Dissecting the Role of EpsE Subdomains in ATPase Activity and Type II Secretion in *Vibrio cholerae*

by

Marcella D. Patrick

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Microbiology and Immunology)
in The University of Michigan
2011

Doctoral Committee:

Assistant Professor Maria B. Sandkvist, Chair
Associate Professor Philip C. Hanna
Associate Professor Billy Tsai
Assistant Professor Eric S. Krukonis

Acknowledgements

A special thanks to my family for all of their love and support-- even if they never quite understood why someone would want to be a student for more than 20 years, they were always behind me in every decision. I can't thank you enough for allowing me to make my own mistakes, learn from them and move forward; lessons that I will carry with me in every aspect of my life.

I would like to thank all of my friends who have accompanied me on this tremendous journey. You've each helped shape me as a person and a scientist. When things were looking dark you were always a beacon of light. And when I started complaining too much you were always there to tell me to just shut up and climb!

And a final thanks to the ladies of the Sandkvist Lab, especially Maria. I wouldn't have made it though graduate school without having such a happy and supportive group of people to collaborate with and rarely does one find people who love food nearly as much as I do!

Table of Contents

List of Figures

Figure

List of Tables

Table

Abstract

The Type II secretion (T2S) apparatus is required for the survival and fitness of *Vibrio cholerae* both in the human host as well as its natural aquatic environment due to its ability to secrete proteins such as cholera toxin and hydrolytic enzymes. This apparatus, which is widely distributed among Gram negative pathogens, is composed of at least 12 proteins and spans the entire cell envelope. The energy to support secretion through the T2S system is provided by an ATPase associated with this complex, EpsE. EpsE is homologous to many ATPases involved with other molecular trafficking machineries and is comprised of three distinct subdomains: the N-terminal domain (NTD), the C-terminal domain (CTD), and a tetracysteine coordinating zinc binding domain (C_M). These studies demonstrate that the mechanism of ATP hydrolysis requires oligomerization of EpsE, which allows for the realignment of specific arginine residues from the NTD, CTD and the CTD of a neighboring subunit to form the active site. Replacement of any of these arginine residues in the active site or of residues within the subunit interface halts the activity of EpsE, suggesting conformational changes associated with the binding, hydrolysis and release of nucleotide must be precise. An additional arginine from the C_M domain is also required to complete the active site, suggesting a role for all three domains in ATP hydrolysis and energizing the T2S machinery. Previous work has shown

EpsE to interact with the T2S apparatus through its NTD with another T2S component, EpsL. The movement of the NTD in association with EpsL likely provides a means of transforming the energy from ATP hydrolysis into mechanical work to support T2S. This work provides further evidence that the C_M domain may be involved in additional protein-protein contacts with either EpsL or another protein associated with the T2S machinery. It is conceivable that the C_M domain may play a direct role in the regulation of ATPase activity as it is in prime position to connect the nucleotide binding pocket to other components of the T2S apparatus through protein-protein interactions.

Chapter 1

Introduction

Vibrio cholerae in the environment

Vibrio cholerae, the agent of the diarrheal disease cholera, is a Gram negative, rod shaped bacterium found in aquatic environments. It is typically attached to aquatic organisms, including the the chitinous surfaces of aquatic zooplankton such as copepods (Huq *et al.*, 1984), egg masses of chironomids, (Halpern *et al.*, 2004), as well as associated with amoeba (Abd *et al.*, 2007, Abd *et al.*, 2005), and algae (Islam *et al.*, 1989) among others. It is thought that attachment to these plankton and other surfaces give the bacteria a survival tool as this likely provides not only protection from the environment, but in many cases the surface polymers could also be used as carbon and nitrogen sources (Huq et al., 1984, Tamplin *et al.*, 1990). Recently, it was discovered that *V. cholerae* secrets a chitin binding protein that mediates attachment to chitinous surfaces (Kirn *et al.*, 2005). Once attached, other secreted proteins can be utilized to promote environmental survival such as the use of hemagglutinin protease to promote attachment and disintegration of the chironomid egg masses

(Halpern *et al.*, 2003) and an exochitinase which breaks down the chitinous surfaces to use as a carbon source (Connell *et al.*, 1998). The interaction of *V. cholerae* with chitin is vastly important not only because it provides a food source and protection from environmental stresses and predators, but it also induces a natural competence, making it easier for the bacterium to take up foreign genetic material from the surrounding environment (reviewed in Vezzulli *et al.*, 2010).

It is still under debate which aquatic reservoir is most important in the transmission of *V. cholerae* to humans, but the abundance of chitinous plankton makes it the current favorite for fostering the persistence of these bacteria (Colwell & Huq, 1994). Studies have shown that, between seasonal occurrences of cholera, the bacteria can be found in a viable but non-culturable state living in biofilms within the aquatic environment (Alam *et al.*, 2006). Recently the O1 biotype was shown to form biofilms that were shed in stool in an animal model (Alam *et al.*, 2007). These bacteria remained infective if transmitted to another host, but quickly lost this infectivity if introduced back into an aquatic environment. This may contribute to epidemic cholera in highly populated areas, where the fecal to oral transmission by means of contaminated water is likely rapid and further highlights the importance of water sanitation to the prevention of cholera outbreaks (Alam *et al*., 2007). Many *Vibrio* species are able to survive in these environmental niches, and currently there are over 200 known serotypes of one of those species, *V. cholerae*, however, of these 200 serotypes, only two are known to cause epidemic cholera; the O1 mentioned above and the O139 serotypes (Faruque *et al.*, 1998). The O1 can further be divided into two

biotypes based on differences in biochemical properties such as polymyxin B sensitivity and production of hemolysin: the classical and El Tor. Of these two biotypes, the classical was likely responsible for the first six pandemics of cholera, and is believed to be the more severe biotype. The seventh and current pandemic, however, is caused by the El Tor variant and an additional, concurrent pandemic, sometimes referred to as the eighth pandemic, is caused by the O139 strain (Faruque *et al*., 1998). Upon further analysis of the O139 serotype, it was found that this strain likely arose from an O1 El Tor strain which picked up exogenous DNA, thus altering its O antigen (Mooi & Bik, 1997) This emphasizes how rapidly *V. cholerae* can adapt by means of horizontal gene transfer. For both pathogenic and non-pathogenic strains of *V. cholerae*, the ability to persist in aquatic environments requires unique strategies, including the extracellular secretion of proteins. These strategies are likely conserved across all Vibrios, and the ability of *V. cholerae*, in particular, to be transformed with exogenous genetic material may play a critical role in its ability to acquire genes responsible for pathogenicity among other functions as were demonstrated for O139 strain.

Pathogenesis

Aside from being the primary reservoir for *V. cholerae*, the aqueous environment may also play a major role in transmission to humans, as the incidences of cholera outbreaks are seemingly controlled by environmental factors (reviewed in Vezzulli *et al.*, 2010). As the temperature rises, warming of the water causes the phytoplankton to bloom, which in turn increases the number

of zooplankton to which *V. cholerae* attaches (Huq *et al*., 1984). The zooplankton can then be taken up by a human through the consumption of contaminated food or water, and once consumed, *V. cholerae* is able to colonize the small intestine. The chitin binding protein mentioned above affords one way for the bacteria to bind to the host's intestinal epithelial lining as its function has been adapted to aid in gut colonization, providing evidence that *V. cholerae* can make use of some if its environmental strategies to thrive in the host environment (Kirn *et al*., 2005, Matson *et al.*, 2007). Additional colonization factors including, toxin co-regulated pilus (TCP) and other TCP accessory colonization factors (ACFs) are required for intestinal colonization, as these factors mediate aggregation and microcolony formation (Faruque *et al*., 1998, Matson *et al*., 2007). Once established in the host, the bacteria can secrete the cholera toxin. Cholera toxin is an AB5 toxin, with the A subunit being cytotoxic and the B pentamer being required for secretion of the toxin out of the bacterium (Hirst *et al.*, 1984, Spangler, 1992). Once secreted, the toxin binds the host GM1 gaglioside where it is internalized and the active form of the A subunit can then act as an ADP-ribosyl transferase. The regulatory subunit of a G protein complex is the target of ADP ribosylation and can thus activate adenylate cyclase which, in turn, stimulates production of cAMP, causing a massive flux of water and electrolytes into the intestinal lumen resulting in the hallmark rice-water stool indicative of the disease cholera (Spangler, 1992). The dehydration resulting from the massive loss of fluids can result in death, which occurs in more than 50% of untreated cases (Matson *et al*., 2007).

Type II Secretion

In both the aquatic and host environments, extracellular secretion of proteins is likely necessary for survival of *V. cholerae*. The Type 2 secretion (T2S) apparatus of *V. cholerae* is utilized in both environments and has been shown to support the secretion of proteins, such as the above noted hemagglutinin protease, chitin binding protein, chitinase and lipase, which are all necessary for the fitness of the bacterium in its given environment. Also, this secretion system plays a direct role in pathogenesis by secreting the major virulence factor cholera toxin. The T2S system (T2SS) is highly conserved across many pathogenic and non-pathogenic Gram negative bacteria (depicted in Fig 1.1) and is composed of at least 12 gene products which assemble into a large protein complex spanning the entire cell envelope (Cianciotto, 2005, Johnson *et al.*, 2006). Following transport across the cytoplasmic membrane through the Sec or Tat systems, the protein to be exported is folded in the periplasm. It is unknown how secreted toxins, hydrolytic enzymes, and other proteins are targeted to the T2S system, but it is likely due to a structural feature rather than sequence motif (Connell *et al.*, 1995, Hirst *et al*., 1984).

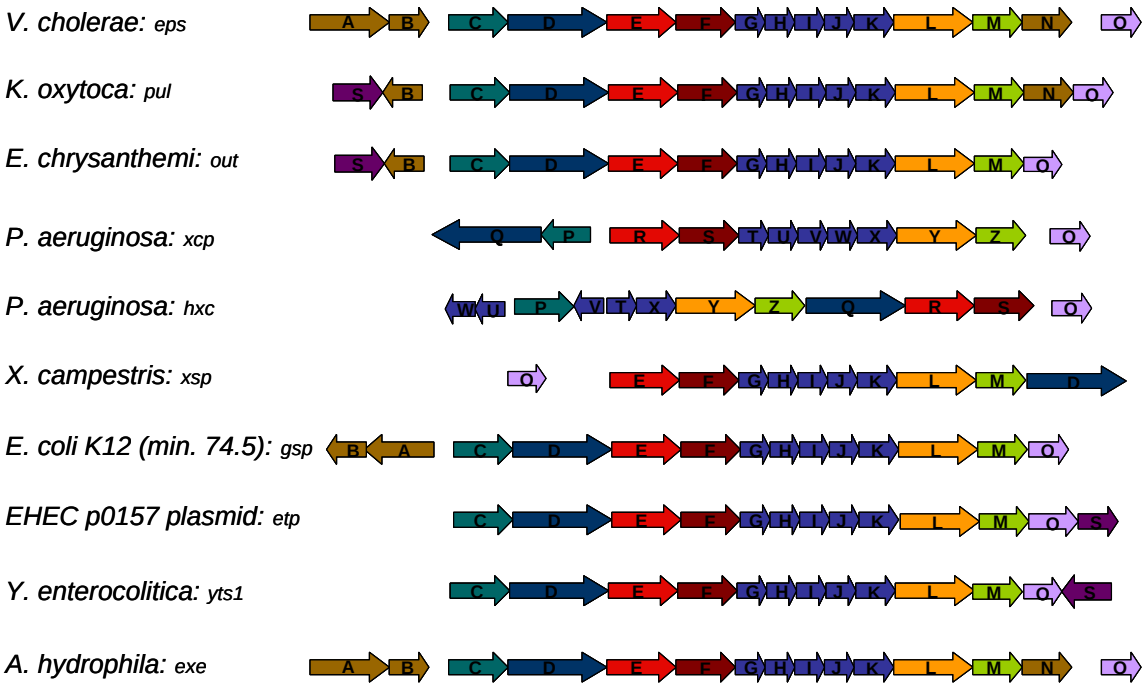


Figure 1.1: T2SS gene alignments.
The genes comprising the T2SSs from individual species are shown schematically. Homologous components are colored similarly; with arrowheads showing orientation.

Once it engages the T2S apparatus, the protein to be secreted is thought to be pushed through the outer membrane secretin comprised of multiple copies of the T2S protein, EpsD. The pushing likely occurs through the polymerization of a pseudopilus structure made up of EpsG-K (Filloux, 2004, Hobbs & Mattick, 1993, Sandkvist, 2001a). This process is energy dependent with the cytoplasmic protein, EpsE providing the energy by means of ATP hydrolysis (Camberg & Sandkvist, 2005). Being cytoplasmic, EpsE engages the T2S apparatus through interactions with EpsL, a member of the T2S inner membrane complex, which

also includes EpsC, EpsF, and EpsM (Fig. 1.2) (Abendroth *et al.*, 2005, Py *et al.*, 2001, Robert *et al.*, 2005, Sandkvist *et al.*, 1995).

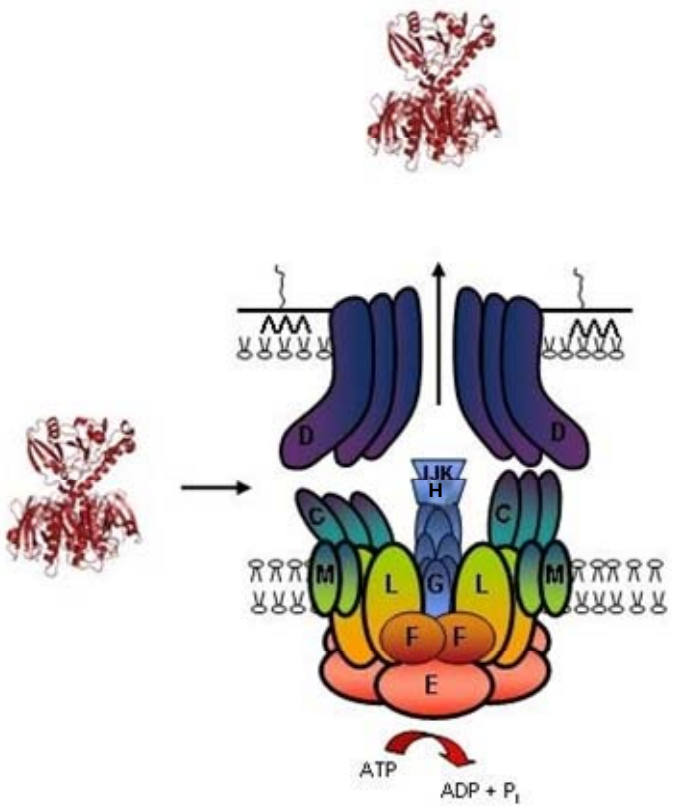


Figure 1.2: Model of the T2S apparatus.
The T2S system is comprised of a cytoplasmic ATPase (EpsE), an inner membrane complex (EpsC, F, L, and M), pseudopilins (EpsG, H, I, J and K) and an outer membrane secretin (EpsD), which assemble into a complex spanning the entire Gram negative cell envelope. The fully folded toxin (shown in red ribbon structure) engages the T2S apparatus in the periplasm to allow for its extracellular translocation.

The ATPase:

T2S is an energy-dependent process. Only one protein encoded by the T2S operon has been characterized to have ATPase activity and that is the cytoplasmic protein, EpsE. EpsE is related to a larger family of secretion ATPases, which are known to be involved in many activities including protein secretion, Type IV pilus (T4P) biogenesis, DNA uptake, and archaeal flagellar assembly (Planet *et al*., 2001), which will be discussed in more detail below. The *V. cholerae* ATPase was one of the earliest T2S proteins to be crystallized, and from that work it was hypothesized to form a hexameric ring structure based on homologies to other characterized RecA-like ATPases (Robien *et al.*, 2003). Later functional work supported this hypothesis by demonstrating that a small fraction of purified GST-tagged EpsE corresponding to the size of a hexamer is capable of hydrolyzing ATP (Camberg & Sandkvist, 2005). It is possible that the GST tag on EpsE played a role in forming the small fraction of active EpsE oligomers, therefore further factors which could similarly induce EpsE's activity were sought out. It was discovered that the ATPase activity of EpsE was stimulated by the association of its N-terminal domain with the cytoplasmic domain of the inner membrane protein, EpsL (Camberg *et al.*, 2007); an interaction which is necessary to link EpsE to the greater T2S complex (Figure 1.2) (Abendroth *et al*., 2005, Possot *et al.*, 2000, Py *et al.*, 1999, Sandkvist *et al*., 1995). EpsE and the cytoplasmic domain of EpsL themselves are monomeric when purified, but they co-crystallize as heterodimers (Abendroth *et al*., 2005). Although EpsE and cytoplasmic domain of EpsL co-purify in a 1:1 ratio, the final

stoichiometry of the EpsE-L complex in the functional T2SS remains to be determined.

The Inner Membrane Complex:

Although proteins secreted through the T2SS engage the apparatus in the periplasm, the majority of the T2S components are found in the inner membrane. EpsL is a bitopic membrane protein with a large cytoplasmic domain, a short membrane-spanning helix, and a smaller periplasmic domain, and it is believed to dimerize in the membrane (Sandkvist *et al.*, 1999). Aside from linking the cytoplasmic EpsE to the membrane, it has been shown to interact with the other inner membrane proteins, EpsC, M, F, and J (Fig. 1.2) (Douet *et al.*, 2004, Lee *et al.*, 2004, Py *et al*., 2001, Robert *et al.,* 2005, Sandkvist *et al.*, 2000, Tsai *et al.*, 2002). The C-terminal periplasmic domain of EpsL was recently crystallized as a dimer (Abendroth *et al.*, 2009a). When comparing the periplasmic EpsL structure, it was found that its closest structural homolog is the periplasmic domain of another T2S inner membrane protein, EpsM (Abendroth *et al.*, 2009a). Interestingly, periplasmic domain of EpsM also forms a dimer (Abendroth *et al.*, 2004b, Johnson *et al.*, 2007); however, the periplasmic domains of EpsL and EpsM are not homologous at the primary sequence level and their dimer arrangements are not the same (Abendroth et al., 2009a).

EpsL and EpsM are known to form a stable complex that protects them from degradation (Possot *et al*., 2000, Py *et al*., 2001, Sandkvist *et al*., 1999), and it has been demonstrated that residues 84 to 99 in the periplasmic domain of

EpsM are necessary to interact with EpsL residues 216 to 296 (Abendroth *et al.*, 2004b, Johnson *et al*., 2007). Unfortunately, structural information is not available for these regions in EpsM and EpsL; thus, it is still unknown how these proteins form a complex and the relative stoichiometry of said complex. When the periplasmic domain of EpsM dimerizes, a hydrophobic cleft with a hydrophilic rim is formed, which may bind a yet undefined ligand or peptide. Because the cleft residues in EpsM are relatively well conserved, the cleft is likely important for secretion (Abendroth *et al*., 2004b).

A third inner membrane protein EpsF has also been shown to make up part of this inner membrane complex (Fig. 1.2). EpsF spans the membrane three times with its N-terminal domain in the cytoplasm and its C-terminal domain in the periplasm (Abendroth *et al.*, 2009b). Only the larger, first cytoplasmic domain has been crystallized and it was found to form crystallographic dimers. The first 55 residues of EpsF were too flexible to crystallize, and are therefore thought to potentially interact with other T2S components (Abendroth *et al*., 2009b). Using yeast two-hybrids and co-immunoprecipitations, Py *et al.* (Py *et al*., 2001) demonstrated the formation of an EpsE, L, F complex in *Erwinia chrysanthemi*, but this was conditional on EpsE and L forming a complex first. Similarly, Arts *et al*. (Arts *et al.*, 2007) demonstrated that the EpsF homolog in *Pseudomonas aeruginosa* is unstable unless in the presence of the EpsE and L proteins. The authors determined the second cytoplasmic domain to be important for this stabilization, which is in contrast to the *E. chrysanthemi* studies; however, it is feasible that both cytoplasmic domains play a role in interaction with other T2S

components (Arts *et al*., 2007, Py *et al*., 2001). The structure of the first cytoplasmic domain of EpsF was recently solved and a careful sequence comparison between the first and second cytoplasmic domains shows many similarities between the two that may translate into structural similarities as well (Abendroth *et al*., 2009b). Interestingly, EpsE binds in a pocket within the EpsL cytoplasmic domain as demonstrated by co-crystallizing the two proteins. EpsE only fills a portion of this binding cleft; making it tempting to speculate that EpsF may also bind within this same pocket in EpsL (Abendroth *et al*., 2005).

The pseudopilins:

The pseudopilins consist of five additional proteins, EpsG-K, which are thought to form a pilus-like structure that can act like a piston, pushing the secreted proteins out of the apparatus (Fig. 1.2) (Filloux *et al.*, 1998, Sandkvist, 2001a). The pseudopilins are termed as such due to their high structural homology to the type IV pilins, with EpsG as the major component and EpsH-K being minor components due to their lower expression levels compared to that of EpsG (Kohler *et al.*, 2004, Korotkov *et al.*, 2009a, Korotkov & Hol, 2008). Based on structural comparisons of the type IV pilins and EpsG-K, it is believed EpsG forms a thin helical filament stabilized by calcium with the H-I-J-K complex forming a tip (Durand *et al.*, 2003, Kohler *et al*., 2004, Korotkov *et al*., 2009a, Korotkov & Hol, 2008). Previous studies have demonstrated that the ATPase EpsE is required for polymerization of the pseudopilus and more recently studies have shown EpsG to interact with EpsL, providing a direct link between the

ATPase and the pseudopilus (Camberg & Sandkvist, 2005, Durand *et al.*, 2005, Gray *et al.*, 2011, Pelicic, 2008).

Connecting the Inner and Outer Membranes:

At least one other T2S protein is associated with the inner membrane complex, EpsC (Fig. 1.2). EpsC is a bitopic inner membrane protein with a periplasmic domain that is divided into two subdomains: the HR domain and either a PDZ or coiled-coil domain. EpsC has been shown to interact and stabilize the inner membrane EpsL-M complex (Gerard-Vincent *et al.*, 2002, Lee *et al*., 2004, Tsai *et al*., 2002). The interaction of EpsC's N-terminal 46 residues with the inner membrane proteins EpsL and M protect them from degradation (Lee *et al*., 2004), and similarly, the presence of the EpsL-M complex seems to be necessary to protect EpsC from degradation and may play a further role in stabilizing EpsC in its proper conformation to support secretion (Lybarger *et al.*, 2009). EpsC homologs are known to be degraded in the absence of outer membrane protein EpsD in *Klebsiella oxytoca* and *P. aeruginosa* (Bleves *et al.*, 1999, Possot *et al.*, 1999), and it was recently demonstrated that the periplasmic subdomains of EpsC and D could be co-purified, indicating a direct interaction between these two T2S components (Korotkov *et al.*, 2006). Korotkov *et al.* (Korotkov *et al*., 2006) further determined that the HR domain is primarily responsible for interaction with EpsD. As EpsC spans the periplasm, it is a nice candidate for substrate recognition and/or gating of the EpsD pore. The second periplasmic subdomain, the PDZ domain, is the most likely candidate to interact

with the secreted proteins, another T2S protein, or possibly both (Korotkov *et al*., 2006). Further research into this region of the protein is ongoing; however it has been shown that the PDZ domain is required for the secretion of most proteins in *E. chrysanthemi* (Bouley *et al.*, 2001).

The Secretin:

The T2S outer membrane protein EpsD, also termed the secretin, is thought to oligomerize into a pore used in the export of secreted proteins out of the bacterium (Fig. 1.2) (Brok *et al.*, 1999). It is part of a larger secretin superfamily that consists of outer membrane proteins required for T2S, T4P biogenesis, Type III secretion, and filamentous phage extrusion. Most secretins require the use of a small lipoprotein termed the pilotin (EpsS) which binds the C-terminal end of EpsD and is required for proper positioning of EpsD in the outer membrane (Hardie *et al.*, 1996, Shevchik *et al.*, 1997). Although many of the secretins that make up the larger family require a pilotin for proper insertion, not all T2SSs have an identified pilotin protein. In fact, only *K. oxytoca* and *E. chrysanthemi* have shown their respective EpsS pilotin to be important for secretin insertion in the outer membrane; however, putative *epsS* genes in other T2S operons (including the *E. coli* p0157 plasmid-encoded T2SS) have been identified (Figure 1.1) (Sandkvist, 2001b, Schmidt *et al.*, 1997). It is possible that the pilotin has not been identified because of limited sequence homology in other T2SSs, or there may be an alternative pathway by which the secretin is inserted in the outer membrane. In support of this latter point, Viarre *et al*. (Viarre *et al.*,

2009) showed that the secretin for the T2SS in *P. aeruginosa*, HxcQ, contains a lipid anchor, which is sufficient to target this protein to the outer membrane and allow for correct insertion. Regardless of the mechanism, once properly inserted, secretins are characterized by their ability to form large heat- and detergent-resistant oligomers. They form rings of 12 to 14 subunits, containing a central channel ranging from 50 to 100 Å in diameter that is occluded by a central plug (Bitter *et al.*, 1998, Chami *et al.*, 2005, Nouwen *et al.*, 2000). The EpsD protein has a well conserved C-terminal domain that is protease resistant and thought to be anchored in the outer membrane by amphipathic transmembrane β-strands (Bitter *et al*., 1998, Guilvout *et al.*, 1999) as well as a more variable N-terminal domain which extends into the periplasm. Recently, Korotkov *et al.* (Korotkov *et al.*, 2009b) crystallized the N-terminal domain of the periplasmic domain of EpsD from ETEC, and showed a region termed the N0:N1 lobe to be easily accessible and thus a good candidate for interacting with other proteins in the periplasm. These interactions may include other T2S components such as EpsC and/or EpsJ (Douet *et al*., 2004, Korotkov *et al*., 2006) as well as proteins targeted to, and ultimately secreted by, the T2SS (Bouley *et al*., 2001, Shevchik *et al.*, 1997).

Stoichiometry of the T2S Complex:

Although much is known about the interactions between the various proteins of the T2S apparatus, the relative stoichiometry of the individual components remains elusive. It is probable that the cytoplasmic component, EpsE exists as a hexamer, while EpsD is found in an oligomer of 12 or more

proteins. Most of the inner membrane components analyzed to date appear to crystallize as dimers, although whether they remain so in a functional apparatus has yet to be determined. As the inner membrane proteins are likely interacting with each other (either transiently or in a stable complex), it is anticipated that heterocomplexes form in the membrane. The relative combinations of each protein in those complexes remain in question, however. Likewise, if EpsC is able to interact with proteins both in the inner membrane and the outer membrane, it seems likely that it would need to conformationally stretch between a 6-fold symmetry in the inner membrane and a larger 12+-fold symmetry in the outer membrane.

Finally, some of the interactions between the T2S proteins are probably transient and it is unknown whether a complex comprised of all T2S components actually exists as a stable entity at any given time. It is still unknown how the various components come together across the entire Gram-negative cell envelope to form a functional apparatus. It is known that EpsC is a key protein in bridging the inner membrane complex to the outer membrane secretin. In addition, Lybarger *et al.* (Lybarger *et al*., 2009) determined that EpsD is required for proper localization and focal assembly of the T2S complex in the membrane. These data, in combination with results from the *K. oxytoca* EpsD homolog, demonstrate that the secretin can focally localize in the outer membrane in the absence of other T2S components (Buddelmeijer *et al.*, 2009) and suggests that EpsD may be the driving force behind T2S apparatus assembly. Furthermore, EpsC appears to be able to interact with EpsD in the absence of any other T2S

components and therefore could direct the EpsL and M proteins (and presumably the rest of the inner membrane proteins) to where the secretin is located (Lybarger *et al*., 2009).

Structural Analysis of EpsE/related homologs

EpsE is part of the larger family of Type II/IV secretion ATPases which is comprised of ATPases and putative ATPases that are similar in structure, function, and sequence. EpsE and the other T2S ATPases display the greatest level of homology to the T4P assembly ATPases and therefore can further be classified in the GspE/PilB subfamily which is primarily made up of ATPases associated with T2S and T4P biogenesis (Planet *et al.*, 2001, Robien *et al*., 2003). Lower level of homology is observed with members of other subfamilies including the T4P retraction ATPases such as PilT from *P. aeruginosa*, while the lowest degree of homology is observed with the Type IV secretion (T4S) ATPases including Hp0525 from *Helicobacter pylori* (Planet *et al*., 2001).

Crystal structures are available for many ATPases of the Type II/IV secretion family, which allows for careful comparison between the different members. Each ATPase in this family has at least two domains termed the amino-terminal domain (NTD) and the carboxy-terminal domain (CTD), which are connected by a flexible linker. The RecA-like core of the CTD contains conserved motifs such as the Walker A and B boxes, as well as the Asp and His boxes that form the nucleotide binding site (Fig. 1.3) (Possot & Pugsley, 1994, Robien *et al*., 2003, Thomsen & Berger, 2008). Given that this domain has the

majority of the universal features that define the secretion ATPases; it is unsurprising that it is also the domain with the overall highest sequence homology.

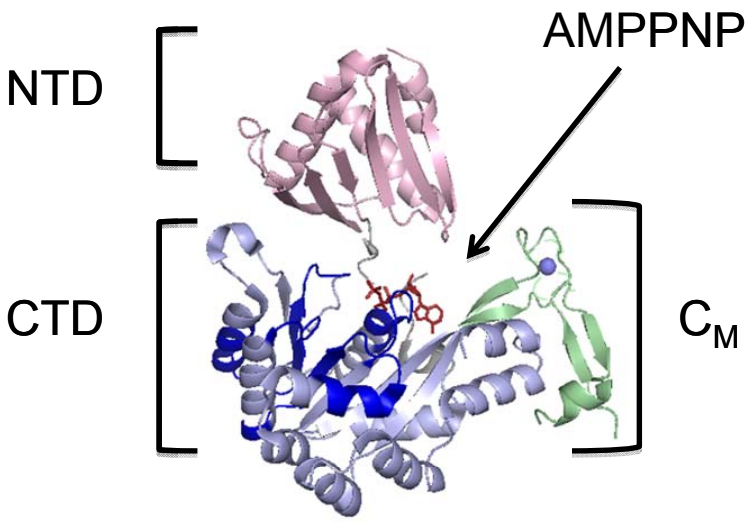


Figure 1.3: The EpsE monomer.
The EpsE monomer with bound AMPPNP is depicted in ribbon structure and colored as follows: the N-terminal domain (NTD) colored pink, the C-terminal domain (CTD) colored blue and the C_M domain colored green. The conserved ATP binding motifs are highlighted in dark blue and the zinc ion is represented as a light blue sphere. Adapted from (Patrick *et al*., 2011).

Even with a conserved ATPase core, these proteins only share 35% sequence identity overall. The NTD of each ATPase has a very low level of sequence homology, but analysis of the crystal structures of this domain show that they are structurally similar. It is likely that this domain imparts system and

species specificity for the different subfamilies of ATPases. EpsE is a nice example of such an occurrence, as previous work has shown the NTD of EpsE co-purifies and co-crystallizes with the T2S inner membrane protein, EpsL (Abendroth *et al*., 2005). This interaction is necessary for EpsE to localize to the larger T2S complex, and may be a common trait among secretion ATPases (Sandkvist *et al*., 1995). Additionally, the NTD may also play a role in regulating ATPase activity through its interactions with other proteins. For example, the T4S ATPase Hp0525 can form a complex with an inhibitory protein, Hp1451 (Hare *et al.*, 2007).

Further analysis of these crystal structures have revealed conformational differences in some of these proteins when in the presence of nucleotides (Hare *et al.*, 2006, Misic *et al.*, 2010, Satyshur *et al.*, 2007, Savvides *et al.*, 2003, Yamagata & Tainer, 2007, Yeo *et al.*, 2000, Robien *et al*., 2003). Analysis of one of the earliest crystal structures, that of the T4S ATPase, Hp0525, suggested this protein closes around the bound nucleotide. Based on this, the authors speculated that this ATPases could act as a gating molecule with the nucleotide state controlling the opening and closing of the protein (Yeo *et al.*, 2000). Later work on AfGspE, an archaeal secretion ATPase from *Archaeoglobus fulgidus*, showed that this may be a universal mechanism, as the NTD of AfGspE was discovered to move as much as 10 degrees with respect to the CTD when nucleotide binds. This work also pin pointed two highly conserved NTD arginine residues that were thought to “clamp down” around the nucleotide, further

suggesting that NTD residues may also directly contribute to the enzymatic activity of the ATPase (Yamagata & Tainer, 2007).

This idea was further substantiated by the identification of two different structural forms of T4P retraction ATPase PilT from *Aquifex aeolicus* (AaPilT) (Satyshur *et al*., 2007). This protein was crystallized as both a hexamer with C6 symmetry as well as a hexamer with C2 symmetry. Closer inspection revealed many conformational changes between the two structures; most notably that of the NTD, which appeared to change position in a given subunit from a wide open to a more closed conformation within the C2 hexamer (Satyshur *et al*., 2007). Later work on the closely related PilT homolog from *P. aeruginosa* (PaPilT) showed that hexamers with C2 symmetry are formed both in the presence and absence of bound nucleotide (Misic et al., 2010). Analysis of both the AaPilT and PaPilT structures revealed that that conformational differences between the different subunits within the C2 hexamer changed not only the position of the NTD, but also seemed to affect the local position of conserved arginine residues in both the NTD, CTD and the neigboring CTD (CTD') (Satyshur *et al*., 2007) as well as the position of two conserved His box histidine residues (Misic *et al*., 2010).

It was previously speculated that a cluster of basic residues spanning the Hp0525 NTD-CTD interface near the nucleotide binding site could have a functional significance for the protein (Savvides *et al*., 2003, Yeo *et al*., 2000). For RecA-like motors, the nucleotide is typically in one monomer, yet, require a residue, usually an arginine dubbed the "arginine finger", from a neighboring

subunit to allow for hydrolysis (Thomsen & Berger, 2008). This phenomenon may explain why RecA-like proteins function as oligomers, as multimerization likely allows the energy from nucleotide hydrolysis to be coupled with conformational changes resulting in mechanical work (Ye *et al.*, 2004). Using the knowledge of RecA's mechanism, coupled with the analysis of the AaPilT C2 hexamer structure, it was proposed that two arginine residues from a neighboring subunit may move in to complete the active site once nucleotide is bound in AaPilT (Satyshur *et al*., 2007). Although not an unusual hypothesis, this was the first time this mechanism was suggested for the Type II/IV family of ATPases (Satyshur *et al*., 2007). Thus, it is likely that the conformational changes associated with ATP binding and hydrolysis proposed above for the secretion ATPases Hp0525, AfGspE, AaPilT and PaPilT could have a substantial effect on the neighboring subunits. This would allow for ATPase activity to be coupled to conformational changes in both the subunits within the hexamer as well as in interactions between the hexamer and the greater protein complex with which they interact (Hare *et al*., 2006, Misic *et al*., 2010, Satyshur *et al*., 2007, Savvides *et al*., 2003, Yamagata & Tainer, 2007, Yeo *et al*., 2000).

While full length EpsE was not amenable to crystallization, the crystal structure was solved in spacegroup $P6_122$ of a truncated form of EpsE lacking a part of the N-terminal domain that interacts with EpsL (Robien *et al*., 2003). In contrast to the ATPases discussed above, EpsE was not crystallized as a hexamer, instead the subunits were found to form a helical filament. Subsequently, EpsE was found to form active oligomers in the presence of the

cytoplasmic domain of EpsL and acidic phospholipids (Camberg *et al*., 2007). In this study, an increase in ATPase activity was detected under the same conditions as those used to promote oligomer formation (Camberg *et al*., 2007, Camberg & Sandkvist, 2005), suggesting a similar mechanism for coupling EpsE's ATPase activity to T2S. Comparison of the liganded and unliganded structures of EpsE did not indicate a significant conformational change due to AMPPNP binding (Robien *et al*., 2003), which may be a consequence attributed to the lack of its hexamerization. We have, however, demonstrated that EpsE, when bound to the cytoplasmic domain of EpsL, must be in its nucleotide-bound form in order to interact with acidic phospholipids in a manner that stimulates its ATPase activity, suggesting that EpsE undergoes a change in conformation in the presence of nucleotide as well (Camberg *et al*., 2007, Robien *et al*., 2003).

EpsE has distinct features compared to some of the other Type II/IV secretion ATPases of known structure, including a unique N-terminal "N1-domain" which is responsible for binding to the cytoplasmic domain of EpsL, and a C-terminal metal-binding "C_M-domain" which is present only in the T2S and T4P assembly ATPases (Camberg *et al*., 2007, Planet *et al*., 2001, Robien *et al*., 2003). These features may contribute to specific functions but may also add to the difficulty in crystallizing hexamers of the T2S ATPases. In the absence of a hexamer, it is therefore of great interest to probe the importance of active site residues and putative interface residues of EpsE to determine if side chains from different subunits contribute to the active site, and, if the inter-subunit interfaces

are similar to those observed in hexameric structures of Type II/IV secretion ATPase homologs.

Zinc coordination in ATPases

As mentioned above, one unique domain found in EpsE is an elbow-shaped loop region, dubbed the C_M domain. This structure contains a $CxxCx_{29}CxxC$ motif which has been shown to coordinate one zinc ion per monomer causing a hairpin turn at the elbow (Camberg & Sandkvist, 2005, Robien *et al*., 2003). This tetracysteine motif is found, albeit with some variability, in most members of the GspE/PilB subfamily of secretion ATPases (with the notable exception of *Xanthomonas campestris* XpsE) along with the T4P assembly ATPases TcpT and PilB, yet is remarkably absent from the T4P retraction ATPase PilT and other members of the PilT/PilU, VirB11 and ComG1 subfamilies (Robien *et al*., 2003). The conservation of this motif suggests an important role for this region to the overall function. Tetracysteine motifs are not limited to T2S ATPases and have been shown to be important for other proteins including zinc finger proteins that participate in DNA binding, the redox regulated chaperone, Hsp33, and the motor protein for Sec-mediated transport, SecA. In Hsp33, the tetracysteine motif is required to regulate dimerization and activity of the protein (Graumann *et al.*, 2001) Under reducing conditions, all four cysteines are reduced and coordinate one zinc ion; however, when exposed to oxidizing conditions, the zinc is released and likely causing significant structural changes and partial activation of chaperone activity. The cysteines residues can then

form pairs of disulfide bonds which allows for dimerization and full activation of the protein (Graf *et al.*, 2004, Graumann *et al*., 2001). For the Sec system, interaction of SecA with SecB requires a conserved domain in SecA which contains three highly conserved cysteines and one residue that is either a cysteine or histidine (Fekkes *et al.*, 1999). Treatment of SecA with a zinc chelator inhibits the binding of SecA to SecB and thus abolishes the SecB-dependent stimulation of SecA's ATPase activity. Further analysis indicated that SecA coordinates one zinc per monomer and this is dependent on the cysteine residues (Fekkes *et al.*, 1999).

Few studies have investigated the role of the elbow to T2S or the tetracysteine motif found therein. A notable one of the *K. oxytoca* genes, demonstrated that replacement of one or more cysteines within the motif severely reduced secretion of pullulanase by the T2S apparatus (Possot & Pugsley, 1997). Further, it was shown that removal of the zinc from purified EpsE with the addition of fourfold molar excess of PMPS reduced, but did not abolish EpsE's ATPase activity *in vitro* (Camberg & Sandkvist, 2005). This may be indicative of the tetracysteine motif playing an indirect role in ATPase activity, and likely is necessary for the overall function of the protein within the T2S complex.

The C_M region terminates at an arginine residue (R441). This residue is highly conserved across the secretion ATPase family and it has been speculated that it could act as a sensor of nucleotide binding state (Robien *et al*., 2003). In EpsE, this residue makes contact at three sites on the bound nucleotide: the

backbone contacts the adenyl group and the side chain forms hydrogen bonds with the ribose moiety (Robien *et al*., 2003), making it tempting to suggest a role for this residue in ATPase function. As this residue connects the elbow with the nucleotide binding site, studies on this region could provide insight into the mechanism of EpsE's ATPase activity and energy transduction through the T2S apparatus.

EpsE interacts with EpsL to facilitate ATP hydrolysis

As mentioned briefly before, EpsE and EpsL interact through the N1 domain of EpsE and the cytoplasmic domain of EpsL. To probe this interaction, EpsE was co-purified with the His-tagged cytoplasmic domain of EpsL and tested for the effects of EpsL on the ATPase activity of EpsE. The presence of EpsL increased ATPase activity 2-fold, which is in agreement to a similar study conducted in *X. campestris* (Camberg *et al*., 2007, Shiue *et al.*, 2006). As EpsL is an inner membrane protein the effect of various phospholipids on the activity of EpsE in complex with EpsL(1-253)His_6 was also examined. Negatively charged phospholipids, specifically cardiolipin (CL), dramatically increased the ATPase activity of EpsE in complex with EpsL(1-253)His_6 and this was modulated in part by an amphipathic helix located at the C-terminal end of EpsL(1-253)His_6 (Camberg *et al*., 2007). Truncating this amphipathic helix reduced interactions with the acidic phospholipid CL and simultaneously reduced the ATPase activity of EpsE-EpsL(1-253)His_6 and also altered protease-sensitivity pattern in the presence of CL. Taken together, these data suggest that the helix is primarily

responsible for interacting with phospholipids. The helix was further analyzed by introducing specific mutations in this region. Replacement of many individual residues within this helix did not significantly alter the ability of the mutant proteins to complement an *epsL* deficient strain of *V. cholerae*; however, the replacement of two specific lysines and two hydrophobic residues compromised the protein's ability to support T2S, indicating these residues are likely playing an important role *in vivo*. Further analysis demonstrated that these mutations diminished EpsL's interaction with CL and reduced the ATPase activity of EpsE, indicating that these residues are required but not sufficient for CL binding and stimulated ATPase activity (Camberg *et al*., 2007).

Significance and scope of these studies

V. cholerae is endemic to southern Asia and parts of Africa, where outbreaks occur seasonally due to increased water temperatures resulting in aquatic plankton blooms and increased numbers of zooplankton, and are usually associated with areas of poor sanitation (Faruque *et al*., 1998). Many times after a natural disaster that includes flooding or other occurrences of standing water and a breakdown in sanitation, major outbreaks of the disease cholera can occur and, without treatment, are devastating. Approximately 5 million people contract cholera annually and if left untreated, many people will die from the disease (Matson *et al*., 2007). During disaster-related outbreaks, many more become infected and die if treatment is not readily available. The standard treatment for this disease is oral rehydration therapy (ORT), which reduces the incidence of

death from >50% to approximately 1% (Matson *et al*., 2007); however, the ability to administer ORT in the areas that need it most is lacking and usually ineffective at eradicating the disease because it fails to treat the source of the *V. cholerae*-infected water. New therapeutic technology is needed to shorten the course and reduce the severity of disease not only to help those who are infected, but also to control the outbreak of the bacteria themselves in order to gain control over the contaminated areas before an outbreak becomes unmanageable.

A potential target for these new therapeutic measures would be the T2SS. This apparatus is not only conserved across *V. cholerae*, including all toxigenic strains, but is also found in the majority of Gram negative pathogens, and could therefore provide therapeutic measures for a wide range of diseases. The energy-providing component, EpsE is required for T2S and is highly conserved in not only the T2S family but also in the T4P biogenesis pathway, among others and, therefore, the understanding of how this protein works could give us broad insight into how to target it for therapeutics.

The overall goal of this study is to characterize the subdomains of EpsE to understand how they contribute to its function specifically as an ATPase and more generally within the T2S apparatus. The work presented here contributes a significant step toward the understanding of how EpsE works and allow us to speculate how it may transduce its energy to support T2S.

When characterizing EpsE, I first wanted to determine the mechanism of ATPase activity. It was previously understood that EpsE could hydrolyze ATP and this required interactions with EpsL and membrane phospholipids; however,

the mechanism of hydrolysis remained elusive. Even though it had been speculated that EpsE needs to hexamerize in order to function, there was no definitive data to support this, nor was the necessity of multimerization to the function of the protein well understood. In Chapter 2, I and co-authors Konstantin Korotkov, Wim G. J. Hol and Maria Sandkvist investigated the contribution of specific arginine residues to ATPase activity as well as analyzed the contribution of the residues in the subunit interface in an attempt to confirm our longstanding hypothesis that EpsE functions as an oligomer. The experiments conducted in Chapter 2 probed the formation of the active site, which brings together arginine residues from multiple subunits and domains. Further, we examined residues within the subunit interface to understand their contribution to oligomerization as well as to the overall function of this protein. Taken together, we were able to show that EpsE must function as an oligomer and our data support the idea that ATP hydrolysis induces dynamic movements which could be directly translated to mechanical force, prompting secretion through the T2S apparatus.

Next I wanted to investigate the contribution of the C_M domain, which contains the zinc-coordinated elbow to EpsE's ATPase activity and T2S in general. Very little is known about this part of the protein, therefore my data represents a significant advance in the understanding of this subdomain and how its role in ATPase activity may be translated specifically to the T2SS. In Chapter 3 I characterized the unique zinc-coordinated elbow that is highly conserved within the GspE/PilB subfamily of secretion ATPases and determined, through

deletion of the majority of the C_M region, that this subdomain is required for T2S. I also present data suggesting that making serine replacements within the tetracysteine motif has negative effects, likely related to defects in the folding of this subdomain, which is detrimental for its function. Further I explored the idea that this subdomain may participate in protein-protein interactions with other components of the T2SS, and speculate these interactions could be directly tied to nucleotide interactions through R441; the only C_M residue that contacts the bound ATP.

Collectively, these data expand the current knowledge of how secretion ATPases function in general, and, in particular, how EpsE may promote secretion through the T2S apparatus. Gaining a better appreciation for the mechanism of ATP hydrolysis will expand the current knowledge of the ATP hydrolysis mechanism and could lead to the development of therapeutics specifically targeting the ATPase, which could halt not only T2S, but as this protein has conservation across its domains, potentially many other ATPases within this family.

Chapter 2

Oligomerization of EpsE coordinates residues from multiple subunits to facilitate ATPase activity

Abstract

EpsE is an ATPase that powers transport of cholera toxin and hydrolytic enzymes through the type II secretion (T2S) apparatus in the Gram negative bacterium, *Vibrio cholerae*. On the basis of structures of homologous Type II/IV secretion ATPases and our biochemical data, we believe that EpsE is active as an oligomer, likely a hexamer, and the binding, hydrolysis and release of nucleotide causes EpsE to undergo dynamic structural changes thus converting chemical energy to mechanical work, ultimately resulting in extracellular secretion. The conformational changes that occur as a consequence of nucleotide binding would realign conserved arginines (R210, R225, R320, R324, R336, and R369) from adjoining domains and subunits to complete the active site around the bound nucleotide. Our data suggests that these arginines are essential for ATP hydrolysis, although their roles in shaping EpsE's active site are varied. Specifically, we have shown that replacements of these arginine

residues abrogate the T2S process due to a reduction of ATPase activity, yet, do not have any measurable effect on nucleotide binding or oligomerization of EpsE. We have further demonstrated that point mutations in the EpsE intersubunit interface also reduce ATPase activity without disrupting oligomerization, strengthening the idea that residues from multiple subunits must precisely interact in order for EpsE to be sufficiently active to support T2S. Our findings suggest that the action of EpsE is similar to that of other Type II/IV secretion ATPase family members and thus these results may be widely applicable to the family as a whole.

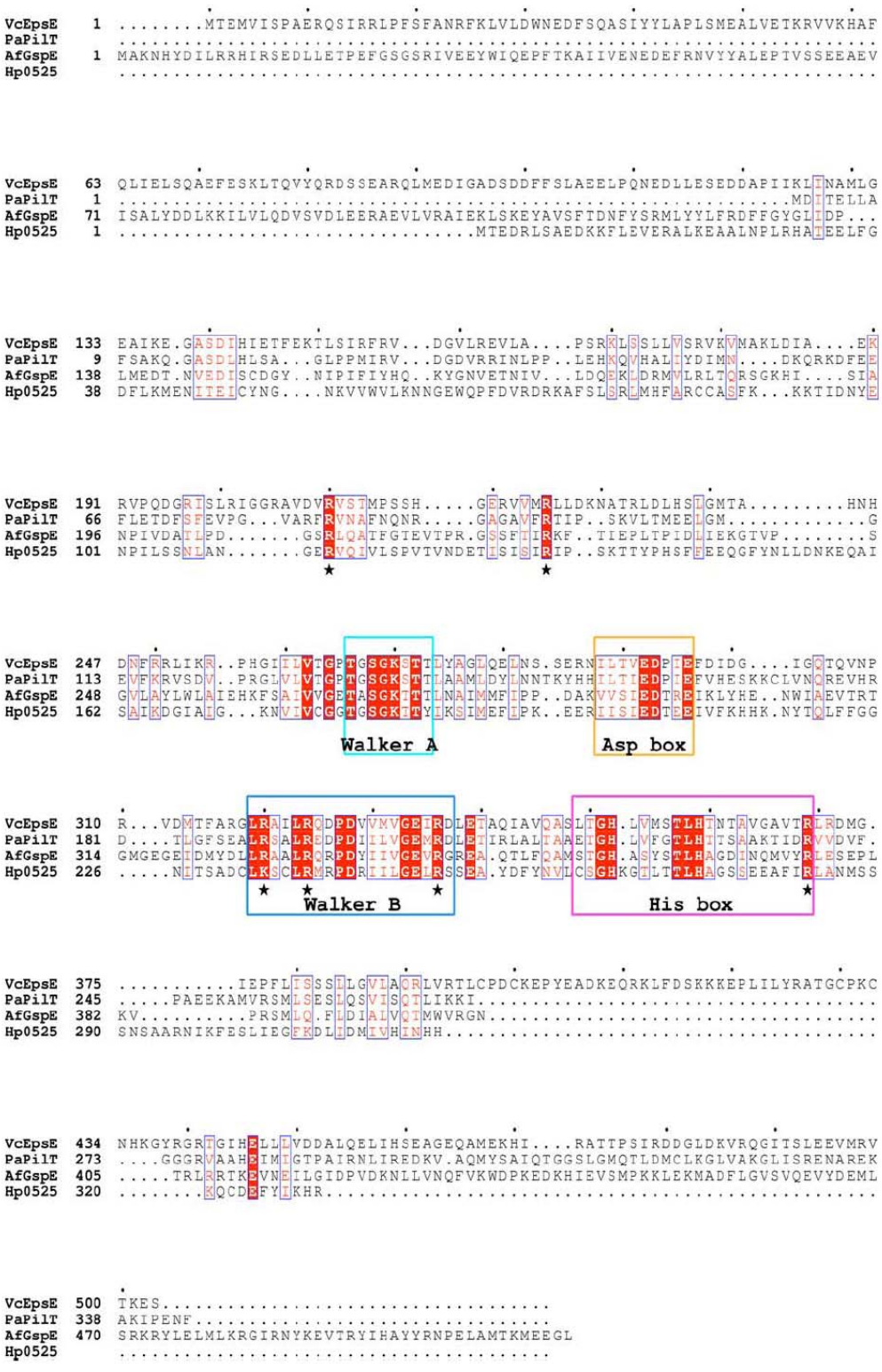
Introduction

The T2S system is composed of at least 12 gene products which assemble into a large protein complex spanning the entire cell envelope of Gram-negative bacteria, including *V. cholerae* (Cianciotto, 2005, Johnson *et al*., 2006). Following transport through the cytoplasmic membrane and folding in the periplasm, toxins and hydrolytic enzymes are targeted to the T2S system and secreted in their native form through a pore in the outer membrane formed by the secretin EpsD. This transport is thought to occur through a piston-like mechanism due to the polymerization of a pseudopilus made up of proteins EpsG-K (Filloux, 2004, Hobbs & Mattick, 1993, Sandkvist, 2001a). The process is energy dependent, requiring the cytoplasmic ATPase, EpsE (Camberg & Sandkvist, 2005). EpsE associates with the T2S system through interactions with EpsL, a protein found in an inner membrane complex also comprised of proteins EpsC, EpsF, and EpsM (Abendroth *et al*., 2005, Lybarger *et al*., 2009, Py *et al*., 2001, Robert *et al*., 2005, Sandkvist *et al*., 1995). The use of an ATPase to provide the energy for assembly and/or substrate translocation is an important feature of not just the T2S machinery, but of several molecular trafficking systems (Filloux, 2004).

EpsE is a member of the large family of Type II/IV secretion ATPases (Planet *et al*., 2001, Robien *et al*., 2003) which includes ATPases involved with T2S, Type IV pilus (T4P) biogenesis and Type IV secretion (T4S) systems and participates in diverse processes such as protein transport, DNA uptake and export, archaeal flagella assembly, and pilus extension or retraction (Craig *et al.*,

2004, Peabody *et al.*, 2003). EpsE and the other T2S ATPases display the greatest level of homology to the T4P assembly ATPases (Planet *et al*., 2001). Lower level of homology is observed with the T4P retraction ATPases as shown for *P. aeruginosa* PilT, while the lowest degree of homology is observed with the T4S ATPases including Hp0525 from *Helicobacter pylori* (Fig. 2.1) (Planet *et al*., 2001).

Figure 2.1: Sequence Alignment.
The *V. cholerae* EpsE sequence was aligned with sequences from *P. aeruginosa* PaPilT, *A. fulgidus* AfGspE, and *H. pylori* Hp0525 using 3D superpositions [Protein structure comparison service SSM at European Bioinformatics Institute (http://www.ebi.ac.uk/msd-srv/ssm), authored by E. Krissinel and K. Henrick]. Identical residues (red boxes) and similar residues (red letters) are primarily concentrated within the nucleotide binding site motifs in the CTD: the Walker A, Walker B, Asp, and His boxes, which are denoted with rectangles colored in cyan, blue, orange, and magenta, respectively (see also Figure 2.2). The EpsE arginine residues examined in this study are labeled by black stars.


Walker A
Asp box
Walker B
His box

The structurally characterized members of this family have two distinct domains: the carboxy-terminal domain (CTD) and the amino-terminal domain (NTD), which are connected through a short flexible linker. Although they have less than 35% overall sequence identity, these proteins share highly homologous sequence motifs in the RecA-like core structure found in the CTD, including a Walker A box encompassing a P-loop structure, and a less well defined Walker B box. Additionally, the CTD includes Asp and His boxes which taken together with the Walker A/B boxes make up the nucleotide binding site (Fig 2.1) (Possot & Pugsley, 1994, Robien *et al*., 2003, Thomsen & Berger, 2008). There is greater size and sequence variation among the NTDs than among the CTDs. These differences likely impart system and species specificity for the different subgroups. Interactions through the NTD could serve to localize the ATPase to a membrane-bound complex as was shown for the N-terminal subdomain N1 of EpsE's NTD which directly interacts with the cytoplasmic membrane protein EpsL within the T2S system (Abendroth *et al*., 2005, Sandkvist *et al*., 1995). Alternatively, NTD interactions with other proteins may help regulate ATPase activity as was seen for Hp0525 in complex with an inhibitory protein, Hp1451 (Hare *et al*., 2007). Regardless of its interaction with other proteins, it should be noted, that the NTD may also contribute directly to the enzymatic activity of these ATPases by providing highly conserved residues to complete the active site (Fig. 2.1) (Misic *et al*., 2010, Robien *et al*., 2003, Satyshur *et al*., 2007, Yamagata & Tainer, 2007).

Structural analyses indicate that significant conformational changes occur in Type II/IV secretion ATPases as they bind nucleotides (Hare *et al*., 2006, Misic *et al*., 2010, Robien *et al*., 2003, Satyshur *et al*., 2007, Savvides *et al*., 2003, Yamagata & Tainer, 2007, Yeo *et al*., 2000). Early work on the T4S ATPase from *H. pylori*, Hp0525, suggested that it may function as a gating molecule at the bacterial inner membrane and regulate its opening and closing by nucleotide binding, hydrolysis and release (Yeo *et al*., 2000). Further analysis indicated that Hp0525 may function as a dynamic hexamer capable of forming a compact, locked structure when in a nucleotide-bound conformation (Savvides *et al*., 2003). More recently, the structure of AfGspE, an ATPase from *Archaeoglobus fulgidus*, showed that the NTD domain moves by 10 degrees with respect to the CTD in the presence of nucleotide, thus allowing for two conserved arginine residues from the NTD to “clamp” around the nucleotide (Yamagata & Tainer, 2007).

The identification of two different structural forms of the T4P retraction ATPase PilT from *Aquifex aeolicus* (AaPilT) provided additional insight into domain movements and subunit realignments which can occur within hexamers of this family of ATPases (Satyshur *et al.*, 2007). In both the ATP and ADP bound forms, AaPilT was observed as hexamers with C6 symmetry. Another crystal structure of AaPilT, at 4.2 Å resolution, revealed a hexamer with approximate C2 symmetry, which contains AaPilT subunits in several different conformations, ranging from a more closed conformation than seen in the AaPilT hexamer with C6 symmetry, to a wide open conformation. In the latter subunits, a large domain

rotation of ~65 degrees compared to the closed form was observed (Satyshur *et al.*, 2007).

More recently, structures of a PilT homolog from *Pseudomonas aeruginosa* (PaPilT) showed that hexamers with C2 symmetry are formed both in the presence and absence of nucleotide (Misic *et al*., 2010). Although the differences between these hexamers are not as drastic as those between the AaPilT hexamer with perfect C6 symmetry and quasi C2 symmetry, the subunit of PaPilT still appears to change from open to closed conformations depending on nucleotide binding. This movement is most clearly demonstrated by looking at the changing positions of two NTD arginines and two "His box" histidines within the three different subunit conformations observed in the hexamer with C2 symmetry. Taken together, the conformational changes associated with ATP binding and hydrolysis proposed for Hp0525, AfGspE, AaPilT and PaPilT could have a substantial effect on the neighboring subunits of the hexamer as well as on interactions between the hexamer and the greater protein complex with which they interact (Hare *et al*., 2006, Misic *et al*., 2010, Satyshur *et al*., 2007, Savvides *et al*., 2003, Yamagata & Tainer, 2007, Yeo *et al.*, 2000).

While full length EpsE was not amenable to crystallization, the crystal structure was solved of a truncated form of EpsE lacking a part of the N-terminal domain that interacts with EpsL (Robien *et al*., 2003). In contrast to the ATPases discussed above, EpsE was not crystallized as a hexamer, instead the subunits were found to form a helical filament, a phenomenon that has also been observed with other hexameric ATPases (Sawaya *et al.*, 1999, Story *et al.*,

1992). The helical nature of EpsE does likely not represent the quaternary structure of EpsE, however, as subsequent analysis showed that EpsE is able to form active oligomers, which are likely hexamers, in the presence of the cytoplasmic domain of EpsL and acidic phospholipids (Camberg *et al*., 2007). A correlation between oligomer formation and increased ATPase activity was demonstrated for EpsE (Camberg *et al*., 2007, Camberg & Sandkvist, 2005), suggesting that the communication between adjoining monomers may be key to coupling nucleotide hydrolysis to substrate trafficking for EpsE as well.

In this study we have modeled EpsE as a hexameric ring with C2 symmetry based on the PaPilT-AMPPCP structure (Misic *et al*., 2010). Based on this model, we have designed experiments to probe the importance of putative active site residues and interface residues of EpsE to determine if side chains from different domains and subunits contribute to the active site, and, if the inter-subunit interface is similar to those observed in hexameric structures of Type II/IV secretion ATPase homologs. In the EpsE model, three pairs of conserved arginines, one pair from the NTD, one from the CTD and one from the CTD of a neighboring subunit (CTD'), appear to realign once nucleotide binds in order to facilitate ATP hydrolysis. The possible involvement of these arginines in ATP binding and hydrolysis was originally proposed by Forest and colleagues, but only a few point mutants were tested for twitching motility *in vivo* and no mutant protein was purified for *in vitro* analyses (Misic *et al*., 2010, Satyshur *et al.,* 2007). In order to determine the role of these arginines in EpsE's function, we have replaced each of the arginines with alanine. Our results indicate that

substitution of each individual arginine abrogates both ATP hydrolysis and the T2S process. Additional mutations within the intersubunit interface similarly disrupts EpsE's function, providing support for the hexamer model and indicating that a precise arrangement of residues from different subunits around the active site is essential for the action of EpsE and other T2S ATPases.

Experimental Procedures

A hexameric ring model of EpsE:

A model was generated using the crystal structure of EpsE (Robien *et al*., 2003) and the available C2 hexameric structure of the homologous PaPilT (Misic *et al*., 2010, Satyshur *et al*., 2007). The structurally conserved core of the EpsE-CTD was superimposed onto CTD's in this hexameric structure (PDB ID: 3JVV). The EpsE-NTD's were added next to the hexameric model to maintain the extensive CTD:NTD' interface observed in the EpsE crystal structure (Robien *et al*., 2003).

Cloning and Expression:

EpsE point mutations R156A, R156D, R210A, R225A, R320A, R324A, D326R, D328R, R336A, and R369A were introduced into the previously constructed plasmid, pET21dEpsE(2-503)-EpsL(1-253)His_6 (Camberg *et al*., 2007) using the QuickChange Site-Directed Mutagenesis kit (Stratagene) as directed. The fragments of *epsE* containing the various mutations were cloned into pMMB384 (Sandkvist *et al.*, 1993) by exchange of an MfeI/BamHI fragment to create the individual pMMBEpsE point mutant plasmids. The pMMB plasmids were then introduced to the *epsE*$^-$ and wild type *V. cholerae* strains through conjugation.

Purification of protein:

The above constructed EpsE(2-503)-EpsL(1-253)His_6 point mutants, with the exception of the one containing D328R, were expressed in *Escherichia coli* BL21(DE3) under IPTG inducing conditions and purified using metal affinity chromatography and gel filtration as described in (Abendroth *et al.*, 2004a, Abendroth *et al*., 2005).

Phospholipids:

E. coli-derived cardiolipin was prepared in 20mM Tris pH 8.0, 150 mM NaCl, 1 mM DTT as described in (Camberg *et al*., 2007).

Protease Secretion Assay:

pMMB plasmids containing the point mutations of *epsE* were expressed in wild type *V. cholerae*, or the *V. cholerae epsE*$^-$ strain in the presence of 10 μM IPTG and tested for secretion of a protease using a modified fluorescence-based assay as described in (Sandkvist *et al.*, 1997). Supernatants from LB-grown overnight cultures were separated from cells by two centrifugation steps and assayed for proteolytic cleavage of N-tert-butoxy-carbonyl-Gln-Ala-Arg-7-amido-4-methyl-coumarin (MCA) (Sigma) as described in (Johnson *et al*., 2007, Sandkvist *et al*., 1997) using excitation and emission wavelengths of 385 nm and 440 nm, respectively. The protease activity was determined by measuring the change in fluorescence over time and comparing to an MCA standard. The amount of MCA generated per minute was normalized by OD_{600} with error

reported as standard error (SEM) of at least three separate samples assayed in triplicate.

ATPase Activity Assays:

ATPase activities of purified proteins were measured using a kinetic assay containing 0.5 μM protein, 5 mM ATP, 5 mM $MgCl_2$, and 125 μM cardiolipin in buffer A (100 mM HEPES pH 8.5, 65 mM NaCl, 5% glycerol). Reactions were incubated at 37ºC for 0, 15, 30, 45, and 60 minutes for EL1 and for 0, 120, 180, 240, and 300 minutes for R210A, R225A, R320A, R324A, R336A, R369A, R156D, R156A, D326R, and BSA and assayed for the release of inorganic phosphate using the BioMol Green reagent as directed. Amount of phosphate was determined by comparing sample absorbance to those of a phosphate standard curve with error reported as standard error (SEM) of at least three separate samples assayed in duplicate.

Liposome Floatation Assay:

Reaction mixtures containing 1 μM protein, 4 mM AMPPNP, 4 mM $MgCl_2$, and 250 μM cardiolipin in buffer A were incubated for 30 minutes at 37ºC, then mixed with 0.8 mL of 20 mM HEPES pH 8.5 containing 85.5% sucrose and layered on the bottom of a 5 mL centrifuge tube. 3 mL of 20 mM HEPES buffer containing 65% sucrose was layered on top, followed by a 1 mL layer of 20 mM HEPES buffer containing 10% sucrose. Samples were centrifuged at 115,000xg (35,000 rpm) for 16 hours at 4ºC using a Sorvall AH650 rotor. Five 1 mL fractions

were collected from the top of each tube and analyzed by SDS-PAGE using the NuPAGE system (Invitrogen) and immunoblotted against EpsE. The top two fractions (# 4-5) represent the cardiolipin-bound protein.

Cross linking*:*

Samples containing 125 μM cardiolipin, 5 mM ATP, 5 mM $MgCl_2$, and 1 μM purified proteins were incubated in buffer A and subjected to the chemical cross linker disthiobis succinimidyl propionate (DSP) at concentrations of 0, 10, 50 and 100 μM as described in (Camberg *et al*., 2007).

Quantification of immunoblots*:*

Densitometry analysis was performed on a Typhoon TRIO (Amersham Biosciences) using ImageQuant TL software. For the liposome binding assay, the % EpsE floated represents the amount of EpsE in F4+F5 divided by the total amount of EpsE in F1, F2, F4 and F5. For quantification of EpsE cross-linking, the amount of monomer remaining in the 100 μM DSP lane was divided by the amount of monomer in the 0 μM DSP lane to obtain the % uncross-linked monomer.

Statistical analysis*:*

A Student's t-test was used for all statistical analyses. Values were considered significant at $p<0.05$.

Thermal Melts*:*

The thermal stability of proteins were measured by fluorescence using Sypro Orange fluorophore as described in (Lo *et al.*, 2004), either without additives, or in the presence of 5 mM ATP and 5 mM $MgCl_2$.

Results

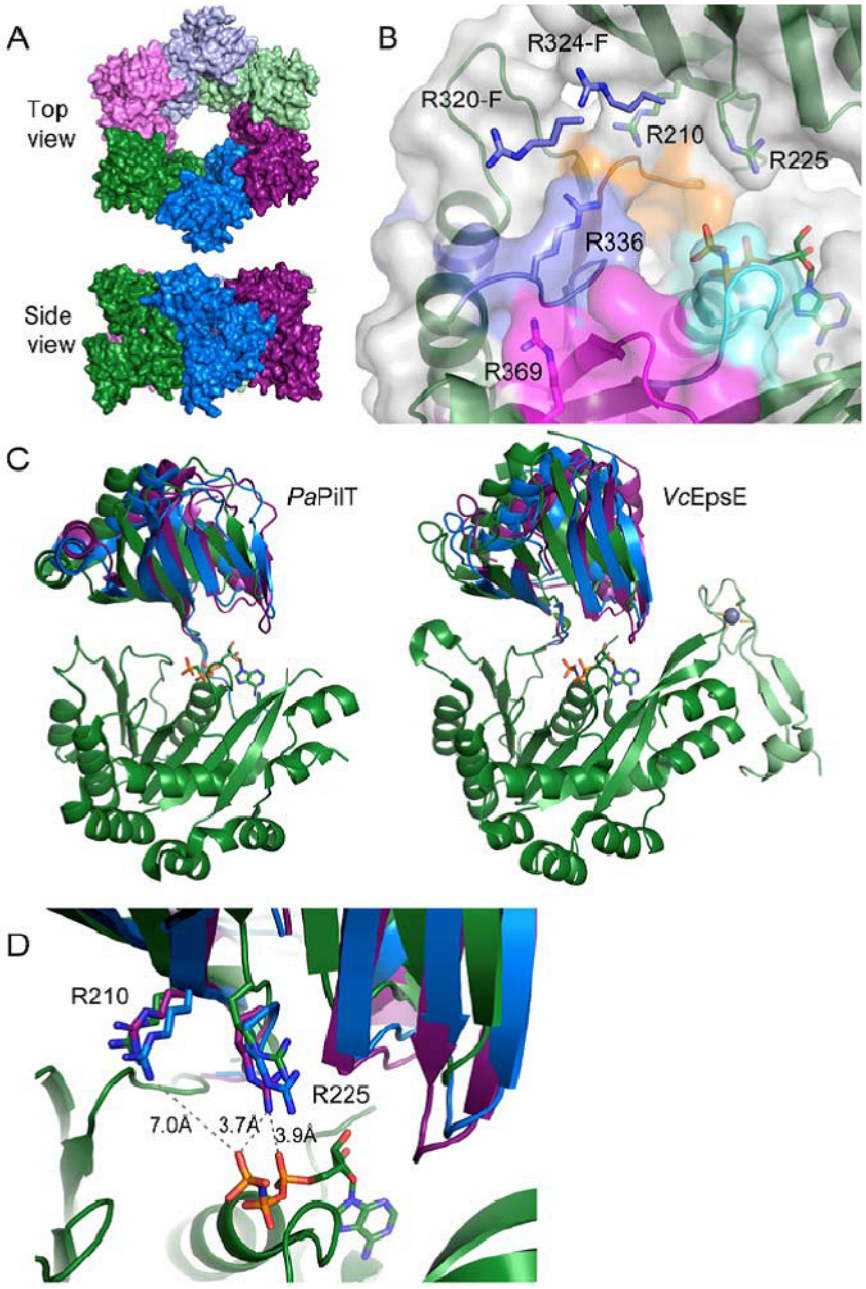
Modeling EpsE as an asymmetric hexamer:

The crystal structure of EpsE without its N-terminal 90 residues was previously solved in spacegroup $P6_122$ where the individual subunits were found in a helical filament; however, based on its structural similarity with the T4S ATPase Hp0525, then the only known hexameric structure of the family, EpsE was modeled as a hexamer (Robien *et al*., 2003). In view of the fact that EpsE is more closely related to PilT than to Hp0525 (Fig. 2.1) (Planet *et al*., 2001), we have updated our working model of the functional hexameric ring of EpsE based on the structure of PaPilT.

In order to identify residues outside the immediate nucleotide binding site that are important for ATP hydrolysis and to test the hypothesis originally put forth by Forest and colleagues that there are conserved arginines that move in to complete the active site once nucleotides bind, we have constructed a C2 hexamer model of EpsE using the structure of the related PaPilT as a template (Fig. 2.2A). The positions of six highly conserved arginines (Fig. 2.1) are conserved in the new EpsE hexameric model with C2 symmetry which emphasizes the idea of a universal function for these residues in the activity of these ATPases. A closer assessment of the two conserved NTD arginines reveals a similar movement toward the bound nucleotide as seen for both AfGspE and PaPilT, further substantiating the hypothesis that these residues may be required for EpsE to hydrolyze ATP (Fig. 2.2B). When examining the four invariant CTD arginine residues relative to the bound nucleotide, we see that

their placement within the conserved motifs would likely allow them to contribute to the completion of the active site, if not directly participate in nucleotide hydrolysis (Figs. 2.2C, 2.2D). In order to determine the functional role of the identified arginines *in vivo* and *in vitro* we constructed EpsE variants with the following mutations: R210A, R225A, R320A, R324A, R336A and R369A.

Figure 2.2: Hexameric model of EpsE.
(A) EpsE modeled after the PaPilT (PDB ID: 3JVV (Misic *et al.*, 2010)) structure which forms hexameric rings with C2 symmetry, i.e. with three different subunit conformations. Subunits A, B and C are shown in dark green, blue and purple, respectively. Subunits D, E and F are shown in light green, blue and purple, respectively, to indicate that subunits A and D, B and E, C and F have identical structures due to the C2 symmetry of the PaPilT hexamer. (B) The nucleotide binding site of EpsE in the hexamer model formed by subunits A and F. The nucleotide shown is AMPPNP bound to the CTD of EpsE (Robien *et al*., 2003). Residues corresponding to the Walker A, Walker B, Asp and His boxes are highlighted in cyan, blue, orange and magenta, respectively. The side chains of functionally important arginine residues are shown in stick representation. R210 and R225 belong to the NTD, and R336 and R369 to the CTD of subunit A whereas R320-F and R324-F belong to CTD' of the neighboring subunit F. (C) The EpsE hexamer model predicts NTD-CTD motions similar to PaPilT structure. Left: subunits A, B and C of PaPilT structure superimposed using CTD only. NTD's of subunits A, B and C are shown in green, blue and purple respectively; only the CTD of subunit A is shown for clarity. On the right: a superposition of subunits A, B and C of the EpsE hexamer model. A zinc-binding C_M domain of EpsE is highlighted in light green with the Zn^{2+} ion shown as a grey sphere and the side chains of coordinating cysteines shown as sticks. (D) A close-up view of nucleotide binding sites in different EpsE subunits shown in (C). Side chains of R210 and R225 are shown as sticks.


A
Top view
Side view
B
R324-F
R320-F
R210
R225
R336
R369
C
PaPilT
VcEpsE
D
R210
R225
7.0Å
3.7Å
3.9Å

Replacement of conserved arginine residues disrupts T2S:

To test the importance of these arginine residues, first, the plasmid-encoded EpsE mutant proteins were analyzed for their ability to complement the secretion defect observed in an *epsE*$^-$ mutant strain of *V. cholerae* (Fig. 2.3A). Wild type EpsE was able to restore secretion; however, none of the six individual point mutants (including R320A) were able to complement the *epsE*$^-$ mutant to a statistically significant level over the vector only control as shown in Figure 2A. Subsequent immunoblotting analysis indicated that the mutant proteins were expressed at levels similar to that of wild type EpsE (Fig. 2.3B) suggesting that the mutations do not affect the stability of EpsE but rather its function.

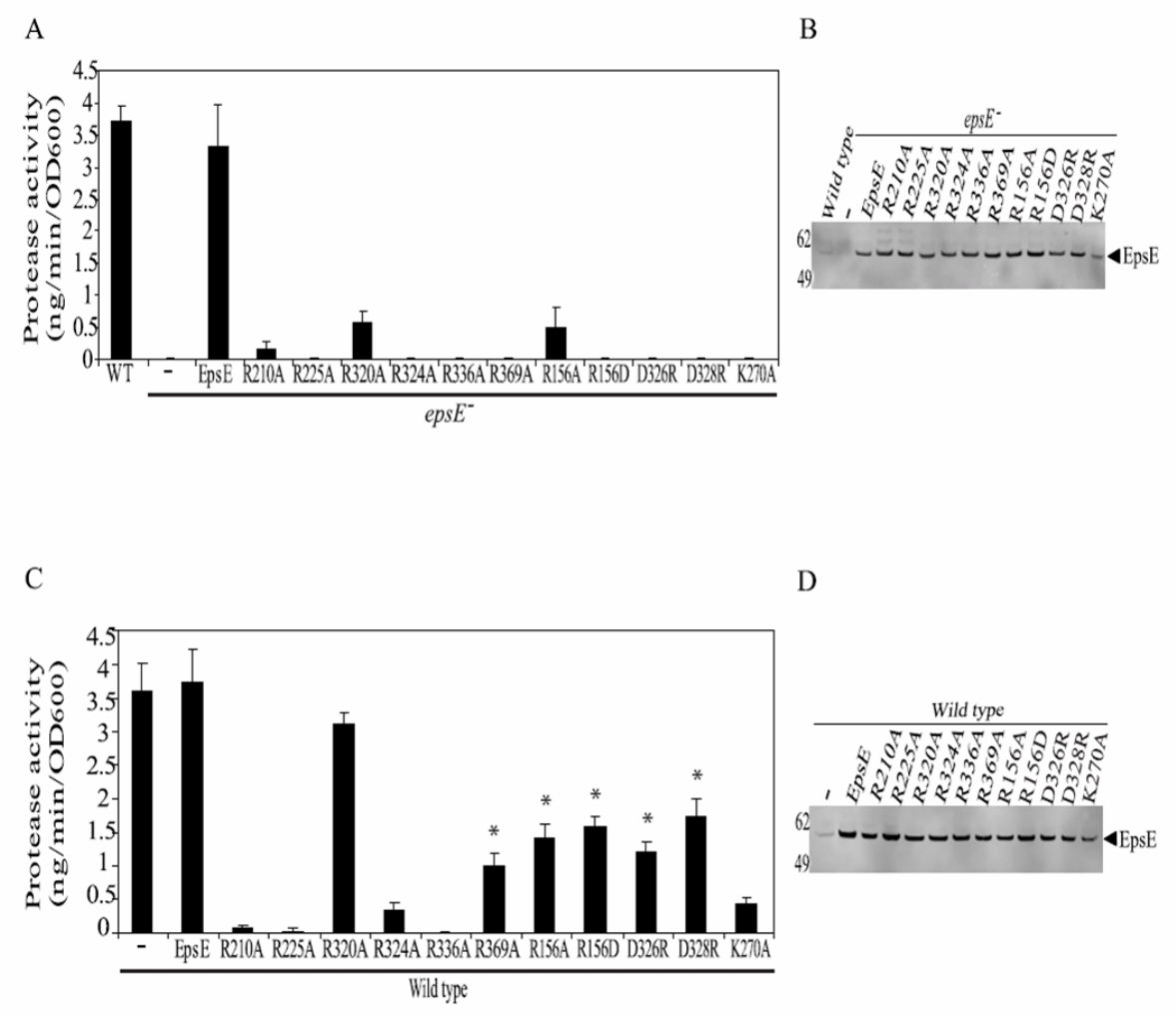


Figure 2.3: EpsE point mutations inhibit protease secretion *in vivo*.
epsE- (A, B) or wild type (C, D) strains of *V. cholerae* containing pMMB67 vector, pMMB EpsE or the plasmid-encoded EpsE with the following point mutations:R210A, R225A, R320A, R324A, R336A, R369A, R156A, R156D, D326R, D328R, or K270A were grown overnight in LB + 10μM IPTG. Culture supernatants and cells were separated by centrifugation. (A, C) Culture supernatants were analyzed for the presence of extracellular protease activity using a fluorescence-based assay with peptide-conjugated MCA as a substrate. The amount of MCA liberated was compared to a standard curve with each sample assayed at least three independent times. The results are presented as ng MCA hydrolyzed per minute, normalized to the OD_{600} of the culture +/- standard error. * indicates a statistically significant difference ($p<0.05$) when compared to both wild type and EpsE-K270A. (B, D) Cell extracts were subjected to SDS-PAGE and immunoblotted with anti-EpsE. Molecular mass markers are indicated on the left and the position of EpsE is indicated by the arrow.

To further verify that the mutations are not grossly affecting the folding of EpsE the mutant proteins were analyzed for their ability to interact with the rest of the T2S apparatus and inhibit secretion in wild type *V. cholerae*. Their effect on secretion was compared to both plasmid-encoded wild type EpsE as well as the nucleotide binding site mutant, EpsE-K270A, which was previously shown to block secretion in wild type *V. cholerae* (Camberg & Sandkvist, 2005, Sandkvist *et al.*, 1995). As shown in Figure 2.3C, all of the mutant EpsE proteins, with the exception of EpsE-R320A, acted in a dominant negative manner when over-expressed in wild type *V. cholerae*. Although the R369A mutant did display trans-dominance over wild type, it did not inhibit secretion as efficiently as the K270A mutant ($p<0.05$). The dominant phenotype could be due to competition between wild type and mutant EpsE for binding to other T2S components such as EpsL, and/or due to the formation of mixed oligomers containing wild type and mutant EpsE.

EpsE point mutations abolish ATPase activity *in vitro:*

Previously we have demonstrated that purified EpsE, when in the presence of EpsL(1-253)His_6 and acidic phospholipids, such as cardiolipin (CL), assembles into larger oligomers and displays ATPase activity (Camberg *et al*., 2007). Therefore, we next tested whether the above-chosen residues are required for hydrolysis of ATP in the presence of EpsL(1-253)His_6 and CL. The point mutations R210A, R225A, R320A, R324A, R336A, and R369A were individually introduced into the previously created EpsE(2-503)-EpsL(1-253)His_6

construct and the mutant proteins were purified by metal affinity chromatography and gel filtration as described (Abendroth *et al*., 2004a, Abendroth *et al*., 2005). The finding that all of these mutants (including EpsE-R320A) could be co-purified with EpsL(1-253)His_6 by these methods indicates that the interaction of EpsE and EpsL was not disrupted as a result of these mutations. As Figure 2.4 demonstrates, each individual point mutation abrogated the ATPase activity of EpsE in the presence of CL more than 50-fold when compared to wild type protein. The rates of ATP hydrolysis for the mutant proteins were not statistically different from that of the BSA control, indicating that all six arginine mutants are inactive.

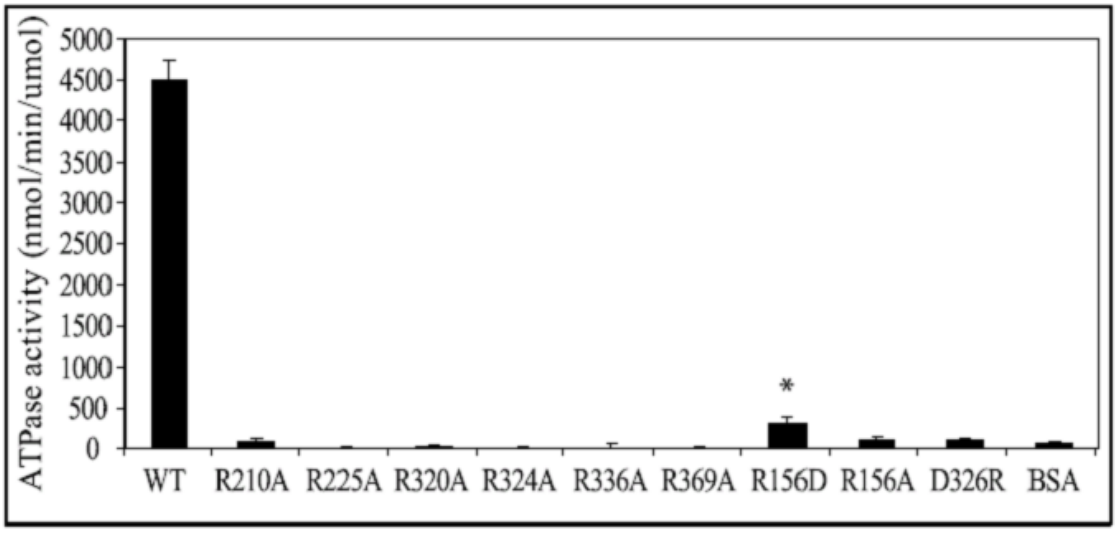


Figure 2.4: EpsE point mutations block CL-stimulated ATPase activity *in vitro*.
CL-stimulated ATPase activity was measured kinetically for wild type EpsE-EpsL(1-253)His_6 over a period of 60 minutes and the individual EpsE mutants with the following mutations: R210A, R225A, R320A, R324A, R336A, R369A, R156D, R156A, D326R or BSA over a period of 300 minutes in reactions containing Mg-ATP and CL. Each protein was assayed at least three separate times and the means +/- standard error are presented. * indicates a statistically significant difference ($p<0.05$) when compared to BSA.

EpsE mutants retain nucleotide binding capabilities:

To determine if the reduced ability to hydrolyze ATP for the mutant EpsE proteins was due to defects in ATP binding we performed thermal melt assays. We determined the temperature at which wild type EpsE, and the individual EpsE point mutants R210A, R225A, R320A, R324A, R336A, and R369A in complex with EpsL(1-253)His_6 melted in the presence or absence of Mg-ATP as a measure of nucleotide binding. As seen in Figure 2.5, each protein showed a significant increase in melting temperature in the presence of Mg-ATP over their respective apo forms, indicating that nucleotide binding increases the stability of each EpsE variant. Furthermore, the wild type and mutant proteins responded with a very similar thermal melt shift to the presence of nucleotide, demonstrating that these proteins have comparable nucleotide binding capabilities. Another important item to note is the fact that the EpsE mutant proteins melted within two degrees of the average melting temperature of wild type protein in both the presence and absence of ATP, suggesting comparable overall stability for these proteins. Taken together with the data from the ATPase assay, the mutations do not appear have a negative effect on ATP binding, but rather on EpsE's ability to hydrolyze ATP.

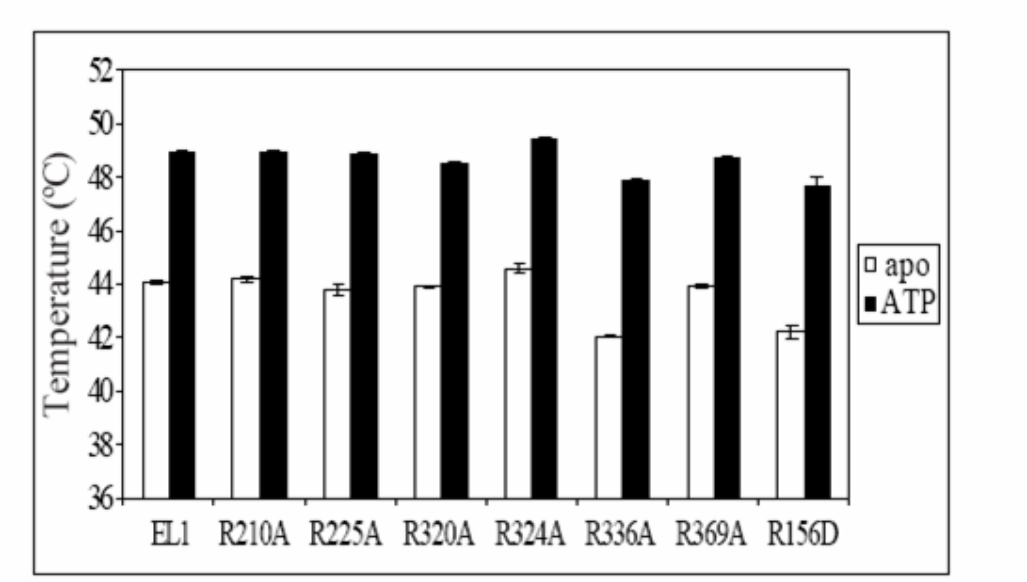


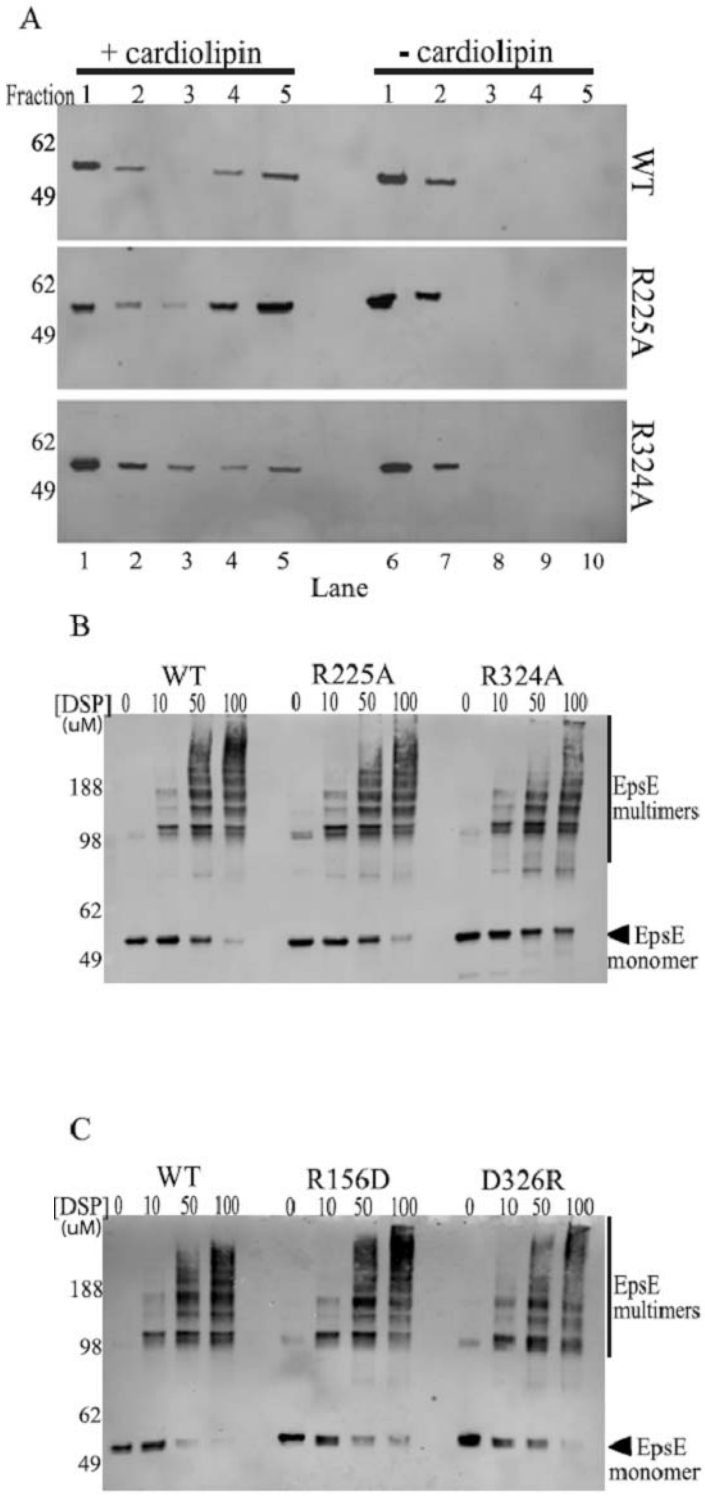
Figure 2.5: EpsE point mutants can bind ATP.
Average thermal melt temperatures for wild type protein, and EpsE variants with the following mutations, R210A, R225A, R320A, R324A, R336A, R369A, and R156D, in the absence or presence of ATP are shown. Melting temperatures for each protein was assayed at least eight times on two separate days and the means +/- standard error are presented.

Interaction with CL is unaltered in EpsE mutants:

To determine whether the loss of activity of the point mutants is solely due to a defect in forming the complete active site or if there is another factor indirectly contributing to the reduced ATP hydrolysis such as altered CL binding, we incubated the purified EpsE-EpsL(1-253)His_6 variants at 37ºC in the presence of Mg-AMPPNP with or without CL. These liposome binding experiments sought to directly measure whether the mutant proteins retained their ability to interact with acidic phospholipids comparable to that of wild type, as failure to interact with CL could account for the loss of ATPase activity of the point mutants. After

incubation the samples were added to an 85.5% sucrose solution in the bottom of a centrifuge tube and a sucrose gradient was layered on top. After centrifugation, samples were taken from the top of the tube with the highest fractions representing those which contained the phospholipids and phospholipids-bound protein. Each fraction was analyzed by SDS-PAGE and immunoblotting. All EpsE mutant proteins in complex with EpsL(1-253)His_6 were able to float to the highest two fractions in the presence of CL, but were retained in the lowest two fractions in the absence of CL, similar to that of wild type EpsE-EpsL(1-253)His_6 (Fig. 2.6A and data not shown). When quantified over at least three individual experiments, each point mutant had a comparable percentage of protein floating up to the highest fractions compared to that of the wild type protein (Table 2.1). Although the CTD' arginines 320 and 324 cluster toward the lower end, they were not statistically different than wild type (Table 2.1). Since EpsL was previously determined to be the main protein involved with the interaction with CL (Camberg *et al*., 2007), it was not surprising that mutations in EpsE far away from the EpsE-EpsL interface are still able to bind phospholipids in the presence of EpsL and confirms that the reduction in ATPase activity is not due to loss of CL binding.

Figure 2.6: EpsE point mutants interact with CL and can form oligomers.
(A) Protein-CL interactions were compared using a liposome floatation assay. Reaction mixtures containing wild type, R225A, or R324A protein +/- CL were subjected to membrane floatation. After centrifugation, five 1mL fractions were collected and analyzed by SDS-PAGE and immunoblotting with anti-EpsE. Molecular mass markers are indicated on the left, fraction number listed across the top, and lane number listed across the bottom. Lanes 4 and 5 are the top fractions in the CL(+) samples and represent liposome-bound protein, while lanes 9 and 10 represent the highest fractions of the CL(-) samples. (B) Reactions containing either wild type EpsE-EpsL(1-253)His_6 or the point mutants EpsE-R225A, or EpsE-R324A protein were cross-linked in the presence of Mg-ATP and CL using increasing concentration of DSP (0, 10, 50, or 100 μM). Samples were analyzed by SDS-PAGE and immunoblotted with anti-EpsE. Molecular weight markers are denoted on the left and the position of monomeric EpsE is indicated with an arrow. (C) Reactions containing either wild type EpsE-EpsL(1-253)His_6 or the point mutant EpsE-R156D, or EpsE-D326R proteins were cross-linked in the presence of Mg-ATP and CL as described in (B). Molecular mass markers are denoted on the left and the position of monomeric EpsE is indicated with an arrow.


A
+ cardiolipin
- cardiolipin
Fraction 1 2 3 4 5 1 2 3 4 5
62
49
WT
R225A
R324A
1 2 3 4 5 6 7 8 9 10
Lane
B
WT
R225A
R324A
[DSP] (uM) 0 10 50 100 0 10 50 100 0 10 50 100
188
98
62
49
EpsE multimers
EpsE monomer
C
WT
R156D
D326R
[DSP] (uM) 0 10 50 100 0 10 50 100 0 10 50 100
188
98
62
49
EpsE multimers
EpsE monomer

EpsE mutants can still form higher molecular weight oligomers*:*

Finally, to determine if the mutant proteins are still capable of forming higher molecular weight complexes in the presence of EpsL(1-253)His_6, ATP and CL, we subjected the proteins to chemical cross linking analysis using disthiobis succinimidyl propionate (DSP) (Camberg *et al.*, 2007). The formation of higher molecular weight species was unchanged for the mutant proteins compared to wild type in the presence of CL when analyzed by SDS-PAGE and immunoblotting (Fig. 2.6B and data not shown). The percentage of monomers remaining uncross-linked was calculated for the highest DSP concentration as shown in Table 2.1. None of the EpsE point mutants have a statistically different amount of uncross-linked monomers remaining when compared to wild type protein indicating that the point mutations introduced into EpsE do not inhibit its ability to form higher molecular weight species. Taken together with all the above mentioned data, it is likely that the substitutions R210A, R225A, R320A, R324A, R336A, and R369A specifically interfere with EpsE's ability to hydrolyze ATP.

Protein	% Uncross-linked monomer	p-value	% Total protein associated with CL	p-value
WT	25 (8.4)		63 (18.3)	
R210A	31 (14)	0.71	77 (5.2)	0.30
R225A	35 (17.7)	0.58	80 (6.3)	0.24
R320A	25 (10)	0.99	38 (13.5)	0.21
R324A	23 (7.5)	0.87	43 (10.7)	0.25
R336A	26 (11.2)	0.44	66 (17.9)	0.89
R369A	41 (9.7)	0.25	62 (6.3)	0.91
R156D	47 (15.5)	0.21	53 (11.4)	0.54
D326R	17 (8.6)	0.60	nt	--

Table 2.1: Quantification of cross-linking and liposome floatation assays. Quantification of at least three independent assays was performed for each protein using ImageQuantTL software. For cross-linking, the amount of monomer remaining in the highest [DSP] lane was compared to that in the no cross-linker lane (+/- SEM). Protein floatation (+/- SEM) was determined by dividing the amount of protein in F4+F5 to the total protein from F1, F2, F4, F5. For both assays, mutant proteins were not statistically different from WT. CL=cardiolipin, nt=not tested.

Mutations in the EpsE subunit interface have a similar phenotype to the arginine mutations:

We have thus far demonstrated that substitutions of conserved arginine residues abolish ATPase activity and suggest that these mutations disrupt the formation of the complete active site, resulting in failure to hydrolyze the nucleotide and inactivation of the entire complex. Further, we have shown that the defect is in the hydrolysis step and not due to inefficient nucleotide binding and hypothesize that oligomerization is a key step in obtaining full activity. In order to strengthen the argument that oligomerization is necessary to form the

complete active site needed for efficient hydrolysis, we introduced point mutations within the subunit-subunit interface. We rationalized based on the EpsE hexamer model that a single mutation within this interface would be unlikely to fully disrupt oligomerization, but may result in altered ability to communicate nucleotide binding in one subunit which could further inhibit the movement of the conserved residues necessary to complete the active site. We chose three residues: R156 in the NTD as well as D326 and D328 in the CTD. These three residues form two salt bridges at the edge of the NTD-CTD' interface of the EpsE filament and also in the C2 hexamer model, and are likely important for properly aligning the subunits with respect to each other in the hexameric ring. The salt bridges are also likely preserved between the active, closed form and the neighboring open form as the distance between these residues does not vary between the different hexamers. R156A, R156D, D326R, and D328R mutations were introduced into EpsE, and the mutant proteins were first tested for complementation and dominance *in vivo* using the protease secretion assay described earlier. As Figures 2.3A and 2.3B demonstrate, all four of the interface mutants failed to complement the *epsE* deficient strain when assayed for protease secretion, yet they were stably overexpressed as verified by immunoblotting. The interface mutants also acted in a dominant negative manner when overexpressed in the wild type strain, indicating that the mutants likewise retain their ability to interact with the T2S complex (Figs. 2.3C and 2.3D). EpsE with the R156A, R156D, D326R, and D328R substitutions showed the same intermediate effect on secretion as the previously discussed R369A,

meaning they were dominant when compared to the plasmid-encoded wild type EpsE ($p<0.05$); however they did not inhibit secretion as efficiently as the K270A mutant ($p<0.05$). This may indicate a trend toward a less severe dominance phenotype for mutations in the interface than those directly participating in the active site.

Next we purified EpsE-R156D and EpsE-D326R in complex with EpsL(1-253)His_6 and tested them in the ATPase and cross linking experiments described above. As predicted, both interface mutants were still able to cross link into higher molecular weight complexes statistically comparable to that of wild type EpsE-EpsL(1-253)His_6 (Fig. 2.6C and Table 2.1). Neither mutant, however, displayed high levels of ATP hydrolysis in the presence of CL (Fig. 2.4). Interestingly, the only point mutant that retained any ATPase activity greater than BSA was EpsE-R156D, however, this was reduced by at least 15-fold when compared to wild type protein and is insufficient to support T2S. Further, the EpsE-R156D mutant was tested in both the liposome binding assay and thermal melt assay which showed that it interacts with CL comparable to that of wild type protein (data not shown and Table 2.1) as well as has a comparable melting temperature to wild type EpsE-EpsL(1-253)His_6 in the absence of nucleotide (Fig. 2.5). Likewise, it also showed an increase in melting temperature over its apo form in the presence of Mg-ATP, which indicates it is still able to bind nucleotide similar to that of the other proteins tested (Fig. 2.5). Overall, these data support the hypothesis that a complete active site requires residues from multiple domains and subunits and the movements of the subunits into their

proper position in response to nucleotide binding and hydrolysis requires oligomerization of the protein.

Discussion

EpsE and other secretion ATPases participate in diverse processes and as a consequence they share low overall sequence homology. What stands out, however, are a few areas of high conservation. Forest and colleagues have mapped the conservation of residues among 27 secretion ATPases and determined that the residues within the active site and at the NTD-CTD' subunit-subunit interface are the most conserved residues across the entire protein (Misic *et al*., 2010). We have shown in this study that mutations in either area of conservation are detrimental to EpsE's ability to hydrolyze ATP thus further supporting the functional importance of these areas in the Type II/IV secretion ATPases. Through *in vivo* methods, we have demonstrated that replacement of each one of six conserved arginines in EpsE inhibits the function of the T2S complex. Additionally, *in vitro* experiments have demonstrated that alanine replacements of these arginines abolish EpsE's ability to efficiently hydrolyze ATP but not to: (i) bind nucleotide (ii) form higher molecular weight oligomers; or (iii) interact with CL. We have further gone on to show that even with an intact active site, EpsE is nonfunctional if the subunit-subunit interface is altered. The finding that replacement of R156 and D326 in the subunit interface has an adverse effect on ATPase activity without preventing oligomerization suggests that altering the interactions in this conserved interface may lead to oligomers that are unable to sense and/or move appropriately when nucleotide binds.

Although the arginine residues from the NTD (R210, R225), CTD (R336, R369), and CTD' (R320, R324) were all found to be required for ATP hydrolysis,

not all of them are within hydrogen bond distance to the nucleotide in our model (Fig 2.2B); therefore, the contribution of each arginine to the enzymatic activity of EpsE may differ. Taking structural information into account, the potential roles of the individual arginines in EpsE's ATPase activity are discussed below.

We have demonstrated in this study that both R225 and R210 are necessary for function both *in vivo* and *in vitro*. These NTD arginines are both highly conserved and likely important for closure of the protein around the bound nucleotide. Modeling of the EpsE hexamer based on the PaPilT structure brings these arginines in the vicinity of the bound nucleotide. As seen in Figure 2.2D, the R225 subunit C from EpsE moves sufficiently close to the γ-phosphate of the bound nucleotide to directly participate in ATP coordination or hydrolysis. The EpsE-R225 equivalent in *Xanthomonas campestris* XpsE (XpsE-R286) has been pinpointed as being an important residue for sensing the bound nucleotide and coordinating the movement of the NTD relative to the CTD (Shiue *et al.*, 2007). The EpsE-R225 and R210 counterparts in both AfGspE and PaPilT have been shown to come close to the bound nucleotide suggesting a conserved function (Shiue *et al*., 2007, Yamagata & Tainer, 2007). Interestingly, the equivalent arginines in PaPilT structure show marked conformational differences in the three observed states (Misic *et al*., 2010). The guanidinium moiety of PaPilT-R97 (the EpsE-R225 equivalent) stays at a quite similar distance of 4.3-5.2 Å from the α- and β- phosphate groups in all three states. In contrast, the guanidinium group of PaPilT-R82 (EpsE-R210) moves dramatically between 3.2 to 9.0 Å distance from the γ-phosphate. It may be argued that R225 serves as a static sensor of

ATP whereas R210 plays a more active role by engaging the γ-phosphate directly and removing it from the active site after hydrolysis. In our model, R210 does not come close to the γ-phosphate to coordinate or assist directly in the hydrolysis of ATP (Fig. 2.2D). This may be attributed to the limitations of our rigid body model. Indeed a change of the arginine side chain conformation could bring the guanidinium group of EpsE-R210 in subunit B close to the γ-phosphate. Also, R210 is within hydrogen bond distance to the carboxylate group of E296 responsible for Mg^{2+} ion coordination. Therefore, an alternative or additional function of R210 might be to stabilize E296 thereby supporting its coordination of Mg^{2+} ion. Regardless of the immediate roles R210 and R225 play in ATP hydrolysis, the observation that these two residues are far from the nucleotide in the initial EpsE structure (Robien *et al*., 2003) but come near the nucleotide bound in our hexameric EpsE model (Fig. 2.2), along with our protease secretion and ATPase data (Figs. 2.3 and 2.4) suggest that EpsE needs to be hexameric in order to perform its function.

The arginines at position 336 and 369 in the CTD of EpsE are found in the Walker B and His box motifs, respectively, and our data demonstrate that both of these residues are required for function of EpsE. In all three subunit conformations of the EpsE hexamer model, EpsE-R336 forms salt bridges with D293 and with the catalytic E334, a residue essential for ATPase activity (Chiang *et al.*, 2008). Therefore, the coordination of EpsE-R336 by both D293 and E334 is likely crucial for correct positioning of the latter for catalysis. A mutation in either R336 or D293 would have negative consequences on function as seen

both in this study and in previous ones looking at the effects of mutating the EpsE-D293 equivalent in PaPilT (Chiang *et al*., 2008) and PulE (Possot & Pugsley, 1994). Interestingly, the equivalent salt bridges in the PaPilT structure are observed only in "active" subunit B. In subunits A and C, a simultaneous shift of a loop containing residues 203-206 along with a change in side chain conformation of PaPilT-E208 (EpsE-E334) brings the carboxylate in contact with His box residue H229 instead. This leads to movement of PaPilT-R206 (EpsE-R336) toward the interface of neighboring CTD'.

EpsE-R369 of the CTD does not appear to make significant contacts with bound nucleotide in our hexamer model, nor in the EpsE crystal structure (Robien *et al*., 2003). Instead, the guanidinium group of EpsE-R369 forms a hydrogen bond with the main chain carbonyl of EpsE-R336 and, therefore, could be important for stabilizing or sensing the conformation of the loop described above. The movement of that loop could be an important element of synchronized structural rearrangements during the ATP hydrolysis cycle. The intra-subunit EpsE-R336 interaction with R369 has not been observed in the PaPilT structure where the EpsE-R369 equivalent (PaPilT-R239) instead makes an inter-subunit contact with PaPilT-E219'. Intriguingly, this salt bridge between PaPilT-R239 and PaPilT-E219' is maintained in all three subunits of the PaPilT C2 hexamer despite the changes in side chain conformations and relative orientation of CTD and CTD'; however, PaPilT-E219 is not conserved in related ATPases (Fig. 2.1). Based on the EpsE hexamer model, residue EpsE-E376 from the neighboring CTD' is a plausible candidate for contacting EpsE-R369.

This fits with our *in vivo* secretion data showing that the EpsE-R369A dominant phenotype shares characteristics with the interface mutants, indicating that the primary role of this residue may be to properly align the subunits, rather than to directly participate in ATP hydrolysis. Further experiments are needed to establish the precise functional role of R369.

Two residues from the CTD', EpsE-R320 and R324, both seem to participate in completing the active site of the neighboring subunit and perhaps support the efficient removal of the γ-phosphate (Fig. 2.2B), although of the two, only R324 is strictly conserved. R324 of CTD' forms a salt bridge with D195 in the NTD of EpsE; however, the importance of this interaction for the stability of the NTD:CTD' interface has not been investigated. While we have demonstrated that R320 is necessary to support secretion through the T2S apparatus of *V. cholerae*, a mutation in the EpsE-R320 equivalent in PaPilT showed only a minor effect on twitching motility (Satyshur *et al*., 2007), further substantiating the idea that the importance of this residue may vary across species.

Whereas both EpsE-R320A and EpsE-R324A are defective in ATP hydrolysis, and thus secretion, only EpsE-R324A has a dominant negative effect on secretion when expressed in a wild type *V. cholerae* strain, suggesting that the defect in EpsE-R320A may be compensated for *in vivo* by the wild type EpsE, possibly through the formation of mixed oligomers. In the presence of wild type EpsE the mutant protein EpsE-R320A may adopt a conformation amenable to ATP hydrolysis. Alternatively, it is possible that not every subunit in the hexamer needs to be active in order to support ATP hydrolysis and secretion;

therefore an EpsE hexamer could function even when it consists of a mixture of wild type EpsE and mutant EpsE-R320A. Interestingly, work investigating the formation of active, mixed hexamers of the AAA+ ATPase ClpX, generated by covalently linking wild type and inactive subunits, concluded that six wild type subunits are not necessary for ClpX's ATPase activity (Martin *et al.*, 2005). The observation that the wild type EpsE alleviates the defective EpsE-R320A mutant *in vivo* further confirms the idea that EpsE is active as an oligomer.

Mutations in the subunit interface were also detrimental to EpsE's function; however, based on our data it is possible that somewhat of a hierarchy exists in the ability of EpsE to compensate for mutations either within the active site or subunit interface, with mutations of residues directly participating in ATP hydrolysis being more severe. This idea is supported by the results from the protease secretion assays in which the trans-dominance phenotype was not nearly as severe for the interface mutants as seen for mutants within the active site indicating that these mutants may be more easily compensated for by the wild type protein. Further, the only mutant protein to retain any appreciable ATPase activity *in vitro* was the interface mutant EpsE-R156D, and although it was not sufficiently active to support secretion on its own *in vivo*, this mutant was able to hydrolyze ATP at a significant level above the BSA control ($p<0.05$).

Based on our findings on the requirement of the conserved arginines for ATP hydrolysis and secretion, and on observations published earlier on conformational differences between nucleotide bound and nucleotide free Hp0525, PilT and AfGspE (Misic *et al*., 2010, Satyshur *et al*., 2007, Savvides *et*

al., 2003, Yamagata & Tainer, 2007), we speculate that the binding of nucleotide to one EpsE subunit likely causes drastic conformational changes which can be relayed to not only the adjoining subunits but also to the larger T2S complex through the interaction of EpsE with the inner membrane protein EpsL. The interactions of EpsL with the major pseudopilin EpsG (Gray *et al*., 2011) may in turn be essential in the formation of the pseudopilus and/or drive the transport of proteins through the EpsD pore in the outer membrane. In addition, understanding how EpsE powers the transport through the T2S system may provide insight into the function of Type II/IV secretion ATPases as a whole due to the high level of structural homology among members of this protein family.

Chapter 3

The zinc-coordinating C_M domain of the ATPase EpsE is required for a functional Type II secretion system in *Vibrio cholerae*

Abstract

Type II secretion (T2S) is an energy-dependent process by which hydrolytic enzymes and toxins are transported across the outer membrane of Gram negative pathogens such as *Vibrio cholerae*. The ATPase, EpsE, is one of twelve T2S proteins and is required to power the export of these proteins through the T2S apparatus. Previous studies have demonstrated EpsE to be active as a dynamic oligomer in which the binding, hydrolysis, and release of nucleotide results in conformational changes within and between EpsE domains and subunits. These structural changes may mechanically transduce the chemical energy of ATP hydrolysis through the T2S apparatus, ultimately resulting in extracellular secretion. One distinct feature found in the GspE/PilB family to which EpsE belongs, is that of a small domain, termed the C_M domain, which coordinates one zinc ion per monomer by means of a tetracysteine motif. Although its role in secretion is not known, it was previously shown in *Klebsiella*

oxytoca that replacement of the cysteines in PulE resulted in loss of pullulanase secretion. I show here that this domain participates in the nucleotide binding pocket through residue R441, as well as contributes to protein-protein interactions between EpsE and another T2S component, EpsL. I further demonstrate that, although the zinc coordination forms a hairpin structure, removal of the zinc does not completely abolish protein-protein interactions between EpsE and EpsL; however, it does inhibit the function of EpsE to support secretion through the T2S apparatus. Taken together, I propose a model in which the ATPase activity of EpsE is regulated by protein-protein interactions within the C_M domain. As this domain is highly similar to that of other GspE/PilB family members, these results may contribute to the understanding of how those ATPases work within other molecular trafficking machineries.

Introduction

Gram negative bacteria have evolved multiple mechanisms to transport toxins and hydrolytic enzymes across their cell envelopes. One of the better studied secretion processes is that of the Type II secretion system (T2SS). This system, made up of 12-15 proteins, is well conserved across many Gram negative pathogens, and as it supports the secretion of toxins, hydrolytic enzymes and other virulence factors, is considered a major virulence mechanism (Cianciotto, 2005, Johnson *et al*., 2006). Much work on the T2SS has been conducted in *Vibrio cholerae*, a Gram negative bacillus known for causing the diarrheal disease, cholera, by means of toxin secretion. Under appropriate conditions, such as those found in the human host, the bacterium produces the toxin which is first transported to the periplasmic space by means of the Sec or Tat pathway. Once in the periplasm, the toxin folds and engages the Type II secretion (T2S) apparatus, which exports the protein to the extracellular milieu using an energy dependant mechanism (Sandkvist, 2001a).

The individual components of the T2SS are well studied, however, their roles in the mechanism of protein secretion is not well understood and the subject of ongoing research. The T2SS forms a large complex spanning the entire cell envelope. Proteins are secreted through a pore in the outer membrane formed by the secretin EpsD using the energy provided the ATPase, EpsE. EpsE associates with the T2S system through interactions with EpsL, a protein found in an inner membrane complex also comprised of proteins EpsC, EpsF, and EpsM (Abendroth *et al.*, 2005, Lybarger *et al*., 2009, Py *et al*., 2001,

Robert *et al*., 2005, Sandkvist *et al*., 1995). The transport of the protein through the secretin is thought to occur through a piston-like mechanism due to the polymerization of a pseudopilus made up of proteins EpsG-K (Filloux, 2004, Hobbs & Mattick, 1993, Sandkvist, 2001a). EpsG has been shown to directly interact with EpsL providing a link between the ATPase and the pseudopilus (Gray *et al*., 2011). Although a more precise model of the secretion process remains elusive, we do know that T2S apparatus shares an evolutionary history with Type IV pili (T4P) (Filloux, 2004, Nunn, 1999) and thus the role of many components, especially that of the ATPase may be conserved across not only the T2S machinery, but the T4P biogenesis machinery as well.

The ATPase, EpsE, is a member of the large family of Type II/IV secretion ATPases, which includes proteins required for diverse functions such as protein secretion, pilus biogenesis, DNA uptake and archaeal flagellar assembly. The members of this family share many conserved features, although their overall sequence homology is only 35%. Secretion ATPases are made up of at least two distinct domains, the N terminal domain (NTD) and the C terminal domain (CTD) which are connected through as short linker (Fig 3.1). Within the CTD are the ATPase specific motifs, including the Walker A and B boxes along with the Asp and His boxes, which come together to form the nucleotide binding site (Patrick *et al.*, 2011, Possot & Pugsley, 1994, Robien *et al*., 2003, Thomsen & Berger, 2008). EpsE may be further grouped into the GspE/PilB subfamily of ATPase, which is primarily comprised of ATPases associated with T2S and T4P (Planet *et al*., 2001). Interestingly, proteins in the GspE/PilB subfamily include

another unique domain, termed the C_M domain, which contains approximately 49 amino acids. Within this domain is a conserved tetracysteine motif which consists of two CxxC motifs, with approximately 29 residues that separate them (Robien *et al.*, 2003). The intervening 29 residues between the second and third cysteines show very low sequence homology to other related proteins; however, structural data shows this domain forms a hairpin elbow-like structure which coordinates a zinc molecule at the bend and is likely conserved across homologous proteins (Fig 3.1) (Camberg & Sandkvist, 2005, Robien *et al*., 2003).

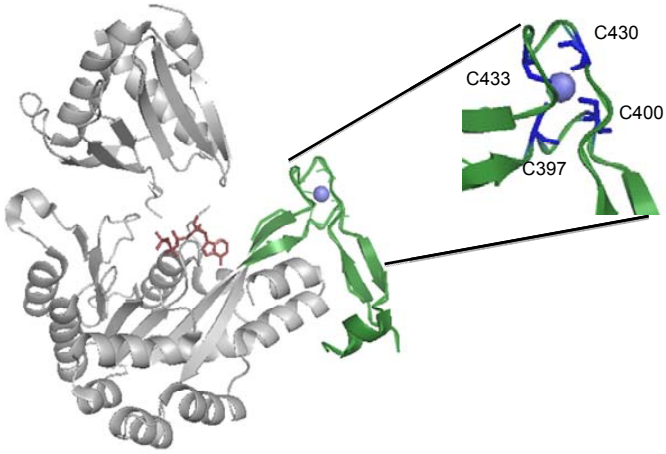


Figure 3.1: The C_M domain coordinates a zinc ion.
An EpsE monomer with bound AMPPNP is shown in gray ribbon structure with the C_M domain colored in green and the side chains of coordinating cysteines are shown as sticks. This is expanded for clarity with the zinc ion depicted as a blue sphere and the tetracysteine motif residues colored blue and labeled. Adapted from (Patrick *et al*., 2011).

These types of zinc-containing domains are generally categorized as zinc fingers. Zinc fingers often participate in DNA binding but they also are an important structural feature of many bacterial proteins (Krishna *et al.*, 2003). For example, they hold proteins in a specific conformational state (which can be either active or inactive), as is seen in Hsp33 (Graumann *et al*., 2001), or stabilize a structure within the protein and allow for protein-protein interactions such as seen with SecA (Fekkes *et al*., 1999). Under reducing conditions, Hsp33 is inactive with its tetracysteine motif coordinating one zinc ion; however when in an oxidized state, the cysteine residues release the zinc causing partial activation. Loss of zinc coordination also leads to disulfide bond formation causing protein dimerization and full activation of this protein (Graf *et al.*, 2004, Graumann *et al*., 2001). The role of zinc to SecA is slightly different. In this case, the zinc coordinated by the cysteine residues provides structural stability within the domain required for another protein, SecB to bind. Loss of SecB binding due to removal of this zinc ion results in loss of SecA's stimulated ATPase activity, and thus its function (Fekkes *et al*., 1999). Further analysis has demonstrated that mutations in basic residues within the zinc- binding domain of SecA likewise compromised SecA-B interaction while keeping the zinc motif intact; indicating that these residues are important for SecA-B complex formation and the zinc is primarily responsible for correctly positioning these residues for interaction (Zhou & Xu, 2003).

In the case of EpsE and its closest homologs, the zinc coordinated by the tetracysteine motifs is likely playing a structural role (Camberg & Sandkvist,

2005, Robien *et al*., 2003); however, the function of the domain in which the zinc is found, remains primarily uncharacterized. I hypothesize that the necessity of the C_M domain to T2S is two-fold. First, I believe that the residue located at the top of the C_M domain near the nucleotide binding site, R441, may participate directly in nucleotide binding and is thus important for overall function (Fig 3.2). Second, I propose that the zinc elbow structure is required for interactions with other components of the T2S, such as EpsL, and these contacts are necessary for a functional T2S apparatus. In this study, I have demonstrated that replacing R441 with either alanine or aspartate is detrimental to the function of the ATPase, but does not grossly affect the interaction of EpsE to EpsL or the ability of EpsE to form multimers. Further, my data suggest that the elbow structure is required for T2S as both the removal of the entire loop as well as mutations which likely cause a loss in zinc coordination inhibit secretion possibly due to a reduced interaction between EpsE and EpsL.

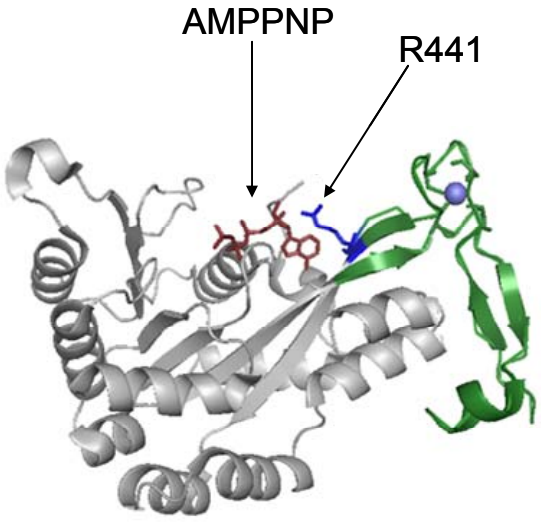


Figure 3.2: The C_M domain terminates at the nucleotide binding site.
A ribbon cartoon depiction of the C-terminal domain of EpsE with bound AMPPNP (colored red). The C_M domain and residue R441 are dark green and blue, respectively. The R441 residue and the nucleotide are indicated with arrows to highlight the position of this residue within the nucleotide binding site. Adapted from (Patrick *et a*l., 2011).

Experimental Procedures

Cloning and Expression:

EpsE point mutations R441A and R441D were introduced into the previously constructed plasmid, pET21dEpsE(2-503)-EpsL(1-253)His_6 (Camberg *et al*., 2007) using the QuickChange Site-Directed Mutagenesis kit (Stratagene) as directed. The fragments of *epsE* containing the various mutations were cloned into pMMB384 (Sandkvist *et al*., 1993) by exchange of an MfeI/BamHI fragment to create the individual pMMBEpsE point mutant plasmids. The pMMB plasmids were then introduced to the *epsE*$^-$ and wild type *V. cholerae* strains through conjugation. To construct the elbow deletions primers were designed to create a series of mutants designated Δloop and ΔloopPro. The Δloop construct has a loop comprised of residues 396-437 deleted, while ΔloopPro replaces the same loop with a Proline.

Purification of protein:

The above constructed EpsE(2-503)-EpsL(1-253)His_6 point mutants were expressed in *Escherichia coli* BL21(DE3) under IPTG inducing conditions and purified using metal affinity chromatography and gel filtration as described in (Abendroth *et al*., 2004a, Abendroth *et al*., 2005).

Phospholipids:

E. coli-derived cardiolipin was prepared in 20mM Tris pH 8.0, 150 mM NaCl, 1 mM DTT as described in (Camberg *et al*., 2007).

Protease Secretion Assay:

pMMB plasmids containing the point mutations of *epsE* were expressed in wild type *V. cholerae*, or the *V. cholerae epsE*⁻ strain in the presence of 10 or 100 μM IPTG and tested for secretion of a protease using a modified fluorescence-based assay as described in (Sandkvist *et al.*, 1997). Supernatants from LB-grown overnight cultures were separated from cells by two centrifugation steps and assayed for proteolytic cleavage of N-tert-butoxy-carbonyl-Gln-Ala-Arg-7-amido-4-methyl-coumarin (MCA) (Sigma) as described in (Johnson *et al*., 2007, Patrick *et al*., 2011, Sandkvist *et al*., 1997) using excitation and emission wavelengths of 385 nm and 440 nm, respectively. The protease activity was determined by measuring the change in fluorescence over time and comparing to an MCA standard. The amount of MCA generated per minute was normalized by OD_{600} with error reported as standard error (SEM) of at least three separate samples assayed in triplicate.

ATPase Activity Assays:

ATPase activities of purified proteins were measured using a kinetic assay containing 0.5 μM protein, 5 mM ATP, 5 mM $MgCl_2$, and 125 μM cardiolipin in buffer A (100 mM HEPES pH 8.5, 65 mM NaCl, 5% glycerol). Reactions were incubated at 37ºC for 0, 15, 30, 45, and 60 minutes for wild type and for 0, 120, 180, 240, and 300 minutes for R441A, R441D and BSA and assayed for the release of inorganic phosphate using the BioMol Green reagent as directed. Amount of phosphate was determined by comparing sample absorbance to those

of a phosphate standard curve with error reported as standard error (SEM) of at least three separate samples assayed in duplicate. For p-hydroxymercuriphenylsulfonic acid (PMPS)-treated assays, aliquots of purified wild type protein were treated with 10-fold molar excess of PMPS, incubated for 5 minutes and assayed as described above.

Liposome Floatation Assay*:*

Reaction mixtures containing 1 μM protein, 4 mM AMPPNP, 4 mM $MgCl_2$, and 250 μM cardiolipin in buffer A were assayed as described in (Patrick *et al.*, 2011). Five 1 mL fractions were collected from the top of each tube and analyzed by SDS-PAGE and immunoblotted against EpsE. The top two fractions (# 4-5) represent the cardiolipin-bound protein.

Cross linking*:*

Samples containing 125 μM cardiolipin, 5 mM ATP, 5 mM $MgCl_2$, and 1 μM purified proteins were incubated in buffer A and subjected to the chemical cross linker disthiobis succinimidyl propionate (DSP) at concentrations of 0, 10, 50 and 100 μM as described in (Camberg *et al*., 2007). For the +PMPS samples, aliquots of purified wild type protein were treated with 10-fold molar excess of PMPS, incubated for 5 minutes and assayed as described above.

Fractionations:

EpsE(2-503)-EpsL(1-253)His_6 point mutants were expressed in *E. coli* BL21(DE3) under IPTG inducing conditions. Bacterial cells were separated by centrifugation and subjected to lysis by French Press. Soluble and insoluble fractions were collected by centrifugation and the soluble fraction subsequently added to Talon metal affinity resin. After incubation, the flow through was collected and the bound protein eluted with imidazole. The fractions were analyzed by SDS-PAGE and immunoblotted against EpsE and EpsL. The ratio of EpsE:EpsL was calculated for three separate assays as described below.

Quantification of immunoblots:

Densitometry analysis was performed on a Typhoon TRIO (Amersham Biosciences) using ImageQuant TL software. For the liposome binding assay, the % EpsE floated represents the amount of EpsE in F4+F5 divided by the total amount of EpsE in F1, F2, F4 and F5. For quantification of EpsE cross-linking, the amount of monomer remaining in the 100 μM DSP lane was divided by the amount of monomer in the 0 μM DSP lane to obtain the % uncross-linked monomer. For the EpsE-EpsL fractionation experiment the ratio of EpsE:EpsL or soluble EpsE:insoluble EpsE was calculated for the individual fractions. The averages of three experiments +/- SEM were calculated.

Statistical analysis*:*

A Student's t-test was used for all statistical analyses. Values were considered significant at $p<0.05$.

Results

Mutations in R441 inhibit T2S *in vivo:*

The R441 residue is found at the top of the C_M domain and makes contact with AMPPNP in the nucleotide bound form of EpsE (Fig 3.2). In order to test the importance of this residue to T2S, I created two different mutants by replacing the arginine with either alanine or aspartate, respectively. To determine whether these point mutants could complement an *epsE*$^-$ *V. cholerae*, I introduced the plasmid-encoded EpsE mutant protein into our *epsE*$^-$ strain. The plasmid-encoded wild type EpsE restored secretion to native levels; however, neither R441A nor R441D could complement the defect (Figure 3.3A). These samples were then immunoblotted against EpsE to demonstrate that both proteins are stably expressed (Figure 3.3B). To further verify that these proteins are not completely misfolded, I also introduced the plasmids containing the mutant proteins into the wild type strain of *V. cholerae* and tested for a dominant negative effect *in vivo*. As Figures 3.3A and 3.3B show, both R441A and R441D reduced the secretion of the wild type strain when stably overexpressed *in vivo*, as compared to an over-expressed wild type EpsE that does not affect secretion.

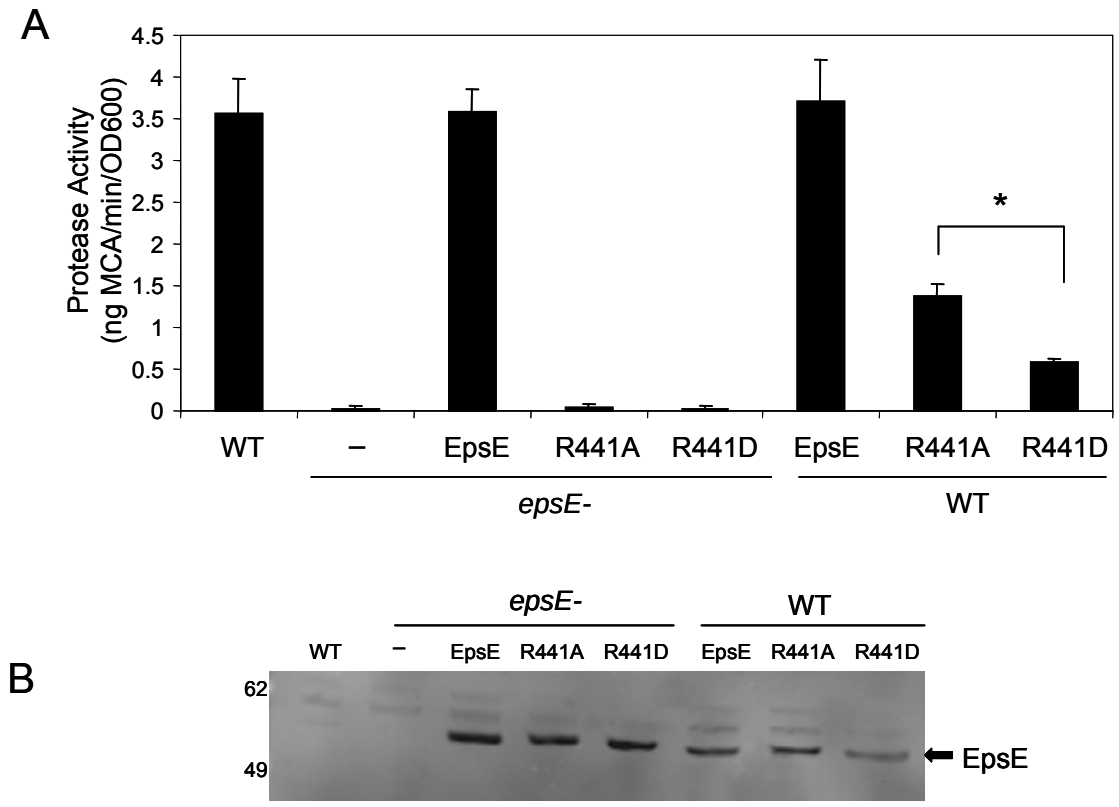


Figure 3.3: EpsE point mutations inhibit protease secretion *in vivo*.
epsE⁻ or wild type strains of *V. cholerae* containing pMMB67 vector, pMMB EpsE or the plasmid-encoded EpsE R441A or R441D were grown overnight in LB + 10μM IPTG. Culture supernatants and cells were separated by centrifugation. (A) Culture supernatants were analyzed for the presence of extracellular protease activity using a fluorescence-based assay with peptide-conjugated MCA as a substrate. The amount of MCA liberated was compared to a standard curve with each sample assayed at least three independent times. The results are presented as ng MCA hydrolyzed per minute, normalized to the OD_{600} of the culture +/- standard error. * indicates a statistically significant difference ($p<0.05$). (B) Cell extracts were subjected to SDS-PAGE and immunoblotted with anti-EpsE. Molecular mass markers are indicated on the left and the position of EpsE is indicated by the arrow.

Both EpsE R441 point mutations abolish ATPase activity *in vitro:*

Previously we have demonstrated that purified EpsE, when in the presence of EpsL(1-253)His_6 and acidic phospholipids, such as cardiolipin (CL), assembles into larger oligomers and displays ATPase activity (Camberg *et al*., 2007, Patrick *et al*., 2011). In order to test whether R441A and R441D have an affect on the hydrolysis of ATP in the presence of EpsL(1-253)His_6 and CL, I purified both mutant proteins as described in (Abendroth *et al*., 2004a, Abendroth *et al*., 2005). Briefly, the point mutations were individually introduced into the previously created EpsE(2-503)-EpsL(1-253)His_6 construct and purified by metal affinity chromatography and gel filtration. Both mutants were co-purified with EpsL(1-253)His_6 in large quantities indicating that the interaction of EpsE and EpsL was not disrupted as a result of these mutations. As Figure 3.4 demonstrates, neither point mutant retained any ATPase activity in the presence of CL when compared to wild type EpsE. The rates of ATP hydrolysis for the mutant proteins were not statistically different from that of the BSA control, indicating that both are completely inactive.

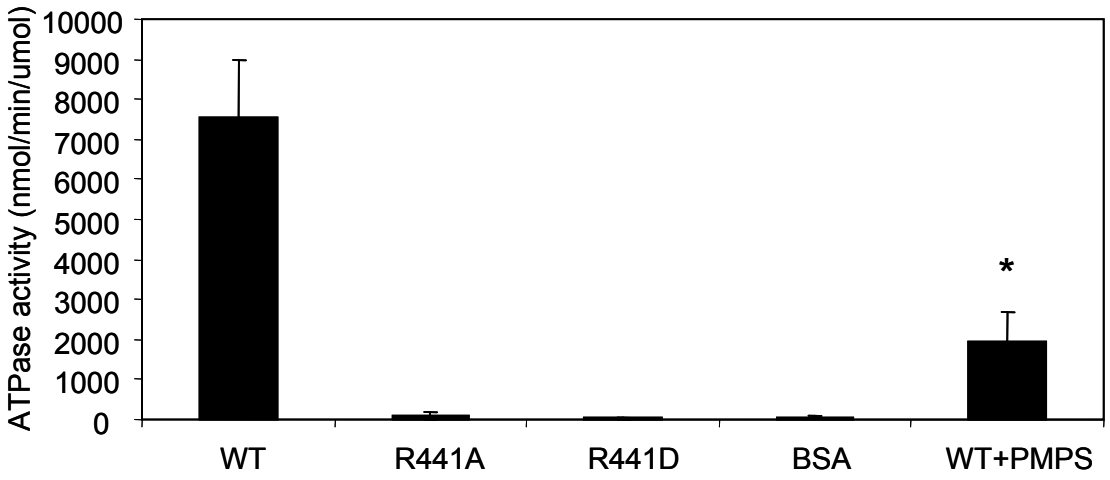


Figure 3.4: Point mutations as well as PMPS treatment of EpsE inhibit ATPase activity *in vitro*.
CL-stimulated ATPase activity was measured kinetically for wild type and PMPS-treated EpsE-EpsL(1-253)His_6 over a period of 60 minutes and the individual EpsE R441A and R441D mutants or BSA over a period of 300 minutes in reactions containing Mg-ATP and cardiolipin. Each protein was assayed at least three separate times and the mean +/- standard error are presented. * indicates a statistically significant difference ($p<0.05$) when compared to BSA.

The R441A and R441D mutants retain liposome binding and multimer forming abilities*:*

To test for the ability of R441A and R441D to form higher molecular weight multimers, I subjected the proteins to chemical cross linking analysis using disthiobis succinimidyl propionate (DSP) (Camberg *et al*., 2007). As seen in Figure 3.5 the formation of higher molecular weight species was unchanged for R441A or R441D compared to wild type in the presence of CL when analyzed by SDS-PAGE and immunoblotted against EpsE.

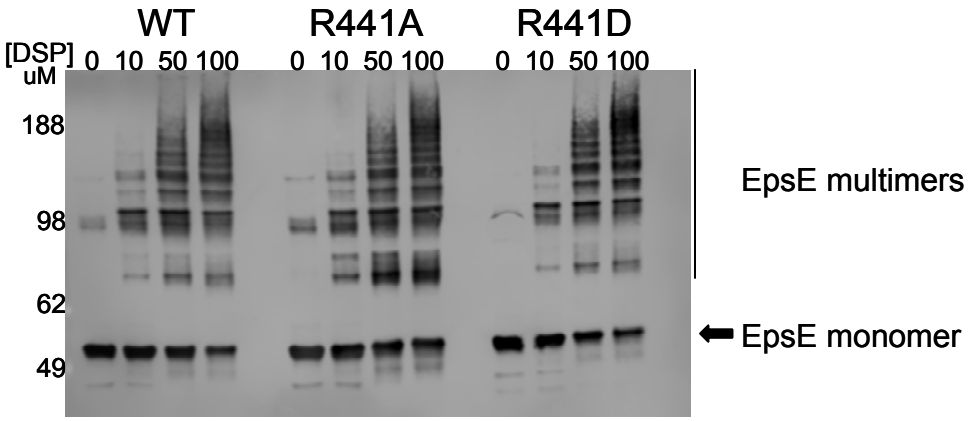


Figure 3.5: EpsE R441 mutants can form oligomers.
Reactions containing either wild type EpsE-EpsL(1-253)His_6 or the point mutants EpsE-R441A, or EpsE-R441D protein were cross-linked in the presence of Mg-ATP and cardiolipin using increasing concentration of DSP (0, 10, 50, or 100 μM). Samples were analyzed by SDS-PAGE and immunoblotted with anti-EpsE. Molecular mass markers are denoted on the left and the position of monomeric EpsE is indicated with an arrow.

To confirm that amount of cross-linked product is similar to that of the wild type protein, the percentage of monomers remaining uncross-linked was calculated for the highest DSP concentration. Neither of the EpsE point mutants had a statistically different amount of uncross-linked monomers remaining when compared to wild type protein indicating that the point mutations introduced into EpsE do not inhibit its ability to form higher molecular weight species (Table 3.1).

Protein	% Uncross-linked monomer	% Total protein associated with CL
WT	56.8 (11.9)	63.1 (10.6)
R441A	66.5 (12.1)	70.8 (3.8)
R441D	29 (0.7)	92.7 (10.6)
+PMPS	38 (9.3)	nt

Table 3.1: Quantification of cross-linking and liposome floatation assays. Quantification of at least three independent assays was performed for each protein using ImageQuantTL software. For cross-linking, the amount of monomer remaining in the highest [DSP] lane was compared to that in the no cross-linker lane (+/- SEM). Protein floatation (+/- SEM) was determined by dividing the amount of protein in F4+F5 to the total protein from F1, F2, F4, F5. For both assays, mutant proteins were not statistically different from WT. CL=cardiolipin, nt=not tested.

Using the floatation assay described in (Patrick *et al*., 2011), I sought to directly ascertain whether the mutant proteins retained their abilities to interact with CL, as a differential interaction with the phospholipids could explain the loss of ATPase activity seen above. As shown in Figure 3.6, both R441A and R441D in complex with EpsL(1-253)His_6 were able to float to the highest two fractions in the presence of CL, but were retained in the lowest two fractions in the absence of CL, similar to that of wild type EpsE-EpsL(1-253)His_6. Although it appears as though the R441D completely floats in the presence of CL, when the amount of protein interacting with CL for both R441A and R441D was quantified over three independent assays, I found that the amounts in the highest fractions are not statistically different than that of the wild type protein (Table 3.1). As the R441 residue is found at the top of the C_M domain and in contact with the nucleotide, I

did not find it surprising that interactions with cardiolipin were unaffected, given that previous work has suggested EpsL to be responsible for the majority of the protein-CL interactions (Camberg *et al*., 2007); however, the visual difference in the amount of protein that floats in the R441D sample may suggest additional interactions either between the C_M domain and CL or the C_M domain and EpsL, which will be discussed further below.

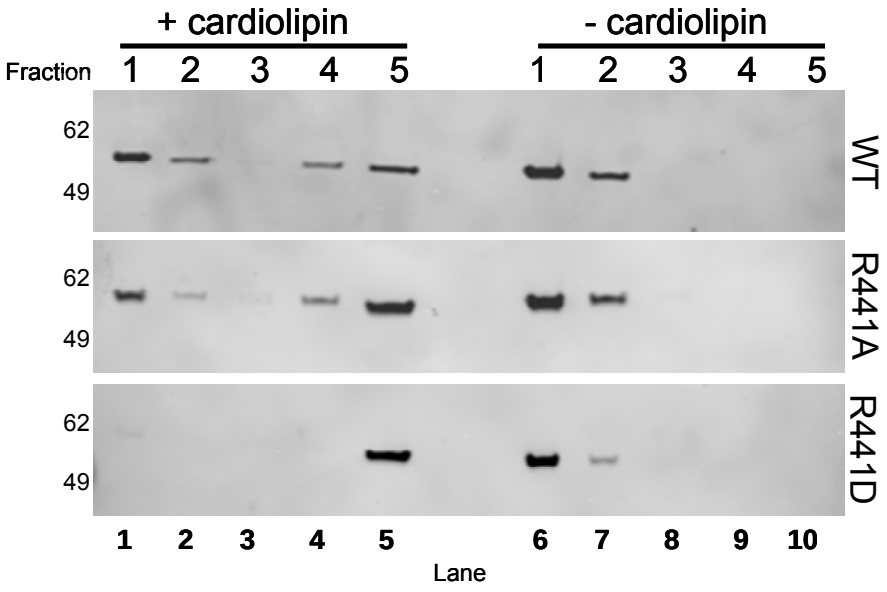


Figure 3.6: EpsE R441 mutants interact with CL.
Protein-CL interactions were compared using a liposome floatation assay. Reaction mixtures containing wild type, R441A, or R441D protein +/- CL were subjected to membrane floatation. After centrifugation, five 1mL fractions were collected and analyzed by SDS-PAGE and immunoblotting with anti-EpsE. Molecular weight markers are indicated on the left, fraction number listed across the top, and lane number listed across the bottom. Lanes 4 and 5 are the top fractions in the CL(+) samples and represent liposome-bound protein, while lanes 9 and 10 represent the highest fractions of the CL(-) samples. Molecular mass markers are denoted on the left and the position of monomeric EpsE is indicated with an arrow.

Removal of the C_M domain disrupts T2S:

As the elbow structure is only found in the GspE/PilB family and not the closely related PilT family, I wanted to investigate whether complete removal of this domain, thus making it more PilT-like, would be tolerated. To test the importance of the entire elbow region, the zinc-containing loop from residues 395-437 was deleted through PCR mutagenesis to create the Δloop construct. As my previous experiments found R441 to be critical for ATP hydrolysis, it was not removed from this construct. In order to preserve the bend that likely occurs through this loop a construct where a proline residue replaces the loop was also made and called ΔloopPro. To test these proteins *in vivo*, the plasmid-encoded EpsE mutant proteins were first analyzed for their ability to complement the secretion defect observed in an *epsE*$^-$ mutant strain of *V. cholerae*. Wild type EpsE was able to restore secretion; however, neither the Δloop nor the ΔloopPro were able to complement the *epsE*$^-$ mutant to a statistically significant level over the vector only control as shown in Figure 3.7A. Subsequent immunoblotting analysis indicated that the mutant proteins were expressed at higher levels than natively expressed EpsE, but somewhat lower levels than the plasmid-encoded wild type protein suggesting that the mutations may slightly affect the stability of EpsE but more likely the greater defect is from the lack of protein function (Figure 3.7B, last two lanes).

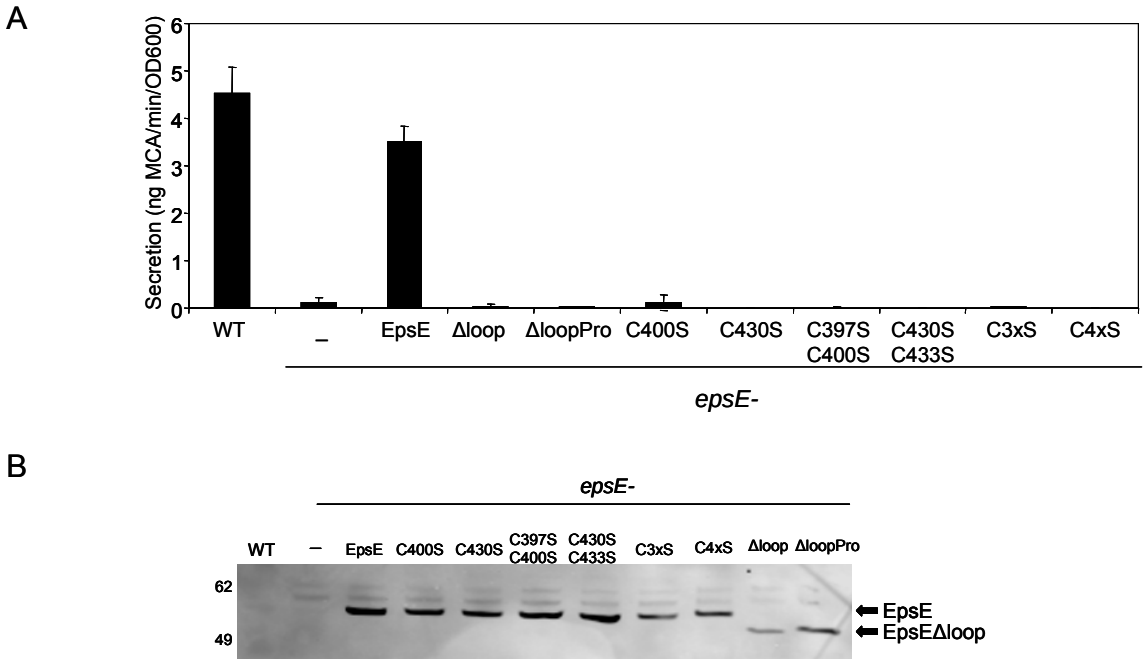


Figure 3.7: Mutations in the C_M domain inhibit protease secretion *in vivo*. An *epsE*$^-$ strain of *V. cholerae* containing pMMB67 vector, pMMB EpsE or the plasmid-encoded EpsE with the following mutations: C400S, C430S, C397SC400S, C430SC433S, C400SC430SC433S (C3xS), C397SC400SC430SC433S (C4xS), Δloop or ΔloopPro were grown overnight in LB + 10μM IPTG. Culture supernatants and cells were separated by centrifugation. (A) Culture supernatants were analyzed for the presence of extracellular protease activity using a fluorescence-based assay with peptide-conjugated MCA as a substrate. The amount of MCA liberated was compared to a standard curve with each sample assayed at least three independent times. The results are presented as ng MCA hydrolyzed per minute, normalized to the OD_{600} of the culture +/- standard error. (B) Cell extracts were subjected to SDS-PAGE and immunoblotted with anti-EpsE. Molecular mass markers are indicated on the left and the position of EpsE is indicated by the arrow.

As these mutations removed 42 amino acids, I wanted to further verify that these constructs did not grossly affect the folding of the mutant EpsE proteins. I therefore analyzed them for their ability to interact with the rest of the T2S

apparatus and inhibit secretion in wild type *V. cholerae*. Their effect on secretion was compared to the plasmid-encoded wild type EpsE. As shown in Figures 3.8A and 3.8B, both the Δloop and the ΔloopPro acted in a dominant negative manner when over-expressed in wild type *V. cholerae*. The dominant phenotype could be due to competition between wild type and mutant EpsE for binding to other T2S components such as EpsL, and/or due to the formation of inactive mixed oligomers containing wild type and mutant EpsE.

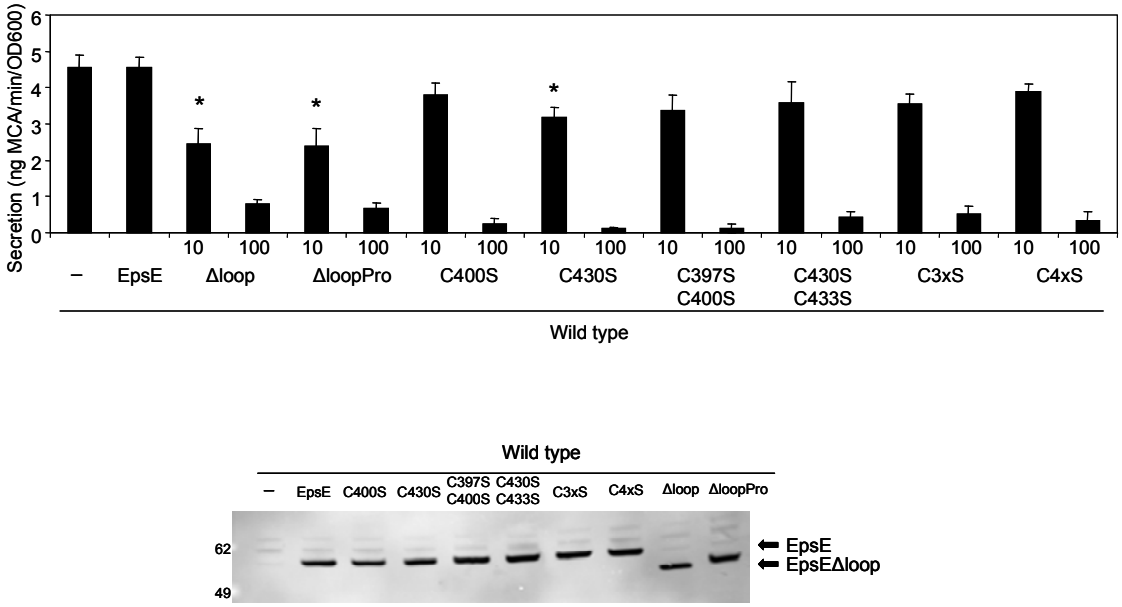


Figure 3.8: Mutations in the C_M domain exhibit a dominant negative effect on secretion *in vivo*.

A wild type strain of *V. cholerae* containing pMMB67 vector, pMMB EpsE or the plasmid-encoded EpsE with the following mutations: C400S, C430S, C397SC400S, C430SC433S, C400SC430SC433S (C3xS), C397SC400SC430SC433S (C4xS), Δloop or ΔloopPro were grown overnight in LB + 10 or 100μM IPTG. Culture supernatants and cells were separated by centrifugation. (A) Culture supernatants were analyzed for the presence of extracellular protease activity using a fluorescence-based assay with peptide-conjugated MCA as a substrate. The amount of MCA liberated was compared to a standard curve with each sample assayed at least three independent times. The results are presented as ng MCA hydrolyzed per minute, normalized to the OD_{600} of the culture +/- standard error. * represents a statistically significant difference ($p<0.05$) compared to the plasmid-encoded wild type EpsE at 10uM IPTG induction. All mutants show a statistically significant difference ($p<0.001$) at 100uM IPTG from that of the plasmid-encoded wild type EpsE. (B) Cell extracts collected from samples induced with 100uM IPTG were subjected to SDS-PAGE and immunoblotted with anti-EpsE. Molecular mass markers are indicated on the left and the position of EpsE is indicated by the arrow.

The C_M domain may interact with EpsL(1-253)His_6:

In order to address the molecular effects of these mutations *in vitro*, I attempted to purify the proteins using the methods described before. Unfortunately I was unable to retrieve high amounts of EpsEΔloopPro, due to what appeared to be a reduced interaction between EpsEΔloopPro and EpsL(1-253)His_6. Previous work has already shown the N1 domain of EpsE to primarily interact with EpsL (Abendroth *et al.*, 2005, Sandkvist *et al*., 1995), so I wanted to further probe this idea that the loop region may play an additional role in protein-protein interactions. I grew BL21(DE3)(EpsE(2-503)ΔloopPro-EpsL(1-253)His_6) under IPTG inducing conditions and fractionated the lysed bacterial cells into their soluble and insoluble fractions. The soluble fraction was passed over Talon resin and the bound protein eluted with imidazole. Although these samples were initially analyzed by SDS-PAGE and silver stain, the non-EpsE/EpsL bands made quantification difficult. As the wild type ratios of EpsE and EpsL of both the stained and immunoblotted gels were comparable, I calculated the relative ratio of wild type and mutant EpsE to EpsL for the soluble, insoluble, flow through and elution immunoblotted fractions. As shown in Figure 3.9 and quantified in Table 3.2, I do see a statistically significant reduction in the ratio of purified E:L in the ΔloopPro, compared to wild type protein suggesting the loop region may be important for EpsE's interaction with EpsL (compare lanes 4 and 8). Also apparent, although not statistically significant, was the accumulation of EpsEΔloopPro in the flow though fraction, suggesting that this protein is soluble but not able to bind efficiently to EpsL under these conditions.

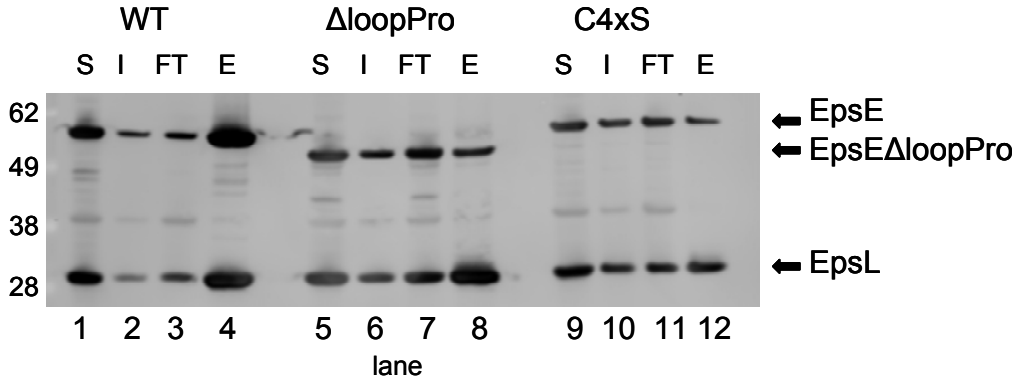


Figure 3.9: EpsE-EpsL interaction may be affected by C_M domain mutations.
E. coli expressing WT, ΔloopPro or C4xS mutants were grown to mid-log phase under IPTG-inducing conditions. Cells were lysed, fractionated into soluble (S) and insouble (I) material. The soluble fraction was applied to Talon resin and the flow though (FT) or elution (E) fractions were collected. Each sample was analyzed by SDS-PAGE and immunoblotted with anti-EpsE and anti-EpsL. Molecular mass markers are denoted on the left and the position of each protein is indicated with an arrow.

Further, I compared the ratio of the soluble EpsE versus the insoluble EpsE. As shown in Figure 3.9 and Table 3.2, the wild type EpsE has a higher portion of protein in the soluble fraction and the ΔloopPro has a much lower, although not statistically significant, ratio. This suggests that the removal of the loop may cause local changes in conformation rendering this protein less soluble than wild type.

	WT		ΔloopPro		C4xS	
Ratio EpsE:EpsL	Soluble	0.82 (0.1)	Soluble	0.74 (0.12)	Soluble	0.67 (0.03)
	Insoluble	1.64 (0.3)	Insoluble	1.01 (0.2)	Insoluble	0.76 (0.15)
	Flow through	0.96 (0.3)	Flow through	1.27 (0.45)	Flow through	0.93 (0.01)
	Elution	0.97 (0.1)	Elution	0.28 (0.11) *	Elution	0.49 (0.16)
Ratio soluble EpsE: insolube EpsE	2.34 (0.87)		1.32 (0.25)		2.46 (0.94)	

Table 3.2: Quantification of fractionation assays. Quantification was performed for proteins in each fraction using ImageQuantTL software. The ratio of EpsE:EpsL (+/- SEM) or the ratio of soluble EpsE: insoluble EpsE (+/- SEM) was calculated and averaged over three independent assays. * denotes a statistically significant difference from that of the wild type fraction ($p < 0.01$).

Mutating the tetracysteine motif inhibits T2S *in vivo*:

Removal of the loop region compromised T2S; however, many residues other than those coordinating the zinc were affected by this deletion. I therefore looked for methods to inhibit zinc coordination without complete removal of the domain. My initial approach to determine the role of zinc coordination in C_M domain was to make a series of point mutations in the tetracysteine motif which binds the zinc ion. In these mutants, the loop region residues would still be present, but would likely have lost the hairpin conformation. I first tested the single mutants: C400S and C430S, the double mutants: C397SC400S and C430SC433S, the triple mutant: C400SC430SC433S (C3xS) and the quadruple mutant: C397SC400SC430SC433S (C4xS) for complementation of an *epsE* deficient strain of *V. cholerae*. As seen in Figure 3.7A, replacement of even one

cysteine residue compromised EpsE's function in T2S. The mutant proteins were stably overexpressed as seen in Figure 3.7B, indicating that removal of the zinc coordinated residues did not grossly affect the overall stability of the protein. I next wanted to investigate the ability of these cysteine mutants to act in a dominant negative manner when introduced into a wild type strain of *V. cholerae*. As seen in Figure 3.8A, the addition of 10uM IPTG was not enough to see a statistically significant decrease in protease secretion; however, increasing the amount of IPTG to 100uM did give the dominant phenotype. When immunoblotted against EpsE, all cysteine mutants were stably overexpressed (Fig 3.8B). These results indicate that replacement of the cysteine residues may introduce some local misfolding which may sterically reduce the ability of the mutant proteins to interact with the wild type T2S components and cause the transdominance phenotype. Alternatively, the incorporation of a few mutant EpsEs into the hexamer may not completely interfere with the function of the hexamer.

Serine replacements of the tetracysteine motif does not grossly affect EpsL(1-253)His_6 interactions:

When the entire elbow was deleted from EpsE, I detected a statistically significant reduction in the ability of EpsE to interact with EpsL. I therefore wanted to investigate the effect of serine replacements in the tetracysteine motif to determine if a similar defect can be detected with these mutant proteins. I decided to first test the C4xS mutant in which all four cysteines were mutated to

serine, hypothesizing that this would show the most drastic phenotype. As Figure 3.9 and Table 3.2 demonstrate the C4xS mutant does seemingly show a slight defect in EpsL binding, although this is not statistically different from that of wild type protein. The ratio of the soluble to insoluble C4xS, however, was comparable to that of wild type EpsE, which may indicate that any local misfolding caused by these mutations did not significantly affect the solubility of the protein. These data suggest that the structure that forms due to the zinc coordination in the elbow domain is important for protein-protein interactions; however, replacement of the cysteines residues does not reduce the interaction of EpsE to EpsL to the same level as complete removal of the loop does.

Chemical removal of zinc inhibits ATPase activity *in vitro*:

Previous studies have demonstrated that the chemical removal of zinc using the reagent PMPS reduced EpsE's ATPase activity by approximately 50% (Camberg & Sandkvist, 2005). For my second approach to determine the role of zinc coordination, I analyzed the affects of zinc removal to EpsE's ATPase activity in the presence of EpsL and CL by performing kinetic ATPase assays on PMPS treated wild type protein. As seen in Figure 3.4, 10-fold molar excess treatment of PMPS abolished ATPase activity *in vitro*. Cross-linking analysis of the PMPS-treated protein did not show any difference in the formation of higher molecular weight product, indicating that the loss of zinc does not grossly affect the ability of EpsE to multerimerize (Figure 3.10 and Table 3.1).

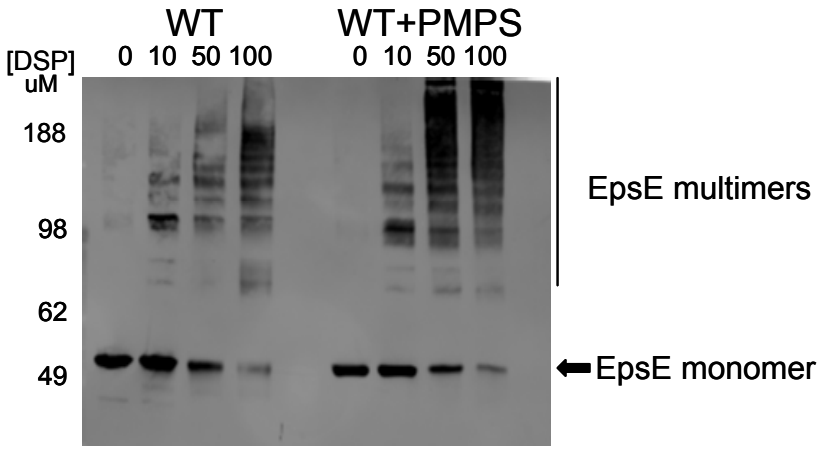


Figure 3.10: PMPS-treated EpsE can form oligomers.
Reactions containing either wild type EpsE-EpsL(1-253)His_6 or wild type protein treated with 10-fold molar excess of PMPS were cross-linked in the presence of Mg-ATP and cardiolipin using increasing concentration of DSP (0, 10, 50, or 100 μM). Samples were analyzed by SDS-PAGE and immunoblotted with anti-EpsE. Molecular mass markers are denoted on the left and the position of monomeric EpsE is indicated with an arrow.

Discussion

The C_M domain of EpsE is highly conserved across the GspE/PilB ATPases, yet little is known about the role this region play in the function of the protein. In this study, I have demonstrated that this region affects the protein-protein interactions either directly or indirectly as removal of the elbow reduced the interaction of EpsE to EpsL. I further showed that replacement of the cysteine residues with serines, either individually or multiple replacements, abolished T2S, but did not have a significant defect on protein-protein interactions. Additionally, I provide data to support the idea that the R441 residue found at the end of the elbow domain likely participates in a functional interaction with the nucleotide, and have shown this residue to be required for ATPase activity *in vitro* and T2S *in vivo*.

The R441 residue likely plays a role in sensing or coordinating the nucleotide in the binding pocket. Previous work has explored the role of other arginine residues to the formation of the active site (Patrick *et al*., 2011) but the residues investigated in that work were limited to those which may interact with the phosphate groups located at one end of the bound nucleotide. This R441 residue is thought to play a role in binding the ribose and adenyl moieties (Figure 3.11) and even though it is highly conserved across the secretion family of ATPases, its role in ATPase activity has not been previously investigated (Robien *et al*., 2003). Here I show that EpsE with either an alanine or aspartate replacement at residue 441 is unable to complement an *epsE*$^-$ strain *in vivo* and abolishes ATPase activity *in vitro*, but does not have a discernable affect on

EpsE multimerization. These data are consistent with what was observed for other arginine mutants within the ATP binding site (Patrick *et al*., 2011) therefore prompting me to conclude that this residue is also participating within the active site.

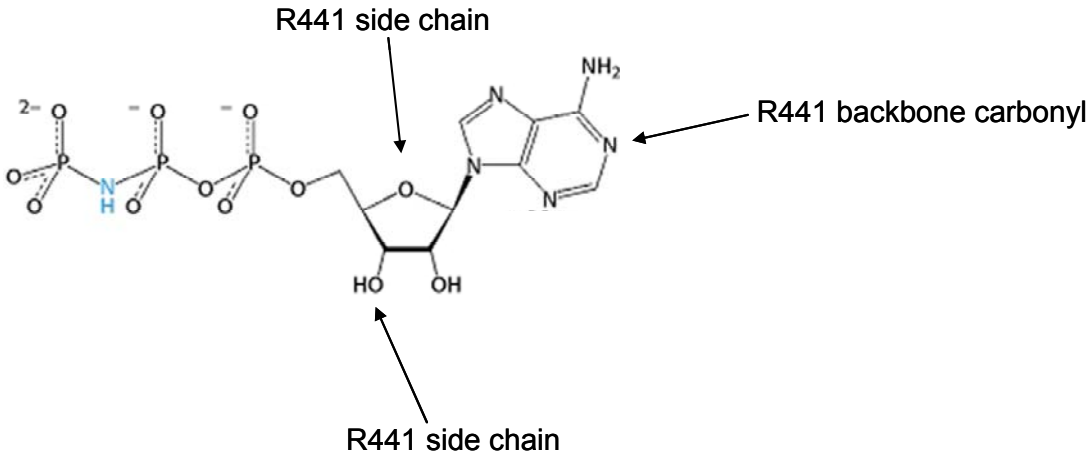


Figure 3.11: Interactions with the bound nucleotide.
Schematic representation showing where the R441 residue of EpsE interacts with bound AMPPNP. Adapted from (Robien *et al*., 2003).

Although the mutation is made in the same residue, I see varied affects when replacing arginine 441 with either alanine or aspartate. Knowing that this residue is found at the top of the C_M domain where the domain terminates and reintegrates with the CTD (Fig 3.2), it is possible that the interaction between the EpsE elbow and EpsL could be affected by modification of this residue. This is supported in part by the fact that the R441D acted more dominant *in vivo* (Fig 3.3), as the aspartate may introduce a charge repulsion with other residues in the

nucleotide binding site, altering its conformation. This change in binding pocket conformation could potentially result in better interactions between EpsE:R441D and EpsL. Similarly, it seems as though the purified R441D protein floated more efficiently in the liposome assay (Fig 3.6), and although this was not statistically significant it is suggestive of a stronger, less variable EpsE:EpsL interaction. As wild type EpsE is dynamic and needs to be in a nucleotide-bound state in order to effectively bind CL, I see some variation in its ability to float with CL. It is possible that this R441D mutation may favor the nucleotide bound conformation, making this protein less dynamic and more amenable to CL floatation. Further experiments are necessary to probe this hypothesis.

I further explored the role of the zinc-coordinated elbow to the *in vivo* function, first by complete removal of the domain and second by mutating the zinc-coordinating residues both individually and in multiples. I found that complete removal of the loop abolished T2S without having any major effect on the stability of the protein (Figs 3.7 and 3.8). I was surprised, however, by my inability to purify this protein in complex with EpsL. Previous work has demonstrated that EpsE and EpsL interact through their N1 and cytoplasmic domains, respectively (Abendroth *et al*., 2005, Sandkvist *et al*., 1995). It was therefore unexpected that removal of a domain for EpsE that likely does not affect the conformation of the distant N1 domain could have such a drastic affect on EpsL interaction.

My data suggests that the chemical removal of the zinc ion abolished the ATPase activity of EpsE *in vitro*. I attempted to replicate this phenomenon *in vivo*

by abolishing zinc coordination through serine replacements within the tetracysteine motif. These replacements, too, failed to complement an *epsE* deficient strain of *V. cholerae*, yet based on my secretion data, they seemed to be more tolerated by the wild type strain of *V. cholerae*, as higher expression levels of the mutant protein were required to see the dominance effect *in vivo* (Figs 3.7 and 3.8). Unlike the loop mutants, however, the complete tetracysteine mutant did not have the same drastic effect on EpsL interactions (Fig 3.9), although this protein presumably has completely lost the ability to coordinate zinc, and thus has likely lost the hairpin structure formed by the metal coordination. Similarly, it appears that removal of the entire elbow causes the protein to lose some solubility, which is not seen with the C4xS mutant. This may contribute to the less severe loss in protein-protein interactions. Having the elbow region intact, but probably not structurally similar to wild type is inadequate to support secretion, but is enough to at least transiently interact with other proteins necessary to support secretion. I can speculate that the 29 intervening residues between the two CxxC motifs are sufficient for EpsE-EpsL interaction; however, in order to transmit a signal to or from the greater T2SS to promote secretion, these residues have to adopt the elbow conformation induced by zinc coordination. This idea is partially supported by the SecA studies in which the actual residues responsible for the binding of SecA and SecB were the basic residues found within the domain itself and not the tetracysteine motif, *per se* (Zhou & Xu, 2003). Purification and further analysis of these cysteine mutants will be necessary to determine their roles in ATPase activity *in vitro*.

Based on my findings here, I can speculate about how the zinc domain may participate in both the function of EpsE as well as generally in T2S. First, I have shown that the R441 residue is a required component in the active site, and as this residue is found at the end of the C_M domain, it likely provides a link between the nucleotide binding site and other proteins in the T2S complex. I can envision that R441 is sensitive to conformational changes in the C_M hairpin. Consequently, removal of the entire loop, replacement of individual cysteines, or binding of another protein such as EpsL to the loop could have substantial effects on EpsE's ability to hydrolyze ATP. I have shown in this work that EpsL likely interacts with the C_M domain, but there is potential for other T2S proteins such as EpsF, to also interact with this domain as well, either transiently or more stably. Work in *Erwinia chrysanthemi* has shown that EpsE, L, and F can form a complex; however, EpsF can only bind when EpsE and EpsL are already in complex with each other (Py *et al.*, 2001). It is therefore tempting to speculate that the elbow may help to orient EpsE and EpsL properly, allowing for the EpsE NTD-EpsL interaction which could then make the C_M domain available for additional interactions with EpsF. As my and others work on EpsE have demonstrated that this protein changes its conformation during nucleotide binding, it seems probable that interactions at the elbow region could, in part, be responsible for regulating this. As EpsE in complex with EpsL has high ATPase activity, it may be reasonable to assume that interactions of another protein at the elbow may provide a way to regulate EpsE function, by stabilizing a conformation that is not amenable to nucleotide binding or hydrolysis. Further

studies are required to answer these questions and will be useful in fine-tuning the model of how the energy produced by EpsE's ATP hydrolysis can be regulated and transmitted through the T2S complex.

Chapter 4

Discussion and significance

The overarching goal of this project was to functionally characterize the subdomains of EpsE in order to differentiate their contribution to EpsE's ability to hydrolyze ATP and support Type II secretion (T2S). Dissecting EpsE has advanced the knowledge in the field as to how it works and I have used my data to construct a model of how EpsE hydrolyzes ATP for energy and how it can then transduce the energy to the greater T2S apparatus. The findings presented in Chapter 2 may be widely applicable to not just the T2S ATPases, but to all members of the Type II/Type IV family. The arginine residues investigated are conserved across multiple EpsE homologs found within this family and may play a role in a universal mechanism of ATP hydrolysis. In contrast, the work from Chapter 3 highlights the role of an important domain unique to T2S-specific ATPases. This subdomain may participate in protein-protein interactions that could be important for system specific regulation of the more universal ATPase mechanism. Overall, my findings will be useful in elucidating the mechanistic similarities and difference between ATPases involved in various macromolecular transport systems.

The role of oligomerization in the ATP hydrolysis mechanism

Previous work detailed the stimulation of EpsE's ATPase activity by means of the cytoplasmic domain of EpsL and the acidic phospholipid, cardiolipin (CL). We speculated based on these studies that EpsE multimerizes, likely into a hexameric ring, and this is aided by the binding of EpsE to EpsL. Further we concluded that the interaction of EpsE to EpsL must induce a conformational change given that the interaction between the EpsE-EpsL complex with CL in part depended on the nucleotide-bound state of EpsE (Camberg *et al*., 2007). Further we found that pre-bound ATP was required in order to see the great increase of ATP hydrolysis in the presence of these factors, leaving us with the unanswered question of what the precise mechanism of ATP hydrolysis could be.

I first tackled the mechanism of nucleotide hydrolysis and the role oligomerization plays in the function of EpsE. As presented in Chapter 2, our work has shown that six conserved arginine residues are required for the hydrolysis of ATP due to their role in completing the active site of EpsE. This required the coming together of multiple subunits and domains as arginines from the NTD, CTD, and CTD' are all essential for hydrolysis. The data from those studies solidified our proposal that EpsE is active as an oligomer, as the requirement for residues from neighboring subunits implies the requirement for oligomerization toward function. Although not all of the residues examined in this work were suggested, based on the structural model of EpsE, to participate directly in the hydrolysis of the nucleotide, we were able to propose functional

roles for each arginine and pin point why a replacement in any of them would be detrimental to the function of the protein.

To further drive home the point that oligomerization is necessary but not sufficient for efficient ATPase activity, we analyzed mutations in the subunit interface. Studying these mutations allowed us to confirm that a mutation that may disrupt the specific alignment of neighboring subunits were just as detrimental as making a mutation directly in the active site. This is apparent when looking at the D326R and R156D mutants. Neither mutant was able to complement an *epsE* deficient strain of *Vibrio cholerae*, yet have a transdominance phenotype in a wild type *V. cholerae* background, indicating a loss of function rather than gross change in protein folding. Interestingly, the R156D mutant retained a very low level of ATPase activity, albeit not enough to secrete the protease, but nonetheless, statistically higher than the inert protein BSA. If one envisions all of the interactions within the subunit interface as being required for the proper alignment of neighboring subunits and thus the proper alignment of residues within the complete binding pocket, the fact that the D326R mutant has no ATPase activity and the R156D mutant retains a low level of ATPase activity suggest that these mutations causes a somewhat imperfect alignment of the subunits thus resulting in insufficient ATPase activity.

The C_M domain

A unique feature that differentiates secretion ATPases in the GspE/PilB subfamily from those in the larger Type II/IV secretion ATPase family is the

additional domain, termed the C_M domain. This domain is found on the outer edge of the EpsE hexamer and is unique in that it coordinates one zinc ion per monomer (Figure 4.1) (Camberg & Sandkvist, 2005, Robien *et al*., 2003). Although largely uncharacterized, this domain has been shown to be important for T2S in *Klebsiella oxytoca*, and may partially aid in ATP hydrolysis in EpsE (Camberg & Sandkvist, 2005, Possot & Pugsley, 1997). A recent study on PilF, a newly characterized ATPase necessary for natural competence in *Thermus thermophilus,* was shown to also contain a C_M domain containing the tetracysteine motif (Rose *et al.*, 2011). As this is a member of the GspE/PilB subfamily of secretion ATPases, it is not surprising that the protein contains this domain; however, they discovered that, as a hexamer, this protein only contains one zinc molecule per complex, in contrast to one per monomer. The authors did not detect any marked defect in the removal of zinc through mutations of the tetracysteine motif on its ATPase activity, which is in opposition to the ~50% reduction we see with chemical removal of zinc from EpsE *in vitro* (Camberg & Sandkvist, 2005). The implications of these observations are not yet understood, but the authors propose that this asymmetrical zinc binding may define one subunit for specific interactions (Rose *et al.*, 2011). As a hexameric EpsE crystal structure has never been obtained and visualized, it may be possible that not all subunits are fully occupied with zinc ions, and this may be common to all members of this family.

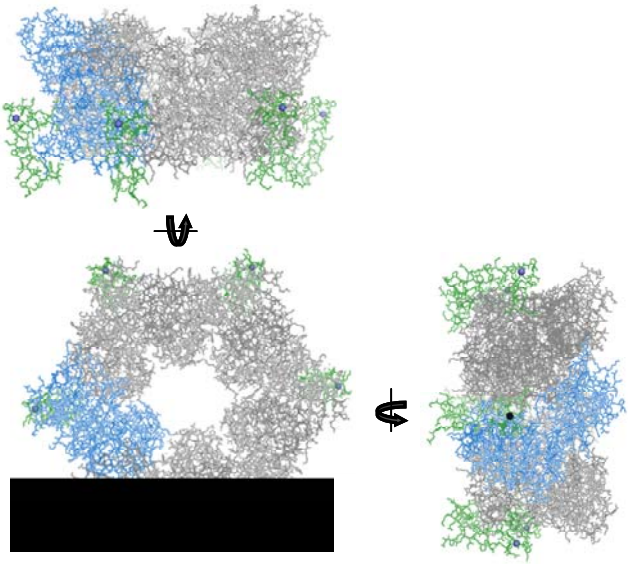


Figure 4.1: The C_M domain is on the outside of the EpsE hexamer.
A hexamer model of EpsE is shown in three different orientations. For clarity, one monomer is colored blue and the zinc-binding (C_M) domain is colored light green. The zinc ion is shown as a light blue sphere.

In my studies, I chose to examine the role of the C_M domain to EpsE's hydrolysis mechanism with the idea that this domain is in prime position to interact with and possibly be regulated by another protein (Fig. 4.1). I hypothesized that the R441 residue found at the end of this domain was required for hydrolysis and could be sensitive to small changes in conformation. Therefore, a protein binding to the C_M domain could have a profound impact on the position of this residue causing substantial effects on EpsE's ATPase activity. My studies in Chapter 3 suggested that the R441 residue does, in fact play a role

in ATP hydrolysis, as replacement of this residue with either alanine or aspartate inhibited T2S as well as abolished EpsE's ATPase activity.

I next asked whether EpsE could function without the zinc elbow region, thinking that homologous proteins, such as PilT, which are missing the C_M domain are still able to hydrolyze nucleotide. I hypothesized that complete removal of this domain may inhibit T2S, but maybe not ATPase activity, unless it is playing a regulatory role. We found that purification of this protein was difficult as the EpsE mutant was not strongly interacting with EpsL(1-253)His_6 and surmised that EpsL may interact at the elbow region. This was unexpected given that I knew EpsL interacts with EpsE through its NTD; however, that NTD interaction does not preclude the idea of other parts of EpsE binding to EpsL as well. In fact, early work on EpsE and EpsL suggested that an additional site of interaction outside of the NTD of EpsE may exist (Sandkvist *et al*., 1995). Taken together, these data suggest a two-step mechanism for EpsE-EpsL interactions whereby the EpsE NTD first interacts with the cytoplasmic domain of EpsL. After this occurs, a secondary interaction between the EpsE C_M domain and a yet undefined region of EpsL may then take place. It is feasible that the EpsE NTD interaction with EpsL is not completely functional, but once the C_M domain engages EpsL, if forms a fully operational complex. This idea will need to be followed up more closely.

I tried to further tease apart the role of the zinc elbow region and probe the idea of EpsL interacting here by making cysteine to serine replacements in the tetracysteine motif. I found that even a single replacement abolished T2S;

however, I have not yet purified these proteins to test the effects of these mutations on the ATPase activity. I know that the potential EpsE-EpsL interaction at this region is not as severely affected by the cysteine mutations as by the complete elbow removal; however, it is probable that the cysteine mutations have lost the ability to coordinate zinc, which one would expect to interfere with the formation or display of the loop. My PMPS-treated EpsE, which chemically removed the zinc, showed a severe decrease in ATPase activity, and based on this, it seems probable that the cysteine to serine mutant proteins will also have lost the zinc ion and therefore will not retain any ATPase activity. Further experiments with the cysteine mutants will help us elucidate the role of this elbow region and may direct me toward additional parts of EpsL or other proteins which may be interacting with this elbow.

One interesting observation regarding the zinc elbow is that it is conserved in the GspE/PilB but not PilT subfamilies of secretion ATPases. Therefore it seems probable that whatever protein is binding to this region is likely T2S specific or may be important for pili polymerization rather than depolymerization as PilB is required to build the pilus and PilT to break it down. It is likely that the specific hairpin fold is important for protein interactions, and the precise sequence may be necessary for species specificity. This is supported by the fact that the sequence homology between the residues of the tetracysteine motif is less than 30%-- a comparable number to that of the NTD, which is known to be required for species specificity (Possot & Pugsley, 1994, Sandkvist *et al*., 1995). Likewise, my data showing that the cysteine mutants have a greater ability to

interact with EpsL than the mutants with the loop region removed, even if this is a non-functional interaction, demonstrates the necessity of the loop residues to protein-protein interactions. It should also be noted that XpsE in *Xanthomonas campestris* do not have this tetracysteine motif; however, other homologous ATPases still have the residues encompassing the C_M domain (Filloux, 2004). Further, the prepillin peptidase that is required for the processing of the pseudopilins in *V. cholerae* also coordinates a zinc with a tetracysteine motif and, although it is unknown whether this is purely coincidence, it should be noted that in *X. campestris*, the peptidase is missing the zinc motif as well, and yet must adopt a functional fold (Filloux, 2004, Robien *et al*., 2003).

A comprehensive model of EpsE's hydrolysis mechanism

While more characterization of the elbow is needed, based on the work presented here, I can speculate on its role in T2S. The placement of the elbow and the orientation of the hairpin strongly suggest that the zinc is playing a structural role, providing a fold necessary to support ATP hydrolysis and a spot for protein-protein interactions. I can speculate that the loss of the hairpin fold due to replacement of cysteines in the tetracysteine motif may result in the movement of the R441 residue out of place in the nucleotide binding site and thus may halt ATPase activity and T2S. As my work has shown R441 to be necessary to support secretion, changes in the elbow that affect its placement within the nucleotide binding pocket may be detrimental. Taking these data together, I can speculate that a protein, such as EpsL, binding to the elbow may

be in prime position to regulate the orientation of the R441 residue and thus control ATPase activity.

With this in mind, I can now envision a model (Figure 4.2) where EpsL interacts with the apo form of EpsE at both its NTD and C_M domains. This would hold EpsE in a completely open form. When nucleotide binds, the NTD of EpsE closes around it, which may alter the interaction between EpsL and the C_M domain. This in turn could allow the R441 residue to interact with the nucleotide and orient it in the binding pocket through its interactions with the ribose and adenyl moieties. The closing of the NTD further allows for the six conserved arginine residues from adjoining domains and subunits to realign and form the active site. Once formed, the γ-phosphate of ATP is hydrolyzed and released which destabilizes the active site due to a loss of the negatively charged Pi. The NTD can then open, which allows for the re-engagement of EpsL with the C_M domain. This may change the conformation of R441, aiding in the release of ADP.

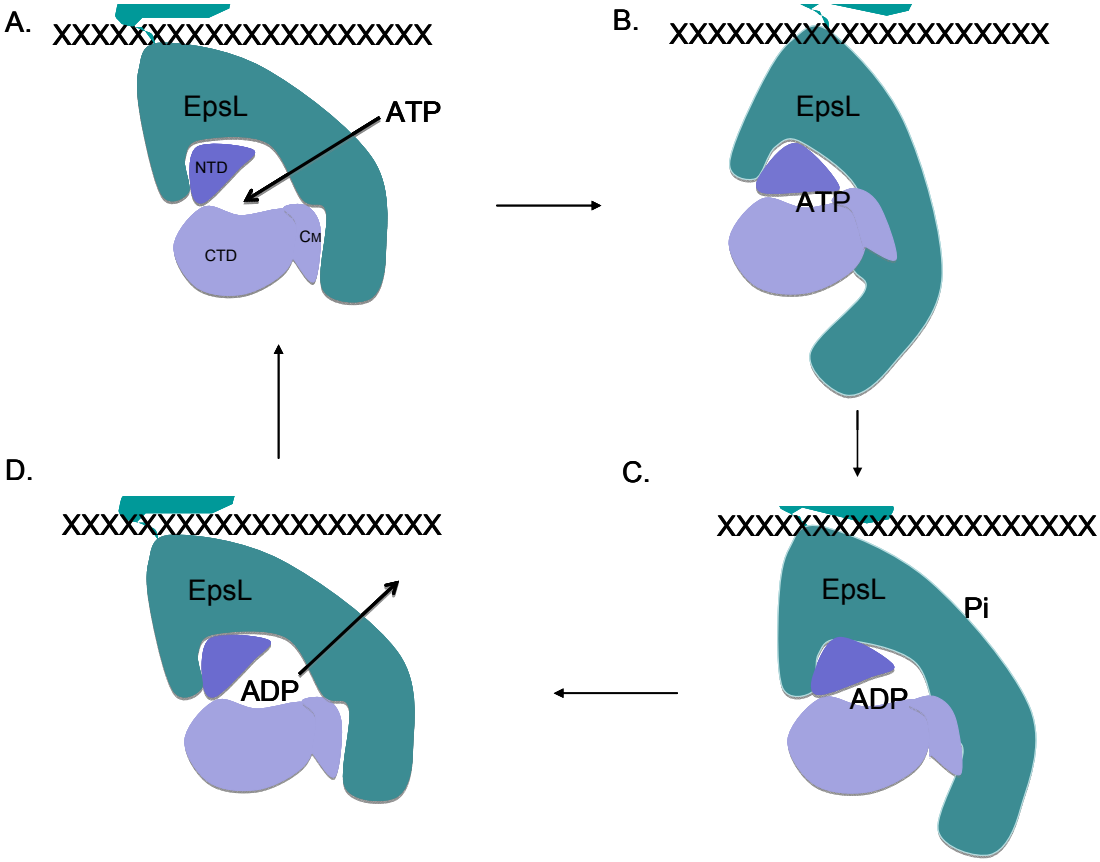


Figure 4.2: A model of ATP hydrolysis by EpsE.
(A) EpsE in apo form (light and dark purple) interacts through its N-terminal domain (NTD) and C_M domains with the cytoplasmic domain of the transmembrane protein EpsL (cyan) at the inner membrane, which allows for ATP to enter the binding site. (B) Bound ATP causes a closing of EpsE's NTD, resulting in movement of EpsL and an altered interaction between EpsE's C_M domain and EpsL. (C) Hydrolysis of ATP and release of Pi destabilizes the active site as Pi is released, causing the NTD to move toward a more open conformation. (D) Once the NTD is in its open form, EpsL re-engages with the C_M domain of EpsE allowing for release of ADP from the binding site.

The dynamic movements which accompany ATP binding, hydrolysis, and release can be translated through the T2S complex by means of EpsL, as the movements of EpsE's NTD likely cause conformational changes in EpsL. EpsL can then transduce those mechanical movements to other T2S proteins that interact with it, most notably to that of the major pseudopilin, EpsG.

This model is based on the idea that the zinc elbow will interact with EpsL. If, in future studies it is found that the elbow is not a site of EpsE-EpsL interaction, the T2S protein EpsF is another candidate for interacting with the C_M domain. EpsF is also found in the T2S inner membrane complex, has a significant cytoplasmic domain, and has been previously suggested to form an ELF complex in *Erwinia chrysanthemi* (Py *et al*., 2001) and *Pseudomonas aeruginosa* (Robert *et al*., 2005). Although studies on EpsF are limited and not much is known about its role in T2S, it is well conserved across the T2S/T4P systems where we find many ATPases of the GspE/PilB family likewise conserved (Planet *et al.*, 2001). It is easy to envision a revised mechanism whereby EpsE and EpsL are in complex through their NTD and cytoplasmic domains, respectively and this complex may then interact with EpsF through the C_M domain of EpsE. When EpsF is bound to the elbow, it may change the conformation of this domain, which in turn may alter the position of the R441 residue in the active site and directly regulate ATP hydrolysis.

In an alternative model, the placement of R441 may allow for it to play a role in sensing the nucleotide binding state, thus transmitting the energy from ATP hydrolysis through the elbow region and to the T2S apparatus by means of

protein interactions. In order to distinguish between these two hypotheses, I would first need to map the interaction sites on both the C_M domain and its interacting partner. Once this is known I could create a series of mutations to try to separate the ATPase activity of EpsE from that of T2S. If the interactions at this elbow are transducing energy rather than playing a regulatory role, one could imagine that the protein could still hydrolyze nucleotide in the absence of interactions at the elbow; however, if the elbow is required for regulation, it may not be possible to separate ATPase activity from protein secretion by this means.

The possibility also exists that the C_M domain interacts with another protein. It is not uncommon for two zinc binding domains to interact, as this has been seen before with the ClpX ATPase (Donaldson *et al.*, 2003). Here it was shown that ClpX could form a trimer of dimers through the strong interactions between two zinc-coordinating domains. Later work suggested that these zinc domains can undergo drastic domain movement due to the ATPase activity of the protein and the repositioning of these domains plays a direct role in protein degradation (Thibault *et al.*, 2006). For EpsE, I do not envision the zinc elbows of adjoining subunits as being able to interact due to their placement on the outer edge of the EpsE hexamer model (Fig 4.1). There is another possible candidate protein, however, that is required for T2S. This is the prepillin peptidase, EpsO, which, as mentioned previously, has the same tetracysteine motif present in its first cytoplasmic domain and could potentially interact with EpsE in a similar fashion to that of ClpX. It could be possible that the prepillin peptidase binds transiently to EpsE at the zinc elbow to process the pseudopilins as this would

position it near the T2S complex. The tetracysteine motif in EpsO is in the first cytoplasmic domain, which does not appear to be necessary for peptidase activity (LaPointe & Taylor, 2000). Although EpsG can be processed in the absence of EpsE, the potential EpsE-EpsO interaction may increase the efficiency of pseudopilin processing. Further, this EpsE-EpsO interaction may be reasonable given the recently discovered direct link between EpsG and EpsL, which is dependant on EpsO processing the major pseudopilin EpsG (Gray *et al*., 2011). It is tempting to speculate that EpsE may transiently hold EpsO in place so it can process EpsG which would then allow EpsG and EpsL to interact and polymerize the pilus. That could also explain why the C_M domain is not found in the retraction ATPase PilT and why the *X. campestris* homologs are both zinc-free.

Future Directions

While this work has greatly enhanced our knowledge about EpsE and how it functions within the T2S apparatus, many questions still remain. An immediate goal is to purify the cysteine mutants to test their activities *in vitro*. It is known that they do not support secretion through the T2S apparatus, however, it is unknown whether that defect is due to the mutant protein's inability to hydrolyze ATP, interact with EpsL or possibly interact with an additional T2S protein. It is expected that these proteins will be amenable to purification and can thus provide insight as to the role the elbow structure plays. Given that the SecA-SecB interaction required basic residues and not the tetracysteine motif,

specifically, and the fact that I do not see the same defect in the C4xS mutant as I do with the elbow deletion mutant, I may be able to additionally use the series of cysteine mutants as well as additional mutants to investigate the role of the specific residues within the loop region.

I would also like to further probe the idea of the differential interactions of R441A and R441D with EpsL and CL and test the hypothesis that the nucleotide state may be playing a role in this difference. A first step in answering this question would be to test the purified proteins for their ability to bind and be stabilized by different nucleotides, such as ATP, ADP, the non-hydrolyzable analog AMPPNP, or with no nucleotide. This would be measured by determining the temperature at which each of these protein-nucleotide complexes melt, with a higher melting temperature indicative of a more stable complex. This could be followed up by determining whether the liposome binding profile changes for either mutant dependant on nucleotide state. These studies could help to fine tune my proposed model and allow for a better appreciation as to how the C_M domain may be tied to the nucleotide state of the protein.

The work presented here is the first time a direct role for the EpsE C_M domain in protein-protein interactions has been suggested. Many future experiments are required to fully characterize the potential of this region to interact with EpsL, some other T2S component, such as EpsF, or the possibility of multiple interactions with different proteins over the course of secretion. One interesting idea is to test for the ability of a suppressor mutation to compensate for a mutation in the elbow region that abolishes protein-protein interactions. To

get at this idea, I am in the process of constructing a plasmid co-expressing EpsE, L and F to which I can then introduce a point mutant of choice into EpsE, transform the plasmid into an *Escherichia coli* mutator strain and then move the plasmid into our *epsE*$^-$ strain of *V. cholerae* to assay for restoration of T2S.

Implications for therapeutic measures

Once a patient exhibits symptoms of cholera, the best chance for survival is to administer oral rehydration therapy. Use of the antibiotics tetracycline and ciprofloxacin have also been found to reduce the severity and shorten the period of disease; however, as antibiotic resistance in the toxigenic strains of *V. cholerae* have increased, the need for a new therapeutic is essential. It also means that a practical place to counteract cholera is before a person contracts it from a contaminated source and requires antibiotics to be administered. Therapeutics halting the transmission of cholera from an infected person and back into the environment and vice versa could prove quite effective in controlling an outbreak, be it a seasonal epidemic or one due to a natural disaster. Components of the T2S system could be ideal candidates for new therapeutics, and an inhibitor directed against ATPase could be beneficial in halting the secretion of not only the cholera toxin itself, but also inhibit the secretion of other proteins from the T2S apparatus, including those necessary for survival in the environment. Based on the work presented here, finding an inhibitor that could bind to or otherwise interfere with the conformational changes associated with ATPase activity could inactivate the whole secretion mechanism. It may also be

possible to exploit the cytoplasmic nature of the ATPase component and use an inhibitor to block the association of EpsE to the rest of the T2S complex. Seeing as these components, especially the ATPase, are also highly conserved in the T4P biogenesis system (which is also necessary for *V. cholerae* colonization) the tools generated from my studies here may be widely applicable at targeting these bacteria from multiple angles as a means of controlling infection and proliferation.

References

Abd, H., A. Saeed, A. Weintraub, G. B. Nair & G. Sandstrom, (2007) Vibrio cholerae O1 strains are facultative intracellular bacteria, able to survive and multiply symbiotically inside the aquatic free-living amoeba Acanthamoeba castellanii. *FEMS Microbiol Ecol* **60**: 33-39.

Abd, H., A. Weintraub & G. Sandstrom, (2005) Intracellular survival and replication of Vibrio cholerae O139 in aquatic free-living amoebae. *Environ Microbiol* **7**: 1003-1008.

Abendroth, J., M. Bagdasarian, M. Sandkvist & W. G. Hol, (2004a) The structure of the cytoplasmic domain of EpsL, an inner membrane component of the type II secretion system of Vibrio cholerae: an unusual member of the actin-like ATPase superfamily. *J Mol Biol* **344**: 619-633.

Abendroth, J., A. C. Kreger & W. G. Hol, (2009a) The dimer formed by the periplasmic domain of EpsL from the Type 2 Secretion System of Vibrio parahaemolyticus. *J Struct Biol* **168**: 313-322.

Abendroth, J., D. D. Mitchell, K. V. Korotkov, T. L. Johnson, A. Kreger, M. Sandkvist & W. G. Hol, (2009b) The three-dimensional structure of the cytoplasmic domains of EpsF from the type 2 secretion system of Vibrio cholerae. *J Struct Biol* **166**: 303-315.

Abendroth, J., P. Murphy, M. Sandkvist, M. Bagdasarian & W. G. Hol, (2005) The X-ray structure of the type II secretion system complex formed by the N-terminal domain of EpsE and the cytoplasmic domain of EpsL of Vibrio cholerae. *J Mol Biol* **348**: 845-855.

Abendroth, J., A. E. Rice, K. McLuskey, M. Bagdasarian & W. G. Hol, (2004b) The crystal structure of the periplasmic domain of the type II secretion system protein EpsM from Vibrio cholerae: the simplest version of the ferredoxin fold. *J Mol Biol* **338**: 585-596.

Alam, M., M. Sultana, G. B. Nair, R. B. Sack, D. A. Sack, A. K. Siddique, A. Ali, A. Huq & R. R. Colwell, (2006) Toxigenic Vibrio cholerae in the aquatic environment of Mathbaria, Bangladesh. *Appl Environ Microbiol* **72**: 2849-2855.

Alam, M., M. Sultana, G. B. Nair, A. K. Siddique, N. A. Hasan, R. B. Sack, D. A. Sack, K. U. Ahmed, A. Sadique, H. Watanabe, C. J. Grim, A. Huq & R. R. Colwell, (2007) Viable but nonculturable Vibrio cholerae O1 in biofilms in the aquatic environment and their role in cholera transmission. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 17801-17806.

Arts, J., A. de Groot, G. Ball, E. Durand, M. El Khattabi, A. Filloux, J. Tommassen & M. Koster, (2007) Interaction domains in the Pseudomonas aeruginosa type II secretory apparatus component XcpS (GspF). *Microbiology* **153**: 1582-1592.

Bitter, W., M. Koster, M. Latijnhouwers, H. de Cock & J. Tommassen, (1998) Formation of oligomeric rings by XcpQ and PilQ, which are involved in

protein transport across the outer membrane of Pseudomonas aeruginosa. *Mol Microbiol* **27**: 209-219.

Bleves, S., M. Gerard-Vincent, A. Lazdunski & A. Filloux, (1999) Structure-function analysis of XcpP, a component involved in general secretory pathway-dependent protein secretion in Pseudomonas aeruginosa. *Journal of bacteriology* **181**: 4012-4019.

Bouley, J., G. Condemine & V. E. Shevchik, (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. *J Mol Biol* **308**: 205-219.

Brok, R., P. Van Gelder, M. Winterhalter, U. Ziese, A. J. Koster, H. de Cock, M. Koster, J. Tommassen & W. Bitter, (1999) The C-terminal domain of the Pseudomonas secretin XcpQ forms oligomeric rings with pore activity. *J Mol Biol* **294**: 1169-1179.

Buddelmeijer, N., M. Krehenbrink, F. Pecorari & A. P. Pugsley, (2009) Type II secretion system secretin PulD localizes in clusters in the Escherichia coli outer membrane. *Journal of bacteriology* **191**: 161-168.

Camberg, J. L., T. L. Johnson, M. Patrick, J. Abendroth, W. G. Hol & M. Sandkvist, (2007) Synergistic stimulation of EpsE ATP hydrolysis by EpsL and acidic phospholipids. *Embo J* **26**: 19-27.

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

Chami, M., I. Guilvout, M. Gregorini, H. W. Remigy, S. A. Muller, M. Valerio, A. Engel, A. P. Pugsley & N. Bayan, (2005) Structural insights into the secretin PulD and its trypsin-resistant core. *The Journal of biological chemistry* **280**: 37732-37741.

Chiang, P., L. M. Sampaleanu, M. Ayers, M. Pahuta, P. L. Howell & L. L. Burrows, (2008) Functional role of conserved residues in the characteristic secretion NTPase motifs of the Pseudomonas aeruginosa type IV pilus motor proteins PilB, PilT and PilU. *Microbiology* **154**: 114-126.

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

Colwell, R. R. & A. Huq, (1994) Environmental reservoir of Vibrio cholerae. The causative agent of cholera. *Ann N Y Acad Sci* **740**: 44-54.

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

Connell, T. D., D. J. Metzger, M. Wang, M. G. Jobling & R. K. Holmes, (1995) Initial studies of the structural signal for extracellular transport of cholera toxin and other proteins recognized by Vibrio cholerae. *Infect Immun* **63**: 4091-4098.

Craig, L., M. E. Pique & J. A. Tainer, (2004) Type IV pilus structure and bacterial pathogenicity. *Nat Rev Microbiol* **2**: 363-378.

Donaldson, L. W., U. Wojtyra & W. A. Houry, (2003) Solution structure of the dimeric zinc binding domain of the chaperone ClpX. *The Journal of biological chemistry* **278**: 48991-48996.

Douet, V., L. Loiseau, F. Barras & B. Py, (2004) Systematic analysis, by the yeast two-hybrid, of protein interaction between components of the type II secretory machinery of Erwinia chrysanthemi. *Res Microbiol* **155**: 71-75.

Durand, E., A. Bernadac, G. Ball, A. Lazdunski, J. N. Sturgis & A. Filloux, (2003) Type II protein secretion in Pseudomonas aeruginosa: the pseudopilus is a multifibrillar and adhesive structure. *Journal of bacteriology* **185**: 2749-2758.

Durand, E., G. Michel, R. Voulhoux, J. Kurner, A. Bernadac & A. Filloux, (2005) XcpX controls biogenesis of the Pseudomonas aeruginosa XcpT-containing pseudopilus. *The Journal of biological chemistry* **280**: 31378-31389.

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

Fekkes, P., J. G. de Wit, A. Boorsma, R. H. Friesen & A. J. Driessen, (1999) Zinc stabilizes the SecB binding site of SecA. *Biochemistry* **38**: 5111-5116.

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

Filloux, A., G. Michel & M. Bally, (1998) GSP-dependent protein secretion in gram-negative bacteria: the Xcp system of Pseudomonas aeruginosa. *FEMS Microbiol Rev* **22**: 177-198.

Gerard-Vincent, M., V. Robert, G. Ball, S. Bleves, G. P. Michel, A. Lazdunski & A. Filloux, (2002) Identification of XcpP domains that confer functionality and specificity to the Pseudomonas aeruginosa type II secretion apparatus. *Mol Microbiol* **44**: 1651-1665.

Graf, P. C., M. Martinez-Yamout, S. VanHaerents, H. Lilie, H. J. Dyson & U. Jakob, (2004) Activation of the redox-regulated chaperone Hsp33 by domain unfolding. *The Journal of biological chemistry* **279**: 20529-20538.

Graumann, J., H. Lilie, X. Tang, K. A. Tucker, J. H. Hoffmann, J. Vijayalakshmi, M. Saper, J. C. Bardwell & U. Jakob, (2001) Activation of the redox-regulated molecular chaperone Hsp33--a two-step mechanism. *Structure* **9**: 377-387.

Gray, M. D., M. Bagdasarian, W. G. Hol & M. Sandkvist, (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. *Mol Microbiol* **79**: 786-798.

Guilvout, I., K. R. Hardie, N. Sauvonnet & A. P. Pugsley, (1999) Genetic dissection of the outer membrane secretin PulD: are there distinct domains for multimerization and secretion specificity? *Journal of bacteriology* **181**: 7212-7220.

Halpern, M., Y. B. Broza, S. Mittler, E. Arakawa & M. Broza, (2004) Chironomid egg masses as a natural reservoir of Vibrio cholerae non-O1 and non-O139 in freshwater habitats. *Microb Ecol* **47**: 341-349.

Halpern, M., H. Gancz, M. Broza & Y. Kashi, (2003) Vibrio cholerae hemagglutinin/protease degrades chironomid egg masses. *Appl Environ Microbiol* **69**: 4200-4204.

Hardie, K. R., S. Lory & A. P. Pugsley, (1996) Insertion of an outer membrane protein in Escherichia coli requires a chaperone-like protein. *EMBO J* **15**: 978-988.

Hare, S., R. Bayliss, C. Baron & G. Waksman, (2006) A large domain swap in the VirB11 ATPase of Brucella suis leaves the hexameric assembly intact. *J Mol Biol* **360**: 56-66.

Hare, S., W. Fischer, R. Williams, L. Terradot, R. Bayliss, R. Haas & G. Waksman, (2007) Identification, structure and mode of action of a new regulator of the Helicobacter pylori HP0525 ATPase. *EMBO J* **26**: 4926-4934.

Hirst, T. R., J. Sanchez, J. B. Kaper, S. J. Hardy & J. Holmgren, (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.

Hobbs, M. & J. S. Mattick, (1993) Common components in the assembly of type 4 fimbriae, DNA transfer systems, filamentous phage and protein-secretion apparatus: a general system for the formation of surface-associated protein complexes. *Mol Microbiol* **10**: 233-243.

Huq, A., P. A. West, E. B. Small, M. I. Huq & R. R. Colwell, (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. *Appl Environ Microbiol* **48**: 420-424.

Islam, M. S., B. S. Drasar & D. J. Bradley, (1989) Attachment of toxigenic Vibrio cholerae 01 to various freshwater plants and survival with a filamentous green alga, Rhizoclonium fontanum. *J Trop Med Hyg* **92**: 396-401.

Johnson, T. L., J. Abendroth, W. G. Hol & M. Sandkvist, (2006) Type II secretion: from structure to function. *FEMS Microbiol Lett* **255**: 175-186.

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

Kirn, T. J., B. A. Jude & R. K. Taylor, (2005) A colonization factor links Vibrio cholerae environmental survival and human infection. *Nature* **438**: 863-866.

Kohler, R., K. Schafer, S. Muller, G. Vignon, K. Diederichs, A. Philippsen, P. Ringler, A. P. Pugsley, A. Engel & W. Welte, (2004) Structure and assembly of the pseudopilin PulG. *Mol Microbiol* **54**: 647-664.

Korotkov, K. V., M. D. Gray, A. Kreger, S. Turley, M. Sandkvist & W. G. Hol, (2009a) Calcium is essential for the major pseudopilin in the type 2 secretion system. *The Journal of biological chemistry* **284**: 25466-25470.

Korotkov, K. V. & W. G. Hol, (2008) Structure of the GspK-GspI-GspJ complex from the enterotoxigenic Escherichia coli type 2 secretion system. *Nat Struct Mol Biol* **15**: 462-468.

Korotkov, K. V., B. Krumm, M. Bagdasarian & W. G. Hol, (2006) Structural and functional studies of EpsC, a crucial component of the type 2 secretion system from Vibrio cholerae. *J Mol Biol* **363**: 311-321.

Korotkov, K. V., E. Pardon, J. Steyaert & W. G. Hol, (2009b) Crystal structure of the N-terminal domain of the secretin GspD from ETEC determined with the assistance of a nanobody. *Structure* **17**: 255-265.

Krishna, S. S., I. Majumdar & N. V. Grishin, (2003) Structural classification of zinc fingers: survey and summary. *Nucleic Acids Res* **31**: 532-550.

LaPointe, C. F. & R. K. Taylor, (2000) The type 4 prepilin peptidases comprise a novel family of aspartic acid proteases. *The Journal of biological chemistry* **275**: 1502-1510.

Lee, H. M., J. R. Chen, H. L. Lee, W. M. Leu, L. Y. Chen & N. T. Hu, (2004) Functional dissection of the XpsN (GspC) protein of the Xanthomonas campestris pv. campestris type II secretion machinery. *Journal of bacteriology* **186**: 2946-2955.

Lo, M. C., A. Aulabaugh, G. Jin, R. Cowling, J. Bard, M. Malamas & G. Ellestad, (2004) Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Anal Biochem* **332**: 153-159.

Lybarger, S. R., T. L. Johnson, M. D. Gray, A. E. Sikora & M. Sandkvist, (2009) Docking and assembly of the type II secretion complex of Vibrio cholerae. *Journal of bacteriology* **191**: 3149-3161.

Martin, A., T. A. Baker & R. T. Sauer, (2005) Rebuilt AAA + motors reveal operating principles for ATP-fuelled machines. *Nature* **437**: 1115-1120.

Matson, J. S., J. H. Withey & V. J. DiRita, (2007) Regulatory networks controlling Vibrio cholerae virulence gene expression. *Infect Immun* **75**: 5542-5549.

Misic, A. M., K. A. Satyshur & K. T. Forest, (2010) P. aeruginosa PilT structures with and without nucleotide reveal a dynamic type IV pilus retraction motor. *J Mol Biol* **400**: 1011-1021.

Mooi, F. R. & E. M. Bik, (1997) The evolution of epidemic Vibrio cholerae strains. *Trends in microbiology* **5**: 161-165.

Nouwen, N., H. Stahlberg, A. P. Pugsley & A. Engel, (2000) Domain structure of secretin PulD revealed by limited proteolysis and electron microscopy. *EMBO J* **19**: 2229-2236.

Nunn, D., (1999) Bacterial type II protein export and pilus biogenesis: more than just homologies? *Trends in cell biology* **9**: 402-408.

Patrick, M., K. V. Korotkov, W. G. Hol & M. Sandkvist, (2011) Oligomerization of EpsE coordinates residues from multiple subunits to facilitate ATPase activity. *The Journal of biological chemistry*.

Peabody, C. R., Y. J. Chung, M. R. Yen, D. Vidal-Ingigliardi, A. P. Pugsley & M. H. Saier, Jr., (2003) Type II protein secretion and its relationship to bacterial type IV pili and archaeal flagella. *Microbiology* **149**: 3051-3072.

Pelicic, V., (2008) Type IV pili: e pluribus unum? *Mol Microbiol* **68**: 827-837.

Planet, P. J., S. C. Kachlany, R. DeSalle & D. H. Figurski, (2001) Phylogeny of genes for secretion NTPases: identification of the widespread tadA subfamily and development of a diagnostic key for gene classification. *Proceedings of the National Academy of Sciences of the United States of America* **98**: 2503-2508.

Possot, O. & A. P. Pugsley, (1994) Molecular characterization of PulE, a protein required for pullulanase secretion. *Mol Microbiol* **12**: 287-299.

Possot, O. M., M. Gerard-Vincent & A. P. Pugsley, (1999) Membrane association and multimerization of secreton component pulC. *Journal of bacteriology* **181**: 4004-4011.

Possot, O. M. & A. P. Pugsley, (1997) The conserved tetracysteine motif in the general secretory pathway component PulE is required for efficient pullulanase secretion. *Gene* **192**: 45-50.

Possot, O. M., G. Vignon, N. Bomchil, F. Ebel & A. P. Pugsley, (2000) Multiple interactions between pullulanase secreton components involved in stabilization and cytoplasmic membrane association of PulE. *Journal of bacteriology* **182**: 2142-2152.

Py, B., L. Loiseau & F. Barras, (1999) Assembly of the type II secretion machinery of Erwinia chrysanthemi: direct interaction and associated conformational change between OutE, the putative ATP-binding component and the membrane protein OutL. *J Mol Biol* **289**: 659-670.

Py, B., L. Loiseau & F. Barras, (2001) An inner membrane platform in the type II secretion machinery of Gram-negative bacteria. *EMBO Rep* **2**: 244-248.

Robert, V., A. Filloux & G. P. Michel, (2005) Subcomplexes from the Xcp secretion system of Pseudomonas aeruginosa. *FEMS Microbiol Lett* **252**: 43-50.

Robien, M. A., B. E. Krumm, M. Sandkvist & W. G. Hol, (2003) Crystal structure of the extracellular protein secretion NTPase EpsE of Vibrio cholerae. *J Mol Biol* **333**: 657-674.

Rose, I., G. Biukovic, P. Aderhold, V. Muller, G. Gruber & B. Averhoff, (2011) Identification and characterization of a unique, zinc-containing transport ATPase essential for natural transformation in Thermus thermophilus HB27. *Extremophiles*.

Sandkvist, M., (2001a) Biology of type II secretion. *Mol Microbiol* **40**: 271-283.

Sandkvist, M., (2001b) Type II secretion and pathogenesis. *Infect Immun* **69**: 3523-3535.

Sandkvist, M., M. Bagdasarian & S. P. Howard, (2000) Characterization of the multimeric Eps complex required for cholera toxin secretion. *Int J Med Microbiol* **290**: 345-350.

Sandkvist, M., M. Bagdasarian, S. P. Howard & V. J. DiRita, (1995) Interaction between the autokinase EpsE and EpsL in the cytoplasmic membrane is

required for extracellular secretion in Vibrio cholerae. *Embo J* **14**: 1664-1673.

Sandkvist, M., L. P. Hough, M. M. Bagdasarian & M. Bagdasarian, (1999) Direct interaction of the EpsL and EpsM proteins of the general secretion apparatus in Vibrio cholerae. *Journal of bacteriology* **181**: 3129-3135.

Sandkvist, M., L. O. Michel, L. P. Hough, V. M. Morales, M. Bagdasarian, M. Koomey, V. J. DiRita & M. Bagdasarian, (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., V. Morales & M. Bagdasarian, (1993) A protein required for secretion of cholera toxin through the outer membrane of Vibrio cholerae. *Gene* **123**: 81-86.

Satyshur, K. A., G. A. Worzalla, L. S. Meyer, E. K. Heiniger, K. G. Aukema, A. M. Misic & K. T. Forest, (2007) Crystal structures of the pilus retraction motor PilT suggest large domain movements and subunit cooperation drive motility. *Structure* **15**: 363-376.

Savvides, S. N., H. J. Yeo, M. R. Beck, F. Blaesing, R. Lurz, E. Lanka, R. Buhrdorf, W. Fischer, R. Haas & G. Waksman, (2003) VirB11 ATPases are dynamic hexameric assemblies: new insights into bacterial type IV secretion. *Embo J* **22**: 1969-1980.

Sawaya, M. R., S. Guo, S. Tabor, C. C. Richardson & T. Ellenberger, (1999) Crystal structure of the helicase domain from the replicative helicase-primase of bacteriophage T7. *Cell* **99**: 167-177.

Schmidt, H., B. Henkel & H. Karch, (1997) A gene cluster closely related to type II secretion pathway operons of gram-negative bacteria is located on the large plasmid of enterohemorrhagic Escherichia coli O157 strains. *FEMS Microbiol Lett* **148**: 265-272.

Shevchik, V. E., J. Robert-Baudouy & G. Condemine, (1997) Specific interaction between OutD, an Erwinia chrysanthemi outer membrane protein of the general secretory pathway, and secreted proteins. *EMBO J* **16**: 3007-3016.

Shiue, S. J., I. L. Chien, N. L. Chan, W. M. Leu & N. T. Hu, (2007) Mutation of a key residue in the type II secretion system ATPase uncouples ATP hydrolysis from protein translocation. *Mol Microbiol* **65**: 401-412.

Shiue, S. J., K. M. Kao, W. M. Leu, L. Y. Chen, N. L. Chan & N. T. Hu, (2006) XpsE oligomerization triggered by ATP binding, not hydrolysis, leads to its association with XpsL. *Embo J* **25**: 1426-1435.

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

Story, R. M., I. T. Weber & T. A. Steitz, (1992) The structure of the E. coli recA protein monomer and polymer. *Nature* **355**: 318-325.

Tamplin, M. L., A. L. Gauzens, A. Huq, D. A. Sack & R. R. Colwell, (1990) Attachment of Vibrio cholerae serogroup O1 to zooplankton and phytoplankton of Bangladesh waters. *Appl Environ Microbiol* **56**: 1977-1980.

Thibault, G., Y. Tsitrin, T. Davidson, A. Gribun & W. A. Houry, (2006) Large nucleotide-dependent movement of the N-terminal domain of the ClpX chaperone. *EMBO J* **25**: 3367-3376.

Thomsen, N. D. & J. M. Berger, (2008) Structural frameworks for considering microbial protein- and nucleic acid-dependent motor ATPases. *Mol Microbiol* **69**: 1071-1090.

Tsai, R. T., W. M. Leu, L. Y. Chen & N. T. Hu, (2002) A reversibly dissociable ternary complex formed by XpsL, XpsM and XpsN of the Xanthomonas campestris pv. campestris type II secretion apparatus. *Biochem J* **367**: 865-871.

Vezzulli, L., C. Pruzzo, A. Huq & R. R. Colwell, (2010) Environmental reservoirs in Vibrio cholerae and their role in cholera. *Environmental Microbiology Reports* **2**: 27-33.

Viarre, V., E. Cascales, G. Ball, G. P. Michel, A. Filloux & R. Voulhoux, (2009) HxcQ liposecretin is self-piloted to the outer membrane by its N-terminal lipid anchor. *The Journal of biological chemistry* **284**: 33815-33823.

Yamagata, A. & J. A. Tainer, (2007) Hexameric structures of the archaeal secretion ATPase GspE and implications for a universal secretion mechanism. *Embo J* **26**: 878-890.

Ye, J., A. R. Osborne, M. Groll & T. A. Rapoport, (2004) RecA-like motor ATPases--lessons from structures. *Biochimica et biophysica acta* **1659**: 1-18.

Yeo, H. J., S. N. Savvides, A. B. Herr, E. Lanka & G. Waksman, (2000) Crystal structure of the hexameric traffic ATPase of the Helicobacter pylori type IV secretion system. *Mol Cell* **6**: 1461-1472.

Zhou, J. & Z. Xu, (2003) Structural determinants of SecB recognition by SecA in bacterial protein translocation. *Nat Struct Biol* **10**: 942-947.