61
    - Introduction 62
    - Results 66
    - Discussion 84
    - Materials and Methods 91

4. Discussion
Function and Structure of VesB 98
The GlyGly-CTERM domain, rhombosortase,
and the T2S system 100
Working Model 104
Future Directions 106

References 113

List of Figures

1.1: The life cycle of *V. cholerae* and the necessary secreted proteins 3

1.2: Model for the T2S system and the cholera toxin (CT) secretion 6

1.3: Schematic diagram of VesB 13

1.4: Crystal structure of GlpG from *E. coli* and the alignment of rhomboid protease (GlpG) and rhombosortase (RssP) from *V. cholerae* 16

1.5: A schematic diagram of the LPXTG/sortase system 20

2.1: Schematic diagrams of VesA, VesB, and VesC 29

2.2: VesB purification by benzamidine-sepharose affinity chromatography 32

2.3: Protease activity of purified VesB 35

2.4: VesB degrades intact proteins 36

2.5: Protease activity and secretion of VesB-S221A and VesB-R32E 38

2.6: Crystal structure of pro-VesB 40

2.7: Similarities between pro-VesB, trypsinogen, thrombin and *Streptomyces griseus* trypsin 42

2.8: Close-up views of the activation site and the Tyr250 region of pro-VesB 45

2.9: Structural homologs of the C-terminal domain of VesB 48

3.1: Alignment of GlyGly-CTERM containing proteins in *V. cholerae* 63

3.2: Secretion profile of VesB in *V. cholerae* 67

3.3: VesB is surface-associated in WT but not in the T2S mutant 69

3.4: Cleavage of VesB by rhombosortase 71

3.5: Native VesB is slightly larger, secreted, and inactive in the rhombosortase mutant 74

3.6: The propeptide is present in VesB and VesC in the rhombosortase mutant..... 77

3.7: VesB is not retained on the cell surface when produced without the GlyGly-CTERM domain 81

3.8: VesB is localized to the inner membrane prior to its engagement with the T2S system..... 83

3.9: Model 86

4.1: Model for VesB biogenesis, translocation, and function 105

4.2: The molecular structure of phospholipids in Gram-negative bacteria 108

4.3: The structure of lipid A of lipopolysaccharides (LPS) from El Tor strains of *V. cholerae*..... 108

List of Tables

2.1. Proteolytic activity of VesB 34

2.2. Data collection and refinement statistics 39

2.3. Structural homologs of the C-terminal domain of VesB 47

3.1: Four of the six GlyGly-CTERM domain containing substrates are present in lower amounts in the supernatant of the rhombosortase mutant 79

3.2: List of primers and plasmids used 92

Abstract

Some proteins carry recognition motifs that allow them to be sorted into different compartments of the cell. In Gram-positive bacteria, proteins with the LPXTG sorting motif are localized to the cell surface where they are attached to peptidoglycan by a membrane bound cysteine protease called sortase. The work presented here entails a novel Gram-negative sorting mechanism that includes the rhombosortase protease and the C-terminal recognition motif, GlyGly-CTERM. GlyGly-CTERM domains contain serines and glycines, followed by a stretch of hydrophobic and positively charged residues. The composition of this domain makes it an ideal candidate to span the membrane and serve as a possible substrate for rhombosortase, a rhomboid-like protease. Rhomboid proteases are intramembrane serine proteases that cleave substrates with a single transmembrane domain that contains helix-destabilizing residues. With this information, we postulated that rhombosortase cleaves and promotes cell surface association of GlyGly-CTERM containing proteins similar to sortase. In order to test this hypothesis, the GlyGly-CTERM containing protein *Vibrio* extracellular serine protease B (VesB) was utilized. VesB contains an N-terminal signal peptide, a protease domain that includes the propeptide and the catalytic triad, and a C-terminal domain with the GlyGly-CTERM extension. VesB is a trypsin-like serine protease that has an apparent preference for arginine and cleaves both peptides and intact proteins.

Using activity assays, western blotting, proteomic analysis and fluorescent microscopy, these studies showed that the type II secretion (T2S) system is responsible for outer membrane translocation of VesB, while rhombosortase promotes its surface retention by cleaving the GlyGly-CTERM domain and perhaps further modifying the newly generated C-terminus of VesB. It was demonstrated that native VesB in a rhombosortase mutant and VesB produced without its GlyGly-CTERM domain were inactive and no longer cell-associated. This latter finding suggests that removal of the N-terminal propeptide and thereby activation of VesB occurs on the cell surface. Taken together, rhombosortase cleaves and promotes cell surface retention of the GlyGly-CTERM domain containing protein, VesB, thereby highlighting a novel protein sorting mechanism in Gram-negative bacteria.

CHAPTER 1:

Introduction

Secreted Proteins and the Life Cycle of *Vibrio cholerae:*

The life-threatening diarrheal disease cholera is caused by the Gram-negative curved bacillus *Vibrio cholerae* and is found in two different biological niches, as an environmental reservoir in brackish waters and as a human pathogen (Nelson et al, 2009; Sack et al, 2004). The persistence of an environmental reservoir is extremely important for the transmission of cholera. In areas where *V. cholerae* is endemic, such as South Asia, outbreaks occur seasonally due to environmental changes and an increase in the *V. cholerae* population (Huq et al, 1984). Isolated outbreaks can occur when clean water, sanitation and health infrastructures are failing, like in Zimbabwe in 2008, where cholera was responsible for over 4,000 deaths (http://www.redcross.org.au/files/Zimbabwe-Report.pdf). In addition, through human travel cholera can be carried anywhere in the world within hours, and if brought to areas that have experienced natural disasters, like the earthquake in Haiti in 2010, cholera can cause massive outbreaks (Grad & Waldor, 2013).

In aquatic environments, *V. cholerae* is associated with phytoplankton, such as algae; zooplankton, including copepods or chironomid egg masses; and

are either free-living or exist as biofilms on chitinous particles (Vezzulli et al, 2008). Environmental factors such as temperature, salinity and concentrations of organic matter in water supports plankton blooms (Huq et al, 1984). Plankton are not only believed to protect *V. cholerae* from environmental stresses, but also serves as a nutrient source for growth, as attachment to live copepods aids in multiplication of *V. cholerae* (Huq et al, 1984). Studies have shown that plankton blooms correlate with an increase in *V. cholerae* population, which can provide enough of an infectious dose to result in clinical cholera if ingested by a human host (Colwell, 1996; Huq et al, 1984).

In both the environment and in the host, the secretion and function of multiple proteins from *V. cholerae* are necessary for its survival and pathogenesis (Nelson et al, 2009). Some of these proteins play roles in both environments (Figure 1.1). Chitin binding protein (GbpA) is involved in the attachment of *V. cholerae* to zooplankton (Kirn et al, 2005). Hemagglutinin protease (HapA) has been shown to degrade the gelatinous matrix of chironomid egg masses, possibly providing a nutrient source to *V. cholerae* (Halpern et al, 2003). Chitinase (ChiA) is responsible for degradation of chitinous material for *V. cholerae* to use as a carbon source (Connell et al, 1998). RbmC and Xds maintain the structure of biofilms in laboratory settings, a finding that can likely be extrapolated to biofilms in the environment (Fong & Yildiz, 2007; Seper et al, 2011).

Once *V. cholerae* enters the human host through consumption of contaminated food or water, it travels to the upper small intestine to colonize epithelial cells (Sack et al, 2004). Again, several secreted proteins help with the progression of disease and continuation of the life cycle. GbpA binds to mucin on intestinal epithelial cells, which enhances attachment of *V. cholerae* (Bhowmick et al, 2008; Kirn et al, 2005). *In vitro* studies suggest that neuraminidase increases the uptake of cholera toxin (CT), a major virulence factor and member of the AB_5 toxin family, by cleaving complex gangliosides into GM_1, the receptor for CT (Griffiths et al, 1986; Haksar et al, 1974; Miller-Podraza et al, 1982).

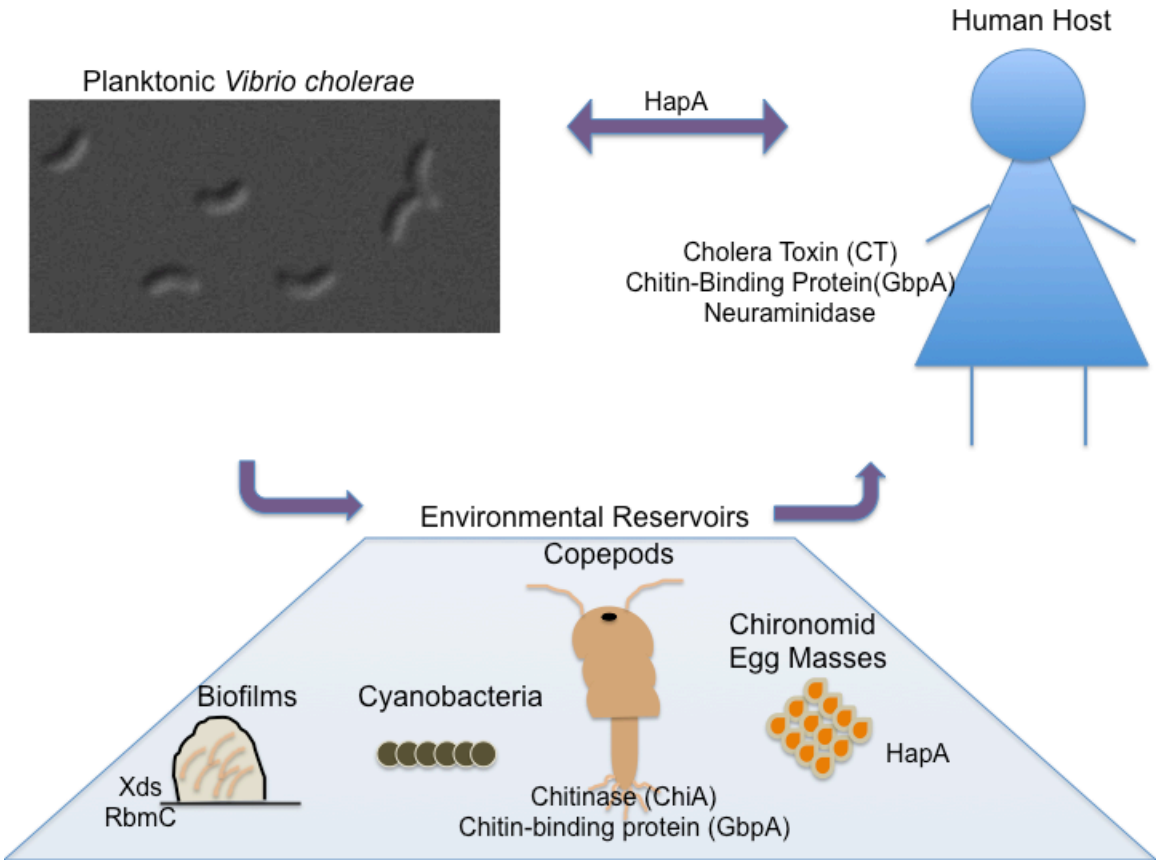

Figure 1.1: The life cycle of *V. cholerae* and the necessary secreted proteins.
This diagram depicts the life cycle of *V. cholerae*, which shuttles between the environment and human host. The known secreted substrates are labeled on the figure based on where they play a role in their respective reservoirs.

The B subunit of CT is able to bind to the exposed GM_1 receptor and subsequently induce endocytosis of the A subunit. Before endocytosis, the A subunit is cleaved or "nicked" at Arg192 by either host intestinal proteases or *Vibrio* proteases, like HapA or *Vibrio* extracellular serine protease A and B (VesA and VesB) (Booth et al, 1984; Lencer et al, 1997; Sikora et al, 2011). The activated A subunit catalyzes ADP ribosylation of the G-protein complex leading to activation of adenylate cyclase and increased levels of cyclic-AMP. This causes rapid efflux of electrolytes and water into the intestinal lumen, which results in watery diarrhea (Nelson et al, 2009; Spangler, 1992). If water and electrolytes are not replenished in patients, they can die within hours (Sack et al, 2004). To complete the life cycle, HapA is suggested to play a role in dissemination by aiding in the detachment of *V. cholerae* from intestinal epithelial cells (Finkelstein et al, 1992). In these cases, patients can shed up to 10^{12} hyperinfectious vibrios per liter and this can be subsequently transmitted to the environment and other human hosts (Nelson et al, 2009).

Type II Secretion (T2S) System:

Several different secretion systems exist in Gram-negative bacteria to translocate proteins, however, all of the secreted factors mentioned above utilize

the T2S system to gain access to the extracellular milieu (Connell et al, 1998; Finkelstein et al, 1992; Gerlach & Hensel, 2007; Halpern et al, 2003; Kirn et al, 2005; Overbye et al, 1993; Sandkvist et al, 1993; Sikora et al, 2011). In *V. cholerae*, the T2S (*eps*) genes are encoded in a single operon which translates into 13 proteins that span the outer membrane, inner membrane, and the periplasm of Gram-negative bacteria (Korotkov et al, 2012). The T2S system can be broken down into four major components; the outer membrane porin (EpsD), the inner membrane platform (EpsC, EpsF, EpsL and EpsM), the ATPase (EpsE), and the pseudopilus (EpsG, EpsH, EpsI, EpsJ and EpsK) (Korotkov et al, 2012) (Figure 1.2). The Eps proteins are well conserved among many other Gram-negative species, including nonpathogenic and pathogenic bacteria of humans, plants and aquatic organisms (Korotkov et al, 2012).

All T2S substrates contain an N-terminal signal peptide that allows them to enter the Sec or Tat system. Cleavage of the signal peptide releases the proteins from the inner membrane to generate folded periplasmic intermediates. Each A and B subunit of CT is generated with its own N-terminal signal peptide. With the help of DsbA, a protein responsible for introducing disulfide bonds, the subunits assemble into the AB_5 toxin to then engage the T2S system for outer membrane translocation (Figure 1.2) (Sandkvist, 2001).

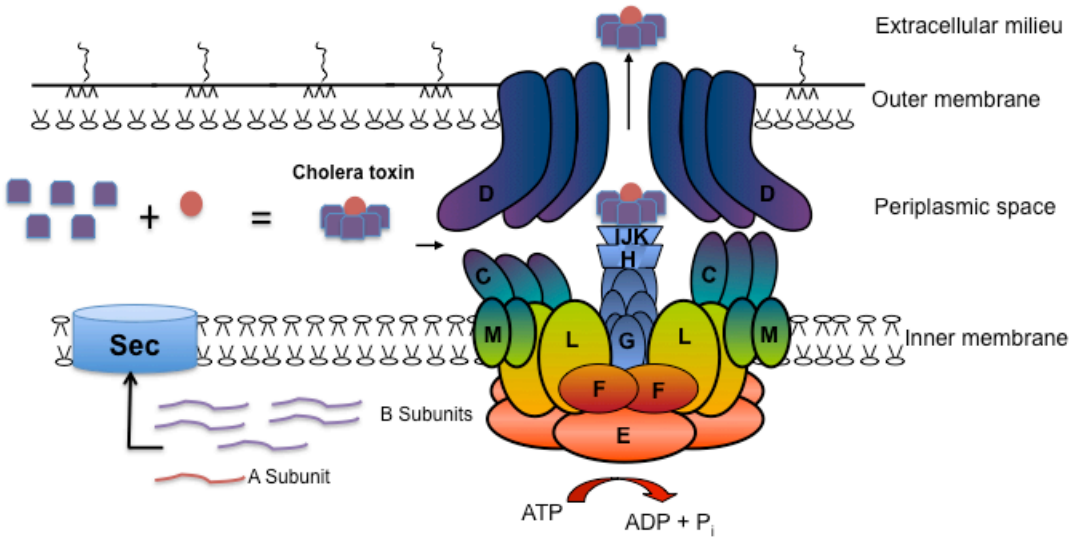

Figure 1.2: Model for the T2S system and cholera toxin (CT) secretion.
The A (red) and B subunits (purple) of CT are generated with a signal peptide that allows for inner membrane translocation via the Sec system. Once in the periplasm, the subunits fold into a multimer that can engage the secretion system to reach its final destination in the extracellular milieu. The T2S system is composed of a porin (EpsD), inner membrane platform (EpsC, F, L and M), the ATPase (EpsE), and the pseudopilus (EpsG, H, I, J and K).

Although much research has been dedicated to understanding the role of each Eps protein and how they interact with each other, the process of substrate recognition and secretion still remains to be fully elucidated. It is believed that the folded T2S substrates gain entry to the apparatus via interaction with EpsD and/or EpsC (Bouley et al, 2001). Once the substrates enter the system and make contact with the pseudopilus, a signal is sent to EpsE, the ATPase, to generate energy from ATP hydrolysis (Camberg et al, 2007; Camberg & Sandkvist, 2005). The hydrolysis of ATP changes the conformation of EpsE (Lu et al, 2013; Patrick et al, 2011), which then may transduce energy to EpsL, the protein that couples EpsG (major component of the pseudopilus) and EpsE (Gray

et al, 2011). This results in extension of the pseudopilus, which may act as a piston to push the T2S substrates through the outer membrane porin to the extracellular milieu (Korotkov et al, 2012).

While many T2S substrates diffuse away from the cell following outer membrane translocation, there are a few that are known to be surface-attached through various mechanisms. Several lipoproteins like pullulanase (PulA), SslE and OmcA from *Klebsiella oxytoca,* enteropathogenic *Escherichia coli* (EPEC) and *Shewenella oneidensis*, respectively, are found to be anchored to the cell surface (Rondelet & Condemine, 2013). Specifically, PulA was shown to be cell surface anchored through an N-terminal fatty acid moiety and is slowly released into the culture supernatant over time. In T2S mutants, PulA is found in the inner membrane due to the presence of aspartic acid at position 2, a residue responsible for keeping lipoproteins in the inner membrane (Pugsley et al, 1990). Mutants of PulA that localized to the periplasmic leaflet of the outer membrane prior to T2S are less efficiently localized to the cell surface than the wild type (WT) PulA, suggesting that the inner membrane intermediate of PulA is needed for efficient translocation to the cell surface by the T2S system (Pugsley & Kornacker, 1991). SslE and OmcA are believed to behave similarly to PulA because of their shared lipoprotein-specific signal peptide. Another mechanism of cell surface retention is seen with heat labile enterotoxin (LT) that engages the T2S system through a periplasmic intermediate similar to CT. Once translocated through the outer membrane it binds to lipopolysaccharides (LPS) on the cell surface (Horstman et al, 2004). Another example, PnlH, a cell-wall degrading

enzyme of *Dickeya dadantii*, is anchored to the cell surface through its noncleavable Tat signal peptide. In a T2S mutant, PnlH is found on the periplasmic side of the outer membrane (Ferrandez & Condemine, 2008). Whether T2S substrates are secreted away from the cell or remain on the surface, the recognition motifs of these substrates by the T2S system are unknown.

T2S proteins may need to fold before engaging the secretion apparatus, suggesting that a structural motif is recognized (Hirst & Holmgren, 1987a). Moreover, the crystal structures of T2S substrates reveal that they are rich in β-sheets, pointing to a possible recognition motif (Korotkov et al, 2012; Sandkvist, 2001). In fact, studies on CT have shown that when each subunit is expressed and made separately, the β-rich B subunit of CT is secreted through the T2S system, whereas the A subunit remains in the periplasm. From this it was concluded that the B subunit contains the recognition motif for T2S and the A subunit comes along with the B subunit during secretion (Hirst et al, 1984). Other studies in various organisms have been done to identify a secretion motif by observing translocation of fusion proteins, where parts of a T2S substrate are fused to a reporter protein. Although some regions have been identified for secretion, they have been far too large to pinpoint specific domains or regions and no linear secretion signal has been identified. It is possible that the secretion signal may consist of different regions that come together in the folded protein (Korotkov et al, 2012).

Many T2S substrates including proteases, lipases, phosphatases and carbohydrate degrading enzymes, as well as toxins have been identified and as indicated above, play important roles in pathogenesis and/or environmental survival, highlighting the importance of the T2S system (Korotkov et al, 2012). In an effort to identify additional *V. cholerae* T2S substrates, our laboratory conducted proteomic analysis on supernatants of WT *V. cholerae* and a T2S mutant. Three related serine proteases were identified and named *Vibrio* extracellular serine protease A/B/C (VesA, VesB and VesC) (Sikora et al, 2011).

Serine Proteases:

There are over 80 families of serine proteases. The chymotrypsin family or S1, contains many serine proteases that function in digestion, blood clotting, fibrinolysis and complement activation (Rawlings & Barrett, 1994; Rawlings et al, 2012). While serine proteases that function in digestion, like trypsin and chymotrypsin, contain a single domain, serine proteases like plasmin, complement component C1r and tissue plasminogen activator contain multiple domains that are N-terminal to the protease domain (Forneris et al, 2012; Hedstrom, 2002; Rawlings & Barrett, 1994; van Zonneveld et al, 1986). These additional domains regulate serine proteases by providing substrate specificity or membrane localization.

Serine proteases are anchored to the membrane in various ways; either generated with an N- or C-terminal transmembrane extension or post-translationally modified with a glycosylphosphatidylinositol (GPI) linkage (Antalis

et al, 2011). Prostasin, a GPI-linked serine protease, is cleaved at its C-terminal transmembrane domain and linked to GPI by GPI transamidase, which results in association with lipid rafts in the plasma membrane (Chen et al, 2001). Studies show that this protein can be released from the membrane via cleavage by GPI-specific phospholipase D (Verghese et al, 2006). Membrane-anchored serine proteases still retain their protease activity without their GPI-anchor or transmembrane domains, suggesting that membrane-anchoring is important for properly localizing these proteins to a specific site but not for their activity (Antalis et al, 2011).

Another mechanism of regulating serine proteases is the presence of additional domains that provide substrate specificity. Two subtilisin-like serine proteases, cucumisin from the plant *Cucumis melo* and *Aeromonas sobris* ASP, are known to contain an immunoglobulin-fold (Ig-fold) domain that interacts with the protease domain or is positioned close to the active site, suggesting a possible mechanism of controlling substrate specificity (Kobayashi et al, 2009; Murayama et al, 2012). Additionally, serine proteases of complement activation contain multiple domains that are involved with protein interactions and domain rearrangements, both of which lead to cleavage of specific substrates (Forneris et al, 2012).

Another way by which serine proteases are regulated involves the binding of Na^+ to augment protease activity. Studies have shown that residues surrounding and including tyrosine at position 225 in thrombin are important for Na^+ ion binding and leads to enhanced activity; a regulation that is relevant to the

blood coagulation serine proteases (Dang & Di Cera, 1996; Di Cera et al, 1995). In contrast, the digestive enzyme trypsin contains a proline at the equivalent tyrosine 225 position of thrombin, which induces an orientation change of Lys224 (thrombin numbering) and therefore the inability to coordinate Na^+. (Guinto et al, 1999).

Regardless of the domain arrangement, subcellular location or manner of regulation, all serine proteases of the chymotrypsin family contain the catalytic triad residues serine, histidine and aspartic acid and are first made as a zymogen with an N-terminal propeptide. The propeptide is cleaved at the activation site, resulting in an active serine protease. Once activated, these proteins can degrade substrates or act in a cascade pathway where one activates another. Serine proteases of the chymotrypsin family are subdivided into three clans based on substrate specificity. A trypsin-like protease cleaves at an arginine or lysine residue. A chymotrypsin-like protease cleaves at a tryptophan, tyrosine or phenylalanine residue. An elastase-like protease cleaves after a small, uncharged residue (Hedstrom, 2002).

The serine proteases of the chymotrypsin subfamily A or S1A are found mostly in eukaryotes and rarely in prokaryotes. Until recently, only serine proteases from the Gram-positive *Streptomyces* genus were present in this subfamily. The trypsin-like proteins SOT and SGT from *S. omiyaenis* and *S. griseus*, respectively, contain a signal peptide and are secreted into the culture supernatant (Uesugi et al, 2008). They preferentially cleave different types of collagen despite sharing 77% sequence identity and similar structures (Read &

James, 1988; Uesugi et al, 2008; Yamane et al, 1995). Studies indicated that two residues of SOT are important for substrate binding. When Tyr71 and Arg72 were mutated, SOT lost the ability to hydrolyze collagen type I/IV. Furthermore, a variant of SGT that was genetically engineered to contain Tyr71 and Arg72 gained the ability to cleave collagen type I/IV like SOT (Uesugi et al, 2011; Uesugi et al, 2009). The structure of SOT reveals that Tyr71 is close to Trp83, which is adjacent to the catalytic residue of the active site, Asp82, suggesting that once substrate binding occurs, a conformation change could result in activation of the enzyme and subsequent cleavage of substrates (Uesugi et al, 2011). In addition to the serine proteases from the *Streptomyces* genus, the recently identified VesA, VesB and VesC from *V. cholerae* have been added to the S1A subfamily. My dissertation work focuses on VesB.

Vibrio extracellular serine protease B:

Vibrio extracellular serine protease B (VesB) contains an N-terminal signal peptide, a protease domain and a C-terminal domain with a GlyGly-CTERM extension (Figure 1.3). Sequence alignment of the protease domain of VesB and trypsinogen reveals 30% homology. The predicted catalytic triad, His-Asp-Ser, and the putative activation site of VesB are evident in these alignments. Additionally, VesB contains the equivalent tyrosine 225 of thrombin rather than proline of trypsin at this position, suggesting that Na+ ion binding may regulate activity. Preliminary analysis of culture supernatants revealed that VesB cleaves synthetic peptides at arginine, suggesting that VesB is a trypsin-like serine

protease (Sikora et al, 2011). At the very C-terminus, VesB contains a GlyGly-CTERM domain, which includes Ser-Gly-Gly-Ser, followed by a stretch of hydrophobic residues and positively charged amino acids. This domain is predicted to be a transmembrane domain.

Previous studies in our laboratory have shown that VesB is not required for infant mouse colonization, as a *vesB* mutant strain (Δ*vesB*) colonized similarly to the WT strain (Sikora et al, 2011). Contrary to our infant mouse data, additional large-scale analysis has implied a role of VesB in colonization or infection in both humans and infant rabbits. Proteomic analysis of bacteria isolated from stool samples of patients infected with *V. cholerae* identified VesB (LaRocque et al, 2008). Also, the expression of *vesB* was upregulated following cultivation of *V. cholerae* in the rabbit ileum loop model (Xu et al, 2003). Furthermore, using transposon insertion site sequencing, a *vesB* mutant was identified to have reduced colonization ability in infant rabbits following oral inoculation (Fu et al, 2013). In this same study, the T2S proteins and rhombosortase, a rhomboid like protease, were also found to be important for colonization (Fu et al, 2013).

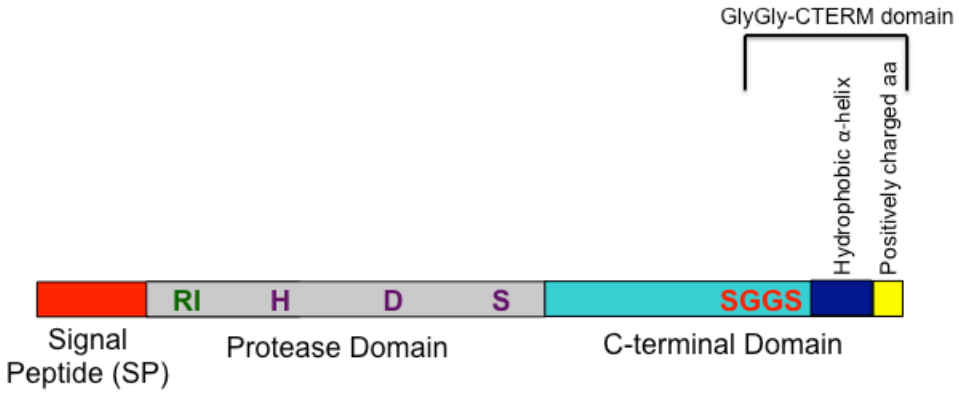

Figure 1.3: Schematic diagram of VesB.
A. VesB contains an N-terminal signal peptide in red, a protease domain in gray containing the activation site (green) and the residues of the catalytic triad (purple), and the C-terminal domain in light blue. At the very C-terminus, VesB contains the GlyGly-CTERM domain made up of two serine and glycines (red), followed by a hydrophobic alpha helix (dark blue) and positively charged residues (yellow).

Rhomboid Proteases:

Rhomboid proteases are one of the most highly conserved families of membrane proteases and found in all three domains of life (Brooks & Lemieux, 2013). They function in many processes like mitochondrial remodeling and growth factor signaling, as well as host cell invasion by apicomplexan parasites (Freeman, 2008). In bacteria, it has been shown that rhomboid protease mutants of *Bacillus subtilis* are altered in cell division and glucose uptake and *Mycobacterium smegmatis* mutants display changes in biofilm formation, colony morphology and susceptibility to antibiotics (Rather, 2013). Even though these findings suggest a possible role for rhomboid proteases in these processes, their mechanisms and substrates are unknown. TatA, the channel-forming protein of the twin-arginine translocation (TAT) system is the only bacterial rhomboid protease substrate known. In *Providencia stuartii*, the rhomboid protease AarA cleaves TatA, which results in formation of the TAT system, a complex involved in the translocation of folded proteins across the inner membrane (Fritsch et al, 2012; Stevenson et al, 2007).

Rhomboid proteases contain six transmembrane domains that fold to form a Ser-His active site and some contain either an N- or C-terminal cytoplasmic domain (Figure 1.4). Although the cytoplasmic domain of the *E. coli* rhomboid

protease, GlpG, does not contribute to the stability or activity of the membrane domain, possible functions of the cytoplasmic domain includes providing substrate specificity, cellular compartmentalization or interacting with the membrane domains of rhomboid proteases (Lazareno-Saez et al, 2013). Crystallography studies on the *E. coli* and *Haemophilus influenzae* GlpG proteins show that the active site is 10 Å below the plane of the lipid bilayer, suggesting that proteolysis of substrates occurs within the membrane (Ben-Shem et al, 2007; Lemieux et al, 2007; Wang et al, 2006; Wu et al, 2006) (Figure 1.4). In agreement with this, rhomboid proteases have been shown to cleave substrates at small residues within transmembrane domains that contain helix-destabilizing residues like prolines and glycines (Akiyama & Maegawa, 2007; Strisovsky et al, 2009).

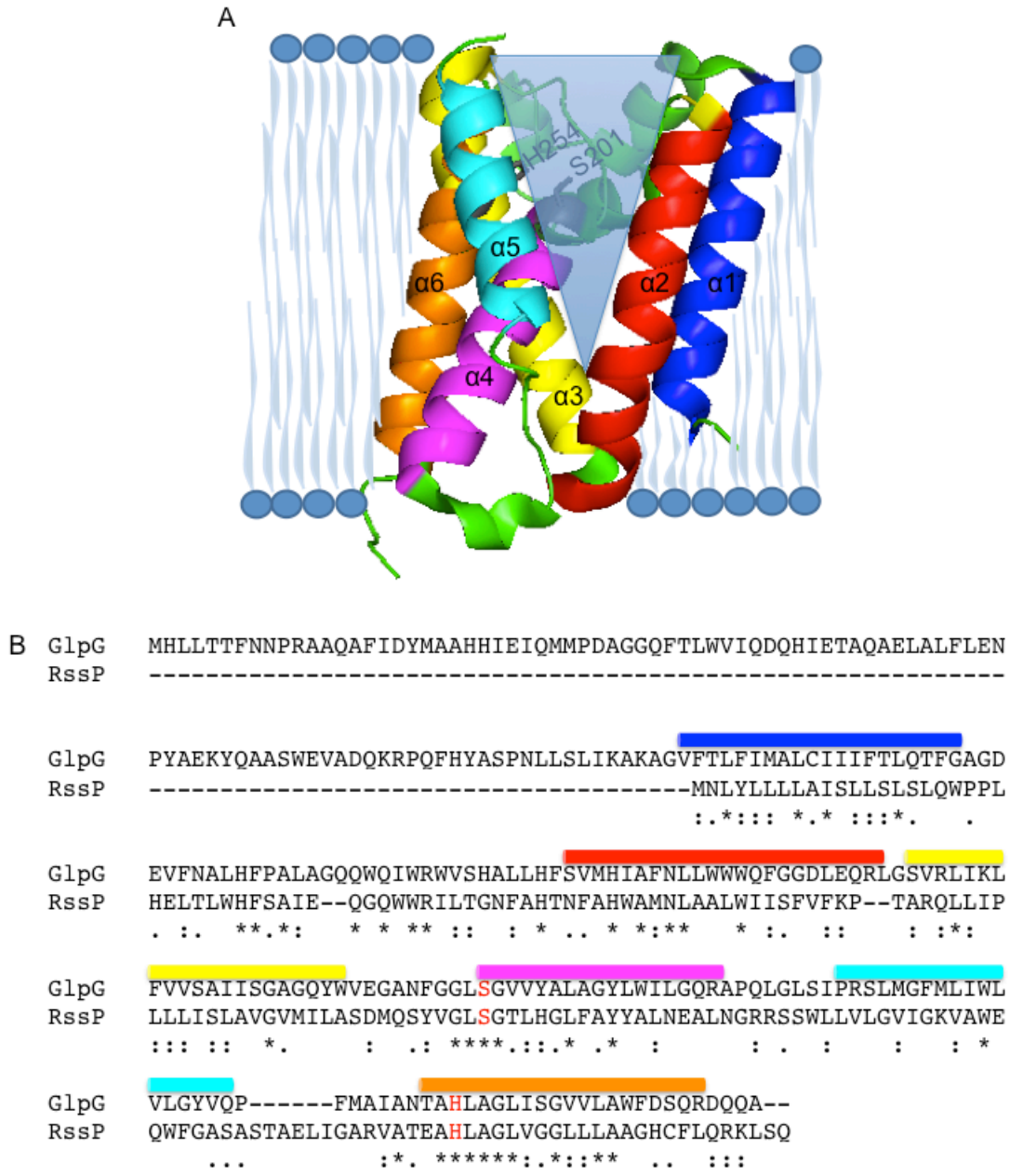

Figure 1.4: Crystal structure of GlpG from *E. coli* and the alignment of rhomboid protease (GlpG) and rhombosortase (RssP) from *V. cholerae*.
A. The crystal structure of GlpG shows six transmembrane helices that form a water channel containing the two Ser-His active site residues. Figure prepared using Pymol, PDB: 2IRV (Ben-Shem et al, 2007).
B. The sequence alignment of GlpG and RssP of *V. cholerae*. The catalytic dyad, Ser-His, is highlighted in red. The putative transmembrane domains are indicated as colored bars based on the transmembrane domains of *E. coli* GlpG.

In an effort to identify the important residues of the recognition motif, the transmembrane domains of native rhomboid protease substrates were fused to reporter proteins. In these studies, the helix-destabilizing residues were shown to be important for cleavage by rhomboid protease, as mutating those residues to more stabilizing residues did not result in cleavage. Conversely, when the transmembrane helix was mutated to contain more destabilizing residues, more substrate was cleaved by rhomboid protease (Akiyama & Maegawa, 2007; Strisovsky et al, 2009). Using the same fusion substrates with various point mutations, the specific residues cleaved by rhomboid protease were identified. The consensus was that rhomboid protease cleaves between amino acids with small side chains like alanines, serines and glycines (Akiyama & Maegawa, 2007; Strisovsky et al, 2009).

Rhombosortase and the GlyGly-CTERM domain:

Recently, a bioinformatics study using partial phylogenetic profiling showed a possible link between a subfamily of rhomboid proteases, called rhombosortases, and proteins containing a GlyGly-CTERM domain. In genomes encoding only one GlyGly-CTERM domain protein, the gene was always found next to the rhombosortase gene; however, in genomes encoding many GlyGly-CTERM domain proteins this arrangement was not maintained. The presence of the GlyGly-CTERM domain containing proteins and rhombosortase is exclusive to Gram-negative bacteria including *Vibrio*, *Shewenella, Acinetobacter, Ralstonia* and *Aeromonas* genera (Haft & Varghese, 2011). In *V. cholerae*, the genome

encodes both a rhomboid protease (GlpG) and rhombosortase protease (RssP). Protein sequence analysis reveals that rhombosortase contains the predicted catalytic dyad and the same topology as rhomboid protease; however, it is missing the N-terminal cytoplasmic domain (Figure 1.4) (Haft & Varghese, 2011). Because of the similarity to rhomboid protease, rhombosortase may also cleave substrates with a transmembrane domain. Due to their possible genomic connection and the tendency of GlyGly-CTERM domain containing proteins to have predicted transmembrane domains, proteins with the GlyGly-CTERM domain may be substrates for rhombosortase.

In the aforementioned bioinformatics study, the transmembrane domains of 219 proteins containing the GlyGly-CTERM domain were analyzed for amino acid content and arrangement. The alignment of these proteins shows the typical length of the GlyGly-CTERM domain is 22 amino acids and consists of a tripartite motif. The arrangement of this domain includes an area rich in glycines and serines, followed by hydrophobic and positively charged residues. With the exception of the glycines and serines, the GlyGly-CTERM domain is similar to the LPXTG sorting motif, which is the target of sortase, a transpeptidase that anchors LPXTG-containing proteins to the surface of Gram-positive bacteria. The similarities between these motifs and the possibility that the rhombosortase may process the GlyGly-CTERM motif were the reasons for the name of this newly identified variant of rhomboid protease (Haft & Varghese, 2011). Furthermore, the hydrophobic region of the GlyGly-CTERM domains deviate from typical transmembrane domains in that they contain more leucines than valines and

isoleucines, indicating that the transmembrane helix may be involved in specific protein-protein interactions with a membrane protein (Haft & Varghese, 2011).

The role of the GlyGly-CTERM domain is not known. Interestingly, a few GlyGly-CTERM domain containing proteins were reported to have dual localization patterns. ExeM, an endonuclease from *S. oneidensis* was identified to be membrane associated using proteomics, but also had culture supernatant activity (Brown et al, 2010; Godeke et al, 2011; Tang et al, 2007). Additionally, the *V. cholerae* homolog of ExeM, Xds was shown to have activity in both the culture supernatant and cells (Newland et al, 1985; Seper et al, 2011). Furthermore, ExeM and Xds were shown to be important for biofilm formation and stability (Godeke et al, 2011; Seper et al, 2011). Collectively, these data suggest that the GlyGly-CTERM domain and rhombosortase may have a role in the dual localization of proteins, potentially highlighting a new protein sorting mechanism in Gram-negative bacteria.

Protein Sorting Mechanisms:

A well-known protein sorting mechanism is the LPXTG/sortase system of Gram-positive bacteria. Sortase is a cysteine protease with a Cys-His active site. The sortase substrates contain an N-terminal signal peptide and a C-terminal sorting signal made up of 32 amino acids consisting of the LPXTG motif, followed by a hydrophobic domain and a positively charged tail (Figure 1.5-1). The peptide bond in between Thr and Gly in the LPXTG motif is cleaved by sortase (Figure 1.5-4). This results in an acyl intermediate with sortase that is resolved by a

nucleophilic attack by the pentaglycines of the lipid II precursor of peptidoglycan (Figure 1.5-5). As peptidoglycan synthesis progresses, the LPXTG substrates are incorporated into the cell wall envelope to generate cell surface-anchored proteins in Gram-positive bacteria (Figure 1.5-6,7) (Marraffini et al, 2006).

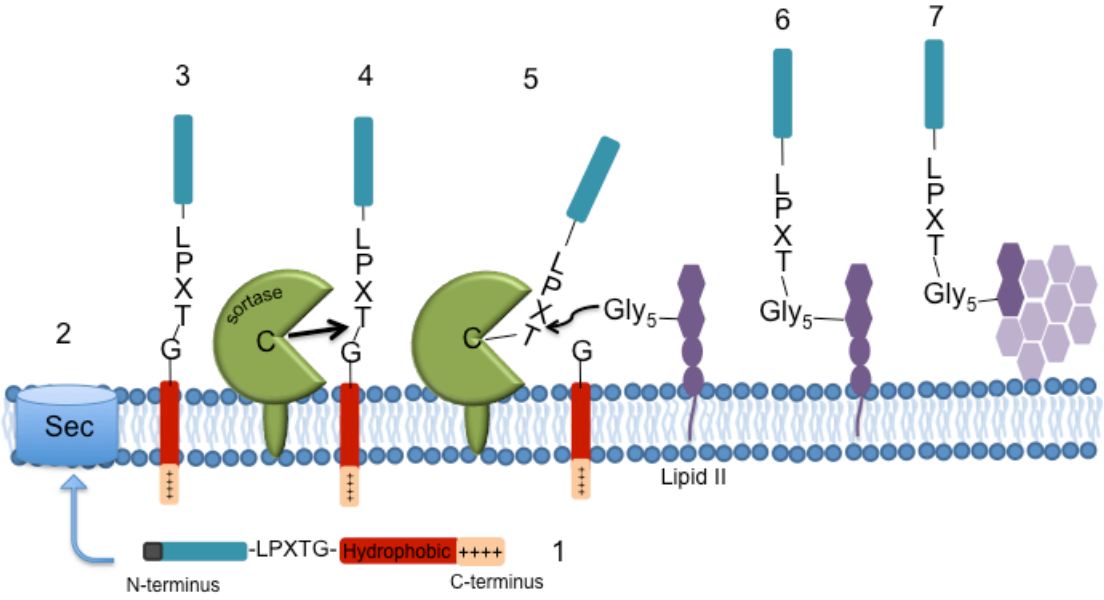

Figure 1.5: A schematic diagram of the LPXTG/sortase system.
The LPXTG substrate (1) is composed of an N-terminal signal peptide (gray) and the C-terminal LPXTG motif. The transmembrane domain (red) and the positively charged residues (+) follow the LPXTG recognition sequence. The signal peptide is cleaved and the protein enters the membrane via the hydrophobic C-terminus (2-3). Sortase cleaves in between the Thr and Gly of the LPXTG motif (4), generating an acyl intermediate (5). The pentaglycines of lipid II acts as a nucleophile and links to Thr (5). This precursor is incorporated into peptidoglycan, resulting in a cell-wall anchored protein (6-7).

Lesser-known protein sorting mechanisms are found in environmental bacteria and archaea and these are termed PGF-CTERM/archaeosortase and PEP-CTERM/exosortase, respectively (Haft et al, 2006; Haft et al, 2012). The PGF-CTERM/archaeosortase system is commonly found in archaea. Proteins

with PGF-CTERM domains have a recognition motif consisting of Pro-Gly-Phe (PGF) and a predicted hydrophobic domain followed by positively charged residues. Most substrates containing the PGF-CTERM domain are glycoproteins (Haft et al, 2012). For example, the S-layer glycoprotein from *Haloferax volcanii* was shown to retain its PGF-CTERM domain in the absence of archaeosortase, suggesting that archaeosortase cleaves the PGF-CTERM domain (Abdul Halim et al, 2013). Furthermore, S-layer glycoprotein was shown to undergo lipid modification once it transverses through the plasma membrane, however, the site of lipid modification is not known (Kandiba et al, 2013; Konrad & Eichler, 2002).

The PEP-CTERM domain is comprised of 24 amino acids and has a similar arrangement to the tripartite motif of the LPXTG, PGF-CTERM and GlyGly-CTERM domains. A bioinformatics study using partial phylogenetic profiling identified EpsH, a protein involved in extracellular polysaccharide biosynthesis, as the potential exosortase enzyme that could cleave the PEP-CTERM domain and assist in sorting. Although the sequence identity of EpsH and sortase are low, they contain the same conserved Cys-His active site residues, suggesting that these enzymes may act in a similar fashion (Haft et al, 2006). Very little is known about where the modifying enzymes cleave and the exact mechanism of cell surface anchoring of their substrates.

Significance and Scope of this Study:

Although the exact role of VesB in *V. cholerae* pathogenesis and environmental survival is unknown, it has been identified in various large-scale analyses to have possible roles in colonization and disease progression (Fu et al, 2013; LaRocque et al, 2008; Xu et al, 2003). Apart from this, the unique domain arrangement consisting of a trypsin-like protease domain and a C-terminal domain containing the GlyGly-CTERM extension has led me to study VesB in detail. The overall goal of my dissertation project is to dissect the protease domain and understand the role of the GlyGly-CTERM domain and its relationship to rhombosortase and T2S.

Until VesB was discovered, the only other bacterial trypsin-like serine proteases from the chymotrypsin subfamily (S1A) under investigation were SGT and SOT. They were shown to cleave different forms of collagen based on their residue content at particular positions and studies focused on their structure revealed a resemblance to trypsin (Read & James, 1988; Uesugi et al, 2008; Uesugi et al, 2011; Uesugi et al, 2009; Yamane et al, 1995). Even though the protease domain of VesB is predicted to have a trypsin-like fold, VesB sets itself apart from SGT and SOT in that it contains the equivalent Tyr250 motif of thrombin, indicating a possible role of Na^+ regulation. Additionally, VesB has a C-terminal domain that SGT and SOT lack.

In the second chapter, I sought to understand the catalytic triad and activation site of the protease domain, as well as determine substrate specificity and the effects of Na^+ on the activity of VesB. Furthermore, my colleague

Konstantin Korotkov, solved the structure of the two-domain VesB protein. Taken together, our combined work revealed that VesB is a trypsin-like protease with a typical trypsin-like fold and a C-terminal Ig-fold. It also demonstrated that the presence of Na^+ does not enhance VesB activity.

I next wanted to study the role of the GlyGly-CTERM domain of VesB. VesB was believed to be fully secreted away from cells, as it was identified in *V. cholerae* culture supernatants. Therefore, the presence of an unusual putative transmembrane motif, the GlyGly-CTERM, motivated us to study this domain in detail along with its connection to the T2S system. My main question dealt with the presence of the putative transmembrane domain and its possible removal for extracellular release. Shortly thereafter, Haft *et. al*. published a paper showing the genetic co-existence of proteins with the GlyGly-CTERM domain and rhombosortase, a subfamily of rhomboid proteases (Haft & Varghese, 2011). Rhomboid proteases are known to cleave substrates with a single transmembrane domain (Akiyama & Maegawa, 2007; Strisovsky et al, 2009). This led to my hypothesis that rhombosortase may cleave the GlyGly-CTERM domain of VesB, resulting in extracellular release. However, after generating antibodies against VesB, I found that most of VesB is cell-associated, the secreted VesB is the same size as the cell-associated form and both forms are active. This led me to therefore predict that rhombosortase cleaves and perhaps modifies the GlyGly-CTERM domain of VesB to keep it on the cell surface, analogous to the sortase/LPXTG system.

The third chapter details my work on the biogenesis of VesB, specifically the relationship between the GlyGly-CTERM domain, rhombosortase, and the T2S system. Using western blot analysis, activity assays and cell-staining with microscopy, I show that VesB is a T2S-dependent cell surface localized protein. Furthermore, I demonstrate that rhombosortase cleaves and possibly modifies the GlyGly-CTERM domain to retain VesB on the cell surface. Also, I show that in the absence of rhombosortase, *V. cholerae's* rhomboid protease, GlpG, can cleave the GlyGly-CTERM domain of VesB and this cleavage leads to an alternative phenotype where VesB is secreted completely from the cell, however, this form of VesB is inactive.

Together, my dissertation work contributes to a better understanding of a multi-modular bacterial serine protease and introduces a novel protein sorting mechanism in Gram-negative bacteria. This project opens an exciting avenue of research for the GlyGly-CTERM/rhombosortase system, which will be discussed in chapter 4. The information learned from these studies can be extrapolated to other GlyGly-CTERM domain containing proteins that may function in pathogenesis or survival of bacteria, and may identify ways to target the GlyGly-CTERM/rhombosortase system for therapeutic treatments.

CHAPTER 2:

Functional and Structural Characterization of *Vibrio cholerae* extracellular serine protease B, VesB

Notes:

A modified version of this chapter was previously published in the Journal of Biological Chemistry (Gadwal et al, 2014).

Konstantin Korotkov and Wim Hol contributed the structural data of VesB.

Abstract:

The chymotrypsin subfamily A of serine proteases consists primarily of eukaryotic proteases, including only a few proteases of bacterial origin. VesB, a newly identified serine protease that is secreted by the type II secretion system in *Vibrio cholerae,* belongs to this subfamily. VesB is likely produced as a zymogen since sequence alignment with trypsinogen identified a putative cleavage site for activation and a catalytic triad, His-Asp-Ser. Using synthetic peptides, VesB efficiently cleaved a trypsin substrate, but not chymotrypsin and elastase substrates. VesB was also capable of cleaving intact proteins such as BSA, casein, IgA and lactoferrin, suggesting that it may be a promiscuous protease with broad specificity. The reversible serine protease inhibitor, benzamidine,

inhibited VesB and served as an immobilized ligand for VesB affinity purification, further indicating its relationship with trypsin-like enzymes. Consistent with this family of serine proteases, N-terminal sequencing implied that the propeptide is removed in the secreted form of VesB. Separate mutagenesis of the activation site and catalytic serine rendered VesB inactive, confirming the importance of these features for activity, but not for secretion. Similar to trypsin but, in contrast to thrombin and other coagulation factors, Na^+ did not stimulate the activity of VesB, despite containing the Tyr250 signature. The crystal structure of catalytically inactive pro-VesB revealed that the protease domain is structurally similar to trypsinogen. The C-terminal domain of VesB was found to adopt an immunoglobulin (Ig) fold that is structurally homologous to Ig folds of other extracellular *Vibrio* proteins. Possible roles of the Ig fold domain in stability, substrate specificity, cell surface association and type II secretion of VesB, the first bacterial multi-domain trypsin-like protease with known structure, are discussed.

Introduction:

Many bacteria require secretion of a variety of proteins for optimal adaptation to the diverse environments they encounter. One of the several bacterial protein secretion machineries is the type II secretion (T2S) system. This large two-membrane spanning assembly is remarkable in that it translocates multidomain and sometimes multimeric proteins in folded form across the outer membrane of Gram-negative bacteria (Douzi et al, 2012; Korotkov et al, 2012;

McLaughlin et al, 2012). In *Vibrio cholerae*, the T2S system is responsible for the secretion of cholera toxin (CT), the major virulence factor, and a large number of enzymes, including proteases into the extracellular milieu (Sandkvist et al, 1993; Sikora et al, 2011)

Proteases are naturally occurring enzymes that are found in all domains of life. One of the largest classes of proteases, serine proteases, contains over eighty families. The S1, or chymotrypsin family consists of structurally homologous endoproteases with conserved His, Asp, and Ser residues that form the catalytic triad of the active site. Most members of this family are secreted as inactive precursors, or zymogens, with an N-terminal propeptide. The propeptide is proteolytically cleaved at a conserved activation site, which induces a conformational change that results in an active enzyme (Hedstrom, 2002; Rawlings & Barrett, 1994).

Although structurally very similar, members of the chymotrypsin family are further subdivided into three groups based on substrate specificity. Trypsin-like proteases cleave at an arginine or lysine residue, chymotrypsin-like proteases act on tryptophan, tyrosine, or phenylalanine, and elastase-like proteases target small, uncharged residues (Hedstrom, 2002). Eukaryotic proteases of the chymotrypsin family are important for many cellular processes. Functionally, they aid in digestion, blood coagulation, complement activation, inflammation and development (Hedstrom, 2002). While the digestive enzymes, trypsin and chymotrypsin, are single domain proteases, proteins like tissue plasminogen activator, plasmin, and the complement component C1r contain multiple

regulatory domains that are N-terminal to the protease domain (Forneris et al, 2012; Rawlings & Barrett, 1994; van Zonneveld et al, 1986). Furthermore, some chymotrypsin family proteases are anchored to membranes by either glycosylphosphatidylinositol (GPI), a single-pass transmembrane domain at the C-terminus, or a signal anchor at the N-terminus (Szabo & Bugge, 2011).

The S1A, or chymotrypsin A subfamily, consists primarily of eukaryotic proteases, however, a few bacterial serine proteases belong to this subfamily according to the peptidase database MEROPS (Rawlings et al, 2013). To date, proteases that are structurally related to trypsin have been identified in *Streptomyces* species (Read & James, 1988; Yamane et al, 1995). These were shown to contain a single protease domain that has collagenolytic activity (Read & James, 1988; Uesugi et al, 2008; Yamane et al, 1995). In addition, we recently identified three related *Vibrio cholerae* extracellular serine proteases, VesA, VesB, and VesC, in a proteomic screen designed to detect proteins secreted via the T2S system (Sikora et al, 2011). Consistent with their reliance on the T2S system for extracellular secretion, when overexpressed, measurable activity of all three proteases has been detected in culture supernatants of *V. cholerae* (Sikora et al, 2011; Syngkon et al, 2010). Furthermore, these proteases have a predicted N-terminal signal peptide that allows them to enter the periplasm via the Sec pathway prior to engaging with the T2S system (Filloux, 2004). In addition, they all consist of an N-terminal protease domain containing a putative activation site, residues that comprise the catalytic triad, and a C-terminal domain that varies in length and includes a GlyGly-CTERM extension of unknown function (Figure 2.1)

(Haft & Varghese, 2011). The domain organization of VesA, VesB and VesC differs from that of most bacterial and eukaryotic trypsin-like proteases in that the additional non-protease domain is positioned at the C-terminus.

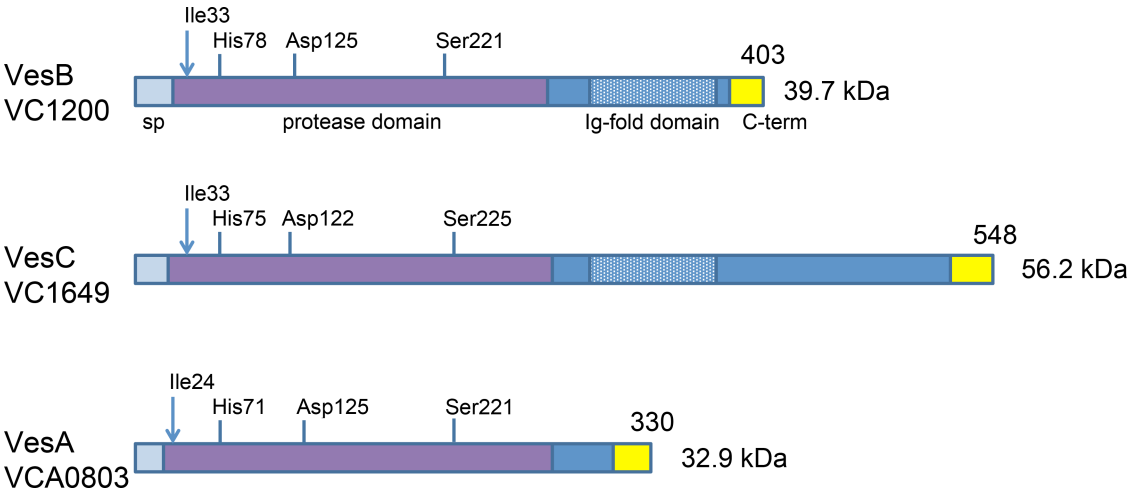

Figure 2.1: Schematic diagrams of VesA, VesB, and VesC.
All three contain an N-terminal signal peptide in light blue, an N-terminal protease domain in purple, a C-terminal domain in dark blue with the highlighted Ig-fold domain in light blue, and the GlyGly-CTERM extension in yellow. The N-terminal protease domain contains the predicted activation site indicated by the arrow and the catalytic triad comprised of His, Asp, and Ser.

VesB, the focus of this study, shares approximately 30% sequence identity with trypsin and other members of the S1A family of serine proteases and displays a similar positioning of its catalytic residues. While the biological role of VesB has yet to be determined, VesB is produced both *in vitro* and *in vivo*. Besides its detection in laboratory-grown cultures, VesB has also been detected in *V. cholerae* that was isolated from stools of patients with clinical cholera, inferring that VesB may contribute to intestinal growth or pathogenesis (LaRocque et al, 2008; Sikora, 2013; Sikora et al, 2011). Also, intestinal *vesB*

gene expression was detected in *V. cholerae* cultivated in a rabbit ileum loop, an experimental model for cholera, further suggesting a possible role in survival or disease (Xu et al, 2003). While VesB may contribute to intestinal growth of *V. cholerae*, it is not the only factor required for intestinal survival as *vesB* inactivation had no negative effect on infant mice colonization (Sikora et al, 2011). Finally, VesB is capable of cleaving the A subunit of CT, a process important for CT activation (Gordon & Leppla, 1994; Sikora et al, 2011).

The rarity of characterized bacterial serine proteases in the chymotrypsin A subfamily, the unique domain arrangement, and the possible role of VesB in intestinal growth or pathogenesis led us to study VesB in detail. Here, we focus on the activity and structure of VesB. We show that the protease domain of pro-VesB contains a trypsin-like fold with an incomplete oxyanion hole, and an additional C-terminal domain with an Ig fold containing seven β-strands and two short α-helices. Furthermore, we show that purified VesB has activity against synthetic peptides, cleaves after Arg independently of Na^+ ions, can be inhibited by serine protease inhibitors, and is produced with a propeptide that is critical for activity but not for extracellular secretion by *V. cholerae.*

Results:

Purification of secreted VesB:

To determine whether benzamidine sepharose could be utilized for affinity purification of native VesB, the reversible serine protease inhibitor benzamidine was analyzed for its ability to inhibit VesB activity. Specifically, VesB was

overexpressed from pMMB67EH-*vesB* in Δ*vesABC,* a *V. cholerae* strain lacking all three serine protease genes *vesA*, *vesB* and *vesC* (Sikora et al, 2011). Culture supernatant was isolated and incubated with increasing amounts of benzamidine, and then subjected to a kinetic protease activity assay using the synthetic peptide *Boc*-Gln-Ala-Arg-7-amido-4-methylcoumarin as a substrate. The results showed that as the concentration of benzamidine increased the activity of VesB decreased, suggesting that benzamidine is binding to VesB and that benzamidine-sepharose can be used for the purification of VesB (Figure 2.2A). To purify VesB, culture supernatant was isolated, concentrated and applied to benzamidine-sepharose chromatography. VesB was eluted with benzamidine and all of the fractions were analyzed by SDS-PAGE followed by GelCode blue and silver staining (Figure 2.2B-C). The GelCode blue and silver stained gels showed that some VesB was found in the flow-through fraction (lane 4), while there was no additional loss of VesB during washing (lanes 5-10). Bound VesB eluted efficiently with 100 mM benzamidine (lanes 11-14) and once benzamidine was removed the activity of VesB could be measured (Figure 2.3). Antibodies were then raised against the purified material and the saved fractions were analyzed by SDS-PAGE and immunoblotting with anti-VesB antiserum (Figure 2.2D).

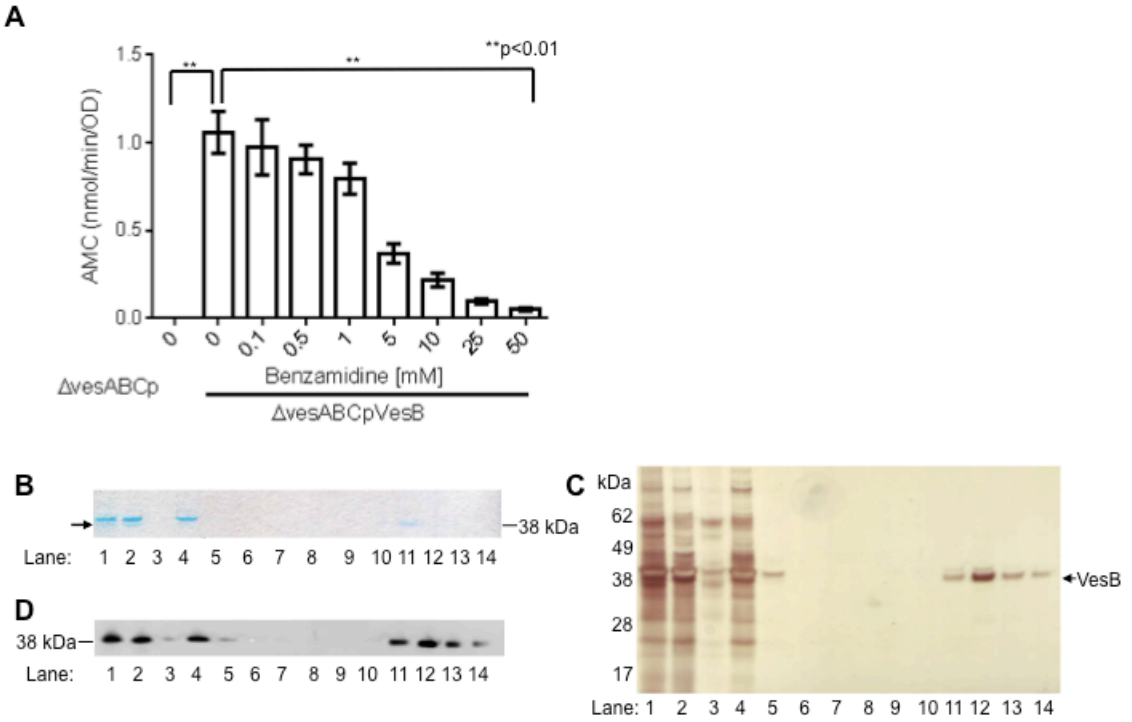

Figure 2.2: VesB purification by benzamidine-sepharose affinity chromatography.

A. The supernatants of Δ*vesABC* strains with pMMB67EH (p) or pVesB were incubated for 10 minutes at 37°C with different concentrations of the serine protease inhibitor, benzamidine, and the protease activity was measured (experiments were done in triplicates and S.E. bars are shown).

B-D. The supernatant from Δ*vesABC* overexpressing VesB (lane 1) was added to 60% saturation of ammonium sulfate and the precipitated material (Lane 2) and supernatant (Lane 3), were separated by centrifugation. The pellet was resuspended and dialyzed in 50 mM Tris-HCl, pH 8.0, 450 mM NaCl. The sample was used for affinity chromatography using a benzamidine-sepharose column. Flow-through fraction was discarded (Lane 4) and the column was washed with 50 mM Tris-HCl, pH 8.0, 450 mM NaCl (Lanes 5-10). VesB was eluted using 100 mM benzamidine (Lanes 11-14). The samples were analyzed by SDS-PAGE and GelCode blue staining (B), silver staining (C) or Western blotting with anti-VesB antibodies (D).

Characterization of purified secreted VesB:

Previously, we have shown that Boc-Gln-Ala-Arg-7-amido-4-methylcoumarin is cleaved when incubated with supernatant from *V. cholerae* overexpressing VesB and that this activity is inhibited by common serine protease inhibitors ((Sikora et al, 2011) and Figure 2.2A). To further determine the substrate specificity and verify that purified VesB is inhibited by serine protease inhibitors, we added purified VesB to different commercially available synthetic peptides that are conjugated to 7-amido-4-methylcoumarin (AMC) (Table 2.1 and Figure 2.3A) and preincubated VesB with different inhibitors (Figure 2.3B). VesB efficiently cleaved the trypsin substrate Boc-Gln-Ala-Arg-AMC, while the plasmin and aminopeptidase substrates, Boc-Glu-Lys-Lys-AMC and Leu-AMC, respectively, were cleaved with greatly reduced efficiency, and no cleavage was observed for the chymotrypsin or elastase substrates (Table 2.1). Measurement of kinetic parameters yielded Vmax and K_M values of 0.137 ± 0.0030 nmol/min and 0.0327 ± 0.0025 mM, respectively, for Boc-Gln-Ala-Arg-AMC (Figure 2.3A). We were unable to determine Vmax and K_M for Boc-Glu-Lys-Lys-AMC and Boc-Leu-AMC as the reactions never approached maximum rates due to the lack of solubility of the substrates at higher concentrations. In addition, we observed inhibition of purified VesB activity by the serine protease inhibitors benzamidine and leupeptin, while VesB was not affected by the presence of the metal chelator, EDTA (Figure 2.3B). We also determined if VesB could degrade intact proteins by incubating various concentrations of VesB with readily available proteins, including BSA, casein, lactoferrin, and IgA. The

samples were analyzed by SDS-PAGE and GelCode blue staining (Figure 2.4). The results suggest that VesB is capable of cleaving several intact protein substrates (Figure 2.4). Taken together, our results suggest that VesB is an extracellular trypsin-like protease that may have a preference for arginine at the P1 position of substrate(s) (using the commonly accepted nomenclature (Schechter & Berger, 1967)) and does not require divalent metal ions for activity. While the natural targets of VesB are unknown, VesB may be a promiscusous protease with broad specificity and multiple targets.

Table 2.1. Proteolytic activity of VesB

Peptide Sequence[a]	**VesB Activity (nmol AMC/min/nmol VesB)**[b]
Suc-Ala-Ala-Ala-AMC	0
Suc (OMe)-Ala-Ala-Pro-Val-AMC	0
Z-Leu-Arg-Gly-Gly-AMC	0
Leu-AMC	0.518 ± 0.124
Suc-Ala-Ala-Pro-Phe-AMC	0
Boc-Glu-Lys-Lys-AMC	0.449 ± 0.060
Boc-Gln-Ala-Arg-AMC	286 ± 34.0
Boc-Gln-Ala-Arg-AMC + 200 mM NaCl	122 ± 21.9 **
Boc-Gln-Ala-Arg-AMC + 400 mM NaCl	137 ± 31.3 **
Boc-Gln-Ala-Arg-AMC + 600 mM NaCl	110 ± 18.9 **
Boc-Gln-Ala-Arg-AMC + 800 mM NaCl	97.2 ± 21.6 **

[a] Peptide sequence of the substrates all conjugated to 7-Amino-4-methylcoumarin (AMC).
[b] Substrate concentration was 0.05 mM. VesB concentration was 0.079 µg/mL when Boc-Gln-Ala-Arg-AMC was analyzed and 7.9 µg/mL for all other substrates. Experiments were done in triplicates and the mean and standard error are shown.
** p values < 0.01: all compared to the sample with no NaCl

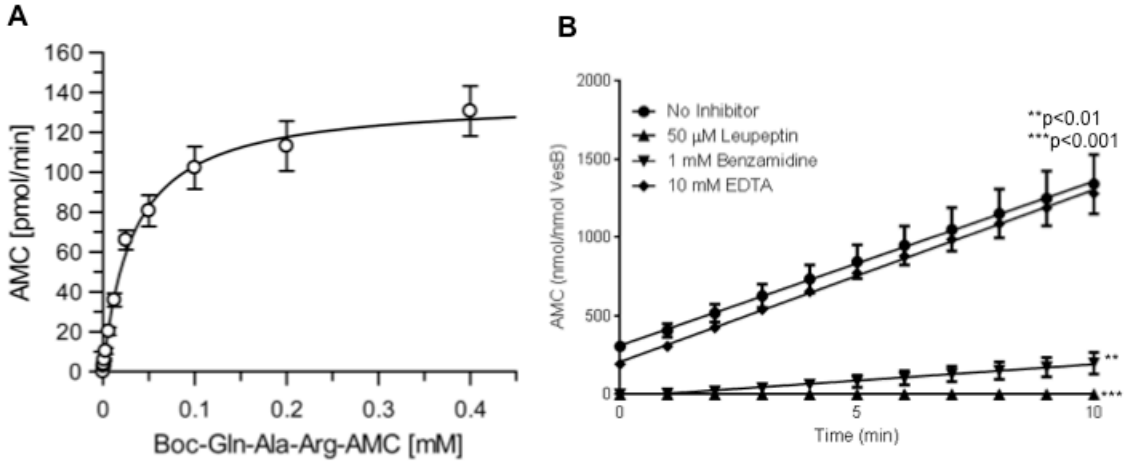

Figure 2.3: Protease activity of purified VesB.

A. VesB activity was measured with different concentrations of Boc-Gln-Ala-Arg-7-amido-4-AMC in 5 mM HEPES (pH 7.5) at 37˚C.

B. Purified VesB (0.08 μg/mL) was incubated with 50 μM leupeptin, 1 mM benzamidine, or 10 mM EDTA for 10 minutes at 37°C. The Boc-Gln-Ala-Arg-7-amido-4-AMC (0.05 mM final concentration) was added and VesB activity was measured. All experiments were done in triplicates and S.E. bars are shown.

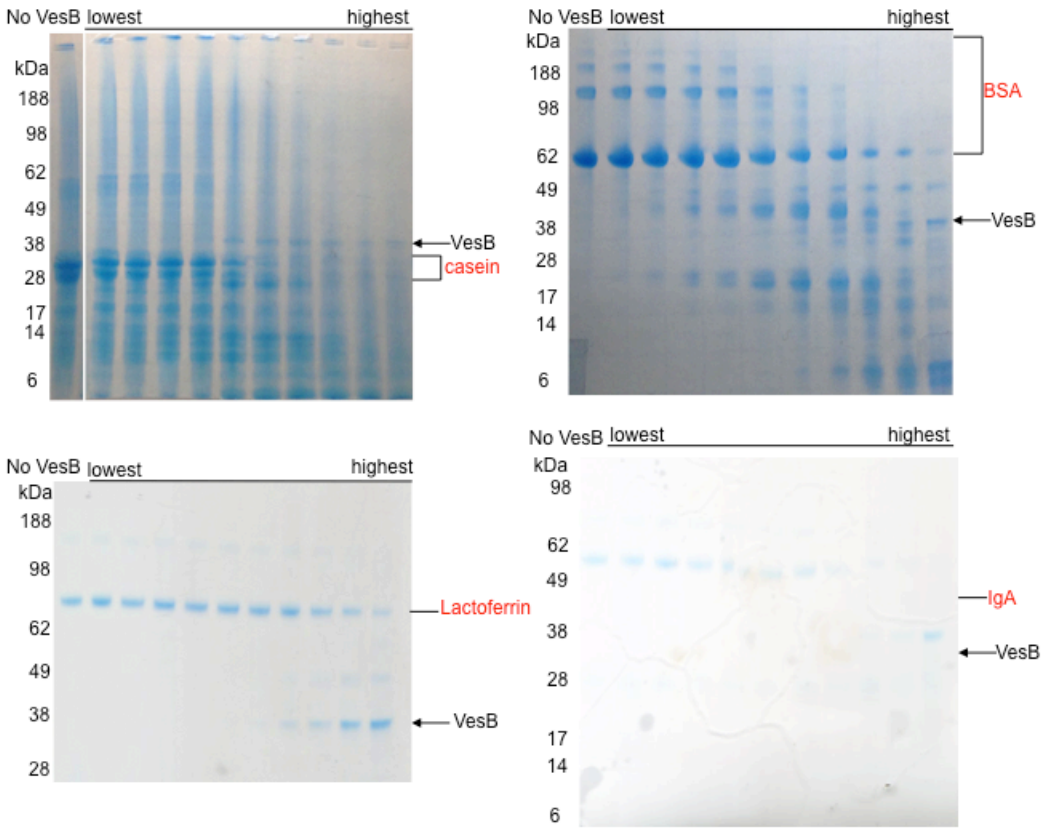

Figure 2.4: VesB degrades intact proteins.
Various concentrations of purified VesB (0-40 μg/mL) were incubating with different substrates (in red) for 1 hour at 37˚C. Samples were analyzed by SDS-PAGE and GelCode Blue staining.

VesB is produced as a zymogen:

The alignment of trypsin, thrombin and the protease domain of VesB revealed two important features of the chymotrypsin family of serine proteases: the catalytic triad and an activation site, with the latter corresponding to residues Arg32-Ile33 in pro-VesB (Figure 2.1). To verify cleavage at the predicted activation site, purified VesB was subjected to N-terminal sequencing by Edman

degradation. The sequence of the first five residues, IINGS, indicated that the propeptide had been removed from VesB and that VesB is cleaved at Arg32Ile33 similar to other trypsin-like proteases. To determine if cleavage of the propeptide is necessary for VesB activity, we mutated the *vesB* gene to replace the activation site residue Arg 32 with Glu. In another variant, the putative catalytic residue Ser 221 was replaced with Ala. The mutant proteins were expressed in the *V. cholerae* strain Δ*vesABC* and supernatants and cells were isolated and subjected to analysis of protease activity (Figure 2.5A) and immunoblotting with VesB antiserum (Figure 2.5B). Both VesB-S221A and VesB-R32E were expressed and secreted in amounts similar to WT VesB (Figure 2.5B). However, VesB-S221A and VesB-R32E were deficient in cleaving the synthetic peptide *Boc*-Gln-Ala-Arg-AMC (Figure 2.5A). Although the immunoblot suggested that a fraction of VesB-R32E is cleaved, the site of cleavage is unknown and does not result in an active protease. Taken together, the activation site and catalytic triad are vitally important for the activity of VesB, while they play no role in VesB secretion.

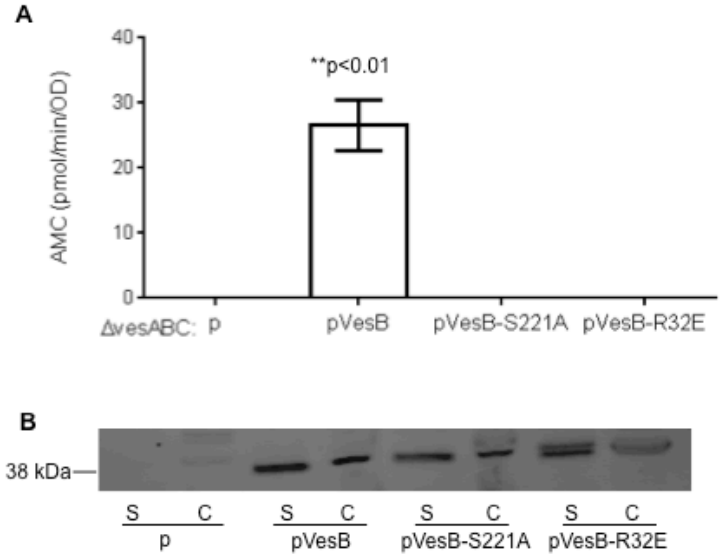

Figure 2.5: Protease activity and secretion of VesB-S221A and VesB-R32E. The *vesB* gene was subjected to site-directed mutagenesis and the obtained *vesB-S221A* and *vesB-R32E* constructs were cloned into pMMB67EH and expression was induced with 1 μM IPTG in the Δ*vesABC* strain.
A. The culture supernatants of Δ*vesABC* strains with pMMB67EH (p), pVesB, pVesB-S221A, and pVesB-R32E were analyzed for protease activity. The data of VesB was compared to the data of the three strains to generate p values.
B. The supernatants (S) and cell extracts (C) were isolated from the bacterial cultures of Δ*vesABC* strains with pMMB67EH, VesB, VesB-S221A, and VesB-R32E and subjected to SDS-PAGE and Western blot analysis using anti-VesB antibodies.

Structure of pro-VesB:

As the yield of VesB purified from *V. cholerae* culture supernatant was not sufficiently high for crystallization, a soluble variant of VesB comprising residues 24-373, but lacking the C-terminal residues 374-403, that contain the C-terminal GlyGly-CTERM extension, was expressed in *E. coli* and purified for crystallization and structure determination. To overcome a putative overexpression toxicity, the catalytic Ser221 of VesB was replaced with Ala. Crystallization trials produced crystals that yielded X-ray diffraction data to 2.4 Å (Table 2.2). The structure was

solved by a combination of molecular replacement and single wavelength diffraction methods using a sodium iodide derivative crystal. The structure was refined to R_{work}/R_{free} of 0.187/0.236 with good geometry (Table 2.2).

Table 2.2. Data collection and refinement statistics

	Native (PDB 4LK4)	NaI derivative[a]
Data collection		
Wavelength (Å)	0.9999	1.5418
Space group	$P4_22_12$	$P4_22_12$
Cell dimensions		
a, *b*, *c* (Å)	121.57, 121.57, 71.31	121.44, 121.44, 71.86
α, β, γ (°)	90, 90, 90	90, 90, 90
Resolution (Å)	46.3–2.40 (2.53–2.40)[b]	33.9–3.20 (3.37–3.20)
R_{sym}	0.065 (0.985)	0.154 (0.683)
$I / \sigma I$	20.4 (2.5)	14.9 (3.6)
Completeness (%)	99.5 (99.9)	99.9 (100)
Multiplicity	6.3 (6.5)	7.7 (7.6)
Refinement		
Resolution (Å)	46.3–2.40	
No. reflections (total / free)	21376 / 1070	
R_{work} / R_{free}	0.187 / 0.236	
No. atoms		
Protein	2465	
Ligand/ion	0	
Water	93	
B-factors		
Protein	60.2	
Water	48.0	
Wilson *B*	59.0	
R.m.s. deviations		
Bond lengths (Å)	0.012	
Bond angles (°)	1.466	
Ramachandran distribution (%)[c]		
Favored	97.5	
Outliers	0.0	

[a] Friedel pairs are treated as different reflections.

[b] Values in parentheses are for the highest-resolution shell.

[c] Calculated using the MolProbity server (http://molprobity.biochem.duke.edu) (Chen et al, 2010).

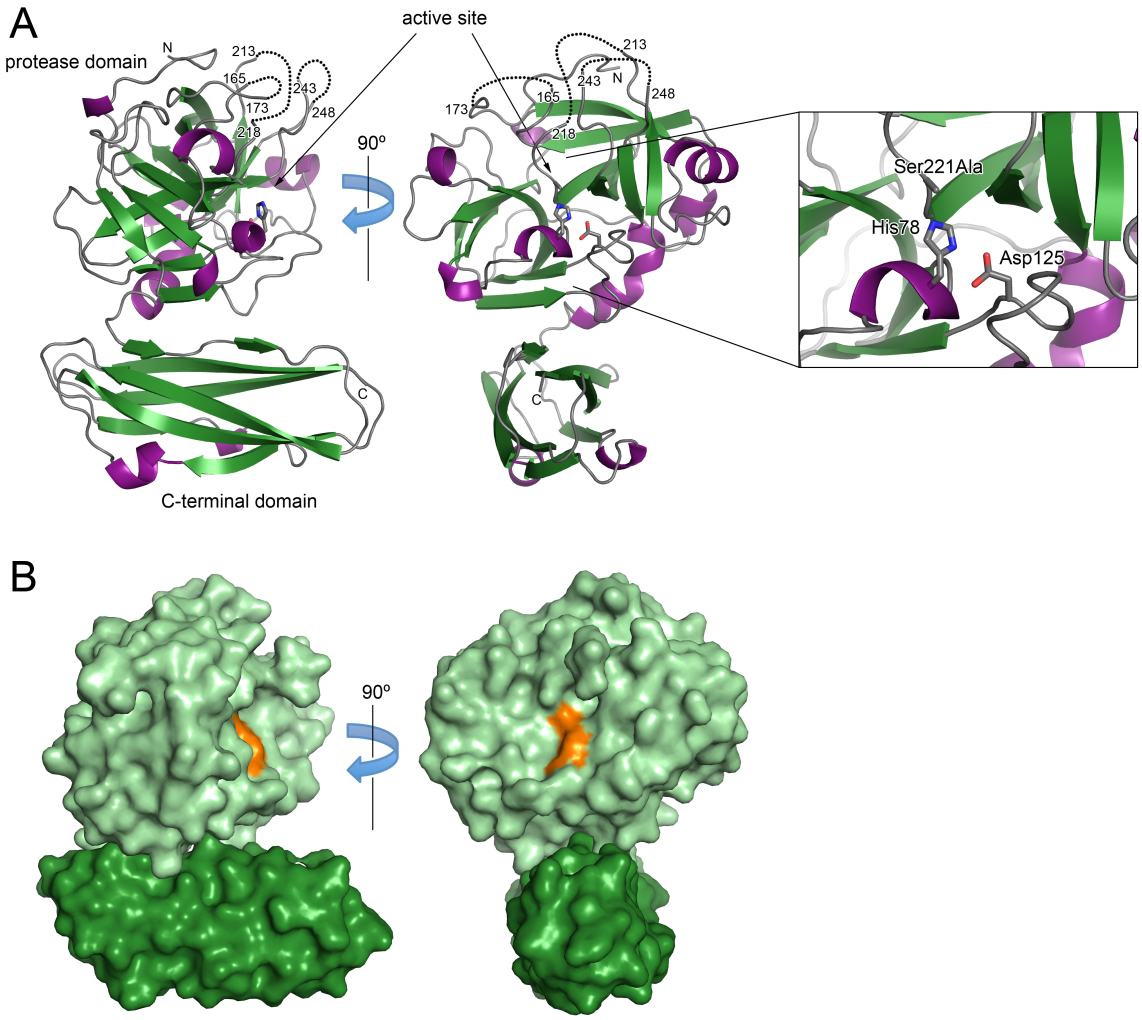

Figure 2.6: Crystal structure of pro-VesB.
A. The crystal structure is shown in ribbon representation with α-helices in purple and β-strands in green. The insert on the right shows the catalytic triad residues Asp125, His78 and Ser221. Note that Ser221 was substituted to Ala to alleviate possible toxicity during protein expression. The disordered loops are indicated by dotted lines.
B. The structure of VesB in surface representation. The protease domain of VesB is shown in light green and the Ig domain in dark green. The catalytic triad residues are colored in orange. Note that the catalytic site is not blocked by the Ig fold domain.

The crystal structure revealed that pro-VesB consists of two domains: an N-terminal protease domain with a typical trypsin/chymotrypsin fold and a C-terminal domain with an Ig fold (Figure 2.6). The structure of the protease domain of pro-VesB can be superimposed onto the bovine trypsinogen structure (PDB 1TGB, (Fehlhammer et al, 1977)) with an r.m.s.d. of 1.7 Å with 34% sequence identity for 198 residues (Figure 2.7). While there are two conserved disulfide bonds, Cys63-Cys79 and Cys190-Cys208, that are clearly present in the crystal structure, a putative Cys217-Cys244 disulfide is not visible as both Cys residues are part of disordered regions in pro-VesB. The protease domain of pro-VesB was captured in an inactive conformation in our structure either because *E. coli* does not express a protease that is capable of cleaving VesB and/or VesB is unable to undergo autocatalysis due to the Ser221Ala substitution. Residues 24-32, corresponding to the N-terminal propeptide sequence, were present in the crystallized protein, although most of the propeptide is disordered in the crystal. Due to the presence of the intact propeptide, the side chain of residue Ile33, which would be the N-terminal residue in the active form of VesB, cannot occupy the hydrophobic pocket lined by residues Val159 and Val180. Instead, this hydrophobic pocket in the current structure is occupied by Ile164, which prevents the correct orientation of Asp220 that, based on homology with trypsin and trypsinogen ((Fehlhammer et al, 1977; Marquart, 1983); PDB 2PTC and 1TGB), should coordinate the N-terminal amino group of Ile33 in active VesB (compare Figure 2.8A and B). These structural features probably contribute to the disorder of three loops in the vicinity of the active site, which include residues 166-172,

214-217 and 244-247 (Figure 2.6A). Consistent with VesB being a trypsin-like protease, the disordered loops in pro-VesB are reminiscent of the disordered loops in the crystal structure of bovine trypsinogen (Fehlhammer et al, 1977).

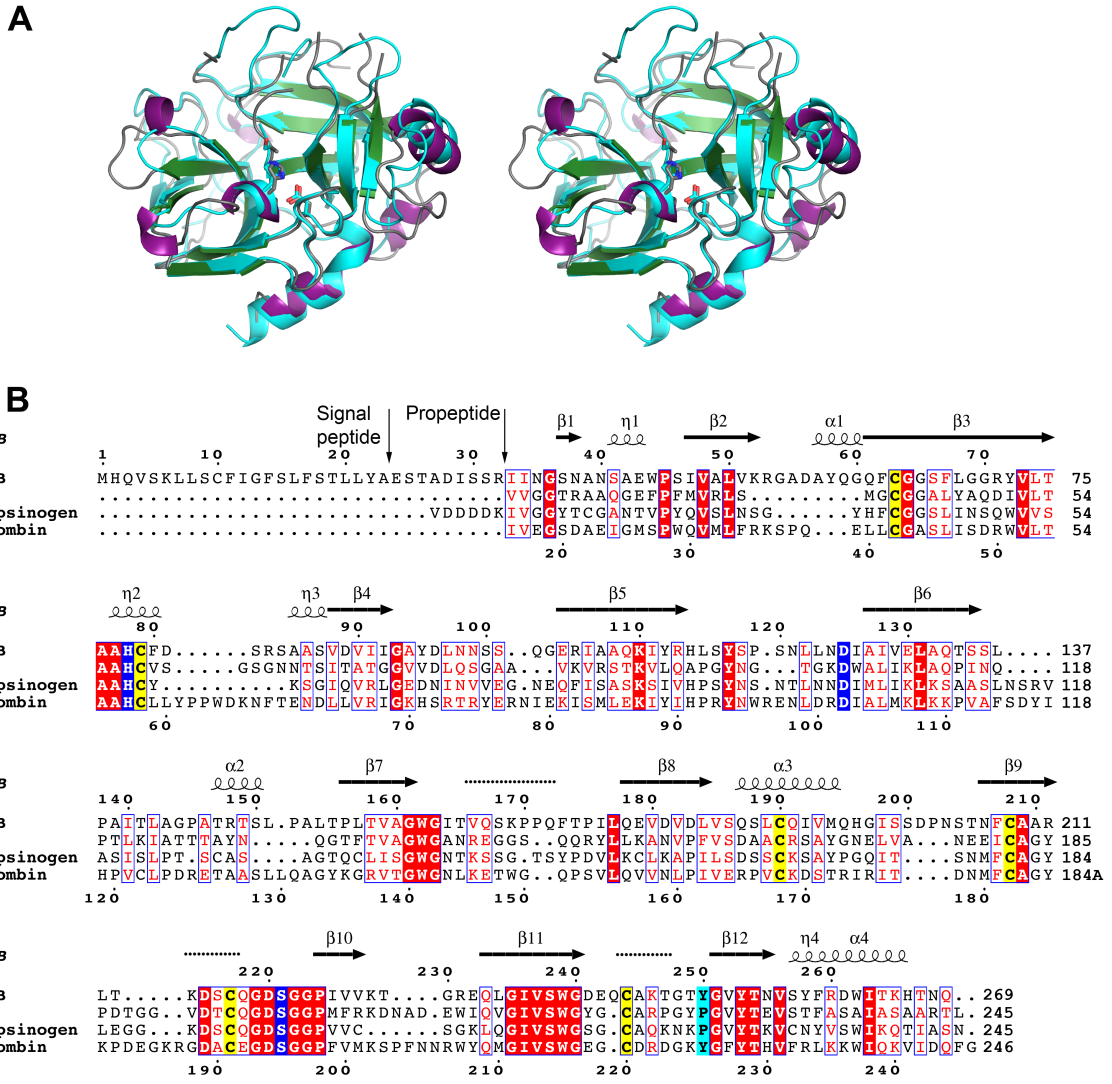

Figure 2.7: Similarities between pro-VesB, trypsinogen, thrombin and *Streptomyces griseus* trypsin.

A. Stereo view of superposition of the protease domain of pro-VesB and trypsinogen (PDB 2PTC, (Marquart, 1983)). VesB is shown in the same colors as in Figure 5A, i.e. with α-helices in purple and β-strands in green; trypsinogen is shown in cyan. The catalytic triad residues of both proteases are shown as sticks.

B. Structure-based sequence alignment of VesB, *S. griseus* trypsin (SGT), bovine trypsinogen and human thrombin (trypsin template numbering for

trypsinogen and thrombin (Fehlhammer et al, 1977)). Catalytic triad residues are highlighted in blue. Tyr250 of VesB, at a position equivalent to that of Tyr225 in thrombin is highlighted in cyan. In thrombin, this residue is involved in Na^+ binding (but neither in SGT, trypsinogen or VesB, see text), Conserved Cys residues that form disulfide bonds are highlighted in yellow, note that the disulfide C217-C244 was invisible in the pro-VesB structure due to flexibility. The three disordered loops in the pro-VesB crystal structure are indicated with dotted lines above the top sequence. The sequence numbers on top are those of VesB, at the bottom those of thrombin. The alignment was rendered using the ESPript server (Gouet et al, 2003).

In thrombin, residues Tyr184, Arg221, and Lys224 were shown to coordinate a Na^+ ion that binds to and allosterically enhances thrombin activity (Dang & Di Cera, 1996; Di Cera et al, 1995). While the catalytic activity of several other coagulation and complement factors is also enhanced by sodium, trypsin is not. In the Na^+-activated members of the trypsin-like family, Tyr 225 (thrombin numbering) plays a critical role. A proline residue at this position changes the orientation of the nearby Lys224 such that its carbonyl O atom can no longer coordinate Na^+ (Dang & Di Cera, 1996). When thrombin residue Tyr225 was replaced with a Pro, as occurs in several trypsin-like proteases which are not Na^+-sensitive, there was a loss in Na^+ activation (Dang & Di Cera, 1996). The sequence alignment of VesB with thrombin shows that the residue corresponding to thrombin's Tyr225 is Tyr250 in VesB (Figure 2.7B) i.e. a residue that is compatible with Na^+ activation. The structural superposition of pro-VesB and thrombin is shown in Figure 7C, specifically highlighting the position of Tyr225 and surrounding residues in thrombin and the corresponding residues in pro-VesB. However, no sodium or other ions are apparent in the pro-VesB electron density (Figure 2.8D). The absence of sodium may be due to the inability of pro-

VesB to coordinate sodium. Similarly, sodium is not present in the zymogen forms of thrombin (Pozzi et al, 2013). Thus, the structure of pro-VesB does not provide insight into the presence of a putative Na^+ binding site in active VesB. To establish whether active VesB can be stimulated by Na^+, we determined the activity of purified VesB in the absence and presence of increasing concentrations of NaCl using Boc-Gln-Ala-Arg-AMC (Table 2.1). The presence of Na^+ did not stimulate VesB, instead we observed a decrease in activity (Table 2.1). When a different substrate with Arg at the P1 position (Z-Gly-Gly-Arg-AMC) was tested, a similar negative effect on the VesB activity was observed (not shown). In addition, we isolated culture supernatant from a *V. cholerae* culture grown in a medium without NaCl and determined the activity of VesB in the absence or presence of 200 mM NaCl. There was no change in VesB activity when 200 mM NaCl was included (data not shown). Taken together, VesB does not appear to be activated by Na^+ in a similar fashion to thrombin.

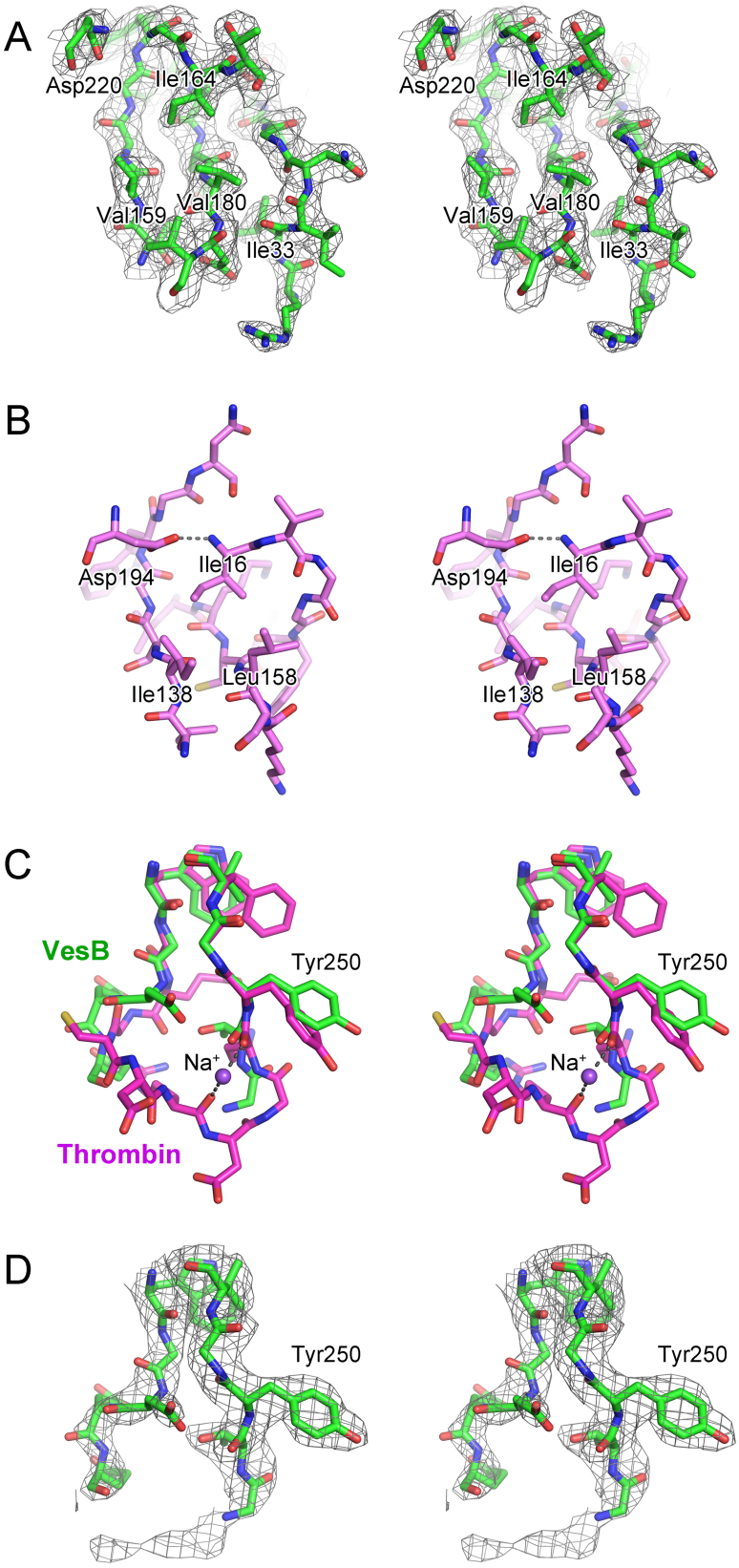

FIGURE 2.8: Close-up views of the activation site and the Tyr250 region of pro-VesB.
A. Stereo view of the activation site of VesB. Selected residues including Ile33 are shown in stick representation and labeled. σ_A-weighted $2F_O$-F_C map is shown as grey mesh (contoured at 1σ).
B. Stereo view of equivalent region in trypsin is shown (PDB ID 2PTC). Note that the N-terminal Ile16 residue occupies the hydrophobic pocket formed by Ile138 and Ile158, and coordinates Asp194.
C. Stereo view of superposition of Na^+ binding loop of thrombin (magenta) and the equivalent loop of pro-VesB (green). Na^+ ion is shown as a purple sphere.
D. The same view as in (C) showing σ_A-weighted $2F_O$-F_C map as grey mesh (contoured at 1σ) of pro-VesB.

The crystal structure of the C-terminal non-protease domain of VesB revealed a domain rich in β-strands that adopts an Ig fold (Figure 2.6). A structural similarity search with the Dali server identified a number of distant homologs with Ig folds (Table 2.3). While the best hit was a domain from a *Porphyromonas gingivalis* protein of unknown function (Z factor = 9.3 and 12 % amino acid sequence identity), homologs of possibly greater interest are domains from bacterial PapD-like chaperones such as SfaE (Z factor = 8.2) and subtilisin-like proteases like cucumisin (Z factor = 7.7). Superpositions of the C-terminal domain of VesB with SfaE and cucumisin domains are shown in Figure 2.9. Interestingly, VesB contains two additional short α-helices compared to the other proteins, and two extra β-strands compared to SfaE. We also observed a degree of structural homology, albeit less closely than with the previous three homologs, between the Ig domain of VesB and domains from several extracellular *Vibrio* proteins, including chitin-binding protein (GbpA), chitinase (ChiA), rugose and biofilm structure modifier (RbmA), and another protease of *V. cholerae* (PrtV) (Table 2.3). The Ig domain of GbpA has been reported to support its *V. cholerae* surface association (Wong et al, 2012), while the Ig domain of ChiA is a carbohydrate binding module (Songsiriritthigul et al, 2008). While we were unable to detect binding of purified VesB to chitin-Sepharose, suggesting that VesB is not a chitin-binding protein (data not shown), we have observed that a large fraction of extracellular VesB is surface associated (Gadwal *et al.*, manuscript in preparation). We have not yet confirmed a role for the Ig domain in this

interaction as the removal of the Ig domain of VesB results in an unstable protein that cannot be detected when expressed in *V. cholerae* (data not shown).

Table 2.3. Structural homologs of the C-terminal domain of VesB

Protein	Z-score[1]	r.m.s.d.	Aligned aa	Sequence identity, %	PDB ID	Reference
Unknown function	9.3	2.5	94	12	2QSV	N/A
SfaE	8.2	2.7	92	5	1L4I	65
Cucumisin	7.7	2.7	91	11	3VTA	57
RbmA	7.0	2.6	88	10	4KKQ	61
GbpA	5.8	3.4	84	11	2XWX	44
ChiA	4.5	3.7	73	11	3B8S	45
PrtV	3.7	2.8	63	17	4L9D	60
KP-43	2.3	2.9	60	7	1WMD	59
ASP	2.0	2.9	59	10	3HJR	58

[1] Z-score from the Dali server

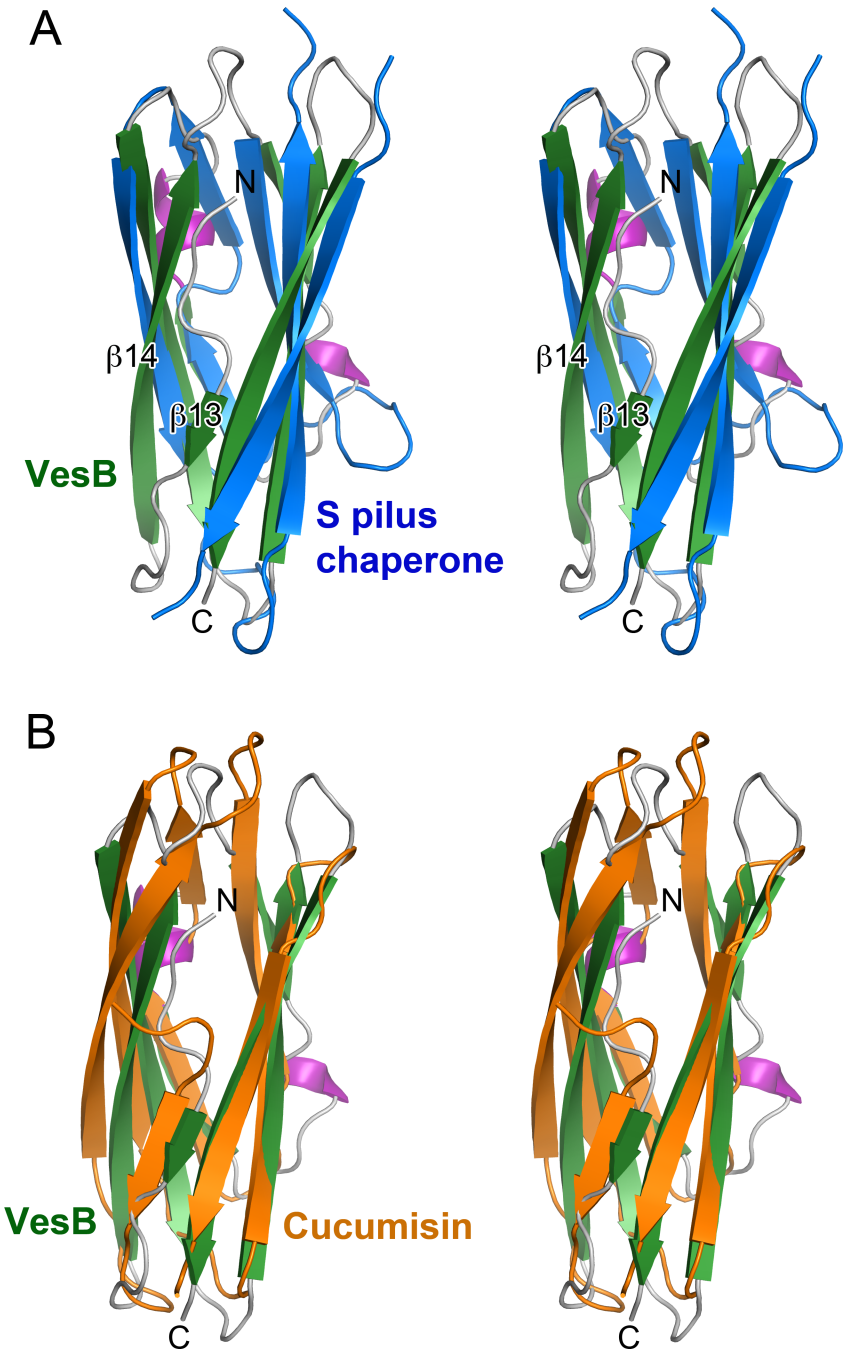

Figure 2.9: Structural homologs of the C-terminal domain of VesB.
A. Stereo view of superposition of the C-terminal domain of VesB (green and purple) and the N-terminal domain of S pilus chaperone (blue) (PDB 1L4I) (Knight et al, 2002). The N- and C-termini of the C-terminal domain of VesB are labeled. Note that the C-terminal domain of VesB has two additional β-strands and two additional short helices (purple).
B. Stereo view of superposition of the C-terminal domains of VesB (green and purple) and cucumisin (orange) (PDB 3VTA) (Murayama et al, 2012). Pro-VesB contains two additional short helices (purple).

Interface between the protease and C-terminal domains:

In the crystals, the two domains of pro-VesB interact with each other with a buried surface area (BSA) of 1220 $Å^2$ (Figure 2.6B). This interface is classified by the PISA server as insignificant (Krissinel & Henrick, 2007), although physiologically relevant interfaces with the same extent of buried surface area have been reported (Korotkov et al, 2011). Hence, it is uncertain whether the two domains of pro-VesB have a different or flexible relative orientation with respect to each other in solution or during secretion via the T2S apparatus.

Discussion:

In this study, we show that the protease domain of VesB is related to trypsin in sequence, structure, and substrate specificity. Upon sequence alignment of trypsin and VesB, the catalytic triad and the activation site of VesB became apparent (Figure 2.7B). Mutating these motifs rendered VesB inactive, suggesting that the catalytic triad and the activation site are essential for activity, but have no impact on extracellular secretion of VesB (Figure 2.5). Our enzymatic studies suggested that VesB may preferentially cleave peptides that have an Arg at the P1 position and showed that VesB is inhibited by the serine protease inhibitors, leupeptin and benzamidine (Figures 2.2 and 2.3, Table 2.1). While the natural targets of VesB remains to be identified, we speculate that VesB may be a promiscuous protease with broad specificity as it is capable of cleaving both peptides and intact proteins (Figure 2.4). Lastly, the crystal structure revealed that the N-terminal protease domain of pro-VesB is structurally

homologous to mammalian trypsin of the S1A subfamily and the C-terminal domain contains an Ig fold (Figures 2.5-2.8).

Bacterial serine proteases belonging to the S1A subfamily were once considered rare and only thought to exist in certain *Streptomyces* species (Rojas, 2002; Uesugi et al, 2009). With the rapidly growing number of sequenced genomes, however, it is becoming increasingly clear that additional bacterial species also carry genes for trypsin-like enzymes. Besides the *V. cholerae vesB*, *vesA* and *vesC*, genes encoding proteases of the chymotrypsin family are also found in other marine Gram-negative bacteria, many belonging to the *Vibrionaceae* and *Shewanellaceae* families. Previously, a study was carried out to determine the nature of acquiring trypsin-encoding genes in bacterial genomes. A phylogenetic tree based on the sequence and structure of trypsins from various mammalian species and *Streptomyces griseus* revealed that these proteases originate from a common ancestor (Rypniewski et al, 1994). A later study performed a similar analysis looking at a larger selection of trypsin-like proteases from a variety of species, but only including a single *Streptomyces* protease as bacterial representative (Rojas, 2002). Based on the data presented in the latter study it is difficult to determine the evolutionary relationship between the trypsin-like proteases. It is not clear whether trypsin encoding genes originated in bacterial cells, and were maintained during subsequent evolutionary steps in eukaryotic species but lost in many bacterial species, or if the genes have been transferred horizontally between species.

A way to distinguish and further analyze the relationship between the serine proteases of the chymotrypsin family is by determining how the activity of these proteases is regulated. For several trypsin-like proteases, which function in blood-coagulation or complement formation, activation by Na^+ plays a well-established role (Dang & Di Cera, 1996; Di Cera et al, 2007). It has been shown that in these proteins the residue equivalent to Tyr225 in thrombin plays a key role in the proper formation of the Na^+-binding site (Dang & Di Cera, 1996). In VesB, Tyr250 is equivalent in position to Tyr225 of thrombin (Figures 2.6B and 2.7C) and hence it would be possible that VesB is activated by sodium ions. However, addition of NaCl does not appear to increase purified VesB activity. In the crystal structure of pro-VesB no density representing a sodium ion at the expected position is evident, although it should be noted that the zymogen form of thrombin also lacks sodium (Figure 2.8D) (Pozzi et al, 2013). These results are not too surprising for two reasons.

First, elegant protein engineering studies on *S. griseus* trypsin (SGT), which contains a Pro at the equivalent position of Tyr225 in thrombin and is not activated by Na^+, have shown that up to 19 residues need to be altered, in addition to the Pro to Tyr change near the Na^+ activation site, to obtain a SGT variant which is fully Na^+-sensitive (Page et al, 2006). Hence, the presence of Tyr at the position equivalent to Tyr225 in thrombin is a necessary but not sufficient requirement for a member of the trypsin-like family to become activated by Na^+ which is also supported by studies on complement mannan-binding lectin-associated serine protease 2 (MASP-2) (Harmat et al, 2004).

Second, VesB is secreted and needs to function in the external milieu like the gut and/or aquatic habitats, sites with highly variable Na^+ ion concentrations. For example, at aquatic environmental sites where *V. cholerae* is frequently isolated, the NaCl concentration varies between 0.2% and 3.0% (Vezzulli et al, 2010). These conditions are quite different from that of blood, where the sodium ion concentration is precisely controlled. Hence, Tyr250 plays most likely a structural role and is not involved in Na^+-activation of this *V. cholerae* protease. Taken together, our findings suggest that a tyrosine in trypsin-like proteases equivalent to thrombin Tyr225 cannot be used as the sole indicator of Na^+ mediated activation.

The C-terminal domain of VesB is structurally similar to Ig folds of PapD-like chaperones, subtilisin-like proteases, and several extracellular proteins from *Vibrio.* The role of the Ig fold in VesB is currently unknown; however, understanding the role of Ig folds in other proteins may shed light on a possible function of the VesB Ig fold. The PapD chaperones reside in the periplasm of Gram-negative bacteria and bind to type 1 pilin subunits via a donor-strand complementation mechanism to target them to the usher protein for outer membrane translocation and pilus assembly on the cell surface (Sauer et al, 2000). The chaperone-mediated stabilization of the pilin subunits is evident in a *papD* deficient strain, in which DegP degrades the pilin subunits (Lindberg et al, 1989; Sauer et al, 2000). The subtilisin-like serine protease cucumisin contains a C-terminal Ig fold domain that interacts with its catalytic domain (Murayama et al, 2012) and in the *Aeromonas sobria* subtilisin-like protein ASP, the Ig fold domain

is located close to the active site; therefore, it has been proposed that this domain is involved in substrate specificity (Kobayashi et al, 2009). The Ig fold domain of yet another subtilisin-like protease, KP-43 from *Bacillus,* is suggested to provide stability, as KP-43 without its Ig fold domain could not be expressed (Nonaka et al, 2004). While the Ig fold domain may similarly provide stability to VesB, as we are unable to detect VesB when expressed without its C-terminal domain in *V. cholerae*, the Ig fold may alternatively, or in addition, provide substrate specificity although it is located relatively far (~20 Å) from the active site (Figure 2.6B).

Other possible roles of the Ig fold domain may relate to the extracellular secretion or surface association of VesB. VesB, GbpA, ChiA, PrtV, and RbmA are extracellular *Vibrio* proteins that contain Ig fold domains (Edwin et al, 2013; Giglio et al, 2013; Songsiriritthigul et al, 2008; Wong et al, 2012). Some of these proteins (VesB, GbpA and ChiA) have been shown to utilize the T2S system for outer membrane translocation (Connell et al, 1998; Kirn et al, 2005; Sikora et al, 2011). While the Ig fold domains may be involved in the transport of VesB and the other extracellular proteins via this translocation system, it is also possible that they assist in surface association of these proteins similar to the Ig fold domain of GbpA, which binds to the bacterial cell surface (Wong et al, 2012). Surface localization of VesB may optimally allow for the immediate uptake and intracellular delivery of peptides generated by VesB mediated proteolysis.

In summary, we have shown that VesB has a structure and specificity profile resembling that of eukaryotic trypsin-like proteases. Additionally, the

structure of the C-terminal domain of VesB revealed an Ig fold domain that may be involved in one or more different functions such as stabilizing the protease domain, co-defining substrate specificity, binding to the bacterial surface, and being part of a yet undefined secretion motif of the T2S system (Korotkov et al, 2012).

Materials and Methods:

Growth Conditions:

All strains were grown on Luria-Bertani (LB) agar (Fisher). Single colonies were inoculated in Luria-Bertani (LB) broth (Fisher) with 100 μg/mL of carbenicillin and 1 or 100 μM isopropyl-β-D-thiogalactopyranoside (IPTG) (Sigma) and grown for 16 h at 37 °C. *Bacterial strains and Plasmids* **-** The *Vibrio cholerae* strain N16961, an El Tor O1 biotype, and the isogenic Δ*vesABC* strain (Sikora et al, 2011) were used for all experiments. VesB was expressed from pMMB67EH-*vesB* (Lybarger et al, 2009), a low copy vector with an IPTG-inducible promoter and ampicillin cassette, and pMMB67EH served as a control plasmid or a cloning vector for additional constructs. The QuikChange Site-Directed Mutagenesis Kit (Agilent) was used to create *vesB-S221A* and *vesB-R32E* with the following primers 5'-TCATGTCAGGGAGATGCTGGTGGCCCAATTGTA-3' (Fwd), 5'-TACAATTGGGCCACCAGCATCTCCCTGACATGA-3' (Rev) and 5'-TCCACAGCAGATATTTCATCTGAAATTATTAATGGTTCGAATGCA-3' (Fwd), 5'-TGCATTCGAACCATTAATAATTTCAGATGAAATATCTGCTGTGGA-3' (Rev),

respectively. Primers were synthesized at IDT Technologies. Restriction enzymes, buffers, and T4 DNA ligase were purchased from New England Biolabs. Plasmid constructs were transformed into *Escherichia coli* MC1061 and a triparental conjugation protocol using the helper strain, MM294, was used to transfer the plasmids into N16961 and Δ*vesABC* strains (Sikora et al, 2007).

Native VesB Purification for Protease Assays:

VesB was overexpressed in *V. cholerae* and culture supernatant was isolated and precipitated with 60% ammonium sulfate for 1 h at 4 °C. The sample was centrifuged at 10,000 ×g for 35 minutes at 4 °C. The pellet was suspended in 20 ml of 50 mM Tris-HCl, pH 8.0, 450 mM NaCl (Buffer A) and residual ammonium sulfate was removed through dialysis against Buffer A. The dialyzed sample was subjected to affinity chromatography on benzamidine-sepharose (GE Healthcare). The flow-through fraction was collected and the column was washed with Buffer A. Then, VesB was eluted using 10 ml of 100 mM benzamidine, 50 mM Tris-HCl, pH 8.0 buffer. Fractions were pooled and the buffer was exchanged to 50 mM Tris-HCl, pH 8.0 using a PD10 column (GE Healthcare). VesB was concentrated by ultrafiltration (30 kDa cut-off; Millipore). Protein concentration was determined with the Bradford assay (BioRad) using bovine serum albumin as a reference. Polyclonal antiserum against VesB was generated by Covance Inc.

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and immunoblotting:

Samples were prepared and analyzed via SDS-PAGE and immunoblotting as described previously (Sikora et al, 2007) with the following modifications. VesB antiserum was incubated in culture supernatants from the Δ*vesABC* strain for 1 h to pre-absorb cross-reactive antibodies prior to incubating with the membrane for 2 h (1:5,000 dilution) and horseradish peroxidase-conjugated goat anti-rabbit IgG (Biorad) was used at 1:20,000 dilution. Membranes were developed using ECL 2 Western blotting reagent (Thermo Scientific) and protein was visualized using a Typhoon Trio variable mode imager system and Image Quant software. SDS-PAGE gels were stained with GelCode Blue (Thermo Scientific) or Silver Stain Kit reagents (Invitrogen).

Protease Assay:

Protease activity in *V. cholerae* culture supernatants was measured using *Boc*-Gln-Ala-Arg-7-amido-4-methylcoumarin as described previously (Johnson et al, 2007). In other assays, supernatants were pre-incubated with benzamidine (Thermo Scientific) at the various concentrations listed for 10 minutes at 37 °C prior to measuring protease activity. Purified VesB was measured for protease activity with various methylcoumarin conjugated peptides (Peptides International) listed in Table 2.1 in a final reaction volume of 100 µl. The peptides, Suc-Ala-Ala-Ala-AMC, Suc-(OMe)-Ala-Ala-Pro-Val-AMC, Z-Leu-Arg-Gly-Gly-AMC, Leu-AMC, Suc-Ala-Ala-Pro-Phe-AMC, *Boc*-Glu-Lys-Lys-AMC, and *Boc*-Gln-Ala-Arg-AMC

are substrates that have been successfully used for elastase, pancreatic elastase, isopeptidase, aminopeptidase, chymotrypsin, plasmin, and trypsin, respectively. All assays were done with 50 mM Tris-HCl, pH 8.0. Where stated, assays were performed in the presence of different concentrations of NaCl. Inhibition assays were done with purified VesB that was pre-incubated for 10 minutes at 37 °C with 50 μM leupeptin (Thermo Scientific), 10 mM EDTA, or 1 mM benzamidine. A change in fluorescence per minute was converted to moles of liberated 7-amido-4-methylcoumarin (AMC) per minute via a standard curve with known amounts of AMC. The results were then normalized by the amount of purified VesB or OD600 of the bacterial cultures. Standard errors were generated from three independent experiments each performed in three technical replicates.

Degradation of Intact Proteins:

VesB concentrations ranging from 0-40 μg/mL were added to 0.781 μM of lactoferrin, 0.162 μM IgA, 0.015 M of casein, and 0.005 M of BSA in a 100 μL reaction volume. Samples were incubated for 1 hr at 37 °C, prepared and analyzed by SDS-PAGE and GelCode Blue staining.

N-terminal Sequencing:

Purified VesB secreted by *V. cholerae* was subjected to SDS-PAGE and transferred to a PVDF membrane using 10 mM CAPS, 10% methanol, pH 11. The membrane was stained with 0.025% Coomassie Brillant Blue R-250, 40%

methanol, 5% acetic acid solution, destained using 40% methanol, 5% acetic acid, washed with water and then air-dried. The VesB band was cut and subjected to automated Edman degradation at the protein facility core at University of Michigan, Ann Arbor.

Statistical Analysis:

t tests were done on samples as indicated. Results yielding a P value of <0.05 were considered statistically significant.

Protein expression and purification for crystallization:

A DNA fragment corresponding to residues 24-373 of VesB was PCR amplified and cloned into a modified pCDF-Duet-1 vector (EMD Millipore). The construct contains an N-terminal *pelB* signal sequence and a C-terminal tobacco etch virus (TEV) protease cleavage site followed by a His_6-tag. To overcome possible difficulties when overexpressing WT pro-VesB in *E. coli*, the Ser221 residue was replaced with Ala using QuickChange mutagenesis protocol (Stratagene). The S221A variant of pro-VesB was expressed in *E. coli* Rosetta2(DE3) cells at 18 ºC for 4 h after induction with 0.5 mM IPTG. The cells were harvested by centrifugation and resuspended in buffer containing 20 mM Tris-HCl pH 7.4, 0.5 M NaCl, 1 mM PMSF, 1 mM benzamidine and 15 mM imidazole. Pro-VesB was purified via a Ni-NTA agarose (Qiagen) column followed by His_6-tag cleavage with TEV protease and a second Ni-NTA agarose chromatography step. The yield of pro-VesB was improved from <0.5 mg to ~1.5

mg per 6 L of culture after addition of 2 mM Fe^{2+} ions to the clarified cell lysate (Magnusdottir et al, 2009). The final size-exclusion chromatography was performed using a Superose6 column (GE Healthcare) in buffer containing 20 mM Tris-HCl pH 7.4 and 0.5 M NaCl. Pro-VesB was concentrated to 3 mg ml^{-1} and flash frozen in liquid nitrogen (Deng et al, 2004).

Crystallization, data collection and structure solution:

Screening for crystallization conditions was performed using Index (Hampton Research) and Wizard I and II (Emerald Bio) screens. Crystallization trials were set-up in vapor diffusion sitting drop format (0.2 μL protein + 0.2 μL crystallization solution) using a Phoenix crystallization robot (Art Robbins Instruments). The initial crystals grew with 0.1 M HEPES pH 7.5, 3.0 M NaCl. The optimized crystals were obtained using a crystallization solution containing 0.1 M Tris-HCl pH 9.0, 3.0 M NaCl (1.0 μL protein + 1.0 μL crystallization solution + 0.5 μL water). Crystals were cryoprotected in crystallization solution supplemented with 20% glycerol and flash-cooled in liquid nitrogen.

A 2.4 Å native dataset was collected at the Advanced Light Source, beamline 8.2.2. In addition, a derivative dataset was obtained in-house using a Saturn 94 CCD detector on a Rigaku Micromax HF-7 rotating anode from a crystal incubated in 0.1 M Tris-HCl pH 9.0, 2.0 M NaCl, 1.0 M NaI. Data were processed using HKL2000 and XDS (Kabsch, 2010; Otwinowski, 1997). The pro-VesB structure was solved by a combination of molecular replacement and single wavelength diffraction methods as implemented in Phaser (McCoy et al, 2007).

The initial molecular replacement solution was found by using the Balbes server (Long et al, 2008) and the structure of MT-SP1/matriptase (PDB 1EAX) as a search model for the protease domain of pro-VesB (Friedrich et al, 2002; Long et al, 2008). 12 iodide sites were found by Phaser during the first round of phasing. After density modification using Parrot, model building using Buccaneer and iterative model improvement using Coot (Cowtan, 2006; Cowtan, 2010; Emsley et al, 2010), the preliminary model was submitted for the second round of phasing, which led to identification of 5 additional (17 total) iodide sites. Subsequent iterative density modification and automated model building led to an almost complete model that included the second domain of VesB. The model was improved using Coot and refined using REFMAC5 and 9 translation/libration/screw (TLS) groups identified by the TLSMD server (Murshudov et al, 2011; Painter & Merritt, 2006). The quality of the model was assessed using Coot and the MolProbity server (Chen et al, 2010). Residues 32-372 are included in the final model, while residues 166-172, 214-217 and 244-247 had poorly defined density and were not modeled. The final model includes 93 water molecules and has good geometry with 97.5% residues in the favorable areas of the Ramachandran plot (Table 2.2). The atomic coordinates and structure factors for pro-VesB (code 4LK4) have been deposited in the Protein Data Bank (http://www.pdb.org/).

Chapter 3:

A novel protein sorting mechanism for cell surface localization in Gram-negative bacteria

Abstract:

Vibrio cholerae and a subset of other Gram-negative bacteria express proteins with C-terminal tripartite transmembrane domains called GlyGly-CTERM, which consist of a motif rich in glycines and serines, followed by a hydrophobic region and positively charged residues. Here we show that VesB, a *V. cholerae* serine protease containing a GlyGly-CTERM domain, requires both the type II secretion system and a rhomboid-like protease for localization at the cell surface. The rhomboid-like protease, rhombosortase, cleaves the GlyGly-CTERM domain and promotes surface retention possibly through transamidation and linkage of VesB to a membrane component. In support of this, cell-associated VesB activity is no longer detected when the rhombosortase gene is inactivated or when VesB is expressed without its GlyGly-CTERM domain. Instead, the GlyGly-CTERM deleted VesB is released to the extracellular milieu, while in the absence of GlyGly-CTERM processing, wild-type VesB is inactive and unstable. We propose that the rhombosortase/GlyGly-CTERM system is the

Gram-negative equivalent of the sortase/LPXTG system in Gram-positive bacteria, which anchors surface proteins to the cell wall through transpeptidation.

Introduction:

The type II secretion (T2S) system is a multi-protein complex used by many Gram-negative bacteria to secrete extracellular proteins (Cianciotto, 2005; Korotkov et al, 2012). Most notably, *Vibrio cholerae*, the causative agent of cholera, uses the T2S system to secrete cholera toxin (Sandkvist et al, 1993). Cholera toxin and other T2S substrates are secreted in a two-step process. First, proteins translocate to the periplasmic compartment via recognition of their signal peptide by the Sec system. Then, the periplasmic secretion intermediates fold and engage with the T2S system for outer membrane translocation. These substrates are then free to diffuse away from the cell (Sandkvist, 2001). However, some substrates that utilize the T2S system have been found to be cell-associated. The heat-labile enterotoxin of enterotoxigenic *Escherichia coli* (ETEC); pullulanase, a lipoprotein from *Klebsiella oxytoca;* the *Dickeya dadantii* pectin lyase, PnlH; and the *Shewenella oneidensis* c-type cytochrome, OmcA, are first transported across the cell envelope by the T2S system and then associate with the cell surface (Donald et al, 2008; Ferrandez & Condemine, 2008; Horstman & Kuehn, 2002; Pugsley et al, 1986; Rondelet & Condemine, 2013).

In addition to cholera toxin, *V. cholerae* transports a number of other proteins, including three serine proteases, VesA, VesB and VesC, across the

outer membrane via the T2S system (Sandkvist et al, 1997; Sikora et al, 2011). VesB, the focus of this paper, is a trypsin-like serine protease that contains an N-terminal signal peptide, a protease domain, and an immunoglobulin-fold domain (Gadwal et al, 2014). The protease domain is 30% homologous to trypsin and includes a typical His-Asp-Ser catalytic triad and activation site. VesB is made as a zymogen and cleavage at the activation site (Arg32-Ile33) results in the removal of the N-terminal propeptide to generate active VesB (Gadwal et al, 2014). VesB, along with five other *V. cholerae* proteins, VesC, VesA, a putative metalloprotease (VCA0065), a hypothetical protein (VC1485), and a nuclease, Xds, also contain a C-terminal domain called GlyGly-CTERM (Haft & Varghese, 2011) (Figure 3.1).

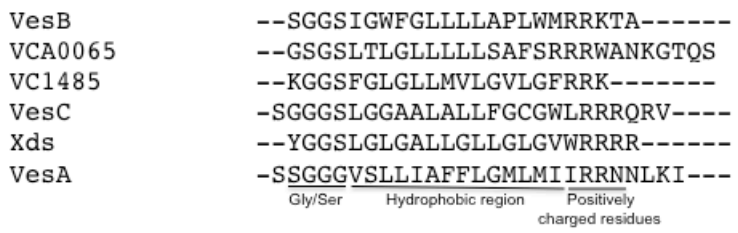

Figure 3.1: Alignment of GlyGly-CTERM containing proteins in *V. cholerae*. The primary sequences of the C-terminal portions of the six proteins in *V. cholerae* containing the GlyGly-CTERM domain are aligned based on the GlyGly-CTERM sequence logo in (Haft & Varghese, 2011). The region rich in glycines and serines, the hydrophobic residues, and the positively charged residues are underlined.

Predicted to be a transmembrane domain, the GlyGly-CTERM has a consensus tripartite motif that contains two prominent glycines surrounded by serines, followed by a hydrophobic helix and positively charged residues (Figure 3.1). Thus far, only a few GlyGly-CTERM proteins have been characterized and

their subcellular location has not been fully determined. While the GlyGly-CTERM containing ExeM is found as an active endonuclease in the culture supernatant of *S. oneidensis* (Godeke et al, 2011), proteomic studies of membranes have identified ExeM as a membrane-associated protein (Brown et al, 2010; Tang et al, 2007). Xds, an ExeM homolog, displays a similar distribution as Xds nuclease activity has been detected in both culture supernatant and cells (Newland et al, 1985; Seper et al, 2011). Although these dual locations may seem contradictory, they may be the result of a sustained, slow, cellular release from the T2S system. Alternatively, the cell surface may be the primary location of these enzymes and membrane-association may involve their GlyGly-CTERM domain. The novel GlyGly-CTERM domain was originally identified bioinformatically in Gram-negative bacteria of the *Vibrio*, *Shewanella*, *Acinetobacter* and *Ralstonia* genera (Haft & Varghese, 2011). Through *in silico* partial phylogenetic profiling, proteins that have the GlyGly-CTERM domain were identified to co-exist with rhombosortases, a subfamily of rhomboid proteases (Haft & Varghese, 2011).

Rhombosortase is homologous to rhomboid proteases, a class of membrane-embedded serine proteases that contain a Ser-His active site located within the lipid bilayer (Ben-Shem et al, 2007; Lemieux et al, 2007; Wang et al, 2006; Wu et al, 2006). Crystallography studies of the rhomboid protease, GlpG, indicate that it is comprised of six transmembrane domains, with a deep cavity within its folded form that contains the active catalytic dyad (Ben-Shem et al, 2007; Lemieux et al, 2007; Wang et al, 2006; Wu et al, 2006). Rhomboid

proteases are nearly ubiquitous and function in a wide range of processes (Freeman, 2008) (Rather, 2013). In bacteria, only one substrate has been identified, the channel-forming TatA of the twin-arginine translocation (TAT) system. The *Providencia stuartii* rhomboid protease, AarA, cleaves and thereby activates TatA, which results in the formation of the TAT system (Fritsch et al, 2012; Stevenson et al, 2007). Rhomboid proteases typically cleave substrates with a single transmembrane domain that contains helix-destabilizing residues, such as prolines and glycines. Cleavage preferably occurs at small residues such as alanine within the transmembrane helix of substrates (Akiyama & Maegawa, 2007; Strisovsky et al, 2009).

The presence of helix destabilizing residues in the putative transmembrane helix of the GlyGly-CTERM domain in VesB and other substrates suggests that they may be cleaved by rhombosortase (Haft & Varghese, 2011). Because of the similarity of the GlyGly-CTERM domain and the Gram positive tripartite C-terminal LPXTG sorting motif that is cleaved and covalently cell wall anchored by sortase in Gram positive cells (Mazmanian et al, 1999; Schneewind et al, 1992), we hypothesize that rhombosortase may similarly cleave and anchor GlyGly-CTERM proteins to the cell envelope of Gram negative bacteria.

In this study, we address the role of rhombosortase, the T2S system, and the GlyGly-CTERM domain on proteins by using VesB as a model protein. Here we show that VesB is distributed in a pattern similar to the nucleases, ExeM and Xds, in that VesB has both secreted and surface-associated activities. Analysis of the localization and activity of VesB in different mutant backgrounds reveals

that rhombosortase cleaves VesB and promotes T2S-dependent cell retention and activity of VesB.

Results:

Cell-associated VesB is active:

Previously, we have shown that *V. cholerae* produces three related extracellular serine proteases, VesA, VesB and VesC, of which VesB contributes to the majority of activity (data not shown, Sikora et al, 2011). While VesB is a type II secreted protease with its N-terminal signal peptide and propeptide removed, a large fraction of VesB remains cell-associated (Gadwal et al, 2014; Sikora et al, 2011). To determine whether the cell-associated form is also active, we expressed plasmid-encoded VesB in a strain lacking the three related *vesA*, *vesB* and *vesC* genes (Δ*vesABC;* [16]). We included two additional strains, one containing an empty vector and one expressing an active site mutant protein, VesB-S221A, to serve as negative controls. Cultures were grown in the presence of 50 μM IPTG in LB for 4 hours, after which culture supernatants and cells were separated by centrifugation and the VesB activity was analyzed using two different approaches. When using the fluorogenic peptide, Boc-Gln-Ala-Arg-AMC, active VesB, but not VesB-S221A, was found in both the culture supernatant and cells (Figure 3.2A). Similarly, when using the fluorophosphonate probe, ActivX TAMRA-FP, which covalently labels the serine residue in the active sites of serine hydrolases (Patricelli et al, 2001), we found that the probe bound to the WT form of VesB both in the culture supernatant and cells as visualized by

SDS-PAGE and fluorescent gel imaging (Figure 3.2B, top panel). While VesB-S221A was recognized by VesB antibodies (Figure 3.2B, bottom panel), it could not be detected with ActivX TAMRA-FP; highlighting the specificity of the probe (Figure 3.2B, top panel). Collectively, the results from the two assays suggest that both the secreted and cell-associated forms of VesB are active. Like other trypsin-like proteases, VesB is produced as a zymogen and removal of the propeptide results in active VesB (Gadwal et al, 2014). The finding that VesB-S221A migrated slower compared to VesB when analyzed by SDS-PAGE and immunoblotting indicated that it still contains the propeptide and suggests that VesB is an autoactivating enzyme (Figure 3.2B, bottom panel).

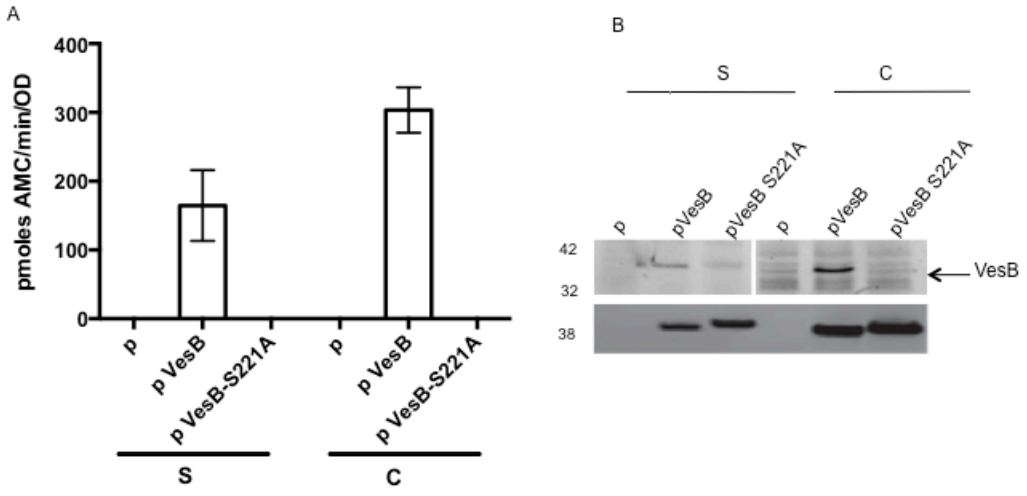

Figure 3.2: Secretion profile of VesB in *V. cholerae*. The *ΔvesABC* strain containing empty vector (p), pVesB, or pVesB-S221A were grown in LB with 100 µg/mL carbenicillin and 50 µM IPTG for four hours. Samples were separated into supernatants (S) and cells (C). The experiments were done in biological triplicates and standard error bars are shown.

A. Secreted and cell-associated VesB are active. Serine protease activity against a fluorogenic peptide, Boc-Gln-Ala-Arg-AMC, was measured in culture supernatants and cells.

B. ActiveX-TAMRA probe binds to both secreted and cell-associated VesB. The culture supernatants and cells were incubated with the ActiveX-TAMRA serine hydrolase probe followed by SDS-PAGE and fluorescent

imaging using the Typhoon imager system (top panel) or subjected to SDS-PAGE and immunoblotting with VesB antibodies (bottom panel).

VesB is cell surface-associated:

Because we detected cell-associated VesB activity, we hypothesized that VesB may be surface-localized in T2S competent cells, but remains internally trapped in T2S mutant cells (Δ*eps*). To test this, we expressed plasmid-encoded VesB in Δ*vesB* and Δ*vesB*Δ*eps* strains and subjected the cells to labeling with anti-VesB antibodies and Alexa Fluor 488 $F(ab')_2$ goat anti-rabbit IgG. In this set of experiments, the cultures were grown in M9 medium supplemented with glucose and casamino acids to avoid autofluorescence observed with LB grown cells. Under this growth condition, the majority of VesB protein and activity was cell-associated in the ΔvesB/pVesB strain; very little VesB was found in the culture supernatant (Figure 3.3AB). While all of VesB remained cell-associated in an inactive form in the Δ*vesB*Δ*eps*/pVesB strain, occasionally a very small amount of inactive VesB can be detected in the culture supernatant (Figure 3.3AB). VesB migrated slower in Δ*vesB*Δ*eps* than Δ*vesB,* perhaps indicating the presence of the N-terminal propeptide in the absence of the T2S system; thus suggesting that activation of VesB may occur after translocation through the outer membrane. For the VesB labeling experiment, the Δ*vesB* and Δ*vesB*Δ*eps* strains with empty vector served as negative controls. After thorough washing, the cells were visualized by fluorescence microscopy (Figure 3.3C). While circumferential fluorescent foci were observed for 94% of Δ*vesB*/pVesB cells, no fluorescent foci were revealed on Δ*vesB*Δ*eps*/pVesB mutant cells. The labeling

efficiency was quantitated by fluorimetry and the fluorescent units were normalized to OD_{600} of the original cultures. The fluorescence measured for Δ*vesB*/pVesB was 4.8-fold higher when compared to Δ*vesB*Δ*eps*/pVesB, suggesting that VesB is surface-localized in WT cells, while it primarily remains intracellular in the Δ*eps* cells (Figure 3.3D). Taken together, our results suggest that outer membrane translocation by the T2S system is required for VesB localization to the surface of WT cells.

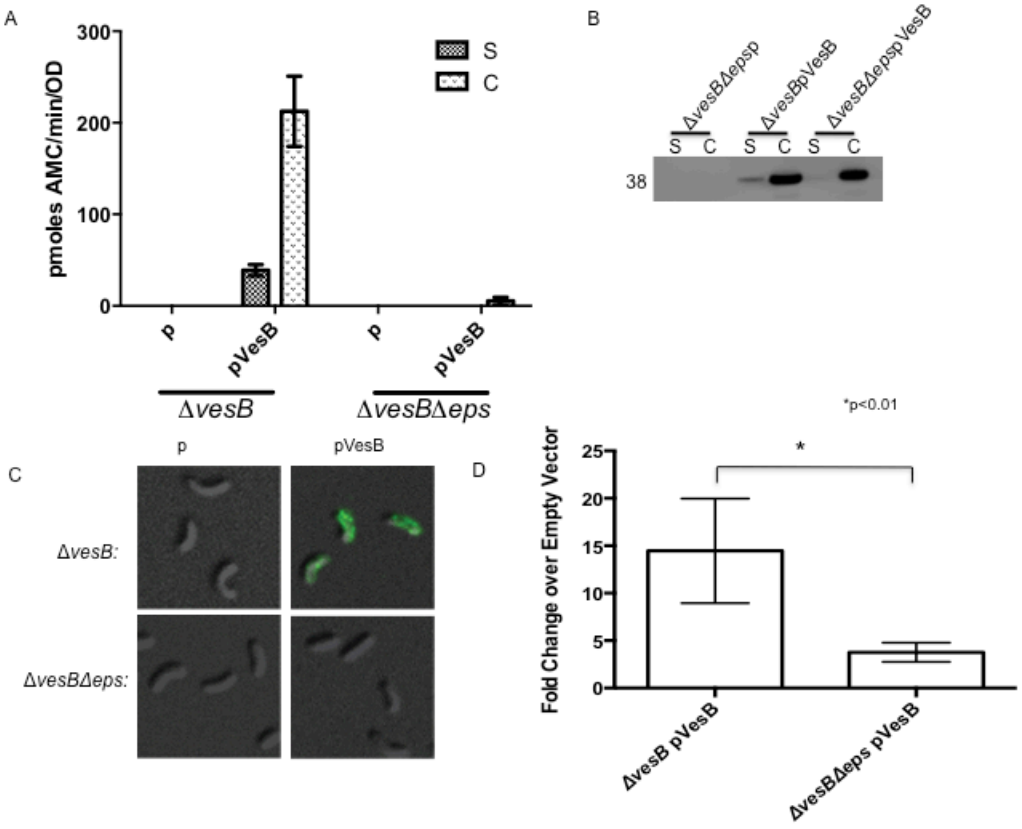

Figure 3.3: VesB is surface-associated in WT but not in the T2S mutant. The Δ*vesB* and Δ*vesB*Δ*eps* strains containing empty vector (p) or pVesB were grown for four hours in M9 media supplemented with glucose and casamino acids in the presence of 100 µg/mL carbenicillin and 50 µM IPTG. All experiments were done in triplicates and standard error bars are shown.

A. The majority of VesB activity is cell-associated in T2S competent cells. Supernatants (S) and whole cells (C) were isolated and analyzed for protease activity using a fluorogenic peptide, Boc-Gln-Ala-Arg-AMC.

B. The majority of VesB is cell-associated. Samples from A were analyzed by SDS-PAGE and western blotting for the protein levels of VesB.
C. Detection of VesB on the surface of T2S competent cells. Cells were washed and incubated with VesB antibodies followed by AlexaFluor 488 antibodies. Labeled cells were visualized by fluorescence microscopy.
D. Quantitative measurements indicate the presence of cell surface VesB in T2S competent cells. Cells that were labeled with VesB antibodies and AlexaFluor 488 antibodies were subjected to quantitative fluorimetry. The fluorescence units were normalized by OD_{600} of the original cultures and fold changes over cells with empty vector were calculated. The experiments were done in triplicates and standard error bars are shown. p values were generated by comparing Δ*vesB* and Δ*vesB*Δ*eps* strains containing pVesB.

Rhombosortase-mediated processing of VesB:

The above experiments showed that the T2S system is responsible for outer membrane translocation of VesB; however, they did not provide an explanation for the relatively large amount of surface retained VesB and why only a small amount of VesB is released to the culture supernatant. Therefore, in the next set of experiments we determined whether surface retention of VesB involves the putative transmembrane domain, GlyGly-CTERM, and/or its possible removal by the rhombosortase protease, RssP. We postulated that by using VesB-S221A, which does not undergo autoactivation (Figure 3.2B), we could ascribe any size changes of VesB to C-terminal processing, possibly involving the removal of the GlyGly-CTERM domain. Cultures of Δ*vesB*/pVesB-S221A, Δ*vesB*Δ*eps/*pVesB-S221A and *rssP::kan/*pVesB-S221A, in which the rhombosortase (*rssP)* gene is inactivated, were grown in the presence of 50 μM IPTG in M9 media supplemented with glucose and casamino acids. Culture supernatants and cells were separated and size differences of VesB-S221A were

analyzed by SDS-PAGE and western blotting (Figure 3.4). We found that VesB-S221A was unstable in the absence of rhombosortase, thus we loaded ten times less of the cell extracts containing a functional rhombosortase (Figure 3.4, lanes 2, 10 and 12) to clearly visualize potential size differences of VesB-S221A in the various mutants.

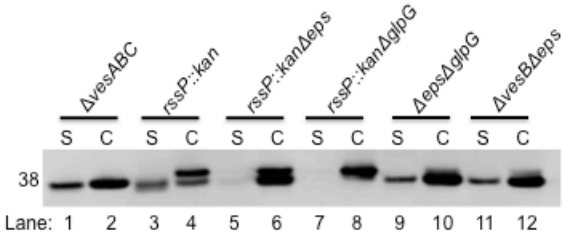

Figure 3.4: Cleavage of VesB by rhombosortase. The indicated mutant strains with pVesB-S221A were grown for four hours in M9 media supplemented with casamino acids, glucose, 100 µg/mL carbenicillin and 50 µM IPTG. Culture supernatant (S) and cells (C) were isolated, matched by equivalent OD_{600} units (except for the cell samples of the *ΔvesABC, ΔglpGΔeps,* and *ΔvesBΔeps* strains which were diluted 1:10) and analyzed by SDS-PAGE and western blotting with VesB antibodies.

As expected, VesB-S221A was processed in the Δ*vesB* and Δ*vesB*Δ*eps* mutant strains, in which rhombosortase is present. When the rhombosortase gene was inactivated (*rssP::kan*), a portion of VesB-S221A migrated slower than VesB-S221A in the Δ*vesB* cells (Figure 3.4, compare lanes 2 and 4). This form of VesB-S221A was exclusively detected in the cell fraction (Figure 3.4, lane 4). Interestingly, a small amount of VesB-S221A was still cleaved and secreted extracellularly (Figure 3.4, lane 3). This cleaved form was no longer released when the T2S system was inactivated in the *rssP::kanΔeps* double mutant (Figure 3.4, lane 5). While these results suggested that rhombosortase likely cleaves off the GlyGly-CTERM domain or some part thereof in the Δ*vesB* and

Δ*vesB*Δ*eps* cells, there was no explanation for the residual processing of VesB-S221A in the absence of rhombosortase. As *V. cholerae* encodes a homologous rhomboid protease, GlpG, with which rhombosortase shares 19% amino acid sequence identity, we hypothesized that GlpG is responsible for cleaving a fraction of VesB-S221A in the *rssP::kan* mutant. To test this, we made a double mutant, where both protease genes *glpG* and *rssP* were inactivated (*rssP::kanΔglpG).* In this background, VesB-S221A remained associated with the cells in an un-processed form suggesting that GlpG is capable of cleaving VesB-S221A, albeit inefficiently, when rhombosortase is absent (Figure 3.4, lane 8). As VesB-S221A was completely cleaved in the *ΔepsΔglpG* double mutant in which the rhombosortase gene remains intact, it suggests that rhombosortase is most likely the protease that cleaves the GlyGly-CTERM domain (Figure 3.4, lanes 9 and 10). The latter result also indicates that rhombosortase-mediated processing of VesB does not require the presence of a functional T2S apparatus and suggests that cleavage normally occurs before VesB engages the T2S system. Taken together, rhombosortase cleaves the GlyGly-CTERM domain of VesB and likely promotes surface retention, as the majority of VesB remains cell-associated and can be detected on the surface of WT cells (Figures 3.3 and 3.4).

C-terminal processing by rhombosortase is a prerequisite for VesB autoactivation:

To establish the functional consequences of C-terminal processing of VesB, we next analyzed WT VesB expressed from its native promoter.

Specifically, we wanted to determine whether N-terminal processing and activation of VesB is dependent on C-terminal modification by rhombosortase. Cultures were grown in LB and culture supernatants and intact cells were analyzed for VesB activity and location by immunoblotting. In both WT and Δ*glpG* strains, VesB was active in the culture supernatants and cells (Figure 3.5A), and the size and distribution of VesB were the same according to immunoblot analysis (Figure 3.5B, lanes 2, 4, 9 and 11), suggesting that GlpG does not play a significant role in the localization or activation of VesB when rhombosortase is present. In the experiment shown in Figure 3.4, VesB-S221A was unstable and detected in two forms when overexpressed in the *rssP::kan* mutant and only the GlpG-cleaved form was found in the culture supernatant. Here, when VesB was expressed from its native promoter in the *rssP::kan* mutant it was only detected in the culture supernatant (Figure 3.5B, compare lanes 3 and 10). VesB in the *rssP::kan* mutant was inactive (Figure 3.5A) and migrated slower than VesB in the WT strain (Figure 3.5B; compare lanes 3 and 4) indicating that the GlpG-cleaved form of VesB does not undergo autoactivation. Furthermore, the VesB species that retained the GlyGly-CTERM extension could not be detected in cells expressing native VesB in the *rssP::kan* mutant (Figure 3.5B, lane 10), again suggesting that un-cleaved VesB is unstable and prone to degradation.

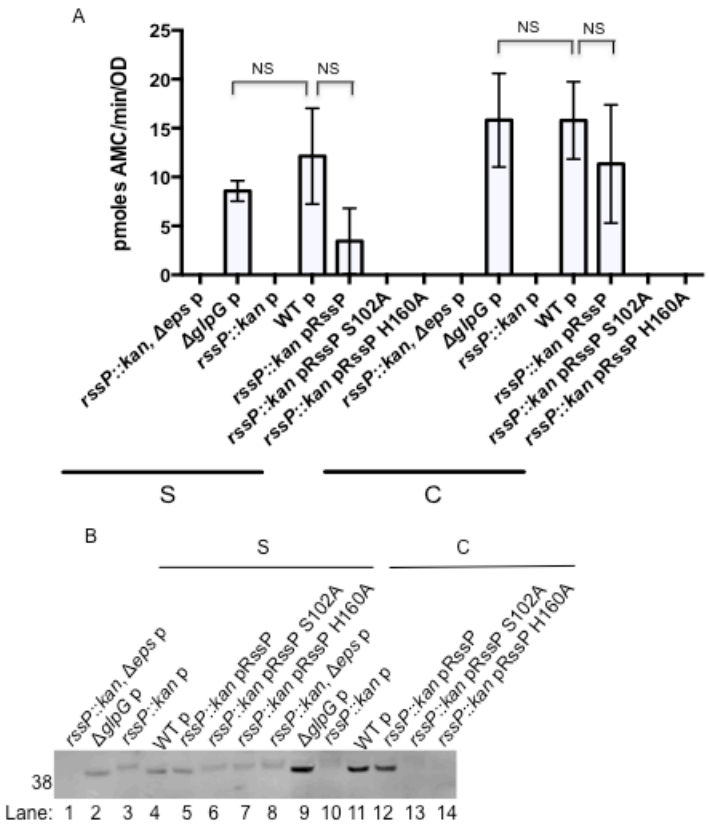

Figure 3.5: Native VesB is slightly larger, secreted, and inactive in the rhombosortase mutant. WT and the indicated mutants containing empty vector (p), pRssP, pRssP-S102A or pRssP-H160A were grown for 4 hours in LB broth with 100 µg/mL carbenicillin and 10 µM IPTG. Culture supernatants (S) and whole cells (C) were isolated and analyzed.

A. VesB is inactive in the rhombosortase mutant. Protease activity of VesB was measured using a fluorogenic peptide, Boc-Gln-Ala-Arg-AMC. All experiments were done in triplicates and standard error bars are shown. The difference in activity of VesB in the WT strain was statistically significant ($p<0.01$**) when compared to the activity of VesB in the other samples except for Δ*glpG*p and *rssp::kan* pRssP supernatants and cells. VesB activity in *rssP::kan/*p strain was not statistically significant when compared to the activity of VesB in the rhombosortase mutant complemented with the point mutants.

B. VesB migrates slower in the supernatant of the rhombosortase mutant. VesB protein in supernatants (S) and cells (C) were analyzed by SDS-PAGE and immunoblotting with VesB antibodies.

When the *rssP::kan* mutant was complemented with the plasmid pRssP, we observed partial restoration of VesB activity in the culture supernatant and cells (Figure 3.5A) and the same migration and localization pattern as that of VesB in the WT strain (Figure 3.5B, lanes 5 and 12). In addition, the overall amount of VesB protein increased (Figure 3.5B, lanes 5 and 12 compared to 3 and 10). No complementation of the *rssP::kan* mutant was apparent in the presence of pRssP-S102A and pRssP-H160A, which express mutant forms of rhombosortase in which the predicted catalytic residues serine and histidine, respectively, are substituted with alanine (Figure 3.5A and B, lanes 6, 7, 13 and 14). Furthermore, VesB was found only in the cells and had no activity in the *rssP::kan*Δ*eps* double mutant, suggesting that VesB found in the supernatant of the *rssP::kan* strain utilized the T2S system for extracellular transport (Figure 3.5A and B, lanes 1 and 8). Taken together, while active VesB is primarily associated with the surface of WT cells, with a smaller amount of active VesB released to the culture supernatant, VesB is only found in the culture supernatant in an inactive larger form in the absence of rhombosortase.

The propeptide of VesB is not removed in the *rssP::kan* mutant:

The larger size and lack of activity of VesB in the culture supernatant of the *rssP::kan* mutant suggested that GlpG-cleaved VesB does not undergo autoactivation. To further test this hypothesis, we took a proteomics approach that involved isolating and concentrating the culture supernatants of the WT and *rssP::kan* strains and subjecting them to off-line SDS-PAGE, gel segmentation,

in-gel digestion with trypsin, and LC/MS/MS. Many proteins were identified using this method, however, we focused on the GlyGly-CTERM domain containing proteins. Peptide mapping gave a VesB sequence coverage of 55% in the WT sample (Figure 3.6A). In the *rssP::kan* sample, which gave a sequence coverage of 57%, the N-terminal VesB peptide, ESTADISSR representing the propeptide of VesB, was present (Figure 3.6A). This peptide was absent in VesB from the WT sample, a finding that was also obtained through N-terminal sequencing of purified VesB (Gadwal et al, 2014). The presence of the propeptide of VesB in the *rssP::kan* mutant explains the larger size and the lack of activity of VesB in this mutant (Figure 3.5). Additionally, similar analysis of another GlyGly-CTERM containing trypsin-like protease, VesC, which resulted in 75 and 62% sequence coverage for the WT and *rssP::kan* mutant, respectively, revealed that the propeptide was also present on VesC when rhombosortase is absent (Figure 3.6B). In summary, C-terminal cleavage by rhombosortase is a prerequisite for N-terminal processing and thereby activation of VesB and VesC.

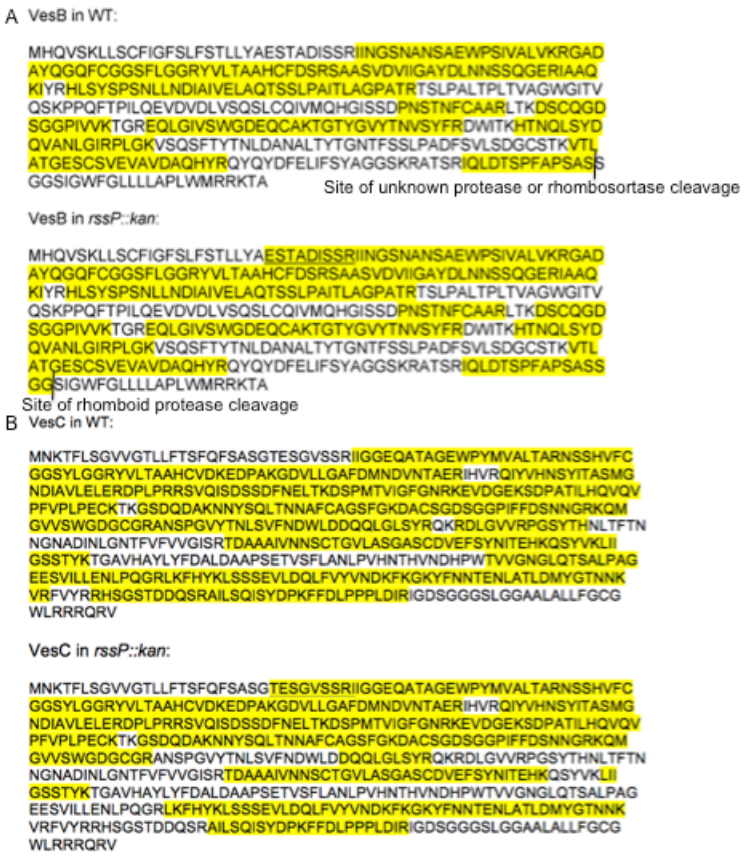

Figure 3.6: The propeptide is present in VesB and VesC in the rhombosortase mutant. Peptide mapping of VesB and VesC by mass spectrometry analysis of culture supernatants of WT and rhombosortase mutant (*rssP::kan*) strains. The yellow areas denote peptide coverage and the underlined sequence represents the N-terminal propeptide.

A. The peptide coverage of VesB.
B. The peptide coverage of VesC.

The lack of peptide coverage for the GlyGly-CTERM domain of VesB in the WT and *rssP::kan* samples supports the results presented in figure 3.4, which showed that VesB is cleaved by rhombosortase in WT cells and partially by GlpG in *rssP::kan* samples. The proteomics data of the *rssP::kan* sample identified a C-terminal VesB peptide, IQLDTSPFAPSASSGG, suggesting that VesB may be cleaved by GlpG at $Gly_{383}Ser_{384}$ resulting in the removal of most of the GlyGly-CTERM domain in the *rssP::kan* mutant (Figure 3.6A). The mechanism by which

VesB is released from the surface in WT cells is not known; however, the detection of the C-terminal peptide, IQLDTSPFAPSAS, in the WT sample suggests that release may occur through cleavage by a protease at Ser_{380} and Ser_{381} (Figure 3.6A). Alternatively, the linking of VesB to the membrane may be inefficient and a fraction of VesB may undergo complete hydrolysis following rhombosortase-mediated cleavage resulting in its release from the surface. If this is the case, Ser_{380}-Ser_{381} may represent the site of rhombosortase processing. The lack of peptide coverage for the GlyGly-CTERM domain in VesC is also suggestive of C-terminal processing in the WT and *rssP::kan* mutant samples (Figure 3.6B). Unfortunately, trypsin cleaves very close to the GlyGly-CTERM domain and may have generated a peptide that was too small to be resolved in these experiments.

Reduction in amounts of other GlyGly-CTERM proteins in the absence of rhombosortase:

The proteomic data were also used to compare the amount of VesB and the other five GlyGly-CTERM containing proteins in the culture supernatant of the WT and *rssP::kan* strains. Spectral counts for each protein were normalized to total number of spectra in each sample and protein size to obtain Normalized Spectral Abundance Factor (NSAF) values for each protein in the two samples (Table 3.1). We hypothesized that the amount of these proteins would be reduced in the culture supernatant of the *rssP::kan* mutant unless some of the other proteins are also cleaved by GlpG in the absence of rhombosortase. The

results showed that the amount of VesC and VCA0065 were significantly reduced by four-fold or more, while VesA and Xds were not detected in the absence of RssP (Table 3.1). Consistent with results from our immunoblotting experiments, VesB was only reduced by a factor of 1.5 in the *rssP::kan* mutant, while the level of VC1485 was increased by a factor of 2. However, the results for VesB and VC1485 were not statistically significant, likely due to GlpG's ability to cleave and thus release VesB and perhaps also VC1485 to the culture supernatant.

Table 3.1: Four of the six GlyGly-CTERM domain containing substrates are present in lower amounts in the supernatant of the rhombosortase mutant.

Protein Name	WT Peptide Coverage (%)	*rssP::kan* Peptide Coverage (%)	NSAF WT	NSAF *rssP::kan*	WT/*rssP::kan**
VCA0065	29	20	0.0023	0.0004	5.4
VesC	67	59	0.0127	0.0026	4.9
Xds	33	0	0.0015	0	-
VesA	11	0	0.0015	0	-
VesB	49	50	0.0023	0.0015	1.5
VC1485	59	67	0.0062	0.0125	0.5

***A ratio of ≤ 4 is considered statistically significant**

The GlyGly-CTERM domain is required for surface retention of VesB:

To determine the functional consequences of expressing VesB without an intact GlyGly-CTERM domain on VesB localization and activity, two VesB variants, VesBΔ5 and VesBΔ30 were constructed. VesBΔ5 has the positively charged residues at the C-terminus removed and VesBΔ30 has the additional Gly/Ser rich motif and hydrophobic region removed. These constructs and VesB

were expressed in Δ*vesABC* cells and at various time points, culture supernatants and cells were collected and analyzed for VesB localization and activity. While VesBΔ5 displayed a slightly higher activity than VesB in the culture supernatant (Figure 3.7A), no cell-associated activity could be detected (Figure 3.7B). No VesBΔ30 activity was detected in any fraction, except for low activity observed in the 8.5h culture supernatant sample (Figure 3.7A). The total combined activity in supernatants and cells of VesBΔ5 and VesBΔ30 was lower than the total activity in supernatants and cells of VesB, suggesting that both VesBΔ5 and VesBΔ30 are unstable. The blots correlated well with the activity assays in that VesB was present in the culture supernatant and cells, while VesBΔ5 was found mostly in the supernatant with a small amount cell-associated, and VesBΔ30 was detected exclusively in the supernatant (Figure 3.7C). The reduction in total amounts of protein and activity of VesBΔ5 and VesBΔ30 in comparison to VesB suggests that without the GlyGly-CTERM domain, VesB is unstable. More importantly, however, is the finding that the presence and the subsequent removal of the GlyGly-CTERM domain by rhombosortase are essential for the cell surface retention of VesB. In summary, the GlyGly-CTERM domain and the processing by rhombosortase are prerequisites for surface retention of VesB.

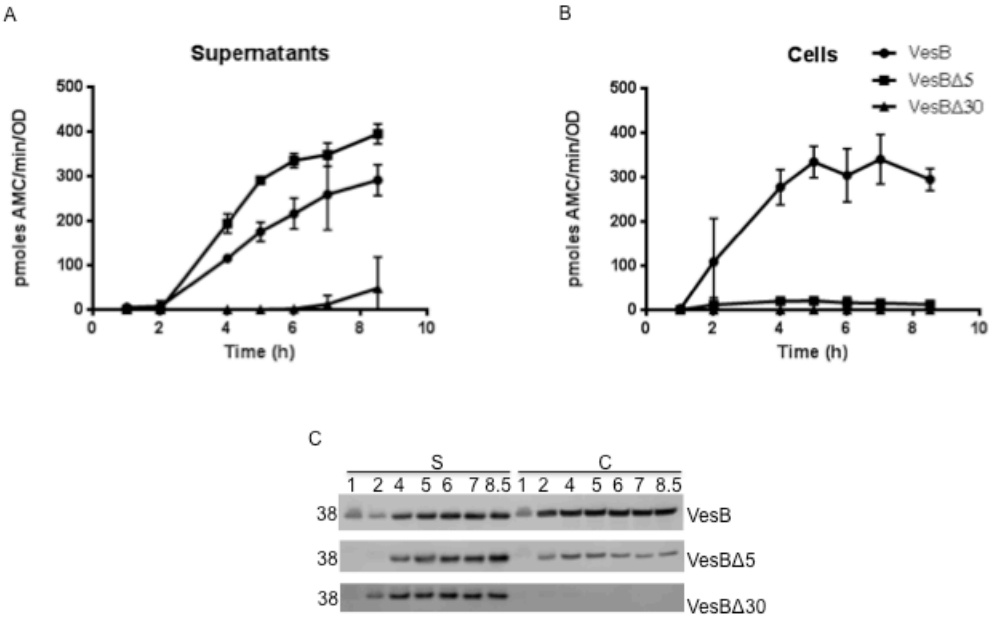

Figure 3.7: VesB is not retained on the cell surface when produced without the GlyGly-CTERM domain. A-C. Cultures of indicated strains expressing VesB or C-terminally deleted VesBΔ5 and VesBΔ30 were grown in LB broth with 100 µg/mL carbenicillin and 40 µM IPTG. Samples were collected at time points indicated and separated into culture supernatant (S) and cells (C). All experiments were done in triplicates and standard error bars are shown. p values were generated by comparing culture supernatants and cells containing VesB to those having VesBΔ5 and VesBΔ30. These were all statistically significant at the 4 - 8.5 hour time points in both supernatants and cells, except for VesB and VesBΔ5 at the 7 hour supernatant time point.

A. VesB and VesBΔ5 are active in culture supernatants. Using the fluorogenic peptide, Boc-Gln-Ala-Arg-AMC, the activity of VesB, VesBΔ5 and VesBΔ30 was measured in culture supernatants.

B. No surface-associated protease activity in cells producing GlyGly-CTERM domain mutants. The activity of VesB, VesBΔ5 and VesBΔ30 was measured in cells as in A.

C. VesB is not cell-associated in the absence of the GlyGly-CTERM domain. The culture supernatants (S) and cells (C) were subjected to SDS-PAGE and immunoblot analysis using VesB antibodies.

Cell-associated VesB can be extracted with Triton X-100:

Next, we determined the localization of VesB prior to engagement with the T2S system. Specifically, we wanted to know whether VesB accumulates as a periplasmic intermediate or a membrane intermediate in a T2S mutant. A common approach is to separate the inner and outer membrane by density centrifugation following cell lysis of bacteria such as *E. coli* and *Salmonella*, however, attempts to separate *V. cholerae* membranes have been unsuccessful. Therefore, we performed subcellular fractionation of WT and Δ*epsD* strains expressing plasmid-encoded Green Fluorescent Protein (pGFP) using a protocol that involves selective extraction of membrane proteins using a non-ionic detergent. The cells (C) were lysed using sucrose/lysozyme/EDTA and sonication followed by centrifugation to separate the soluble fraction (containing periplasmic and cytoplasmic proteins, P/C) from the membrane fraction. The membrane fraction was then treated with Triton X-100 and the samples were centrifuged to separate the triton soluble fraction (inner membrane proteins, IM) from insoluble material (outer membrane proteins, OM). The fractions were analyzed using SDS-PAGE and immunoblotting. Cytoplasmic GFP, inner membrane protein EpsL, and outer membrane protein OmpU were used as fractionation controls, as they are found in the soluble fraction, Triton X-100 soluble fraction, and Triton X-100 insoluble fraction, respectively (Figure 3.8A). VesB was mostly found in the Triton X-100 soluble fraction in the T2S mutant (Figure 3.8A), suggesting that VesB likely enters the T2S system as an inner membrane intermediate and may never enter the periplasmic compartment as a

soluble protein. Surprisingly, VesB was also predominantly detected in the Triton X-100 soluble fraction in WT cells suggesting that VesB remains tethered to the inner membrane, while exposed on the cell surface. Alternatively, VesB is localized in the outer membrane of WT cells in a Triton X-100 extractable form. The retention in the inner (or outer) membrane of rhombosortase-cleaved VesB may be due to additional post-translational modification, possibly through transamidation, of the new VesB C-terminus.

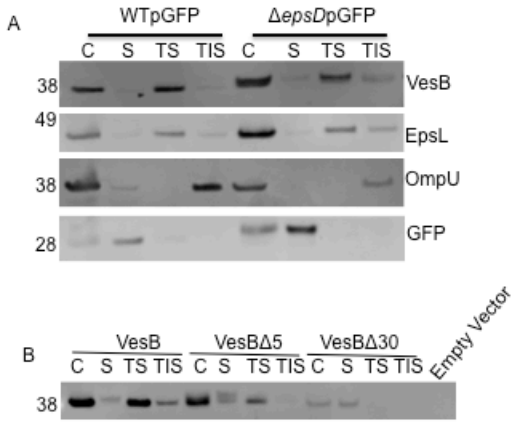

Figure 3.8: VesB is localized to the inner membrane prior to its engagement with the T2S system. Subcellular fractionation of cells was done as described in Material and Methods. Cells (C) were separated into a soluble fraction containing periplasmic and cytoplasmic proteins (P/C) and membrane fraction. The membrane was further fractionated into Triton X-100 soluble and insoluble fractions containing inner (IM) and outer membrane (OM) proteins, respectively.

A. Subcellular fractionation of WT and T2S mutant cells. Cells of WT and *epsD* mutant strains expressing plasmid encoded GFP were grown for four hours in LB with 100 µg/mL carbenicillin and 10 µM IPTG. Samples were analyzed by SDS-PAGE and immunoblotting with antibodies directed against VesB, GFP, EpsL, and OmpU.

B. Mutants of the GlyGly-CTERM domain are found in the soluble fraction. Subcellular fractions of a T2S mutant expressing WT and GlyGly-CTERM modified VesB. Samples of *ΔepsΔvesB* carrying pVesB and pVesBΔ5 were diluted 1:10 while the empty vector and pVesBΔ30 samples were undiluted and analyzed by SDS-PAGE and immunoblotting with VesB antiserum.

As VesBΔ30 and VesBΔ5 were predominantly released to the extracellular space and none or only a small amount of protein was detected in the cell samples, respectively (Figure 3.7), we predicted that these VesB variants are primarily secreted via a soluble periplasmic intermediate similar to cholera toxin (Hirst & Holmgren, 1987b). To test this, the T2S mutant (Δ*epsΔvesB)* cells overexpressing VesB, VesBΔ5, and VesBΔ30 were subjected to subcellular fractionation. As shown in figure 3.8B, VesB was mostly found in the Triton X-100 soluble fraction in the Δ*epsΔvesB/*pVesB strain. In contrast, VesBΔ5 was found in the soluble fraction containing periplasmic and cytoplasmic proteins and the Triton X-100 soluble fraction, while VesBΔ30 was found exclusively in the soluble fraction (Figure 3.8B). Taken together with findings presented above, these results suggest that VesB is capable of engaging the T2S apparatus whether it is tethered to the membrane or soluble in the periplasm. They also suggest that the location of VesB prior to its engagement with the T2S system dictates the final location of VesB. When VesB is transported by the T2S system from a membrane location, it will primarily remain surface-associated, while transport from a periplasmic location leads to release of VesB to the extracellular environment.

Discussion:

In this study, we used the trypsin-like serine protease VesB as a model protein to determine the relationship between the GlyGly-CTERM domain, rhombosortase, and the T2S system in *V. cholerae*. T2S substrates translocate

through the outer membrane and are either released into the extracellular milieu or retained on the cell surface through a variety of mechanisms. We show here that the majority of VesB is surface-associated; however, depending on growth conditions, various amounts of the surface retained VesB are released to the extracellular space (Figures 3.2 and 3.3). While many T2S substrates are found as soluble periplasmic secretion intermediates prior to their engagement with the T2S system, VesB is likely found in the inner membrane or possibly in a Triton X-100 extractable form in the outer membrane (Figure 3.8). Rhombosortase cleaves off the GlyGly-CTERM domain and may further modify the newly generated C-terminus in order to retain VesB on the cell surface once transported through the T2S system (Figures 3.3-3.5). When rhombosortase is absent, GlpG inefficiently cleaves the GlyGly-CTERM domain, however, in this case there is likely no additional post-translational modification and VesB is released to the extracellular milieu (Figures 3.4 and 3.5). Additionally, rhombosortase-mediated C-terminal processing subsequently leads to VesB activation, while cleavage by GlpG results in an inactive VesB that retains its N-terminal propeptide (Figure 3.6).

Taking all of the data together, we have assembled a model for successful delivery of VesB to the cell surface of *V. cholerae* that involves its GlyGly-CTERM domain, rhombosortase and the T2S system (Figure 3.9). VesB is synthesized in the cytoplasm with an N-terminal signal peptide and a C-terminal GlyGly-CTERM domain (1). VesB enters the Sec system via the signal peptide and, as the protein is translocated through the Sec system, the GlyGly-CTERM

domain is positioned in the inner membrane (2). The signal peptide is cleaved off (3) and VesB folds (4). Rhombosortase cleaves the GlyGly-CTERM domain (5) and possibly further modifies the new VesB C-terminus to prevent its release to the periplasmic compartment (6). VesB enters the T2S system to be delivered to the cell surface (7). Under some growth conditions, a fraction of VesB is then released from the cells with an as yet unknown mechanism. In the absence of rhombosortase, GlpG can cleave the GlyGly-CTERM domain and VesB engages with the T2S system for delivery to the extracellular milieu without going through a surface intermediate. The cleavage of VesB by GlpG is inefficient, however, and occurs only when rhombosortase is absent. Why VesB is preferentially cleaved by rhombosortase is not known; but it is possible that rhombosortase has a greater affinity than GlpG for VesB.

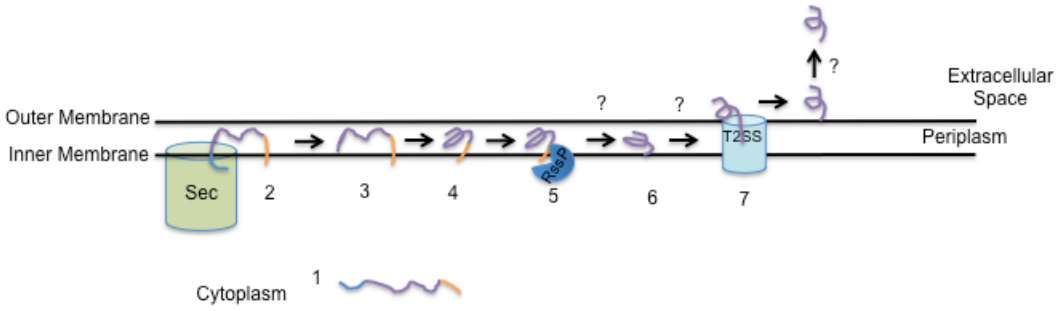

Figure 3.9: Model. VesB is produced in the cytoplasm with an N-terminal signal peptide (blue) and a C-terminal GlyGly-CTERM domain (orange) (1). VesB is directed to the Sec machinery via the signal peptide. As the protein is pulled into the periplasm via the Sec system, the GlyGly-CTERM domain is positioned in the inner membrane (2). The signal peptide is cleaved off (3) and the protein folds (4). Rhombosortase (RssP) binds and cleaves the GlyGly-CTERM domain (5). The newly generated C-terminus of VesB is post-translationally modified to maintain VesB in the membrane (6). VesB is recognized by the T2S system and translocated through the outer membrane to the cell surface (7). The modified C-terminus remains associated with the inner membrane or outer membrane thus preventing VesB from being released to the extracellular milieu. A fraction of surface-associated VesB is released through an additional processing event.

Initially, it was hypothesized that rhombosortase may cleave and thus release proteins from the cell (Haft & Varghese, 2011). We show here that although rhombosortase cleaves VesB, it also appears to promote its surface-retention, and therefore, the rhombosortase/GlyGly-CTERM system can be compared to the sortase/LPXTG system of Gram-positive organisms and the archaeosortase/PGF-CTERM system of archaea. In all of these systems, substrates contain a C-terminal tripartite domain consisting of a recognition motif (GlyGly, LPXTG or PGF) followed by a hydrophobic region and positively charged residues (Abdul Halim et al, 2013; Schneewind et al, 1993). In *Staphylococcus aureus*, sortase cleaves staphylococcal protein A (SPA) a LPXTG substrate, between the Thr and Gly residues, resulting in an acyl-enzyme intermediate that is resolved by lipid II, a precursor of peptidoglycan (PG), instead of H_2O (Marraffini et al, 2006). This results in PG-anchoring of SPA (Marraffini et al, 2006). While VesB may also undergo C-terminal transpeptidation rather than complete hydrolysis, the surface retained VesB is likely not attached to PG as it is of similar molecular weight as the surface released species, and if it were covalently attached to the PG it would not enter the separation gel when subjected to SDS-PAGE. VesB may, however, contain a short C-terminal lipid moiety that allows for membrane retention. The extraction of VesB from the cell envelope with Triton X-100 is supportive of this possibility (Figure 3.8). In the archaea system, archaeosortase appears to cleave or otherwise modify the PGF-CTERM domain of the S-layer glycoprotein. While it is

not known where archaeosortase cleaves or mediates transamidation to lipids, the S-layer glycoprotein is lipid modified at an unknown site (Abdul Halim et al, 2013) and can be extracted using Triton X-100. Further studies will be done to elucidate the mechanism by which VesB is retained on the *V. cholerae* cell surface.

Under some conditions, we observe more VesB in the supernatant than others (for example, in LB vs. M9 medium containing casamino acids and glucose; Figures 3.2 and 3.3). Further experiments will be done to determine exactly how VesB is released into the supernatant. Recently, cell wall attached sortase substrate SPA was found to be released into the culture supernatant by the removal of sugars from PG by LytN, a murine hydrolase, and polypeptide release by LytM, a pentaglycyl-endopeptidase (Becker et al, 2014). Additionally, in eukaryotes, glycosylphosphatidylinositol (GPI)-anchored proteins, like prostasin, can be released from the cell surface by lipid targeting enzymes like phospholipase C (Chen et al, 2001). VesB could similarly undergo an additional and perhaps regulated step of processing by another enzyme that allows it to be released from the cell just like SPA and prostasin. Proteomic analysis of WT culture supernatant suggests that release of VesB involves proteolysis.

The dual locations of VesB pose an interesting question of why proteins have the GlyGly-CTERM extension for surface association only to have a portion of the proteins released to the extracellular space. While it is possible that the release of surface retained VesB only occurs under certain laboratory growth conditions, the release may be of physiological importance. VesB and other

GlyGly-CTERM proteins, which are mostly hydrolytic enzymes, may be of benefit to the individual bacterium when retained on the surface where they can break down macromolecules and generate nutrients for immediate cellular uptake. On the other hand, in the context of a biofilm, the released form of VesB and the other GlyGly-CTERM proteins could benefit the community of bacteria. While the role of VesB and other GlyGly-CTERM proteins has only recently received attention, inactivation of the *vesB* gene results in reduced fitness in *V. cholerae* intestinal colonization (Fu et al, 2013), VC1485 may be essential for growth (Chao et al, 2013) and deletion of the *xds* gene yields unstructured and compact biofilms and also a slight deficiency in cellular detachment from biofilm (Seper et al, 2011). Another GlyGly-CTERM protein, ExeM, has been shown to promote stable biofilm formation of the environmental species *S. oneidensis* (Godeke et al, 2011).

In future studies, we will also investigate when and where the N-terminal propeptide is removed. Preliminary results suggest that VesB as well as VesC, another trypsin-like serine protease with a GlyGly-CTERM extension, may autoactivate once they are delivered to the cell surface. The lack of a surface intermediate for GlpG-cleaved VesB and VesBΔ30 may explain why they are inactive (Figures 3.5 and 3.6). The means by which VesB and VesC are maintained as zymogens until they reach the exterior of the cell may be a built-in mechanism to prevent premature activation of these serine proteases while intracellular, thus avoiding uncontrolled proteolysis that could otherwise be harmful to the cells.

In summary, VesB has a novel GlyGly-CTERM domain and utilizes rhombosortase and the T2S system to be correctly processed, translocated across the cell envelope, and delivered to the cell surface of *V. cholerae*. The rhombosortase/GlyGly-CTERM system offers a new alternative method of surface association that differs from previously observed mechanisms. PnlH from *D. dadantii* possesses a non-cleavable TAT specific signal peptide that is needed for its outer membrane attachment (Ferrandez & Condemine, 2008), heat-labile enterotoxin from ETEC is retained on the cell surface via an interaction with lipopolysaccharides (Horstman & Kuehn, 2002) and the lipidated N-terminus of pullulanase keeps it surface-associated in *K. oxytoca* (Pugsley et al, 1986). Furthermore, the newly acquired knowledge of the GlyGly-CTERM domain of VesB provides insight to the previous findings of ExeM and Xds from *S. oneidensis* and *V. cholerae*, respectively. Both of these GlyGly-CTERM containing proteins are detected in culture supernatants; however, they are also cell-associated, and based on our findings, they are likely retained on the cell surface in a process that involves their GlyGly-CTERM domains and rhombosortase (Brown et al, 2010; Godeke et al, 2011; Newland et al, 1985; Seper et al, 2011; Tang et al, 2007). Consistent with our suggestion that GlyGly-CTERM containing proteins are primarily localized to the cell surface, the four most abundant GlyGly-CTERM proteins, VesC, VC1485, VesB and VC0065, were recently detected in *V. cholerae* outer membrane vesicles by mass spectrometry (Altindis et al, 2014). In conclusion, the rhombosortase/GlyGly-

CTERM system presents a new sorting mechanism in *V. cholerae* and other Gram-negative bacteria.

Materials and Methods:

Bacterial strains and Plasmids:

The *V. cholerae* El Tor O1 strain, N16961, and the isogenic Δ*eps* (Sikora et al, 2007), Δ*epsD* (Sikora et al, 2009), Δ*vesB*, Δ*vesAC,* and Δ*vesABC* (Sikora et al, 2011) mutants were used in this study. All plasmids and primers are listed in Table 3.2. All polymerase chain reactions (PCR), cloning and restriction enzyme digestions were done with Phusion Polymerase, T4 DNA ligase and restriction enzymes from New England Biolabs and primers that were synthesized at IDT Technologies. pCRScript (Stratagene) and pMMB67EH constructs were transformed into *E.coli* MC1061 and pCVD442 constructs into SY327λpir. Triparental conjugation was performed with a helper strain, MM294 to transfer plasmids into N16961 and its isogenic mutants (Taylor et al, 1989).

A kanamycin insertional rhombosortase mutant (*rssP::kan)* was created by amplifying the rhombosortase (*rssP*) gene (VC1981) from *V. cholerae* N16961 chromosomal DNA using rssP primers. The fragment was ligated into a high copy plasmid, pCRscript. A kanamycin cassette was amplified from pK18mobsacB using kan primers containing *BclI* restriction enzyme sites and cloned into the native *BclI* site in *rssP*. The *rssP::kan* fragment was moved into the suicide plasmid, pCVD442. The rhomboid protease gene from *V. cholerae,*

glpG (VC0099) was deleted from the chromosome by amplifying 500 base pair regions upstream and downstream from *glpG* followed by overlap extension PCR of the two fragments to generate a 1.0 kbp fragment that was cloned into pCVD442.

Table 3.2: List of primers and plasmids used.

Gene or Plasmids	Genotypes	Forward Primer 5' to 3'	Reverse Primer 5' to 3'	Relevant Restriction Enzymes sites
pCRscrpt	Cloning vector, Ampr			
pMMB67EH	*Ptac promoter, ampR*			
pCVD442	*ori R6K mobRP4 sacB* Ampr			
pK18mobsacB	*ori-pMB1 oriT* (RP4) *sacB lacZ* Kan			
rssP		GCGGGATCCGAGTGGGCGCTTTTGTTC	TCCTGTTTATTGCGATAGTTTACG	SacI and SalI
kan		GCGTGATCACCGGAATTGCCAGCT	CGCTGATCATCAGAAGAACTCGTC	BclI
glpG - upstream fragment		GCGTCTAGATACATATTGTGGATCGAAAATCAC	AAATGCTAGTCCTTAACTCGCTTCGATGGG	SacI
glpG - downstream fragment		GGACTAGCATTTCCTCCTTTCTTGCGCCGC	CGCCCCGGGTGGTTCTGGCTGTAATGCCGTCGT	SalI
vesBΔ5		(Lybarger et al, 2009)	CGCCTGCAGTCACATCCACAATGGAGCCAAAA	EcoRI and PstI
vesBΔ30		(Lybarger et al, 2009)	CGCCTGCAGTCACGAAGTATCCAGTTGAATACGA	EcoRI and PstI
vesB-S221A		TCATGTCAGGGAGATGCTGGTGGCCCAATTGTA	TACAATTGGGCCACCAGCATCTCCCTGACATGA	
rssP-S102A - upstream		pmmb02: GAAATGAGCTGTTGACAATTAATC	CAGTCTTATGTTGGGCTCGCCAGCAGTTGG	SacI
rssP-S102A - downstream		GTTGGGCTCGCCGGCACCTTGCATGGTCTG	pmmb01: CTGTTTTATCAGACCGCTTCTGCG	SalI
rssP-H160A - upstream		pmmb02: GAAATGAGCTGTTGACAATTAATC	CGAGTAGCGACGGAAGCCGCTTTGGCCGGT	SacI
rssP-H160A - downstream		ACGGAAGCCGCTTTGGCCGGTTTAGTCGGC	pmmb01: CTGTTTTATCAGACCGCTTCTGCG	SalI

For complementation, *rssP* was amplified and cloned into pMMB67EH, a low copy vector with an IPTG inducible promoter and ampicillin cassette to generate pRssP. Using primers with base pair changes and the pRssP plasmid as template, fragments with *rssP*-S102A and *rssP*-H160A were made and cloned into pMMB67EH. *vesBΔ5* and *vesBΔ30* were amplified from pCRscript carrying *vesB* using newly created reverse primers with the forward primer originally used to clone *vesB* (Lybarger et al, 2009). PCR fragments were cloned into pMMB67EH. Quikchange Site Directed Mutagenesis Kit (Agilent) was used to create a mutation in pVesB to generate pVesB-S221A (Gadwal et al, 2014).

Growth Conditions:

Strains were either grown on Luria-Bertani (LB) agar/broth (Fisher) or M9 media containing 0.4% glucose and 0.4% casamino acids with 100 μg/mL of carbenicillin (Sigma) when plasmids were present.

Triton X-100 extraction:

Cultures were centrifuged for 10 min at 16,000 × g to separate culture supernatants and cells. Cell pellets were resuspended and treated with Tris/sucrose, EDTA, and lysozyme as described previously (Johnson et al, 2007). Then, 10 μg/ml of DNase and 500 μl of 50 mM Tris-HCl pH 8.0 were added and the samples were sonicated using 30% amplitude for 10 sec with 1 sec pulses. Lysed cells were centrifuged at 178,000 × g for 1 hour at 4°C to separate the soluble fraction (containing periplasmic and cytoplasmic proteins)

from the membrane fraction. The membrane pellet was then resuspended in 0.25% Triton X-100, 10 mM $MgCl_2$ in 50 mM Tris-HCl pH 8.0, incubated on ice for 30 min, and then centrifuged at 178,000 × g for 1 hour at 4°C. Following centrifugation, the supernatant contained the Triton X-100 extracted proteins. The pellet containing the Triton X-100 insoluble proteins was resuspended in 0.25% Triton X-100, 10 mM $MgCl_2$ in 50 mM Tris-HCl pH 8.0.

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and immunoblotting:

Samples were prepared and analyzed by SDS-PAGE and immunoblotting as described previously (Sikora et al, 2007). Polyclonal antiserum against VesB (Gadwal et al, 2014) was incubated with culture supernatant from the Δ*vesABC* mutant for 1 hour to pre-absorb cross-reactive antibodies prior to incubating with the nitrocellulose membrane for two hours (1:5,000). EpsL and OmpU (gift from K. Klose) antibodies were used at 1:20,000 and anti-GFP rabbit IgG (Molecular Probes) was used at 1:10,000 and incubated for two hours. Horseradish peroxidase-conjugated goat anti-rabbit immunoglobulin G (Bio-Rad) used at 1:20,000 was incubated with the membrane for one hour. Membranes were developed using ECL 2 Western blotting reagent (Thermo Scientific) and visualized using a Typhoon Trio variable mode imager system and Image Quant software.

Protease Assay:

V. cholerae supernatants and whole cells were measured for protease activity using *N-tert*-butoxycarbonyl-Gln-Ala-Arg-7-amido-4-methylcoumarin as described previously (Gadwal et al, 2014). A change in fluorescence per minute was calculated and converted to moles of methylcoumarin (AMC) cleaved per minute via a standard curve with known concentrations of AMC. The rate of AMC generation was normalized by OD_{600} of the cultures.

Serine hydrolase Labeling:

To 25 µl of filtered culture supernatants and washed cells, 0.5 µl of ActivX Serine Hydrolase Probe (Pierce) was added and incubated at room temperature for 30 minutes. Samples were visualized by SDS-PAGE and fluorescent imaging using the Typhoon imager system.

Cell Labeling and Microscopy:

Cells were washed, blocked with 2% BSA and incubated with 1:1000 of VesB antiserum that was pre-incubated with Δ*vesABC* cells to remove cross-reactive antibodies. Following incubation with 1:1000 of Alexa Fluor 488 $F(ab')_2$ goat anti-rabbit immunoglobulin G (Invitrogen) and washing, fluorescence was measured (Ex 495 nm/ Em 519 nm). These samples were also visualized by fluorescent microscopy using a Nikon Eclipse 90i fluorescence microscope equipped with a Nikon Plan Apo VC 100× (1.4 numerical aperture) oil immersion

objective and a CoolSNAP$_{HQ}^2$ digital camera as previously described (Lybarger et al, 2009).

Proteomics/GeLC/MS Analysis:

Culture supernatants of WT and *rssP::kan* strains were prepared as described with the following modifications (Johnson et al, 2013). Culture supernatants from three independent experiments were combined and 20µg of protein was processed by SDS-PAGE. The gel was stained with InstantBlue (Expedeon) and excised into ten equal sized segments using a grid pattern. Gel segments were digested with a ProGest robot (DigiLab) with the following protocol: washed with 25mM ammonium bicarbonate followed by acetonitrile, reduced with 10mM dithiothreitol at 60°C followed by alkylation with 50mM iodoacetamide at RT, digested with trypsin (Promega) at 37°C for 4h, quenched with formic acid, and the supernatant was analyzed directly without further processing.

Each gel digest was analyzed by nano LC/MS/MS with a Waters NanoAcquity HPLC system interfaced to a ThermoFisher LTQ Orbitrap Velos Pro. Peptides were loaded on a trapping column and eluted over a 75µm analytical column at 350nL/min; both columns were packed with Jupiter Proteo resin (Phenomenex). The mass spectrometer was operated in data-dependent mode, with MS performed in the Orbitrap at 60,000 FWHM resolution and MS/MS performed in the LTQ. The fifteen most abundant ions were selected for MS/MS. Data were searched using a local copy of Mascot with the following

parameters: Enzyme: Trypsin; Database: Swissprot *V. cholerae* El Tor N16961 (concatenated forward and reverse plus common contaminants); Fixed modifications: Carbamidomethyl (C); Variable modifications: Oxidation (M), Acetyl (N-term), Pyro-Glu (N-term Q), Deamidation (N,Q); Mass values: Monoisotopic; Peptide Mass Tolerance: 10 ppm; Fragment Mass Tolerance: 0.6 Da; Max Missed Cleavages: 2. Mascot DAT files were parsed into the Scaffold algorithm for validation, filtering and to create a non-redundant list per sample. Data were filtered using protein and peptide thresholds of 99% and 95%, respectively, and requiring at least two unique peptides per protein. The complete mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository (Vizcaino et al, 2013) with the dataset identifier PXD000896. Data for the six GlyGly-CTERM proteins are shown in Table 3.1.

Statistical Analysis:

t tests were done on WT and mutant samples as indicated. Results yielding a P value of <0.05 were considered statistically significant.

CHAPTER 4:

Discussion

The goal of my dissertation work was to characterize the substrate specificity, as well as the catalytic triad and activation site of the protease domain of *Vibrio cholerae* extracellular serine protease B, VesB. Additionally, I sought to identify the roles of rhombosortase, the type II secretion (T2S) system, and the GlyGly-CTERM domain in surface localization of VesB. This chapter encapsulates key findings, the implications and future directions of my work.

Function and Structure of VesB:

In chapter 2, I purified VesB from culture supernatants to determine substrate specificity with synthetic peptides and assess VesB's ability to degrade intact proteins. VesB cleaves at arginine residues in synthetic peptides and is capable of cleaving proteins like casein, BSA, lactoferrin and IgA. Together, this suggests that VesB has a substrate specificity similar to trypsin and may be a general degradative enzyme that cleaves intact proteins. These findings indicate that a possible role of VesB is to provide nutrients for *V. cholerae* growth like other T2S substrates, HapA and ChiA

(Connell et al, 1998; Halpern et al, 2003). However, in the small intestine nutrients are in excess, perhaps indicating that VesB may not be involved in nutrient acquisition but may instead have a specialized role in infection or colonization. Additionally, VesB could function in the context of a biofilm, where it may degrade proteins and provide nutrients to the biofilm.

The role of Tyr250 in regulating VesB activity was investigated by measuring protease activity in the presence of increasing concentrations of NaCl. Despite the presence of this residue, Na^+ did not enhance VesB activity but rather slightly reduced the activity, suggesting that VesB may function in both niches of *V. cholerae*, the human intestine and environmental reservoirs where the Na^+ concentrations vary. In other experiments, we showed that the propeptide is removed in active VesB and mutagenesis of the activation site and the catalytic triad separately, led to undetectable VesB activity while the outer membrane translocation of VesB was unaffected. Taken together, these data suggest that the catalytic triad and removal of the propeptide are required for VesB activity.

The structure of an inactive VesB (VesB-S221A) without the GlyGly-CTERM domain was determined by our colleagues Drs. Konstantin Korotkov and Wim Hol. The protease domain contains a trypsin-like fold that can be superimposed on trypsinogen. The C-terminal domain has an immunoglobulin (Ig)-fold composed of β-sheets and α-helices. The superposition of the C-terminal domain of VesB on cucumisin and SfaE, two structural homologs, revealed that VesB contained two additional α-helices and β-sheets. The

significance of this finding is, at the moment, unknown. In chapter 2, I discuss the possible functions of the C-terminal Ig-fold domain in VesB. While the Ig-fold domain may convey stability, as VesB without the C-terminal domain cannot be detected (data not shown), it could also function as a T2S recognition signal. A role in substrate specificity is also a possibility, as the Ig-fold may confer higher affinity or specificity for a subset of substrates. We also proposed that the Ig-fold domain might be involved in cell surface binding. While the synthesis of a VesB variant without the GlyGly-CTERM domain is no longer cell associated (chapter 3), suggesting a requirement for this domain in cell association, the Ig-fold domain may also contribute to cell surface binding. Taken together, VesB is a trypsin-like serine protease that cleaves a wide range of substrates and the modular domains of VesB contain a trypsinogen-like protease domain and a C-terminal Ig-fold domain.

The GlyGly-CTERM domain, rhombosortase, and the T2S system:

In chapter 2, I focused on secreted VesB. While VesB was originally identified as a T2S substrate (Sikora et al, 2011), a large portion of VesB was found associated with cells (chapter 3). Therefore, I focused on the cell-associated form of VesB. VesB contains a C-terminal GlyGly-CTERM extension, which is composed of a stretch of glycines/serines, followed by hydrophobic and positively charged residues, therefore inferring that this domain possibly spans the cytoplasmic membrane (Haft & Varghese, 2011). Furthermore, in Gram-negative bacteria that encode a single GlyGly-CTERM containing protein, next to

its gene, is a gene encoding rhombosortase, a rhomboid-like protease, suggesting a possible functional relationship between the two (Haft & Varghese, 2011). In fact, rhomboid proteases are intramembrane serine proteases that are known to cleave single transmembrane domain proteins (Akiyama & Maegawa, 2007; Strisovsky et al, 2009).

My initial hypothesis was that VesB is cleaved at the GlyGly-CTERM domain by rhombosortase, thereby releasing it into the culture supernatant. However, in chapter 3 I showed that both the cell-associated and secreted forms of VesB have their GlyGly-CTERM domain removed. Furthermore, the cell-associated form of VesB is surface-localized and active. To gain a better understanding of the cell surface VesB material, subcellular fractionation was conducted and the results showed that VesB is extracted with Triton X-100 in WT cells. This suggests that VesB may undergo lipid modification by rhombosortase or another enzyme, and this allows for surface retention through lipid-lipid interaction. Alternatively, VesB may be attached to the cell surface through a non-covalent interaction with a membrane protein, which is Triton X-100 extractable. These findings led to the hypothesis that the GlyGly-CTERM domain and rhombosortase play a role in localizing VesB to the cell surface. The two pieces of evidence that support this are 1) when VesB is expressed without the GlyGly-CTERM domain, it is no longer cell-associated, and 2) in a rhombosortase mutant, native VesB is not detected in cell samples. In fact, in the latter case, the amount of VesB is reduced and only found in the culture supernatant due to rhomboid protease-mediated cleavage of VesB, suggesting

that rhombosortase may control the stability of VesB. Rhombosortase-mediated cleavage may result in the correct positioning of the C-terminal Ig-fold domain resulting in cell surface retention and/or stability of VesB. Taken together, chapter 3 demonstrates a novel sorting mechanism of Gram-negative bacteria, where the presence and subsequent removal of the GlyGly-CTERM domain of VesB by rhombosortase promotes cell surface retention possible through a lipid moeity.

While studying rhombosortase, VesB and its GlyGly-CTERM domain in chapter 3, I made additional key discoveries related to the GlyGly-CTERM/rhombosortase system. First, I discovered that cleavage of the GlyGly-CTERM domain of VesB by rhomboid protease resulted in a phenotype different than cleavage by rhombosortase. When I focused on the C-terminal processing, I saw that VesB is cleaved by rhombosortase in WT, as well as T2S and rhomboid protease mutant cells. However, in the absence of rhombosortase, rhomboid protease cleaves VesB, which is then released to the culture supernatant. Additionally, when rhombosortase cleaves the C-terminus, VesB is active, while rhomboid protease-mediated cleavage results in inactive VesB. The reason for this lack of activity is due to the propeptide still remaining in rhomboid protease-cleaved VesB. This unexpected, yet exciting finding led to the hypothesis that the activation of VesB occurs at the cell surface. In support of this, the fully secreted VesB variant without its GlyGly-CTERM domain has only a low, barely detectable activity.

Another key finding identified through this work is the ratio of VesB in culture supernatants and cells when grown under different conditions. In LB, overexpression of VesB leads to an even split of 50% of VesB present in the supernatant and 50% cell-associated. However, when grown in M9 media supplemented with casamino acids and glucose, overexpression of VesB leads to a majority of cell–associated VesB, with only a small amount released into the culture supernatant. While the ratio of VesB distribution differs, VesB is active under both conditions. The release of VesB from the cell surface may, therefore, have to do with an additional processing event that only occurs under certain growth conditions. Enzyme-mediated release has also been reported for cell surface anchored proteins in Gram-positive bacteria and GPI-anchored proteins in eukaryotes (Becker et al, 2014; Verghese et al, 2006).

This chapter highlights the significance of the GlyGly-CTERM extension in VesB. Other proteins like Xds and ExeM that contain a GlyGly-CTERM domain have also been detected as extracellular and cell associated proteins (Brown et al, 2010; Godeke et al, 2011; Newland et al, 1985; Seper et al, 2011; Tang et al, 2007). The exact role of the GlyGly-CTERM domain in these proteins was not investigated; however, proteomic analysis of *V. cholerae* culture supernatants showed that Xds is not detected in the absence of rhombosortase (Chapter 3). Therefore, the information learned from our studies of VesB can be extrapolated to these nucleases. Interestingly, Xds and ExeM were both shown to be involved in biofilm formation and maintenance (Godeke et al, 2011; Seper et al, 2011). This suggests that in addition to Xds and ExeM, VesB and other GlyGly-CTERM

domain containing enzymes may also be active in biofilms. It is possible that during planktonic growth, cell surface-retained VesB primarily benefits the individual cell by degrading substrates for immediate uptake of nutrients, while the released form of VesB may dominate during biofilm growth where it may be beneficial to the entire community.

Working Model:

Taking all of the experimental data and hypotheses together, I present a working model for the mechanism of VesB biogenesis, translocation, and function (Figure 4.1). VesB is made in the cytoplasm with its N-terminal signal peptide (blue) and propeptide (green), followed by the protease domain and C-terminal domain (purple) with the GlyGly-CTERM extension (orange) (1). The protein enters the Sec system via the signal peptide and the GlyGly-CTERM domain may be placed into the inner membrane (2). The signal peptide is cleaved off (3) and the protein folds (4). Rhombosortase (RssP) cleaves the GlyGly-CTERM domain (4) and a post-translational modification like lipidation may occur (indicated by the blue asterisk) (5). VesB may stay linked to the inner membrane, and the β-sheet rich C-terminal domain of VesB may be recognized by the T2S system to gain entry (6). VesB translocates through the outer membrane (7) and is delivered to the cell surface, where it remains possibly linked by a lipid moiety (8). Once VesB is on the surface it may self-activate, leading to the removal of the propeptide (9). While the cell surface may be the final location of VesB, there are growth conditions where an additional

processing event either from a secreted or intracellular enzyme might lead to the release of VesB from the cell surface (9 and 10). Both the secreted and cell-surface associated VesB are proteolytically active and may generate smaller peptides and/or amino acids that could be used to benefit the community of cells in a biofilm or be taken up immediately by the individual cell to use as nutrients.

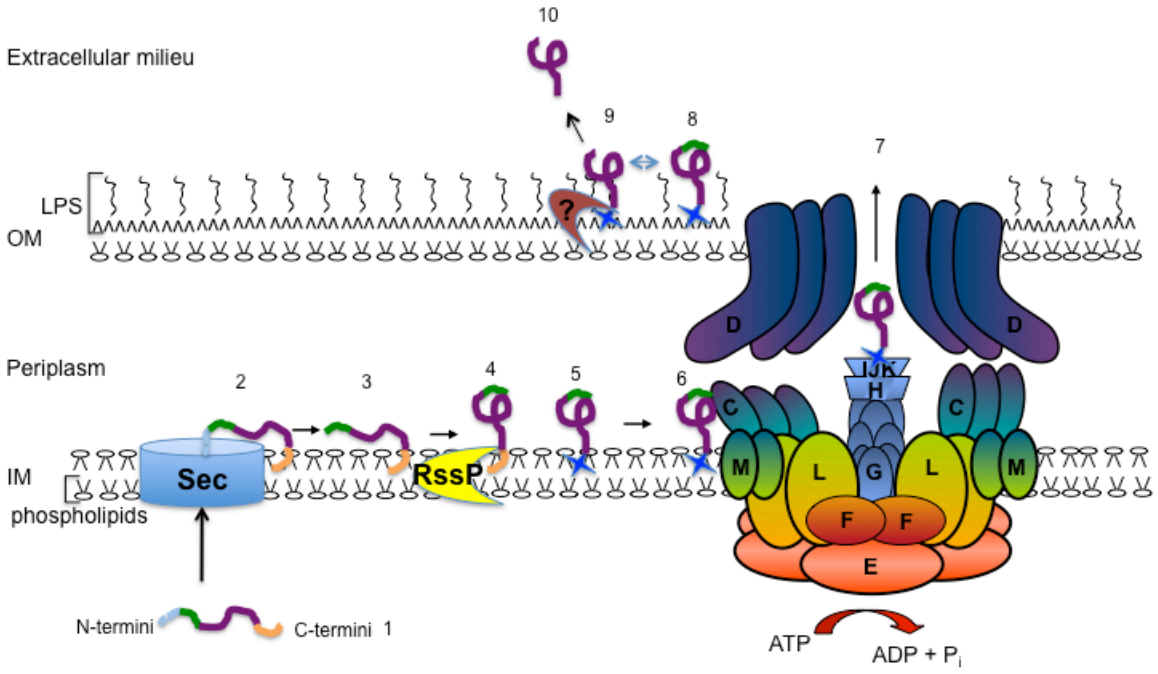

Figure 4.1: Model for VesB biogenesis, translocation, and function.
1. VesB is made in cytoplasm with a signal peptide (blue), propeptide (green), protease domain and C-terminal domain (purple) with the GlyGly-CTERM domain (orange). 2. VesB is translocated through the inner membrane by recognition of the signal peptide by the Sec system. The GlyGly-CTERM may be placed into the inner membrane. 3. The signal peptide is cleaved. 4. The protein folds and rhombosortase cleaves the GlyGly-CTERM domain. 5. Post-translational modification, such as lipidation, may occur by rhombosortase or another enzyme. 6. The C-terminal domain of VesB may be recognized by the T2S system. 7. VesB enters the T2S system to be translocated through the outer membrane. 8. The lipid moiety might allow for cell surface-associated VesB. 9. VesB may self-activate and the propeptide is removed. 10. An additional processing event may occur under certain growth conditions to allow for release of VesB into the extracellular space. Both the extracellular and cell-associated forms of VesB may degrade substrates to be used in biofilms or by the individual cell.

Future Directions:

While the new information here about the GlyGly-CTERM/rhombosortase system has contributed to the overall knowledge of proteins with the GlyGly-CTERM sorting motif, additional questions remain to be answered. First, I want to identify the modification that occurs during processing of VesB by rhombosortase that is responsible for cell surface retention. In the LPXTG/sortase system, a nucleophilic attack by lipid II, a precursor of peptidoglycan, allows for covalent linking of the substrate to peptidoglycan (Marraffini et al, 2006). While we know that rhombosortase cleaves the GlyGly-CTERM domain, the nucleophile that resolves the acyl-enzyme VesB/rhombosortase intermediate remains to be elucidated. Because VesB is present on the cell surface and can be extracted with Triton X-100 like the archeal PGF-CTERM containing S-layer glycoprotein, which is lipid modified (Kandiba et al, 2013; Konrad & Eichler, 2002), this nucleophile may be a lipid. In Gram-negative bacteria, the inner and outer membranes are composed of proteins, lipoproteins, lipopolysaccharides (LPS) and phospholipids (Raetz et al, 2007). This leads to the hypothesis that the potential lipid moiety of VesB may originate from phospholipids or LPS in the *V. cholerae* cell envelope. Specifically, during synthesis of phospholipids or LPS, a precursor molecule may complete the nucleophilic attack, which would result in VesB being incorporated into either phospholipid or LPS.

In *E. coli* and many other Gram-negative bacteria, the inner membrane is a phospholipid bilayer while the outer membrane is a heterologous bilayer consisting of phospholipids and LPS. Phospholipids are relatively small

molecules, weighing less than 1,000 Daltons and are composed of hydrophobic fatty acid chains linked to glycerol 3-phosphate and hydrophilic head groups. In bacteria, phosphatidylethanolamine (PE), phosphatidylglycerol (PG), and cardiolipin (CL) are the major phospholipids of the inner and outer membranes (Figure 4.2). Phosphatidylethanolamine (PE) contains ethanolamine as its head group and the amine can either be positively charged under low pH or uncharged under neutral pH (Cronan, 2003). LPS consists of a lipid A tail (~1,800 Daltons) and polysaccharides that contain the inner core, outer core, and the O-antigen. The lipid A tail makes up the outer leaflet of the outer membrane (Raetz et al, 2007). In the El Tor strains of *V. cholerae*, lipid A of LPS contains either a glycine or diglycine that confers resistance to the antimicrobial peptide, polymyxin B (Hankins et al, 2012) (Figure 4.3). The glycine or diglycines are added onto the lipid A tail in a process involving three *V. cholerae* proteins, AlmG, F and E. Insertional mutants of *almG, F* and *E* synthesize LPS without the glycines and expression of this operon in a classical strain of *V. cholerae*, results in LPS with glycines (Hankins et al, 2012). Knowing this information, I propose that the molecule that resolves the acyl-enzyme VesB/rhombosortase intermediate, may be either the amine group of phosphatidylethanolamine or the glycine of lipid A in LPS. Because of the relatively small sizes of phospholipids and lipid A, the addition of either of these molecules would not significantly change the molecular weight of VesB. This is consistent with our findings, which showed that the surface-bound and extracellular forms of VesB migrate similarly when analyzed by SDS-PAGE.

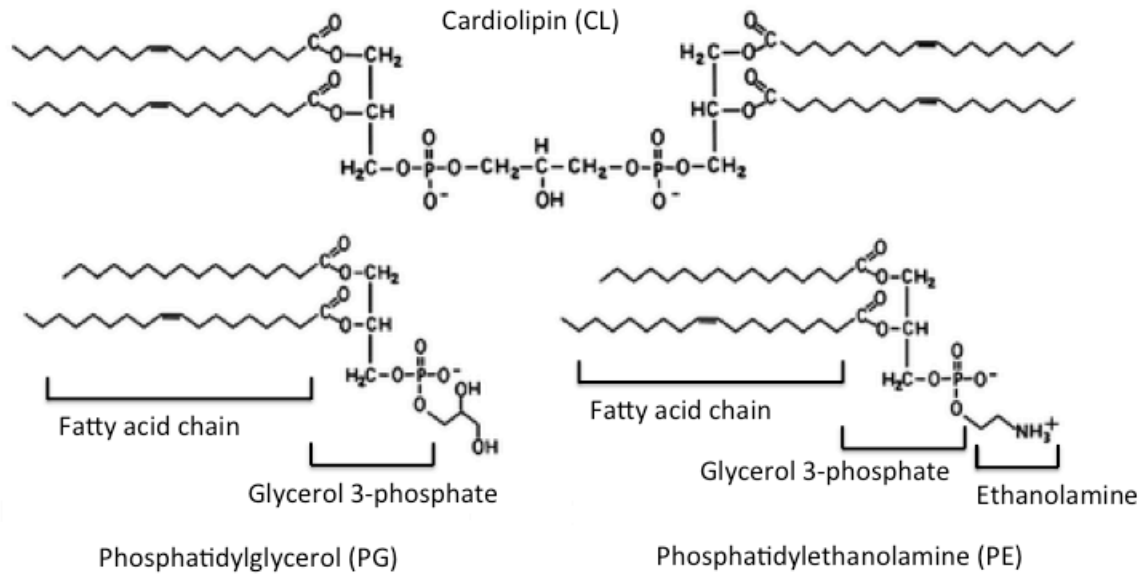

Figure 4.2: The molecular structure of phospholipids in Gram-negative bacteria.
The three major phospholipids of Gram-negative bacteria are cardiolipin, phosphatidyglycerol, and phosphatidylethanolamine. In general, all contain a hydrophobic fatty acid chain, which is then linked to glycerol 3-phosphate and the hydrophilic head group. Taken from (Epand & Epand, 2009).

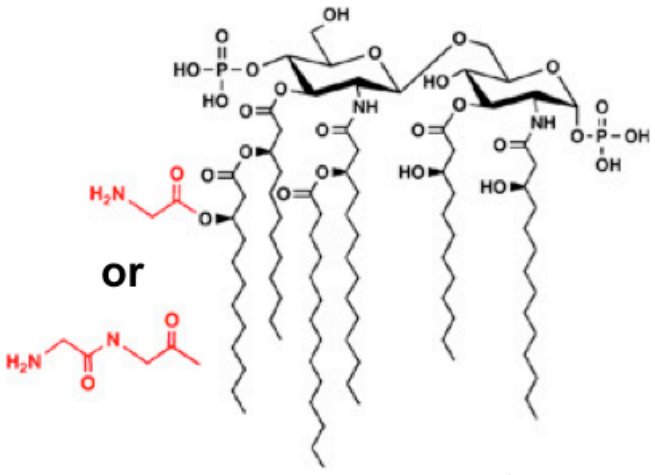

Figure 4.3: The structure of lipid A of lipopolysaccharides (LPS) from El Tor strains of *V. cholerae*.
LPS is composed of a lipid A tail which makes up the outer leaflet of the outer membrane. The lipid A of *V. cholerae* contains either a single or double glycine (in red) attached to the hexa-acylated chain. Taken from (Hankins et al, 2012).

To test the first possibility that the nucleophile is the amine of phosphatidylethanolamine, I will take *V. cholerae* cells overexpressing VesB and treat them with commercially available phospholipases that target specific sites within the phospholipid molecule. After incubation, the supernatant and cells will be separated and analyzed by SDS-PAGE and immunoblotting to see whether VesB is released from cells. Several techniques will be used to determine if the nucleophile is the glycine of lipid A in LPS. In the El Tor biotype of *V. cholerae*, VesB is on the cell surface, however the location of VesB in the classical strains of *V. cholerae,* which lack the glycine(s) in lipid A (Hankins et al, 2012), is not known. Experiments will be done to determine the distribution of VesB in classical strains such as O395. If active VesB is not detected on the cell surface of O395, further studies will be done to determine if the glycines of lipid A are responsible for the nucleophilic attack. *V. cholerae* strains of the El Tor biotype lacking the *almE-G* genes will be constructed and the distribution of VesB will be determined by SDS-PAGE and immunoblotting as well as fluorescence microscopy (described in Chapter 3). If the glycine is not responsible for the nucleophilic attack, another component of lipid A could serve the same purpose. For example, in various Gram-negative bacteria, PmrC adds phosphoethanolamine to lipid A (Beceiro et al, 2011; Kim et al, 2006; Lee et al, 2004), thus suggesting ethanlamine could complete the cleavage reaction by rhombosortase. In parallel experiments, VesB will be extracted from the cell surface, subjected to proteolysis using different enzymes and analyzed by mass spectrometry to determine the nature of the C-terminal modification.

In the case of GPI-anchored serine proteases, the C-terminal transmembrane domain is first cleaved off and then the newly generated C-terminus is linked to ethanolamine of GPI by GPI transamidase (Antalis et al, 2011; Chen et al, 2001). As discussed above, rhombosortase may act similarly to GPI transamidase. However, it is also possible that the cleavage and modification of VesB could happen by two different enzymes. A transposon mutagenesis screen will be constructed to identify the modifying enzyme.

In chapter 3, I demonstrate that both the extracellular and surface attached VesB have their GlyGly-CTERM domain removed. The cell-associated form is retained on the surface possibly through a lipid modification. Therefore, I want to determine whether an additional cleavage step resulting in the removal of the modification allows for VesB to be released from the surface into the extracellular space. If the above experiments show that the modification is due to a lipid moiety, mutants of *V. cholerae* lacking phospholipases or lipases will be made and the distribution of VesB in the supernatant and cells will be observed via SDS-PAGE and immunoblotting. If a phospholipase or lipase is identified to process VesB and release it into the supernatant, studies will be done on the regulation of this protein under different growth conditions. Alternatively, if VesB is cleaved by rhombosortase and then modified by another enzyme, the regulation of this additional enzyme could be different in various growth conditions. This can give insight into the differences of VesB distribution when cells are grown in LB and M9 media.

Further experiments will be done to elucidate how rhomboid protease and rhombosortase function. The VesB experiments done in chapter 3 showed that C-terminal processing by rhomboid protease only happens in the absence of rhombosortase and results in secreted inactive VesB, whereas rhombosortase-mediated cleavage results in cell-associated active VesB. The functional differences of these two enzymes may be due to structural differences. First, rhomboid protease contains an additional N-terminal cytoplasmic domain that may affect the mechanism by which it processes substrates. Second, although rhombosortase appears to display the same membrane topology as rhomboid protease, the composition of their active sites and surrounding areas may differ. One way to address the diference in activity of these proteases is to swap helices of both proteins or add the N-terminal domain of rhomboid protease to rhombosortase and observe the activity and distribution of VesB.

The experiments in chapter 3 suggest that rhombosortase cleaves the GlyGly-CTERM domain whether the T2S system is present or not, suggesting that removal of the GlyGly-CTERM domain precedes outer membrane translocation. Furthermore, subcellular fractionation shows that VesB may enter the T2S system through an inner membrane intermediate similar to PulA, the lipoprotein from *Klebsiella oxytoca* (Pugsley & Kornacker, 1991; Pugsley et al, 1990) rather than a periplasmic intermediate. It is believed that T2S substrates enter the T2S system via EpsD and/or EpsC whether they are soluble periplasmic intermediates or tethered to the inner or outer membrane (Korotkov et al, 2012). Further studies will be done to elucidate the mechanism of how lipid-

modified VesB enters the T2S system from the inner membrane position. This may however be difficult if the entry is rapid. Lipoproteins and other molecules that have fatty acid chains are moved throughout the cell with the help of proteins that protect the lipid moieties (Okuda & Tokuda, 2011; Raetz et al, 2007). VesB may move through the T2S system with its lipid modification protected. Further studies will be done to identify if a component of the T2S system or VesB itself may protect the lipid moiety.

Conclusions:

The work presented here is the first to demonstrate the substrate specificity, structure, and importance of the catalytic triad and activation site for a bacterial two-modular trypsin-like serine protease. Apart from the extensively studied LPXTG/sortase of Gram-positive bacteria and the lesser-known PGF-CTERM/archeosortase systems, the GlyGly-CTERM/rhombosortase system, with the assistance of T2S, is a newly described sorting mechanism of cell-surface proteins in Gram-negative bacteria. The information on the sorting mechanism of VesB is applicable to other GlyGly-CTERM containing proteases like Xds and ExeM that have been shown to play a role in biofilms (Godeke et al, 2011; Seper et al, 2011). Biofilms are important to maintain the survival of bacteria in host organisms or environmental reservoirs, thus imposing a threat to human health. Therefore, the GlyGly-CTERM/rhombosortase system presents a novel target for therapeutics to prevent disease.

References

Abdul Halim MF, Pfeiffer F, Zou J, Frisch A, Haft D, Wu S, Tolic N, Brewer H, Payne SH, Pasa-Tolic L, Pohlschroder M (2013) Haloferax volcanii archaeosortase is required for motility, mating, and C-terminal processing of the S-layer glycoprotein. *Molecular microbiology* **88:** 1164-1175

Akiyama Y, Maegawa S (2007) Sequence features of substrates required for cleavage by GlpG, an Escherichia coli rhomboid protease. *Molecular microbiology* **64:** 1028-1037

Altindis E, Fu Y, Mekalanos JJ (2014) Proteomic analysis of Vibrio cholerae outer membrane vesicles. *Proceedings of the National Academy of Sciences of the United States of America*

Antalis TM, Bugge TH, Wu Q (2011) Membrane-anchored serine proteases in health and disease. *Progress in molecular biology and translational science* **99:** 1-50

Beceiro A, Llobet E, Aranda J, Bengoechea JA, Doumith M, Hornsey M, Dhanji H, Chart H, Bou G, Livermore DM, Woodford N (2011) Phosphoethanolamine modification of lipid A in colistin-resistant variants of Acinetobacter baumannii mediated by the pmrAB two-component regulatory system. *Antimicrobial agents and chemotherapy* **55:** 3370-3379

Becker S, Frankel MB, Schneewind O, Missiakas D (2014) Release of protein A from the cell wall of Staphylococcus aureus. *Proceedings of the National Academy of Sciences of the United States of America* **111:** 1574-1579

Ben-Shem A, Fass D, Bibi E (2007) Structural basis for intramembrane proteolysis by rhomboid serine proteases. *Proceedings of the National Academy of Sciences of the United States of America* **104:** 462-466

Bhowmick R, Ghosal A, Das B, Koley H, Saha DR, Ganguly S, Nandy RK, Bhadra RK, Chatterjee NS (2008) Intestinal adherence of Vibrio cholerae involves a coordinated interaction between colonization factor GbpA and mucin. *Infection and immunity* **76:** 4968-4977

Booth BA, Boesman-Finkelstein M, Finkelstein RA (1984) Vibrio cholerae hemagglutinin/protease nicks cholera enterotoxin. *Infection and immunity* **45:** 558-560

Bouley J, Condemine G, Shevchik VE (2001) The PDZ domain of OutC and the N-terminal region of OutD determine the secretion specificity of the type II out pathway of Erwinia chrysanthemi. *Journal of molecular biology* **308:** 205-219

Brooks CL, Lemieux MJ (2013) Untangling structure-function relationships in the rhomboid family of intramembrane proteases. *Biochimica et biophysica acta* **1828:** 2862-2872

Brown RN, Romine MF, Schepmoes AA, Smith RD, Lipton MS (2010) Mapping the subcellular proteome of Shewanella oneidensis MR-1 using sarkosyl-based fractionation and LC-MS/MS protein identification. *Journal of proteome research* **9:** 4454-4463

Camberg JL, Johnson TL, Patrick M, Abendroth J, Hol WG, Sandkvist M (2007) Synergistic stimulation of EpsE ATP hydrolysis by EpsL and acidic phospholipids. *The EMBO journal* **26:** 19-27

Camberg JL, Sandkvist M (2005) Molecular analysis of the Vibrio cholerae type II secretion ATPase EpsE. *Journal of bacteriology* **187:** 249-256

Chao MC, Pritchard JR, Zhang YJ, Rubin EJ, Livny J, Davis BM, Waldor MK (2013) High-resolution definition of the Vibrio cholerae essential gene set with hidden Markov model-based analyses of transposon-insertion sequencing data. *Nucleic acids research* **41:** 9033-9048

Chen LM, Skinner ML, Kauffman SW, Chao J, Chao L, Thaler CD, Chai KX (2001) Prostasin is a glycosylphosphatidylinositol-anchored active serine protease. *The Journal of biological chemistry* **276:** 21434-21442

Chen VB, Arendall WB, 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta crystallographica Section D, Biological crystallography* **66:** 12-21

Cianciotto NP (2005) Type II secretion: a protein secretion system for all seasons. *Trends in microbiology* **13:** 581-588

Colwell RR (1996) Global climate and infectious disease: the cholera paradigm. *Science* **274:** 2025-2031

Connell TD, Metzger DJ, Lynch J, Folster JP (1998) Endochitinase is transported to the extracellular milieu by the eps-encoded general secretory pathway of Vibrio cholerae. *Journal of bacteriology* **180:** 5591-5600

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

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

Cronan JE (2003) Bacterial membrane lipids: where do we stand? *Annual review of microbiology* **57:** 203-224

Dang QD, Di Cera E (1996) Residue 225 determines the Na(+)-induced allosteric regulation of catalytic activity in serine proteases. *Proceedings of the National Academy of Sciences of the United States of America* **93:** 10653-10656

Deng J, Davies DR, Wisedchaisri G, Wu M, Hol WG, Mehlin C (2004) An improved protocol for rapid freezing of protein samples for long-term storage. *Acta crystallographica Section D, Biological crystallography* **60:** 203-204

Di Cera E, Guinto ER, Vindigni A, Dang QD, Ayala YM, Wuyi M, Tulinsky A (1995) The Na+ binding site of thrombin. *The Journal of biological chemistry* **270:** 22089-22092

Di Cera E, Page MJ, Bah A, Bush-Pelc LA, Garvey LC (2007) Thrombin allostery. *Physical chemistry chemical physics : PCCP* **9:** 1291-1306

Donald JW, Hicks MG, Richardson DJ, Palmer T (2008) The c-type cytochrome OmcA localizes to the outer membrane upon heterologous expression in Escherichia coli. *Journal of bacteriology* **190:** 5127-5131

Douzi B, Filloux A, Voulhoux R (2012) On the path to uncover the bacterial type II secretion system. *Philosophical transactions of the Royal Society of London Series B, Biological sciences* **367:** 1059-1072

Edwin A, Rompikuntal P, Bjorn E, Stier G, Wai SN, Sauer-Eriksson AE (2013) Calcium binding by the PKD1 domain regulates interdomain flexibility in Vibrio cholerae metalloprotease PrtV. *FEBS open bio* **3:** 263-270

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

Epand RM, Epand RF (2009) Domains in bacterial membranes and the action of antimicrobial agents. *Molecular bioSystems* **5:** 580-587

Fehlhammer H, Bode W, Huber R (1977) Crystal structure of bovine trypsinogen at 1-8 A resolution. II. Crystallographic refinement, refined crystal structure and comparison with bovine trypsin. *Journal of molecular biology* **111:** 415-438

Ferrandez Y, Condemine G (2008) Novel mechanism of outer membrane targeting of proteins in Gram-negative bacteria. *Molecular microbiology* **69:** 1349-1357

Filloux A (2004) The underlying mechanisms of type II protein secretion. *Biochimica et biophysica acta* **1694:** 163-179

Finkelstein RA, Boesman-Finkelstein M, Chang Y, Hase CC (1992) Vibrio cholerae hemagglutinin/protease, colonial variation, virulence, and detachment. *Infection and immunity* **60:** 472-478

Fong JC, Yildiz FH (2007) The rbmBCDEF gene cluster modulates development of rugose colony morphology and biofilm formation in Vibrio cholerae. *Journal of bacteriology* **189:** 2319-2330

Forneris F, Wu J, Gros P (2012) The modular serine proteases of the complement cascade. *Current opinion in structural biology* **22:** 333-341

Freeman M (2008) Rhomboid proteases and their biological functions. *Annual review of genetics* **42:** 191-210

Friedrich R, Fuentes-Prior P, Ong E, Coombs G, Hunter M, Oehler R, Pierson D, Gonzalez R, Huber R, Bode W, Madison EL (2002) Catalytic domain structures of MT-SP1/matriptase, a matrix-degrading transmembrane serine proteinase. *The Journal of biological chemistry* **277:** 2160-2168

Fritsch MJ, Krehenbrink M, Tarry MJ, Berks BC, Palmer T (2012) Processing by rhomboid protease is required for Providencia stuartii TatA to interact with TatC and to form functional homo-oligomeric complexes. *Molecular microbiology* **84:** 1108-1123

Fu Y, Waldor MK, Mekalanos JJ (2013) Tn-Seq analysis of vibrio cholerae intestinal colonization reveals a role for T6SS-mediated antibacterial activity in the host. *Cell host & microbe* **14:** 652-663

Gadwal S, Korotkov KV, Delarosa JR, Hol WG, Sandkvist M (2014) Functional and structural characterization of Vibrio cholerae extracellular serine protease B, VesB. *The Journal of biological chemistry*

Gerlach RG, Hensel M (2007) Protein secretion systems and adhesins: the molecular armory of Gram-negative pathogens. *International journal of medical microbiology : IJMM* **297:** 401-415

Giglio KM, Fong JC, Yildiz FH, Sondermann H (2013) Structural basis for biofilm formation via the Vibrio cholerae matrix protein RbmA. *Journal of bacteriology* **195:** 3277-3286

Godeke J, Heun M, Bubendorfer S, Paul K, Thormann KM (2011) Roles of two Shewanella oneidensis MR-1 extracellular endonucleases. *Applied and environmental microbiology* **77:** 5342-5351

Gordon VM, Leppla SH (1994) Proteolytic activation of bacterial toxins: role of bacterial and host cell proteases. *Infection and immunity* **62:** 333-340

Gouet P, Robert X, Courcelle E (2003) ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic acids research* **31:** 3320-3323

Grad YH, Waldor MK (2013) Deciphering the origins and tracking the evolution of cholera epidemics with whole-genome-based molecular epidemiology. *mBio* **4:** e00670-00613

Gray MD, Bagdasarian M, Hol WG, Sandkvist M (2011) In vivo cross-linking of EpsG to EpsL suggests a role for EpsL as an ATPase-pseudopilin coupling protein in the Type II secretion system of Vibrio cholerae. *Molecular microbiology* **79:** 786-798

Griffiths SL, Finkelstein RA, Critchley DR (1986) Characterization of the receptor for cholera toxin and Escherichia coli heat-labile toxin in rabbit intestinal brush borders. *The Biochemical journal* **238:** 313-322

Guinto ER, Caccia S, Rose T, Futterer K, Waksman G, Di Cera E (1999) Unexpected crucial role of residue 225 in serine proteases. *Proceedings of the National Academy of Sciences of the United States of America* **96:** 1852-1857

Haft DH, Paulsen IT, Ward N, Selengut JD (2006) Exopolysaccharide-associated protein sorting in environmental organisms: the PEP-CTERM/EpsH system. Application of a novel phylogenetic profiling heuristic. *BMC biology* **4:** 29

Haft DH, Payne SH, Selengut JD (2012) Archaeosortases and exosortases are widely distributed systems linking membrane transit with posttranslational modification. *Journal of bacteriology* **194:** 36-48

Haft DH, Varghese N (2011) GlyGly-CTERM and rhombosortase: a C-terminal protein processing signal in a many-to-one pairing with a rhomboid family intramembrane serine protease. *PloS one* **6:** e28886

Haksar A, Maudsley DV, Peron FG (1974) Neuraminidase treatment of adrenal cells increases their response to cholera enterotoxin. *Nature* **251:** 514-515

Halpern M, Gancz H, Broza M, Kashi Y (2003) Vibrio cholerae hemagglutinin/protease degrades chironomid egg masses. *Applied and environmental microbiology* **69:** 4200-4204

Hankins JV, Madsen JA, Giles DK, Brodbelt JS, Trent MS (2012) Amino acid addition to Vibrio cholerae LPS establishes a link between surface remodeling in gram-positive and gram-negative bacteria. *Proceedings of the National Academy of Sciences of the United States of America* **109:** 8722-8727

Harmat V, Gal P, Kardos J, Szilagyi K, Ambrus G, Vegh B, Naray-Szabo G, Zavodszky P (2004) The structure of MBL-associated serine protease-2 reveals that identical substrate specificities of C1s and MASP-2 are realized through different sets of enzyme-substrate interactions. *Journal of molecular biology* **342:** 1533-1546

Hedstrom L (2002) An overview of serine proteases. *Current protocols in protein science / editorial board, John E Coligan [et al]* **Chapter 21:** Unit 21 10

Hirst TR, Holmgren J (1987a) Conformation of protein secreted across bacterial outer membranes: a study of enterotoxin translocation from Vibrio cholerae. *Proceedings of the National Academy of Sciences of the United States of America* **84:** 7418-7422

Hirst TR, Holmgren J (1987b) Transient entry of enterotoxin subunits into the periplasm occurs during their secretion from Vibrio cholerae. *Journal of bacteriology* **169:** 1037-1045

Hirst TR, Sanchez J, Kaper JB, Hardy SJ, Holmgren J (1984) Mechanism of toxin secretion by Vibrio cholerae investigated in strains harboring plasmids that encode heat-labile enterotoxins of Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America* **81:** 7752-7756

Horstman AL, Bauman SJ, Kuehn MJ (2004) Lipopolysaccharide 3-deoxy-D-manno-octulosonic acid (Kdo) core determines bacterial association of secreted toxins. *The Journal of biological chemistry* **279:** 8070-8075

Horstman AL, Kuehn MJ (2002) Bacterial surface association of heat-labile enterotoxin through lipopolysaccharide after secretion via the general secretory pathway. *The Journal of biological chemistry* **277:** 32538-32545

Huq A, West PA, Small EB, Huq MI, Colwell RR (1984) Influence of water temperature, salinity, and pH on survival and growth of toxigenic Vibrio cholerae serovar 01 associated with live copepods in laboratory microcosms. *Applied and environmental microbiology* **48:** 420-424

Johnson TL, Scott ME, Sandkvist M (2007) Mapping critical interactive sites within the periplasmic domain of the Vibrio cholerae type II secretion protein EpsM. *Journal of bacteriology* **189:** 9082-9089

Johnson TL, Sikora AE, Zielke RA, Sandkvist M (2013) Fluorescence microscopy and proteomics to investigate subcellular localization, assembly, and function of the type II secretion system. *Methods in molecular biology* **966:** 157-172

Kabsch W (2010) Xds. *Acta crystallographica Section D, Biological crystallography* **66:** 125-132

Kandiba L, Guan Z, Eichler J (2013) Lipid modification gives rise to two distinct Haloferax volcanii S-layer glycoprotein populations. *Biochimica et biophysica acta* **1828:** 938-943

Kim SH, Jia W, Parreira VR, Bishop RE, Gyles CL (2006) Phosphoethanolamine substitution in the lipid A of Escherichia coli O157 : H7 and its association with PmrC. *Microbiology* **152:** 657-666

Kirn TJ, Jude BA, Taylor RK (2005) A colonization factor links Vibrio cholerae environmental survival and human infection. *Nature* **438:** 863-866

Knight SD, Choudhury D, Hultgren S, Pinkner J, Stojanoff V, Thompson A (2002) Structure of the S pilus periplasmic chaperone SfaE at 2.2 A resolution. *Acta crystallographica Section D, Biological crystallography* **58:** 1016-1022

Kobayashi H, Utsunomiya H, Yamanaka H, Sei Y, Katunuma N, Okamoto K, Tsuge H (2009) Structural basis for the kexin-like serine protease from Aeromonas sobria as sepsis-causing factor. *The Journal of biological chemistry* **284:** 27655-27663

Konrad Z, Eichler J (2002) Lipid modification of proteins in Archaea: attachment of a mevalonic acid-based lipid moiety to the surface-layer glycoprotein of Haloferax volcanii follows protein translocation. *The Biochemical journal* **366:** 959-964

Korotkov KV, Johnson TL, Jobling MG, Pruneda J, Pardon E, Heroux A, Turley S, Steyaert J, Holmes RK, Sandkvist M, Hol WG (2011) Structural and functional studies on the interaction of GspC and GspD in the type II secretion system. *PLoS pathogens* **7:** e1002228

Korotkov KV, Sandkvist M, Hol WG (2012) The type II secretion system: biogenesis, molecular architecture and mechanism. *Nature reviews Microbiology* **10:** 336-351

Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *Journal of molecular biology* **372:** 774-797

LaRocque RC, Krastins B, Harris JB, Lebrun LM, Parker KC, Chase M, Ryan ET, Qadri F, Sarracino D, Calderwood SB (2008) Proteomic analysis of Vibrio cholerae in human stool. *Infection and immunity* **76:** 4145-4151

Lazareno-Saez C, Arutyunova E, Coquelle N, Lemieux MJ (2013) Domain swapping in the cytoplasmic domain of the Escherichia coli rhomboid protease. *Journal of molecular biology* **425:** 1127-1142

Lee H, Hsu FF, Turk J, Groisman EA (2004) The PmrA-regulated pmrC gene mediates phosphoethanolamine modification of lipid A and polymyxin resistance in Salmonella enterica. *Journal of bacteriology* **186:** 4124-4133

Lemieux MJ, Fischer SJ, Cherney MM, Bateman KS, James MN (2007) The crystal structure of the rhomboid peptidase from Haemophilus influenzae provides insight into intramembrane proteolysis. *Proceedings of the National Academy of Sciences of the United States of America* **104:** 750-754

Lencer WI, Constable C, Moe S, Rufo PA, Wolf A, Jobling MG, Ruston SP, Madara JL, Holmes RK, Hirst TR (1997) Proteolytic activation of cholera toxin and Escherichia coli labile toxin by entry into host epithelial cells. Signal transduction by a protease-resistant toxin variant. *The Journal of biological chemistry* **272:** 15562-15568

Lindberg F, Tennent JM, Hultgren SJ, Lund B, Normark S (1989) PapD, a periplasmic transport protein in P-pilus biogenesis. *Journal of bacteriology* **171:** 6052-6058

Long F, Vagin AA, Young P, Murshudov GN (2008) BALBES: a molecular-replacement pipeline. *Acta crystallographica Section D, Biological crystallography* **64:** 125-132

Lu C, Turley S, Marionni ST, Park YJ, Lee KK, Patrick M, Shah R, Sandkvist M, Bush MF, Hol WG (2013) Hexamers of the type II secretion ATPase GspE from Vibrio cholerae with increased ATPase activity. *Structure* **21:** 1707-1717

Lybarger SR, Johnson TL, Gray MD, Sikora AE, Sandkvist M (2009) Docking and assembly of the type II secretion complex of Vibrio cholerae. *Journal of bacteriology* **191:** 3149-3161

Magnusdottir A, Johansson I, Dahlgren LG, Nordlund P, Berglund H (2009) Enabling IMAC purification of low abundance recombinant proteins from E. coli lysates. *Nature methods* **6:** 477-478

Marquart M, Walter, J., Deisenhofer, J., Bode, W., Huber, R. (1983) The geometry of the reactive site and of the peptide groups in trypsin, trypsinogen and its complexes with inhibitors. *Acta Crystallographica Structural Science* **39:** 480-490

Marraffini LA, Dedent AC, Schneewind O (2006) Sortases and the art of anchoring proteins to the envelopes of gram-positive bacteria. *Microbiology and molecular biology reviews : MMBR* **70:** 192-221

Mazmanian SK, Liu G, Ton-That H, Schneewind O (1999) Staphylococcus aureus sortase, an enzyme that anchors surface proteins to the cell wall. *Science* **285:** 760-763

McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007) Phaser crystallographic software. *Journal of applied crystallography* **40:** 658-674

McLaughlin LS, Haft RJ, Forest KT (2012) Structural insights into the Type II secretion nanomachine. *Current opinion in structural biology* **22:** 208-216

Miller-Podraza H, Bradley RM, Fishman PH (1982) Biosynthesis and localization of gangliosides in cultured cells. *Biochemistry* **21:** 3260-3265

Murayama K, Kato-Murayama M, Hosaka T, Sotokawauchi A, Yokoyama S, Arima K, Shirouzu M (2012) Crystal structure of cucumisin, a subtilisin-like endoprotease from Cucumis melo L. *Journal of molecular biology* **423:** 386-396

Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta crystallographica Section D, Biological crystallography* **67:** 355-367

Nelson EJ, Harris JB, Morris JG, Jr., Calderwood SB, Camilli A (2009) Cholera transmission: the host, pathogen and bacteriophage dynamic. *Nature reviews Microbiology* **7:** 693-702

Newland JW, Green BA, Foulds J, Holmes RK (1985) Cloning of extracellular DNase and construction of a DNase-negative strain of Vibrio cholerae. *Infection and immunity* **47:** 691-696

Nonaka T, Fujihashi M, Kita A, Saeki K, Ito S, Horikoshi K, Miki K (2004) The crystal structure of an oxidatively stable subtilisin-like alkaline serine protease, KP-43, with a C-terminal beta-barrel domain. *The Journal of biological chemistry* **279:** 47344-47351

Okuda S, Tokuda H (2011) Lipoprotein sorting in bacteria. *Annual review of microbiology* **65:** 239-259

Otwinowski ZaWM (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276:** 307-326

Overbye LJ, Sandkvist M, Bagdasarian M (1993) Genes required for extracellular secretion of enterotoxin are clustered in Vibrio cholerae. *Gene* **132:** 101-106

Page MJ, Bleackley MR, Wong S, MacGillivray RT, Di Cera E (2006) Conversion of trypsin into a Na(+)-activated enzyme. *Biochemistry* **45:** 2987-2993

Painter J, Merritt EA (2006) Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta crystallographica Section D, Biological crystallography* **62:** 439-450

Patricelli MP, Giang DK, Stamp LM, Burbaum JJ (2001) Direct visualization of serine hydrolase activities in complex proteomes using fluorescent active site-directed probes. *Proteomics* **1:** 1067-1071

Patrick M, Korotkov KV, Hol WG, Sandkvist M (2011) Oligomerization of EpsE coordinates residues from multiple subunits to facilitate ATPase activity. *The Journal of biological chemistry* **286:** 10378-10386

Pozzi N, Chen Z, Gohara DW, Niu W, Heyduk T, Di Cera E (2013) Crystal structure of prothrombin reveals conformational flexibility and mechanism of activation. *The Journal of biological chemistry* **288:** 22734-22744

Pugsley AP, Chapon C, Schwartz M (1986) Extracellular pullulanase of Klebsiella pneumoniae is a lipoprotein. *Journal of bacteriology* **166:** 1083-1088

Pugsley AP, Kornacker MG (1991) Secretion of the cell surface lipoprotein pullulanase in Escherichia coli. Cooperation or competition between the specific secretion pathway and the lipoprotein sorting pathway. *The Journal of biological chemistry* **266:** 13640-13645

Pugsley AP, Kornacker MG, Ryter A (1990) Analysis of the subcellular location of pullulanase produced by Escherichia coli carrying the pulA gene from Klebsiella pneumoniae strain UNF5023. *Molecular microbiology* **4:** 59-72

Raetz CR, Reynolds CM, Trent MS, Bishop RE (2007) Lipid A modification systems in gram-negative bacteria. *Annual review of biochemistry* **76:** 295-329

Rather P (2013) Role of rhomboid proteases in bacteria. *Biochimica et biophysica acta*

Rawlings ND, Barrett AJ (1994) Families of serine peptidases. *Methods Enzymol* **244:** 19-61

Rawlings ND, Barrett AJ, Bateman A (2012) MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic acids research* **40:** D343-350

Rawlings ND, Waller M, Barrett AJ, Bateman A (2013) MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic acids research*

Read RJ, James MN (1988) Refined crystal structure of Streptomyces griseus trypsin at 1.7 A resolution. *Journal of molecular biology* **200:** 523-551

Rojas A, Doolittle, Russell F. (2002) The Occurrence of Type S1A Serine Proteases in Sponge and Jelly. *Journal of Molecular Evolution* **55:** 790-794

Rondelet A, Condemine G (2013) Type II secretion: the substrates that won't go away. *Research in microbiology* **164:** 556-561

Rypniewski WR, Perrakis A, Vorgias CE, Wilson KS (1994) Evolutionary divergence and conservation of trypsin. *Protein engineering* **7:** 57-64

Sack DA, Sack RB, Nair GB, Siddique AK (2004) Cholera. *Lancet* **363:** 223-233

Sandkvist M (2001) Biology of type II secretion. *Molecular microbiology* **40:** 271-283

Sandkvist M, Michel LO, Hough LP, Morales VM, Bagdasarian M, Koomey M, DiRita VJ, Bagdasarian M (1997) General secretion pathway (eps) genes required for toxin secretion and outer membrane biogenesis in Vibrio cholerae. *Journal of bacteriology* **179:** 6994-7003

Sandkvist M, Morales V, Bagdasarian M (1993) A protein required for secretion of cholera toxin through the outer membrane of Vibrio cholerae. *Gene* **123:** 81-86

Sauer FG, Knight SD, Waksman, Gj, Hultgren SJ (2000) PapD-like chaperones and pilus biogenesis. *Seminars in cell & developmental biology* **11:** 27-34

Schechter I, Berger A (1967) On the size of the active site in proteases. I. Papain. *Biochemical and biophysical research communications* **27:** 157-162

Schneewind O, Mihaylova-Petkov D, Model P (1993) Cell wall sorting signals in surface proteins of gram-positive bacteria. *The EMBO journal* **12:** 4803-4811

Schneewind O, Model P, Fischetti VA (1992) Sorting of protein A to the staphylococcal cell wall. *Cell* **70:** 267-281

Seper A, Fengler VH, Roier S, Wolinski H, Kohlwein SD, Bishop AL, Camilli A, Reidl J, Schild S (2011) Extracellular nucleases and extracellular DNA play important roles in Vibrio cholerae biofilm formation. *Molecular microbiology* **82:** 1015-1037

Sikora AE (2013) Proteins secreted via the type II secretion system: smart strategies of Vibrio cholerae to maintain fitness in different ecological niches. *PLoS pathogens* **9:** e1003126

Sikora AE, Beyhan S, Bagdasarian M, Yildiz FH, Sandkvist M (2009) Cell envelope perturbation induces oxidative stress and changes in iron homeostasis in Vibrio cholerae. *Journal of bacteriology* **191:** 5398-5408

Sikora AE, Lybarger SR, Sandkvist M (2007) Compromised outer membrane integrity in Vibrio cholerae Type II secretion mutants. *Journal of bacteriology* **189:** 8484-8495

Sikora AE, Zielke RA, Lawrence DA, Andrews PC, Sandkvist M (2011) Proteomic analysis of the Vibrio cholerae type II secretome reveals new proteins, including three related serine proteases. *The Journal of biological chemistry* **286:** 16555-16566

Songsiriritthigul C, Pantoom S, Aguda AH, Robinson RC, Suginta W (2008) Crystal structures of Vibrio harveyi chitinase A complexed with chitooligosaccharides: implications for the catalytic mechanism. *Journal of structural biology* **162:** 491-499

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

Stevenson LG, Strisovsky K, Clemmer KM, Bhatt S, Freeman M, Rather PN (2007) Rhomboid protease AarA mediates quorum-sensing in Providencia stuartii by activating TatA of the twin-arginine translocase. *Proceedings of the National Academy of Sciences of the United States of America* **104:** 1003-1008

Strisovsky K, Sharpe HJ, Freeman M (2009) Sequence-specific intramembrane proteolysis: identification of a recognition motif in rhomboid substrates. *Molecular cell* **36:** 1048-1059

Syngkon A, Elluri S, Koley H, Rompikuntal PK, Saha DR, Chakrabarti MK, Bhadra RK, Wai SN, Pal A (2010) Studies on a novel serine protease of a DeltahapADeltaprtV Vibrio cholerae O1 strain and its role in hemorrhagic response in the rabbit ileal loop model. *PloS one* **5**

Szabo R, Bugge TH (2011) Membrane-anchored serine proteases in vertebrate cell and developmental biology. *Annual review of cell and developmental biology* **27:** 213-235

Tang X, Yi W, Munske GR, Adhikari DP, Zakharova NL, Bruce JE (2007) Profiling the membrane proteome of Shewanella oneidensis MR-1 with new affinity labeling probes. *Journal of proteome research* **6:** 724-734

Taylor RK, Manoil C, Mekalanos JJ (1989) Broad-host-range vectors for delivery of TnphoA: use in genetic analysis of secreted virulence determinants of Vibrio cholerae. *Journal of bacteriology* **171:** 1870-1878

Uesugi Y, Arima J, Usuki H, Iwabuchi M, Hatanaka T (2008) Two bacterial collagenolytic serine proteases have different topological specificities. *Biochimica et biophysica acta* **1784:** 716-726

Uesugi Y, Usuki H, Arima J, Iwabuchi M, Hatanaka T (2011) Molecular dissection of Streptomyces trypsin on substrate recognition. *Biochimica et biophysica acta* **1814:** 1295-1304

Uesugi Y, Usuki H, Iwabuchi M, Hatanaka T (2009) The role of Tyr71 in Streptomyces trypsin on the recognition mechanism of structural protein substrates. *The FEBS journal* **276:** 5634-5646

van Zonneveld AJ, Veerman H, MacDonald ME, van Mourik JA, Pannekoek H (1986) Structure and function of human tissue-type plasminogen activator (t-PA). *Journal of cellular biochemistry* **32:** 169-178

Verghese GM, Gutknecht MF, Caughey GH (2006) Prostasin regulates epithelial monolayer function: cell-specific Gpld1-mediated secretion and functional role for GPI anchor. *American journal of physiology Cell physiology* **291:** C1258-1270

Vezzulli L, Guzman CA, Colwell RR, Pruzzo C (2008) Dual role colonization factors connecting Vibrio cholerae's lifestyles in human and aquatic environments open new perspectives for combating infectious diseases. *Current opinion in biotechnology* **19:** 254-259

Vezzulli L, Pruzzo C, Huq A, Colwell RR (2010) Environmental reservoirs of Vibrio cholerae and their role in cholera. *Environmental microbiology reports* **2:** 27-33

Vizcaino JA, Cote RG, Csordas A, Dianes JA, Fabregat A, Foster JM, Griss J, Alpi E, Birim M, Contell J, O'Kelly G, Schoenegger A, Ovelleiro D, Perez-Riverol Y, Reisinger F, Rios D, Wang R, Hermjakob H (2013) The PRoteomics IDEntifications (PRIDE) database and associated tools: status in 2013. *Nucleic acids research* **41:** D1063-1069

Wang Y, Zhang Y, Ha Y (2006) Crystal structure of a rhomboid family intramembrane protease. *Nature* **444:** 179-180

Wong E, Vaaje-Kolstad G, Ghosh A, Hurtado-Guerrero R, Konarev PV, Ibrahim AF, Svergun DI, Eijsink VG, Chatterjee NS, van Aalten DM (2012) The Vibrio cholerae colonization factor GbpA possesses a modular structure that governs binding to different host surfaces. *PLoS pathogens* **8:** e1002373

Wu Z, Yan N, Feng L, Oberstein A, Yan H, Baker RP, Gu L, Jeffrey PD, Urban S, Shi Y (2006) Structural analysis of a rhomboid family intramembrane protease reveals a gating mechanism for substrate entry. *Nature structural & molecular biology* **13:** 1084-1091

Xu Q, Dziejman M, Mekalanos JJ (2003) Determination of the transcriptome of Vibrio cholerae during intraintestinal growth and midexponential phase in vitro. *Proceedings of the National Academy of Sciences of the United States of America* **100:** 1286-1291

Yamane T, Iwasaki A, Suzuki A, Ashida T, Kawata Y (1995) Crystal structure of Streptomyces erythraeus trypsin at 1.9 A resolution. *Journal of biochemistry* **118:** 882-89