Investigation of factors required for assembly of the GspD secretin in enterotoxigenic *Escherichia coli* and *Vibrio* species

A thesis submitted to the College of Graduate studies
and research in partial fulfillment of the requirements for the
Degree of Doctor of Philosophy in the Department of Microbiology and Immunology
University of Saskatchewan
Saskatoon, SK, Canada

By Timothy Strozen

PERMISSION TO USE

In presenting this thesis/dissertation in partial fulfillment of the requirements for the Degree of Doctor of Philosophy from the University of Saskatchewan, I agree that the libraries of this university may make it freely available for inspection. I further agree that permission for copying of this thesis/dissertation in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis/dissertation work or, in their absence, by the Head of the department or the Dean of the college in which my thesis work was done. It is understood that any copying or publication or use of this thesis/dissertation or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis/dissertation.

Requests for permission to copy or to make other uses of materials in this thesis/dissertation in whole or part should be addressed to:

Head of the Department of Microbiology and Immunology
University of Saskatchewan
Saskatoon, Saskatchewan S7N 5E5
Canada

OR

Dean
College of Graduate Studies and Research
University of Saskatchewan
107 Administration Place
Saskatoon, Saskatchewan S7N 5E5
Canada

ABSTRACT

The type II secretion system (T2SS) is a major virulence factor in Gram-negative bacteria due to the involvement of this system in the pathogenesis of numerous species. This specialized macromolecular apparatus spans the cell envelope and functions to translocate folded proteins located in the periplasm across the outer membrane. In the pathogens enterotoxigenic *Escherichia coli* (ETEC) and *Vibrio cholerae*, the T2SS is integral for infection of the host because it serves as the conduit for secretion of the heat-labile enterotoxin (LT) and the cholera toxin respectively.

The objective of this study was to assess the assembly and function of the T2SSs in ETEC and *Vibrio cholerae* by genetic, physiological and molecular biological techniques. *In silico* analysis of the ETEC H10407 genome identified two T2SSs designated alpha and beta. The $T2SS_{\alpha}$ is not assembled in ETEC under standard laboratory conditions. Replacement of the cryptic endogenous $gspA_{\alpha}$ and $gspC_{\alpha}$ promoters with inducible ones resulted in assembly of the $T2SS_{\alpha}$, however the system remained incapable of secreting LT. Under laboratory conditions the $GspD_{\beta}$ secretin of the $T2SS_{\beta}$ was readily detectable and expression of a functional $T2SS_{\beta}$ was required for secretion of LT into culture supernatant.

The requirement for T2SS accessory proteins GspA, GspB, GspS and LeoA in $GspD_{\beta}$ secretin assembly was investigated. The hypothetical lipoprotein YghG (renamed $GspS_{\beta}$) encoded in the $T2SS_{\beta}$ operon of ETEC was characterized as a pilotin protein required for localization of the secretin protein $GspD_{\beta}$ to the outer membrane, since in the absence of the lipoprotein $GspD_{\beta}$ was localized to the inner membrane in monomeric form and mostly degraded. The hypothetical virulence factor LeoA was shown to be not required for $GspD_{\beta}$ secretin assembly or LT secretion. In *Vibrio cholerae*, the peptidoglycan-remodeling complex GspAB was shown to be partially required for $GspD^{Vc}$ secretin assembly, suggesting the possibility that another protein, possibly the ETEC YghG homologue, is also required for $GspD^{Vc}$ secretin assembly in *Vibrio* species.

Lastly, a novel application of the bacterial two hybrid technique was developed that enabled screening for interactions with a protein of interest with proteins encoded by the entire coding capacity of the ETEC H10407 genome. In particular, a library composed of 6.2×10^5 plasmids comprised of overlapping ETEC genomic fragments was encoded in the bacterial two-hybrid pTRG vector and was screened for interactions with members of the ETEC T2SS.

ACKNOWLEDGEMENTS

I would like to thank my supervisor Dr. Peter Howard for his guidance and mentorship in my work. I would also like to thank Mary Woodsworth, Li Gang, and Qibing Wang for their technical help in the laboratory.

I gratefully acknowledge my wife Rachael for her infinite amount of support and patience in waiting for me to complete this study. In addition I would like to thank my family and friends for their words of encouragement through this process.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

.

LIST OF ABBREVIATIONS

AA	Amino acid
AIEC	Adherent-invasive *Escherichia coli*
Amp	Ampicillin
APEC	Avian pathogenic *Escherichia coli*
B2H	Bacterial II hybrid
EAEC	Enteroaggregative *Escherichia coli*
EHEC	Enterohemorrhagic *Escherichia coli*
ELISA	Enzyme-linked immunosorbent assay
EPEC	Enteropathogenic *Escherichia coli*
ETEC	Enterotoxigenic *Escherichia coli*
ExPEC	Extra-intestinal pathogenic *Escherichia coli*
PCR	Polymerase chain reaction
Cam	Chloramphenicol
CAYE	Casamino acids-yeast extract
CBB	Coomassie Brilliant Blue
D	Dalton
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
GSP	General secretion pathway
HNS	Histone-like nucleoid structuring protein
IPTG	Isopropyl β-D-1-thiogalactopyranoside
kb	Kilobase pairs
Kan	Kanamycin
LB	Luria-Bertani
LPS	Lipopolysaccharide
LT	Heat-labile enterotoxin
LTa	Heat-labile enterotoxin subunit A
LTb	Heat-labile enterotoxin subunit B
PVDF	Polyvinylidene fluoride

1. INTRODUCTION

1.1 Gram-negative gastrointestinal pathogens

The accessibility of the gastrointestinal (GI) tract has made it an attractive niche for bacteria as evidenced by the wide variety of species that comprise the normal flora of the human intestine. It is estimated that the human intestine contains 10-100 trillion bacteria belonging to 15,000 to 36,000 species (Frank *et. al*., 2007; Xu *et. al*., 2007). Species of the normal flora have a symbiotic relationship with the host whereby the host provides a stable environment and a constant supply of nutrients, and the flora prevent colonization by pathogens (Hudault *et. al.*, 2001; Schamberger *et. al*., 2004). Alteration of intestinal homeostasis between the host and flora can lead to the development of chronic intestinal inflammatory disorders including Crohn's disease, inflammatory bowel disease and ulcerative colitis thereby highlighting the importance of the host-flora interaction (Mizoguchi *et. al*., 2003; Sadlack *et. al*., 1993; Rath *et. al*., 1996). In many instances, homeostasis in the intestine is changed by the introduction of an intestinal pathogen. Intestinal pathogens have evolved and acquired specific mechanisms that enable these species to colonize and dominate the environment, in the process causing disease. Gram-negative pathogens utilize an arsenal of virulence factors to give them a competitive advantage over the normal flora of the intestine. Specialized appendages for adherence, toxin secretion, toxin injection among others are utilized by pathogens to proliferate in the intestinal environment and rid itself of its competitors.

In several Gram-negative species, the type II secretion system (T2SS) is utilized by the pathogen to secrete a diarrheagenic toxin into the lumen of the host intestine. It is hypothesized that the competitive advantage of toxin secretion is that by causing diarrhea, the pathogen is able to flush out its competitors from the intestine from which it remains attached. Therefore the T2SS is considered a major virulence factor of several Gram-negative species including the enterotoxigenic *Escherichia coli* that secretes the heat-labile enterotoxin, *Vibrio cholerae* that secretes the cholera toxin and *Aeromonas hydrophila* that secretes the aerolysin toxin.

1.1.1 Enterotoxigenic *Escherichia coli*

Escherichia coli (*E. coli*) serve as integral members of the gut microbiota that are therefore important in maintenance of healthy gastrointestinal tract function (Tenaillon *et. al.*, 2010). However, some strains have evolved or acquired genetic factors that enable the

bacterium to disseminate into a variety of niches in the host, in the process causing disease. From urinary tract infections caused by uropathogenic *E. coli* to colonization of the lower intestine by enterotoxigenic *E. coli* and systemic infections generated by enteroinvasive *E coli*, members of this species have developed their own niches when colonizing the host.

The enterotoxigenic *E. coli* (ETEC) are one of six diarrheagenic *E. coli* pathotypes (Black, 1993), capable of colonizing the small intestine of humans and animals. In humans, ETEC is most common cause of traveller's diarrhea (Coster, *et. al*., 2007) and in areas of the world where ETEC is endemic, is estimated to cause 200 million diarrheal episodes and approximately 400,000 deaths of children under the age of five annually (Wenneras and Erling, 2004; Qadri *et. al.*, 2005). In endemic areas, ETEC is found mainly in surface waters (Ohno *et al.*, 1997), leading to contamination of food and water sources. In animals, ETEC is a major cause of secretory diarrhea in pigs (Berberov *et. al.*, 2004) therefore posing a significant threat to the farming industry.

ETEC is transmitted via a fecal-oral route and cause an infection of the epithelial layer of the small intestine, thereby leading to the development of severe diarrhea and bacterial shedding. To colonize the small bowel, ETEC produce a wide variety of colonization factors that facilitate adherence of the bacterium to the intestinal mucosa (reviewed by Qadri *et. al.*, 2005) and a variety of toxins that when secreted by ETEC, are internalized by the intestinal epithelial cells thereby causing diarrhea. In human ETEC, 25 colonization factors including adhesins and fimbriae have been characterized. Those most commonly expressed in diarrheagenic strains include CFA/1, CS1, CS2, CS3, CS4, CS5, CS6, CS7, CS14, CS17 and CS21 (Reviewed by Walker *et. al*., 2007) of which several have been the target of vaccines (Svennerholm *et al*., 1988; Tobias *et al*., 2010).

Human ETEC also elicit a variety of toxins that when secreted from the cell generate a variety of pathogenic effects upon the intestinal epithelial cells of the host. Of these toxins are two enterotoxins known as the heat-labile enterotoxin (LT) and the heat-stable enterotoxin (ST). Human ETEC strains will produce one or both of these enterotoxins, with the prevalence of strains expressing one or both toxins varying around the world. For example, of 1,220 children in India with diarrhea symptoms caused by ETEC, 27% of ETEC strains were positive for both LT and ST, 41% were positive for LT only and 32% were positive for ST only (Qadri *et. al*., 2005).

ETEC can secrete two structurally and functionally distinct heat-stable enterotoxins, STa and STb (Kennedy *et. al*., 1984) of which only STa performs a role in disease in humans. ST is synthesized as a pre-pro-protein of 72 amino acids that during export is processed to generate an active toxin of 18 to 19 amino acids (2 kDa). The toxin is transported across the inner membrane by the Sec system, three disulfide bonds are generated by disulfide bond isomerase (DsbA), the pro-region is removed and the pre-protein is exported through the TolC outer membrane transporter (Tamanaka *et. al*., 1998). Extracellular STa binds and stimulates the guanylate cyclase receptor (GC-C) of the intestinal epithelium, resulting in intracellular accumulation of cGMP, phosphorylation of CFTR and secretion of chloride ions and water, thereby causing diarrhea (Forte *et. al.*, 1992; reviewed by Turner *et. al*., 2006).

LT is a member of the AB_5 family of enterotoxins, composed of five 11.6 kDa B subunits (LTb) and one 28 kDa A subunit (LTa) that are assembled into an 84 kDa protein (Reviewed by Spangler, 1992). The holotoxin is secreted by the T2SS across the outer membrane and binds to the GM1-ganglioside receptor on the surface of the intestinal epithelium. Once internalized, LTa activates adenylate cyclase by ribosylating the stimulatory G protein. This activation leads to an increased concentration of cellular cAMP that ultimately results in the loss of water by the cell, thereby causing diarrhea (Reviewed by Nataro and Kaper, 1998). The ETEC pathotype, characterized by strain H10407 that was isolated from a patient with severe cholera-like disease in Bangladesh (Evans *et. al*., 1975), contains two chromosomally encoded type II secretion systems (T2SS) designated as T2SSα and T2SSβ, both of which have been implicated as responsible for secretion of the heat labile enterotoxin (LT) (Horstman and Kuehn, 2002; Tauschek *et. al.*, 2002).

1.1.2 *Vibrio cholerae* and other pathogenic *Vibrio* species

Vibrio cholerae (*V. cholerae*) is a non-invasive Gram-negative species with a single polar flagellum. It is classified in the γ subdivision of the *Proteobacteriaceae* family and is comprised of more than 200 serogroups (Chatterjee and Chaudhuri, 2003) of which only two strains O1 and O139 cause the disease cholera. Annually 100,000 cholera cases resulting in 20000-30000 deaths are reported to the WHO with an estimated 90% of cases that are unreported (Sanchez and Holmgren, 2005). The natural resevoir for *V. cholerae* are bodies of water including brackish water and estuarine environments whereby it has been found in

association with various aquatic species including; cyanobacteria (Islam *et al.*, 1989), diatoms (Martin and Bianchi, 1980) and blue crab (Huq *et al*., 1986), among others. The disease is rare in industrialized nations but remains a significant threat mostly to young children in developing nations (Deen *et. al*., 2008). Seven cholera pandemics have spread throughout the world since 1817, the latest being caused of the El Tor biotype (Sack *et. al*., 2004; Faruque *et. al*., 1998) which contains the CTX (cholera toxin) prophage that encodes the major virulence factor of *V. cholerae*, the cholera toxin.

The *ctxAB* genes encoding the cholera toxin are encoded within the genome of a filamentous bacteriophage CTXφ that lysogenizes *V. cholerae* at specific integration sites thereby converting non-pathogenic strains to toxigenic ones (Waldor and Mekalanos, 1996). Following ingestion by the host, *Vibrio* travel to the small intestine where they form micro-colonies. Construction of micro-colonies is mediated by lateral interactions between long fibers of the toxin-coregulated pili (TCP) (Taylor *et. al.*, 1987; Kirn *et al*., 2000). This bundle-forming pilus is essential for intestinal colonization, serves as the receptor for CTXφ attachment and is expressed concurrently with *ctxAB* (Taylor *et. al*., 1987; Waldor and Mekalanos, 1996)

The cholera toxin (CT), like LT, is a member of the AB_5-type toxins composed of a pentameric ring of B subunits and an active A subunit (Zhang *et. al.*, 1995). Similar to LT, the B-subunit binds to the GM1-ganglioside of the intestinal epithelium whereby the A subunit is internalized, ultimately resulting in development of diarrhea (reviewed by De Haan and Hirst, 2004). Interestingly, both CTXφ and CT are actively secreted from the cell by the T2SS (Sandkvist *et. al*., 1997; Davis *et. al*., 2000)

Vibrio species *V. vulnificus* and *V. parahaemolyticus* are invasive pathogenic species that are capable of generating a systemic infection. Similar to *V. cholerae*, these species are also waterborne pathogens whose main routes of transmission into humans occur by ingestion of raw contaminated seafood and contamination of wounds. *V. vulnificus* is the more dangerous pathogen of the two species since a hallmark of a septicemic *V. vulnificus* infection is rapid progression that results in mortality rates as high as 75% in as little as 24 hours post-infection (Hlady and Klontz, 1996; Strom and Paranjpye, 2000). The main virulence factors of *V. vulnificus* include the T2SS secreted proteins hemolysin/cytolysin (Kreger and Lockwood, 1981) and metalloprotease (Smith and Merkel, 1982) (Paranjpye *et. al.*, 1998), and the RtxA toxin (Rhee *et. al*., 2001). In *V. parahaemolyticus*, the main virulence factors are a thermostable

hemolysin (TDH) and a thermostable-related hemolysin (TRH) (Honda and Iida, 1993; Shirai *et. al*., 1990).

1.2 The Type II secretion system

The type II secretion system (T2SS), previously known as the main terminal branch of the general secretory pathway (GSP), is a large complex composed of 12-16 proteins that span both the inner and outer membranes of Gram-negative bacteria and is utilized by many Gram-negative bacteria to translocate folded proteins across the outer membrane from the periplasm to the extracellular milieu.

T2SS substrates reach the periplasm by protein translocation across the inner membrane via the secretion (Sec) or twin-arginine translocation (Tat) pathways. In the Sec pathway, pre-proteins that contain an N-terminal signal peptide are processed by the Sec machinery whereby the signal peptide is cleaved and proteins are translocated into the periplasm (Pugsley *et. al*., 1991) through successive rounds of binding between SecA and the preprotein followed by ATPase hydrolysis that is used to translocate the preprotein in a stepwise fashion through the export pore composed of the heterotrimeric complex SecYEG (Tomkiewicz *et. al.*, 2008). In the Tat pathway, folded substrates complexed with their chaperones bind to the inner membrane receptor complex composed of TatB and TatC (Bolhuis *et. al*., 2001) via an N-terminal signal peptide (Cline *et. al*., 2001; Alami *et. al*., 2003). This interaction initiates formation of an inner membrane complex composed only of TatA that forms the active translocon (Cline and Theg, 2007; Dabney-Smith *et. al.*, 2006) thereby leading to substrate translocation (Yahr and Wickner, 2001).

The T2SS has been identified in many Gram-negative species and consequently genes and proteins of these systems have been given their own designations as a result. For instance the T2SS of *Aeromonas hydrophila* is referred to as the extracellular secretion system (Exe) and proteins are designated as ExeA, ExeB etc. For the purpose of clarity, all proteins and genes will be referred to in the general secretory pathway designation used for the T2SS in *E. coli* strains and the abbreviation of the species name will be denoted in superscript, for example, ExeA will be referred to as $GspA^{Ah}$. In the absence of superscript, the protein or gene is designated as from *E. coli*.

1.2.1 Genetic organization of T2SS operons

In most species, the genes that encode the T2SS proteins are arranged in a major operon composed of genes *gspC,D,E,F,G,H,I,J,K,L,M,N* and *O*, and in some cases a minor operon composed of *gspA* and *gspB* or an independently encoded *gspS* (Fig. 1). The promoter for the operon is usually encoded upstream of *gspC* for the major operon and upstream of *gspA* for the minor operon. In the operons that encode the $T2SS_{\alpha}$ of ETEC, divergent promoters are located adjacent to one another that control expression of *gspAB* and *gspC-O* operons. In the operon that encodes the $T2SS_{\beta}$ of ETEC, the promoter is located upstream of *yghJ* instead of being encoded upstream of *gspC*.

1.2.2 Structure of the T2SS macromolecular complex

The T2SS is a multisubunit complex that spans both the inner and outer membranes in Gram-negative species. The number and identity of members of the T2SS vary among species as 12-16 proteins have been observed to comprise the system. The identity of T2SS proteins and their proposed functions in several Gram-negative species are provided in Table 1. As is evident from this table, not all species encode the same complement of T2SS proteins. For instance, the ETEC $T2SS_{\beta}$ does not encode GspAB or GspN homologues. Likewise, the ETEC $T2SS_{\alpha}$ does not encode a GspS or GspN homologue but does encode a GspAB homologue.

Subcellular localization studies of the T2SS components suggest that the system is composed of four parts: an inner membrane platform, a periplasmic pseudopilus, an outer membrane complex and the cytoplasmic secretion ATPase (Fig. 2). The inner membrane platform complex is composed of proteins GspC, F, L and M and forms the core of the system from which interactions with the pseudopilus, the secretin and the cytoplasmic ATPase are generated. Proteins GspC, L and M each have one transmembrane domain and form a complex in the inner membrane that is required for protection from degradation (Michel *et. al*., 1998; Sandkvist *et. al*., 1999; Robert *et. al*., 2005; Lybarger *et. al*., 2009). In *Erwinia chrsyanthemi* and *Pseudomonas aeruginosa*, inner membrane complexes composed of GspE, GspF, GspL and GspM have been identified directly by co-purification and yeast two-hybrid analysis (Py *et. al*., 2001; Robert *et. al.*, 2005). GspF contains three transmembrane helicies and has two cytoplasmic domains with ~28% sequence identity (Abendroth *et. al*., 2009). GspL contains a cytosolic domain, a transmembrane helix and a periplasmic domain and functions in recruiting

FIG. 1. **Genetic organization of the T2SS operons in *V. cholerae, A. hydrophila, K. oxytoca* and *E. chrysanthemi* species.** Note that the alpha and beta T2SS operons in *E. coli* are a new designation proposed in this study. Homologous genes are designated by colour, promoters are identified by arrows.

Table 1. Components of the T2SS of various species

Function	ETEC (alpha)	ETEC (Beta)	*Vibrio cholerae*	*Aeromonas hydrophila*	*Klebsiella oxytoca*	*Erwinia chrysanthemi*
Peptidoglycan binding	$GspA_{\alpha}$		$GspA^{EpsA}$	$GspA^{ExeA}$		
Peptidoglycan binding	$GspB_{\alpha}$		$GspB^{EpsB}$	$GspB^{ExeB}$	$GspB^{PulB}$	$GspB^{OutB}$
IM platform, secetin interaction	$GspC_{\alpha}$	$GspC_{\beta}$	$GspC^{EpsC}$	$GspC^{ExeC}$	$GspC^{PulC}$	$GspC^{OutC}$
OM secretin	$GspD_{\alpha}$	$GspD_{\beta}$	$GspD^{EpsD}$	$GspD^{ExeD}$	$GspD^{PulD}$	$GspD^{OutD}$
ATPase	$GspE_{\alpha}$	$GspE_{\beta}$	$GspE^{EpsE}$	$GspE^{ExeE}$	$GspE^{PulE}$	$GspE^{OutE}$
IM platform	$GspF_{\alpha}$	$GspF_{\beta}$	$GspF^{EpsF}$	$GspF^{ExeF}$	$GspF^{PulF}$	$GspF^{OutF}$
Major pseudopilin	$GspG_{\alpha}$	$GspG_{\beta}$	$GspG^{EpsG}$	$GspG^{ExeG}$	$GspG^{PulG}$	$GspG^{OutG}$
Minor pseudopilin	$GspH_{\alpha}$	$GspH_{\beta}$	$GspH^{EpsH}$	$GspH^{ExeH}$	$GspH^{PulH}$	$GspH^{OutH}$
Minor pseudopilin	$GspI_{\alpha}$	$GspI_{\beta}$	$GspI^{EpsI}$	$GspI^{ExeI}$	$GspI^{PulI}$	$GspI^{OutI}$
Minor pseudopilin	$GspJ_{\alpha}$	$GspJ_{\beta}$	$GspJ^{EpsJ}$	$GspJ^{ExeJ}$	$GspJ^{PulJ}$	$GspJ^{OutJ}$
Minor pseudopilin	$GspK_{\alpha}$	$GspK_{\beta}$	$GspK^{EpsK}$	$GspK^{ExeK}$	$GspK^{PulK}$	$GspK^{OutK}$
IM platform	$GspL_{\alpha}$	$GspL_{\beta}$	$GspL^{EpsL}$	$GspL^{ExeL}$	$GspL^{PulL}$	$GspL^{OutL}$
IM platform	$GspM_{\alpha}$	$GspM_{\beta}$	$GspM^{EpsM}$	$GspM^{ExeM}$	$GspM^{PulM}$	$GspM^{OutM}$
Unknown			$GspN^{EpsN}$	$GspN^{ExeN}$	$GspN^{PulN}$	$GspN^{OutN}$
Prepilin peptidase	$GspO_{\alpha}$	PppA	VcpD	TapD	$GspO^{PulO}$	$GspO^{OutO}$
Pilotin		YghG ($GspS_{\beta}$)	YghG ($GspS^{EpsS}$)		$GspS^{PulS}$	$GspS^{OutS}$

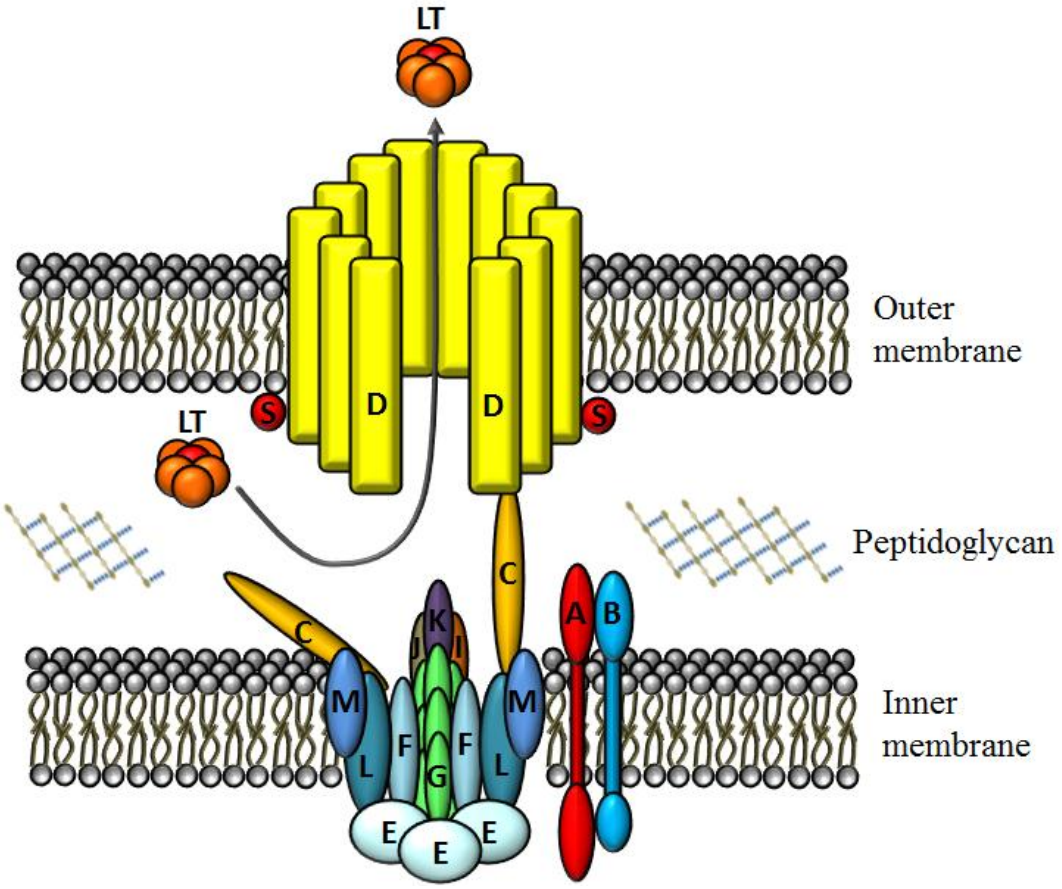


FIG. 2. **Schematic diagram of the T2SS.** The macromolecular T2SS spans the inner and outer membranes of the cell envelope. Subcellular localization studies of the core T2SS components suggest that the system is composed of four parts: an inner membrane platform (GspC, F, L, M), a periplasmic pseudopilus (GspG, H, I, J, K) , an outer membrane complex (GspD) and the cytoplasmic secretion ATPase (GspE). Accessory proteins that are not highly conserved in T2SSs include the GspAB complex as shown in the inner membrane, the GspS pilotin that may remain associated with the secretin following secretin assembly and GspN whose function is not known (not shown). Note that the prepilin peptidase (GspO/PppA/TapD) is also not shown.

the cytoplasmic ATPase GspE to the inner membrane (Sandkvist *et. al.*, 1995). A highly conserved interface that exists between GspE and GspL suggests that conformational changes generated by GspE upon ATP hydrolysis might be translated to the rest of the system through GspL (Sandkvist *et. al.*, 1995; Patrick *et. al*., 2011).

GspC consists of a small cytoplasmic domain, a transmembrane helix and two periplasmic domains designated as the homology region (HR) domain and the PDZ (post synaptic/ Drosophila/ zonula occludens-1 protein) domain that in some species is replaced by a coiled-coil domain. The function of the PDZ domain has not been clarified however deletion of this domain causes the *Erwinia chrysanthemi* T2SS to be selectively functional by allowing secretion of three exoproteins but not enabling the secretion of several pectate lyases (Bouley *et. al*., 2001). As a result, the PDZ domain of GspC may regulate T2SS substrate specificity. The HR domain of GspC is involved in mediating an interaction with the N0 domain of GspD (Korotkov *et. al.*, 2011B). In this manner, GspC acts as a tether between the outer membrane complex composed of GspD and the inner membrane complex (Gerard-Vincent *et. al*., 2002; Lee *et. al.*, 2004; Possot *et. al.*, 1999). The stoichiometry of proteins that comprise the inner membrane complex is not known but has been suggested to contain an equal number of GspL, GspM and GspC proteins. Since GspE and GspL interact with a ratio of 1:1 (Abendroth *et. al*., 2005) and GspE forms a hexamer as described below, an inner membrane complex comprised of six subunits of GspL, GspM, GspC and GspE may exist with an unknown number of GspF subunits (Korotkov *et. al.*, 2011A). However, the 1:1 complex observed in crystal structures composed of GspC and GspD together with the electron microscopy evidence that GspD forms a dodecamer suggest that GspC may also form a dodecamer (Chami *et. al.*, 2005; Reichow *et. al.*, 2010). In a study by Korotkov, 2011A, it was suggested that due to space constraints, every second subunit in the dodecameric ring of GspD may form an interaction with the HR domain of GspC, thereby suggesting a hexameric ring of GspC proteins exist which that would theoretically be consistent with the stoichiometry of the inner membrane complex.

GspE is a Zn-containing Type II/IV secretion ATPase that likely forms hexamers at the interface with the inner membrane (Camberg *et. al.*, 2007; Patrick *et. al*., 2011) since similar ATPase proteins HP0525 of *Helicobacter pylori* and PilT of the type IV pilus have been shown to form hexamers by crystallographic studies (Yeo *et. al.*, 2000; Forest *et. al.*, 2004). The function of the ATPase is integral to the function of the T2SS since in its absence the T2SS is

not functional (Possot *et. al*., 2000). The mechanism in which GspE powers the T2SS is unknown, it may couple energy derived from ATP hydrolysis to drive assembly/disassembly of the pseudopilus since GspE interacts with GspL which in turn interacts with the major pseudopilin GspG (Gray *et. al*., 2011). The function of the ATPase GspE is not the only source of energy required for T2SS function however, since in some species including the Aeromonads, energy derived from the proton motive force is also required (Letellier *et. al*., 1997).

The periplasmic pseudopilus is composed of five proteins GspG, H, I, J and GspK and is designated as a pseudopilus due to the shared sequence identity of this group of proteins with pilins of the Type IV pilus (Nunn and Lory, 1992 and 1993; Pugsley, 1993). The most abundant protein in the pseudopilus is GspG (the major pseudopilin) whereas the other pseudopilins are referred to as minor ones (Nunn and Lory, 1993; Pugsley, 1993b: Hu *et. al.*, 2002). All pseudopilins contain an N-terminal signal peptide that localizes the protein for translocation across the inner membrane by the Sec system through the signal recognition particle (SRP) pathway (Francetic *et. al.*, 2007; Arts *et. al*., 2007). Specialized prepilin peptidases such as GspO or in some species a T4P prepilin peptidase, function in cleavage of the signal peptide and methylation of the N-terminal amino acid, a function necessary for pilus assembly (Lapointe and Taylor, 2000). In addition to the signal peptide, pseudopilins also contain a long N-terminal α-helix, a variable region and a C-terminal semi-conserved β-sheet. The crystal structure of GspG identified the pseudopilus as right-handed, with the N-terminal α-helix located in the core of the structure (Kohler *et. al.*, 2004). A conserved glutamic acid residue at position 5 (with the exception of GspK) is hypothesized to be involved in an interaction with the N-terminal amino group of the next subunit in the structure (Campos *et. al.*, 2010; Craig *et. al.*, 2006; Korotkov and Hol, 2008). The tip of the pseudopilus is likely composed of a trimer of GspK, GspI and GspJ in a triangular arrangement with GspK located at the top because GspK contains a large C-terminal alpha-helical domain instead of a semi-conserved β-sheet structure required for addition of pseudopilins (Campos *et. al.*, 2010; Douzi *et. al.*, 2009; Korotkov and Hol, 2008). In this way, GspK in complex with GspI and GspJ may form the "arrow head" of the pseudopilus whereby the large C-terminal alpha-helical domain of GspK may interact with the outer membrane secretin or substrates of the T2SS (Forest, 2008). Interestingly, when $GspG^{Pa}$ was overexpressed in *P. aeruginosa*, a type II surface-exposed pseudopilus structure

was observed and could be assembled by its endogenous Xcp T2SS, the Hxc T2SS (another T2SS encoded in *P. aeruginosa*) and the type IV pilus assembly system (Durand *et. al.*, 2003). However, if GspK was overproduced together with GspG, shorter pseudopilin structures were observed (Durand *et. al.*, 2005). Lastly, the function of GspI may be in initiation of pseudopilin assembly since in the absence of GspI, assembly of surface-exposed pseudopili by overproduction of GspG was severely negatively affected (Sauvonnet *et. al.*, 2000).

The final complex that comprises an assembled T2SS resides in the outer membrane. The major protein in the outer membrane is a megadalton-sized complex termed the secretin that is composed of 12-14 subunits of the GspD protein (Chen *et. al.*, 1996; Hardie *et. al.*, 1996; Sandkvist 2001). This complex is a remarkably stable macromolecular structure that is both heat and detergent resistant and requires treatment with organic solvents such as phenol for dissociation (Hardie *et. al.*, 1996; Ast *et. al.*, 2002). The T2SS secretin is a member of a family of outer membrane secretin transporters that are present in other cell envelope systems including the type III secretion system (T3SS), toxin co-regulated pili, type IV pilus, type IV bundle-forming pili, and filamentous phage (Burkhardt *et. al.*, 2011; Collins *et. al.*, 2003; Jain *et. al.*, 2011; Korotkov *et. al.*, 2011; Opalka *et. al.*, 2003; Schraidt and Marlovitz, 2011). In each system, the secretin functions as the outer membrane pore through which proteins or macromolecular complexes are translocated.

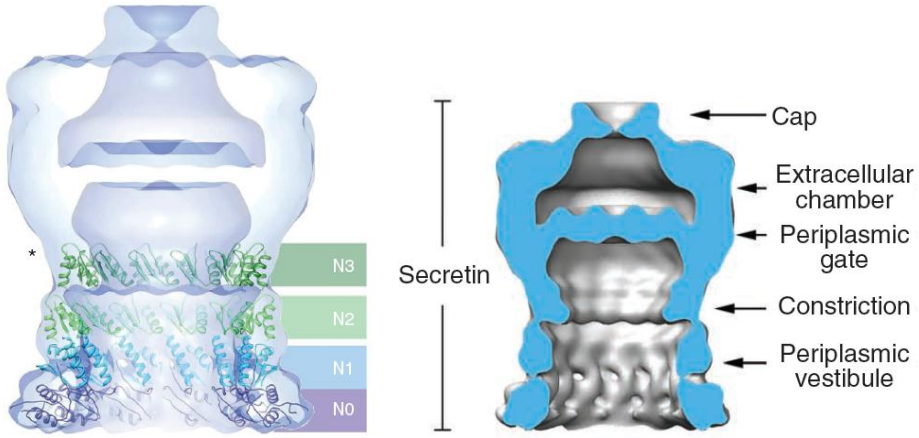
GspD proteins contain several domains that include; the Sec-dependent N-terminal signal peptide that targets the protein to the periplasm, followed by four N-terminal domains (N0, N1, N2 and N3), a conserved C-terminal domain and an S-domain recognized by the pilotin (Reviewed by Korotkov *et. al.*, 2011). The N0-N1 domain of GspD is thought to interact with the homology region (HR) of GspC as stated above (Korotkov *et. al.*, 2011). The conserved C-terminal domain forms the β-barrel in the outer membrane that creates the pore through which substrates or macromolecular complexes are passed through (Chami *et. al.*, 2005; Guilvout *et. al.*, 1999). Cryo-electron microscopy of the *V. cholerae* $GspD^{Vc}$ secretin revealed a cylindrical structure with 12-fold symmetry with a diameter of 155 Å and length of 200 Å (Reichow *et. al.*, 2010) (Fig. 3). The outer membrane surface is relatively smooth whereas the periplasmic domain exhibits three concentric rings. A cross-section of the secretin identified a large periplasmic vestibule composed of the N0, N1 and N2 domains with an opening of 70 Å that narrows to 55 Å approximately two-thirds of the way into the pore. This

constriction site is composed of the N3 domain of GspD. There also exists a small extracellular chamber with an extracellular gate opening of only 10 Å (Fig. 3).

Assembly of the secretin and function of the T2SS require the activity of additional accessory proteins, the identity of which are species-dependent. In some systems, localization of the secretin in the outer membrane requires the function of a small lipoprotein that serves as a pilotin to direct the secretin to the outer membrane and protect the multimer from degradation. To date, two highly similar pilotin lipoproteins involved in localization and protection of a T2SS secretin have been elucidated, these include $GspS^{Ko}$ (PulS) of *Klebsiella oxytoca* (Hardie *et. al.*, 1996; Nouwen *et. al.*, 1999) and $GspS^{Ec}$ (OutS) of *Erwinia chrysanthemi* (Shevchik and Condemine, 1998). In other T2SSs, notably that of *Aeromonas* species *A. hydrophila* and *A. salmonicida*, the peptidoglycan-binding and ATPase functions of the inner membrane complex GspAB are absolutely required for secretin assembly (Ast *et. al.*, 2002; Howard *et. al.*, 2006; Jahagirdar and Howard, 1994; Li and Howard, 2010). In several *Vibrio* species *V. cholerae*, *V. vulnificus* and *V. parahaemolyticus*, the absence of the GspAB complex significantly decreases but does not eliminate the assembly of the secretin, indicating that other factors, possibly including an unidentified pilotin are involved in its assembly (Chapter 5; Strozen *et. al.*, 2011). In *Erwinia chrysanthemi* and *Klebsiella oxytoca* that contain GspB and GspS proteins, GspS is required for assembly and protection of the GspD secretin multimer from degradation (Hardie *et. al.*, 1996A; Shevchik and Condemine, 1998; Hardie *et. al.*, 1996B) and GspB is required for secretion of pectinases by *E. chrysanthemi* (Pugsley *et. al.*, 1990; Condemine and Shevchik, 2000). Lastly, the inner membrane GTPase LeoA (labile enterotoxin output A) (Brown and Hardwidge, 2007) could also be involved in assembly and/or function of the T2SS since it is required for efficient secretion of LT in H10407 (Fleckenstein *et. al.*, 2000). As a result, several accessory proteins and combinations of accessory proteins are involved in secretin assembly, the identity of which are species-dependent.

Once localized to the outer membrane, the T2SS secretin like the Type IV pilin secretin PilQ of *P. aeruginosa*, likely requires the function of the β-barrel assembly machinery (the BAM complex) for assembly of the multimeric secretin in the outer membrane (reviewed by Knowles *et. al.*, 2009; Voulhoux *et. al.*, 2003). Interestingly however, assembly of the *K. oxytoca* T2SS secretin upon expression of $GspD^{Ko}$ in *E. coli* does not require the activity of the

A. B.



C.

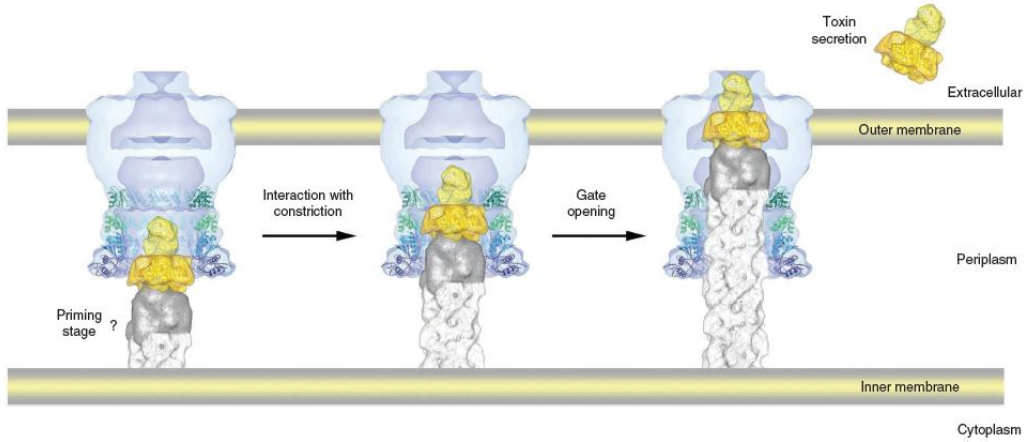


FIG. 3 Structure of the GspD^{Vc} secretin of *V. cholerae* proposed in a study by Reichow *et al.*, 2010. (A) A 12-member ring model of the GspD^{Vc} N-terminal periplasmic domains N0, N1, N2 and N3 were placed into a GspD^{Vc} density map. (B) The location of structures in the secretin protein. (C) A proposed piston-driven model of protein secretion. The figure at left is the secretin in its closed state, when substrates of the T2SS such as the cholera toxin interact with the pseudopilus or the secretin itself, the pseudopilus would extend (middle figure) to push the toxin into the periplasmic vestibule. By pushing the toxin to the vestibule, extension of the pseudopilus would push the toxin through the constriction site created by the N3 domain which in turn generates the conformational changes required for the toxin to pass through the periplasmic gate and enter the extracellular chamber of the secretin (figure on the right), ultimately leading to release of the toxin. These figures from Reichow *et al.*, (2010) were reprinted with permission from the Nature Publishing Group.

BAM complex (Collin *et. al.*, 2007), and thereby serves as the only example of an outer membrane protein that supposedly does not require the BAM complex for assembly.

1.2.3 Mechanism of T2SS secretion

The precise series of events that take place during extrusion of substrates through the T2SS are not well characterized. The current theory suggests that T2SS substrates bind to the periplasmic domains of GspD, GspC or the pseudopilus tip and in doing so the ATPase activity of GspE is activated, thereby providing energy to power extension of the pseudopilus by adding subunits to the pseudopilus. The pseudopilus would then function in a piston-like manner by pushing substrates through the pore created by the secretin (Shevchik *et. al.*, 1997; Hobbs and Mattick, 1993). In a series of studies by Reichow *et. al.*, (2010; 2011), a model of T2SS function was described based on structural data obtained by cryoelectron microscopy and X-ray crystallography of the *V. cholerae* T2SS secretin $GspD^{Vc}$. As shown in Fig. 3 parts B and C, the piston-model of protein extrusion suggests that when the substrate (cholera toxin) binds to the psuedopilus or the secretin, the pseudopilus extends to push the substrate into the periplasmic vestibule of the secretin. Further extension of the pseudopilus pushes the substrate past the constriction site created by the N3 periplasmic domain of the secretin, ultimately leading to conformational changes in the structure of the secretin that results in transfer of the substrate to the extracellular chamber and subsequent secretion.

1.3 The type II secretion systems of *Escherichia coli*

In several *E. coli* pathotypes the T2SS has been shown to be required for pathogenesis of human and animal hosts, thereby signifying this system as a major virulence factor in this species. Examples of the requirement for the T2SS in *E. coli* pathogenesis include; colonization of the mouse small intestine by ETEC strain H10407 is facilitated by T2SS secretion of LT (Tauschek *et. al.*, 2002), intimate adherence of EHEC strain O157:H7 to epithelial cells is promoted by T2SS secretory proteins including the metalloprotease StcE (Grys *et. al.*, 2005) and YodA (Ho *et. al.*, 2008), colonization of the urinary tract and persistence in the mouse bladder requires a T2SS in UPEC (Kulkarni *et. al.*, 2009), the T2SS in EPEC strain E2348/69 is

required for virulence, and *E. coli* O157:H7 strains cured of the T2SS-encoding pO157 plasmid did not colonize the bovine terminal rectal mucosa as efficiently as wild-type (Sheng *et. al.*, 2006).

In addition to the substrates secreted by the *E. coli* T2SS mentioned above, the *E. coli* T2SS of UPEC is also involved in secretion of an adhesin protein. The Dr fimbriae are a family of adhesin proteins expressed on the surface of UPEC strains that are involved in attachment to urinary tract cells (Le Bouguenec *et. al.*, 2001). The fimbriae are composed of repeating monomers of the DraE subunit assembled by the chaperone/usher pathway with a minor subunit DraD that functions as the tip adhesin (Anderson *et. al.*, 2004). Surprisingly the T2SS of UPEC was shown to secrete the DraD tip adhesin since in the absence of the chaperone/usher pathway, the fimbriae composed of DraE is not formed but DraD remains secreted. Inactivation of the T2SS rendered the cell incapable of secreting DraD (Zalewska-Piatek *et. al*., 2008).

Unbeknownst to the scientific community, many *E. coli* pathotypes actually encode two T2SSs that have been designated as $T2SS_{\alpha}$ and $T2SS_{\beta}$ in this thesis (Fig. 1). The presence of more than one T2SS in *E. coli* is not surprising given that the existence of multiple T2SSs has been described in other species including the Yts1 and Yts2 systems of *Yersinia enterolitica* 8081 and the Xcp and Hxc systems of *Pseudomonas aeruginosa*. An in-depth analysis of the $T2SS_{\alpha}$ and $T2SS_{\beta}$ systems of ETEC is given in chapter 3.

1.4 The T2SS of *Vibrio cholerae*

The operons that encode the T2SS of *V. cholerae* include a major operon *gspC-N* and a minor operon *gspAB* that are not located adjacent to one another as in the $T2SS_{\alpha}$ operon of *E. coli* (Fig. 1). The T2SS has been shown as capable of secreting a variety of proteins including a lipase and protease (Strozen *et. al.*, 2011), an endochitinase, hemagglutinin-protease and neuraminidase (Ali *et. al.*, 2000; Sandkvist *et. al.*, 1997) with the most notable substrate of the *V. cholerae* T2SS being the cholera toxin (Sandkvist *et. al.*, 1997) responsible for the diarrheagenic phenotype of *V. cholerae*. Interestingly, *V. cholerae* is also capable of secreting the structurally related ETEC heat-labile enterotoxin B-subunit (Sandkvist 1997). An in-depth discussion of the *V. cholera* T2SS is provided in chapter 5.

1.5 The heat-labile enterotoxin and cholera toxin

The major virulence factors of ETEC and *V. cholerae* are the heat-labile enterotoxin (LT) and cholera toxin (CT) respectively. These toxins are similar in both amino acid sequence and structure and are part of the AB_5 family of enterotoxins in which the holotoxin is composed of one A and five B subunits. LT is considered a heat-labile enterotoxin due to the ability to inactivate the toxin by incubation in temperatures above 60^{o}C (Giannella and Mann, 2003). The cholera toxin A subunit (CT-A) is denatured irreversibly at temperatures greater than 51^{o}C, whereas the cholera toxin B subunit (CT-B) pentamer will dissociate at a temperature of 74^{o}C (Goins and Freire, 1988) yet remains stable in 1% SDS (Gill, 1976). The LT-A, LT-B, CT-A and CT-B subunits are translocated across the inner membrane via the Sec machinery into the periplasm where they are assembled into the AB_5 holotoxin that requires the activity of the disulfide oxidoreductase protein DsbA to form disulfide bonds (Yu *et. al.,* 1992).

CT is secreted from the cell by the sole T2SS of *V. cholerae* (Sandkvist *et. al*., 1997) whereas in ETEC, both the $T2SS_{\alpha}$ and $T2SS_{\beta}$ have been shown as capable of secreting LT. The $T2SS_{\alpha}$ functions in secretion of LT that remains cell-associated by binding to LPS on the extracellular surface of the bacterium via its B-subunit, and is released upon formation of outer membrane vesicles (Horstman and Keuhn, 2002; Horstman *et. al.,* 2004). The $T2SS_{\beta}$ has been shown to secrete soluble LT into culture supernatant since in the absence of $GspD_{\beta}$ (Taushek *et. al.*, 2002), $GspE_{\beta}$ and $GspM_{\beta}$ (Dorsey *et. al.*, 2006), LT toxin was not detected in culture supernatant.

Once secreted, CT in soluble form and LT in soluble and vesicle-associated forms bind to the GM-1 ganglioside on the surface of the intestinal epithelium and are endocytosed by lipid rafts. The toxin is then trafficked in a retrograde manner to the Golgi apparatus and ER. Once inside the cytosol, CT-A and LT-A permanently activate adenylate cyclase by ribosylating the stimulatory G protein. Permanent activation of adenylate cyclase leads to an increased concentration of cellular cAMP that translates into the loss of water by the cell, thereby causing diarrhea (Nataro and Kaper, 1998).

1.6 Lipoproteins and the Lol system

Lipoproteins are a diverse group of proteins that perform integral functions in cellular metabolism including the construction and maintenance of cell surface structures, substrate transport, drug efflux and a protein sorting function that includes the β-barrel proteins as

substrates (Bernadac *et. al.*, 1998; Clavel *et. al.*, 1998; Ehrmann *et. al.*, 1998; Nikaido, 1998). *In silico* analysis of the *E. coli* K-12 proteome suggested that *E. coli* encodes at least 90 lipoproteins, the majority of which perform unknown functions and are are likely localized to the inner leaflet of the outer membrane (Tokuda *et. al.*, 2007). To date, three essential lipoproteins have been identified, these include; BamD that is a component of the BAM complex (Wu *et. al.*, 2005), LptE that is involved in the transport of LPS to the surface of the cell envelope (Bos *et. al.*, 2004) and LolB that is a component of the Lol (localization of lipoproteins) pathway required for lipoprotein sorting as described below (Matsuyama *et. al.*, 1997).

Lipoproteins encode an N-terminal signal peptide, which targets the protein for transport across the inner membrane by the Sec system, followed by a lipobox domain (Fig. 4). The lipobox domain is a four amino acid consensus domain of the sequence L-A/S-G/A-C that is required for lipoprotein maturation. The maturation process is performed at the outer leaflet of the inner membrane in three steps. In the first step, diacylglycerol is added to the Cys residue of the lipobox motif by phosphatidylglycerol/prolipoprotein diacylglyceryl transferase (Lgt). Addition of the lipid moiety to the N-terminal Cys residue of the mature lipoprotein allows the lipoprotein to be anchored in the inner or outer membrane. Next, the signal peptide is cleaved by signal peptidase II (LspA) and in the final step, the Cys residue is aminoacylated by the protein phospholipid/apolipoprotein transacylase (Lnt).

Once processed, lipoproteins are sorted to the inner or outer membrane by the Lol system according to the identity of the amino acids located at the +2 and +3 positions of the mature lipoprotein (directly adjacent to the lipidated Cys residue). In *E. coli*, Asp residues at positions 2 (Yamaguchi *et. al.*, 1988) and 3 (Seydel *et. al.*, 1999) ensure retention of the lipoprotein in the inner membrane by evading the LolCDE complex responsible for lipoprotein transport to the outer membrane (Fig. 5). Most other combinations of amino acids at positions 2 and 3 result in complex formation of the lipoprotein with the ATP-binding cassette transporter (ABC transporter) LolCDE that catalyzes the release of the N-terminal lipid moiety of the lipoprotein from the outer leaflet of the inner membrane (Yakushi *et. al.*, 2000). Once removed from the inner membrane, the lipoprotein is transported across the periplasm via its hydrophobic lipid moiety on its N-terminus by the chaperone LolA (Matsuyama *et. al.*, 1995). When shuttled to the outer membrane, the lolA-lipoprotein complex interacts with the outer

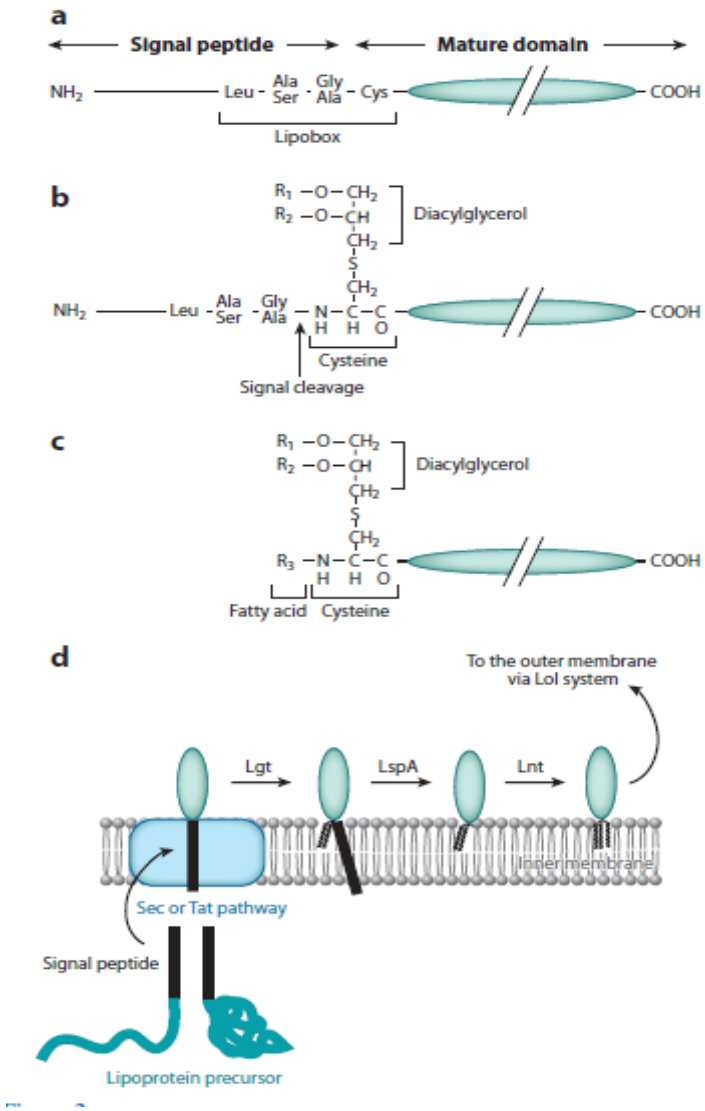


FIG. 4. The lipoprotein maturation process in Gram-negative bacteria. (a) The N-terminus of lipoproteins consist of a signal peptide and a four amino acid lipobox of the consensus sequence L-A/S-G/A-C. (b) Lipoprotein maturation involves addition of a diacylglycerol moiety to the sulfur atom of the conserved cysteine via a thioether linkage by the protein Lgt, cleavage of the signal peptide by LspA and (c) addition of a fatty-acid to the amino group of the N-terminal cysteine by Lnt. (d) A schematic diagram of the lipoprotein processing in the inner membrane is given. This figure from Okuda and Tokuda, (2011) was reprinted with permission from the journal Annual Reviews in Microbiology.

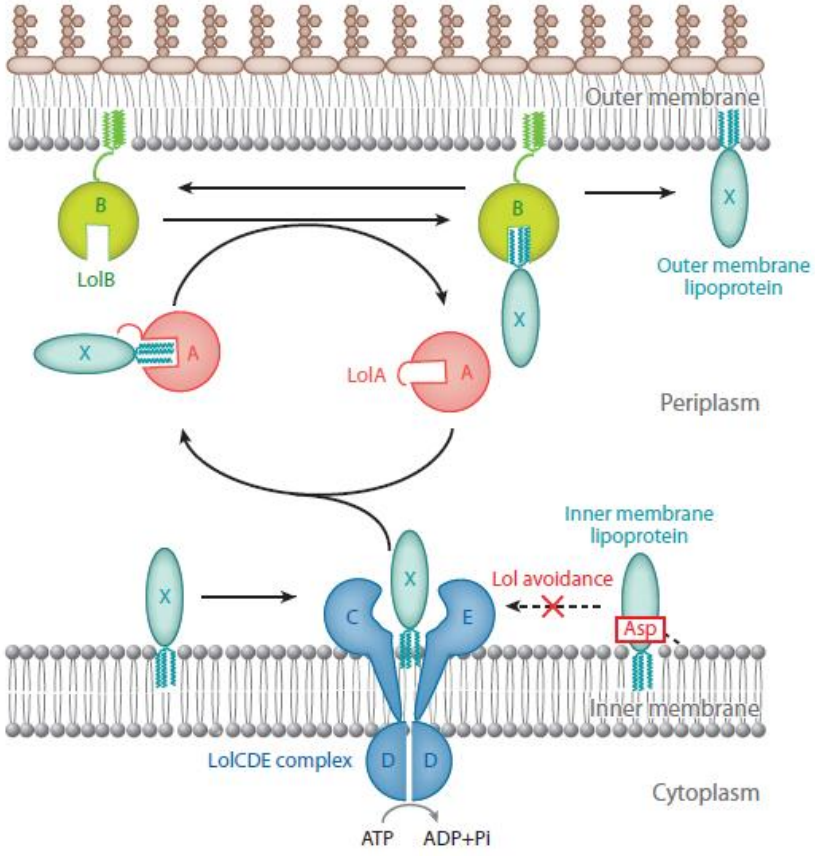


FIG. 5. Lipoprotein sorting by the localization of lipoprotein (Lol) system. Lipoproteins processed by Lgt/LspA/Lnt are localized to the inner membrane via the N-terminal lipid moiety. If the +2 (and +3) amino acid of the mature lipoprotein is Asp, the protein avoids the Lol system. Lipoproteins destined for localization to the outer membrane are bound by the ABC transporter LolCDE complex that functions in catalyzing removal of the hydrophobic N-terminal lipid from the outer leaflet of the inner membrane and transfer to the periplasmic chaperone LolA. The LolA-lipoprotein is shuttled to the outer membrane whereupon the hydrophobic lipid moiety of the lipoprotein is transferred to LolB and inserted into the inner leaflet of the outer membrane. This figure from Okuda and Tokuda, (2011) was reprinted with permission from the journal Annual Reviews in Microbiology.

membrane lipoprotein LolB (Matsuyama *et. al.*, 1997) whereby a "mouth to mouth" transfer of the hydrophobic lipid moiety of the lipoprotein is transferred from LolA to LolB which displays a higher affinity for the hydrophobic lipid than LolA (Okuda and Tokuda, 2009). Once transferred to LolB, the lipoprotein is localized to the outer membrane via insertion of the N-terminal lipid modification in the outer membrane (reviewed by Tokuda and Matsuyama, 2004).

1.7 The pilotin family of outer membrane lipoproteins

In many cell envelope systems that utilize an outer membrane secretin, assembly of the secretin in the outer membrane requires the function of an outer membrane lipoprotein termed the pilotin to localize the components of the secretin to the outer membrane for assembly. The pilotin proteins are relatively small (approximately 150 amino acids) and have been shown to recognize the C-terminal S-domain of the secretin monomer (Hardie *et. al.*, 1996; Daefler *et. al.*, 1997, Okon *et. al.*, 2008). To date, several pilotins have been identified as required for GspD protein localization in the outer membrane, these include; $GspS^{Ko}$ required for localization of T2SS $GspD^{Ko}$ of *Klebsiella oxytoca* (Hardie *et. al.*, 1996; Nouwen *et. al.*, 1999), $GspS^{Ec}$ required for localization of T2SS $GspD^{Ec}$ of *Erwinia chrysanthemi* (Shevchik and Condemine, 1998), MxiM required for localization of the T3SS MxiD in *Shigella flexneri* (Schuch an Maurelli, 1999), YscW for assembly of the T3SS YscC secretin in *Yersinia enterocolitica* (Burghout *et. al.*, 2004) and SciN for assembly of the T6SS SciD secretin in EAEC (Aschtgen *et. al.*, 2008).

Secretin proteins exhibit significant sequence conservation with exception of the C-terminal variable S-domain recognized by the pilotin protein (Burghout *et. al.*, 2004; Daefler *et. al.*, 1997; Okon *et. al.*, 2008). Due to the heterogeneous nature of the pilotin binding site of secretins, pilotin proteins themselves exhibit a large degree of heterogeneity, thereby making *in silico* identification of these proteins difficult. As a result, pilotin proteins are classified as such based on the similarity of the function of the protein and not on amino acid similarity. Consistent with the sequence diversity of pilotin proteins, the structures of pilotins are also diverse. For instance, the EHEC pilotin EtpO (a member of the PulS family) is not similar in structure to pilotins MxiM of the T3SS or PilF of the T4P (Lario *et. al.*, 2005; Izore *et. al.*, 2011; Kim *et. al.*, 2006; Koo *et. al.*, 2008; Trindade *et. al.*, 2008).

Since outer membrane-localized lipoproteins that include pilotin proteins require the function of the Lol system for localization, it is hypothesized that the pilotin and its bound secretin monomer are also localized to the outer membrane by the Lol system (Okon *et. al.*, 2008) (Fig. 6). Involvement of the Lol system in secretin assembly would suggest that a ternary complex composed of LolA-pilotin-secretin monomer would be assembled in the periplasm, the complex would move to the outer membrane whereupon the pilotin-secretin monomer complex would be transferred to LolB thereby resulting in insertion of the N-terminal hydrophobic lipid moiety of the pilotin into the inner leaflet of the outer membrane. Since the pilotin is bound to the secretin monomer, the secretin would also be localized to the inner leaflet of the outer membrane. In a recent study by Collin *et. al.*, 2011, the *Klebsiella oxytoca* T2SS secretin $GspD^{Ko}$ could not be localized in the outer membrane of *E. coli* if a mutant LolA protein incapable of lipoprotein transfer to LolB was expressed, thereby identifying the requirement for the Lol system in secretin assembly. Whether or not the pilotin remains associated with the secretin monomer following assembly of the secretin has not been elucidated, although in a study by Nouwen *et. al.*, 1999, $GspS^{Ko}$ was thought to co-purify with $GspD^{Ko}$ when co-expressed in *E. coli*.

1.8 The GspAB complex

Previous studies of the 60 kDa $GspA^{Ah}$ and 25 kDa $GspB^{Ah}$ (Fig. 2) proteins in *A. hydrophila* demonstrated the requirement for co-ordinated expression of both $GspA^{Ah}$ and $GspB^{Ah}$ such that deletion of either protein prevented detection of the other (Jahagirdar and Howard., 1994). $GspA^{Ah}$ and $GspB^{Ah}$ span the inner membrane once and form a large heteromultimeric complex (Howard *et. al.*, 1996; Jahagirdar and Howard., 1994) that interacts with peptidoglycan (Li and Howard, 2010) and has been demonstrated to be required for localization and multimerization of the GspD secretin multimer in the outer membrane of *A. hydrophila* (Ast *et. al.*, 2002). In addition, $GspA^{Ah}$ contains a novel cytoplasmic ATPase domain (Schoenhofen *et. al.*, 2005) and a periplasmic peptidoglycan-binding domain that are both required for assembly of the secretin multimer (Howard *et. al.*, 2006; Li and Howard, 2010). Therefore, the GspAB complex is hypothesized to reorganize the peptidoglycan matrix, a function presumably necessitated by the 50 kDa size constraint imposed by the peptidoglycan mesh (Demchick and Kock, 1996) for assembly of the T2SS by binding to peptidoglycan and forming a multimeric structure (Li and Howard, 2010; Li *et. al.*, 2011). In *Aeromonas* species

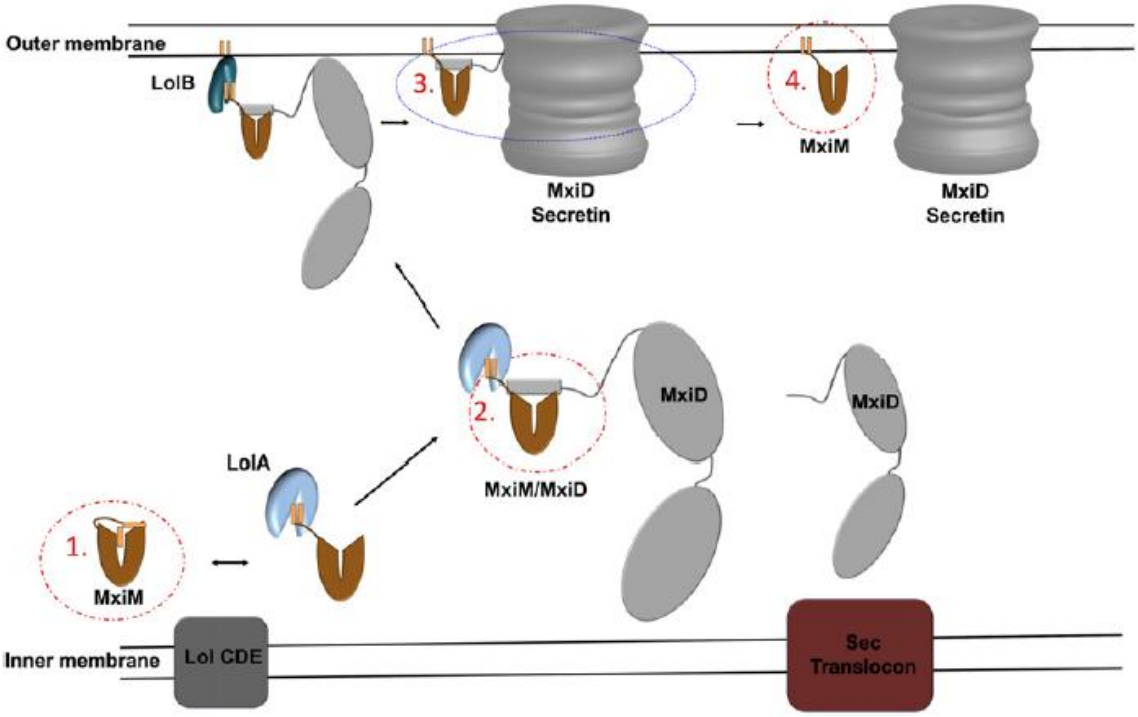


FIG. 6. Localization of the secretin to the outer membrane by the pilotin likely involves the Lol system. In this figure reproduced from Okon *et. al*., 2008, the T3SS pilotin of *Shigella flexneri* is bound by LolA in the periplasm. The interaction with LolA reveals a binding site within MxiM specific for the C-terminal S-domain of the secretin protein MxiD. A ternary complex composed of LolA-MxiM-MxiD is constructed and shuttled to the outer membrane whereby MxiM is transferred to LolB and inserted into the inner leaflet of the outer membrane, in doing so MxiD is also localized to the outer membrane and forms a secretin multimer. This figure was reprinted with permission from the Elsevier publishing group.

this function is absolutely required for assembly of the secretin such that deletion of GspAB^{Ah} results in no secretin multimer observed. These data do not disprove the existence of a pilotin protein that may exist in Aeromonads but instead that the peptidoglycan-binding activity of GspAB^{Ah} is integral to assembly of the GspD^{Ah} secretin in Aeromonads. An interesting conundrum exists for the requirement for proteins GspA, GspB and GspS in assembly and function of the T2SS in Gram-negative species. As evident in Table 1, in the T2SS of many species identified thus far, the presence of *gspS* and *gspA* is mutually exclusive and in no species identified thus far has *gspA* been encoded without *gspB*, but species have been identified that encode *gspB* with *gspS* instead of *gspA*. Due to the mutual exclusivity of GspA and GspS, one would expect these proteins to perform similar functions and as such exhibit extensive amino acid similarity. However, these proteins are not at all similar since GspS is an outer membrane lipoprotein and GspA is part of an inner membrane complex with a cytoplasmic ATPase domain and a periplasmic peptidoglycan-binding domain. What is similar though is that both proteins have been shown to be absolutely required for assembly of the T2SS secretin multimer in species that encode these proteins.

1.9 Purpose of this dissertation

Previous studies in Dr. Howard's laboratory had identified the requirement for GspAB in the assembly and function of the T2SS in *A. hydrophila*. I wanted to further extend this analysis by determining the requirement for GspAB in another *Aeromonas* species *A. salmonicida* to determine if the requirement for GspAB in secretin assembly is a property specific to the *A. hydrophila* species or the *Aeromonas* genus. In addition, since *Vibrio* species also encode a GspAB complex, I wanted to investigate the requirement for these proteins in assembly and function of the T2SS in the major human pathogen *Vibrio cholerae* and extend the analysis to include *Vibrio* species *V. parahaemolyticus* and *V. vulnificus*.

Another main focus of my research has been to investigate the T2SS in an enterotoxigenic *E. coli* species. A great deal of investigation described in the literature has dealt with the Type III secretion system that is utilized by *E. coli* species to adhere to the intestinal cell wall and translocate toxic effector proteins into cells of the intestinal epithelium. However, there is a surprising lack of knowledge concerning the T2SS in *E. coli* even though the T2SS has been characterized as a major virulence factor in this and other species. To this end, I have

investigated several aspects of the virulence mechanism of the enterotoxigenic *E. coli* strain H10407 that include; analysis of the mode in which the heat-labile enterotoxin is secreted by the cell, the involvement of the putative virulence factor LeoA (labile enterotoxin output A) in secretion of the heat-labile enterotoxin and characterization of the function of YghG (renamed $GspS_{\beta}$) as a pilotin protein required for assembly of the ETEC $GspD_{\beta}$ secretin in the outer membrane.

Lastly, a minor focus of my study has been to extrapolate the widely utilized bacterial II hybrid method by developing a library of target library clones to be used in the technique with the hypothesis that this library would enable efficient detection of interactions involving T2SS proteins and the entire proteome of ETEC strain H10407.

2. EXPERIMENTAL PROCEDURES

2.1 Bacterial strains, primers, plasmids and culture conditions

A list of the strains used in this study is provided in Table 2. *E. coli* H10407 was routinely grown in CAYE media [2% Casamino acids, 0.15% yeast extract, 43 mM NaCl, 50 mM K_2HPO_4, 0.25% glucose, 0.1% trace salts (0.415 M $MgSO_4$, 40 mM $MnCl_2$, 30 mM $FeCl_3$)] until mid-exponential phase of growth at 37^{o}C. *E. coli* DH5α was cultured in Luria-Bertani media (LB). *A. salmonicida* was grown in LB with Davis salts (Miller, 1972) at 22^{o}C until mid-exponential phase of growth. *A. hydrophila* was grown in LB supplemented with 30 mM Na_2HPO_4, 30 mM K_2HPO_4, 16.5 mM NaH_2PO_4, 16.5 mM KH_2PO_4, 0.75 mM $(NH_4)_2SO_4$, 0.4 mM $MgSO_4$ pH 7.0 at 30^{o}C until mid-log phase. *Vibrio* species were grown in LB at a temperature of 37^{o}C for *V. cholerae* and 30^{o}C for *V. vulnificus* and *V. parahaemolyticus* until mid to late logarithmic phase of growth. Antibiotics were used at the following concentrations: rifampicin (rif), 50µg/mL; kanamycin (kan), 50µg/mL; chloramphenicol (cat), 1.25 µg/mL and streptomycin (str), 20µg/mL, unless specified otherwise.

2.2 Purification of genomic and plasmid DNA, the polymerase chain reaction method

Minipreparations of plasmid DNA were conducted by use of the alkaline lysis method (Sambrook *et. al.,* 1989). Midipreparations of plasmid DNA were isolated using a Midiprep kit (Qiagen). Genomic DNA was isolated by use of the DNeasy kit (Qiagen). DNA from PCR amplifications, restriction digestions and ligations was purified by use of the QIAquick DNA

TABLE 2: List of strains used in this study

Strains	relevant phenotype/genotype and description	Source or reference
Escherichia coli		
H10407	wild-type ETEC serotype O78:H11, LT+ ST+	Evans et al., 1975
H10407 *gspD*$_{\alpha}$	H10407 *gspD*$_{\alpha}$ mutant	This study
H10407 *gspD*$_{\beta}$	H10407 *gspD*$_{\beta}$ mutant	This study
H10407 *yghJ*	H10407 *yghJ* mutant	This study
H10407 *pppA*	H10407 *pppA* mutant	This study
H10407 *yghG*	H10407 *yghG* mutant	This study
H10407 *eltAB*	H10407 *eltAB* mutant	Dorsey et al., 2006
TGS48	Kan^R,Cm^R; H10407 Pbad-*gspAB*$_{\alpha}$, Ptac-*gspC-O*$_{\alpha}$	This study
TGS49	Kan^R,Cm^R; H10407 Ptac-*gspAB*$_{\alpha}$, Pbad- *gspC-O*$_{\alpha}$	This study
TGS59	Kan^R,Cm^R; H10407 Ptac-*gspAB*$_{\alpha}$, Pbad- *gspC-O*$_{\alpha}$, *gspD*$_{\beta}$	This study
TGS60	Kan^R,Cm^R; H10407 Pbad-*gspAB*$_{\alpha}$, Ptac-*gspC-O*$_{\alpha}$, *gspD*$_{\beta}$	This study
DH5α	F$^-$ φ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169 *deoR recA1 endA1 hsdR17* (r_K^- m_K^-) *phoA supE44* λ$^-$ *thi-1 gyrA96* Δ*relA*	New England Biolabs
SM10λpir	*thi-1 thr leu tonA lacY supE recA*::RP4-2-Tc::Mu;Kan^R	Miller et al., 1988
S17-1	*recA thi pro hsdR*, RP4::2-Tc::Mu::KmTn7 λpir	Simon et al., 1983
XL1-Blue	recA1 lac endA1 gyrA96 thi-1 hsdR17 supE44 relA1[F' proAB lacIq ZΔM15 Tn10]	Bullock et al., 1987
Bacteriomatch reporter	Kan^R; Δ*mcrA* 183 Δ(*mcrCB-hsdSMR-mrr*)173 *endA1 hisB supE44 thi1 recA1 gyrA96 relA1 lac*[F'*laqI*^{q} *HIS3 aadA*]	Stratagene
A. hydrophila		
Ah65	Wild-type	Howard lab
C5.84	Ah65 *gspA*::Tn5-751	Jahagirdar and Howard, 1994
V. cholerae		
Bah2	E7946 derivative, TCP+, Δ*TLC* Δ*attRS* Δ*CTX*Φ Δ*RTX*	Pearson et al., 1993
Bah2-R	Rif^R mutant of Bah-2	Howard laboratory
Bah2-*gspA*	*gspA* mutant of Bah-2R	Howard laboratory
TRH7000	El Tor O1 biotype, *thy* Hg^R *ctxAB*	Overbye et al., 1993
TRH7000 *gspA*	*gspA* mutant of TRH7000	Howard laboratory
TRH7000 *gspD*	*gspD* mutant of TRH7000; Kan^R	Lybarger et al., 2009
V. vulnificus		
V. vul. 67181283		Saskatchewan Disease Control Lab
VvR	Rif^R mutant of *V. vulnificus* 67181283	Howard laboratory
Vv-*gspA*	*gspA::*kan mutant of VvR	Howard laboratory
V. parahaemolyticus		
V.para US32052027		Saskatchewan Disease Control Lab
VpR	Rif^R mutant of *V. parahemolyticus* US32052027	Howard laboratory
Vp-*gspA*	*gspA::*kan mutant of VpR	Howard laboratory
A. salmonicida		
As449	wild-type,	Kay et al., 1981
AsR	As449; Rif^R	Howard laboratory
As-*gspA*	AsR *gspA*::*kan*	Howard laboratory

purification kit (Qiagen). Polymerase chain reactions (PCR) were conducted with use of the Phusion DNA polymerase (New England Biolabs).

2.3 Plasmid construction

A list of plasmids and primers used in this study are provided in Tables 3 and 4 respectively. Recombinant plasmids constructed in this study were designated with the prefix TS as given in Table 3. The *yghG* open reading frame (ORF) was amplified from the H10407 genome with primers US420 and US398 for cloning into vector pBAD322C to construct complementation plasmid pBAD322C/*yghG*. Primers US420 and US400 were used to amplify *yghG* and introduce a 10 amino acid myc tag at the C-terminus of YghG when cloned into pBAD322C (to construct plasmid pBAD322C/*yghG*-myc). Overlapping PCR was used to create substitution mutations in the *yghG* ORF. The YghG C25A mutation was generated by overlapping PCR using templates created by primer sets US459/US462 and US460/US461 for amplification. Likewise, amplification with templates created with primers US459/US464 and primers US460/US463 was used to create the YghG A26D S27D substitution mutation.

The *leoA* open reading frame (ORF) was amplified from the H10407 genome with primers US263 and US258 for cloning into vector pBAD322C to construct complementation plasmid pBAD322C/*leoA* (TS51). An N-terminal Myc-tag was incorporated onto LeoA by amplification of H10407 genomic DNA with primers US265/US258 (TS53) and a C-terminal Myc tag added by amplification with primers US263/US264 (TS52).

Plasmids TS18 and TS19 that encode the periplasmic domains of $GspD_{\alpha}$ (amino acids 26-349) or $GspD_{\beta}$ (amino acids 39-364) that when expressed contain an N-terminal 6x Histidine tag were constructed by amplification of the H10407 genome with primers sets US161/US162 for $gspD_{\alpha}$ and US163/US164 for $gspD_{\beta}$ and cloning into plasmid pET47B with restriction enzymes *SacII* and *XhoI*. Plasmid TS76 that encodes *eltB* in plasmid pMAL-p4X was constructed with an amplicon generated by amplification with primers US280/281.

A set of vectors used for screening against the bacterial two hybrid library consisted of the periplasmic domains of various T2SS proteins and LeoA fused to the cI repressor protein in the vector pBT. A portion of the $gspD_{\alpha}$ gene that encodes amino acids 26 to 349 was amplified with primers US216 and and US162 and cloned into pBT to create plasmid TS32. The portion of the $gspD_{\beta}$ gene that encodes amino acids 39 to 364 was amplified with primers US217 and

TABLE 3: List of pasmids used in this study

Plasmids	TS-designation	Relevant description phenotype/genotype	Reference
pMMB/*gspA*^{epsA}	TS1	Cam^R; deletion of C-terminal half of *gspB*^{Vc} from pMMB/*gspAB*^{Vc}	This work
pTS12	TS12	Amp^R; pGEM-T containing *gspCD*$_\alpha$::FRT-*kan*-FRT[b]	This study
pET47B/*gspD*$_\alpha$$^{26\text{-}324}$	TS18	Kan^R; pET47B/*gspD*$_\alpha$ encoding amino acids 26-324 of GspD$_\alpha$ with N-terminal His-tag	This study
pET47B/*gspD*$_\beta$$^{39\text{-}364}$	TS19	Kan^R; pET47B/*gspD*$_\beta$ encoding amino acids 39-364 of GspD$_\beta$ and N-terminal His-tag	This study
pBT/*gspD*$_\alpha$	TS32	Cam^R; encodes amino acids 26-324 of GspD$_\alpha$	This study
pBT/*gspD*$_\beta$	TS33	Cam^R; encodes amino acids 39-364 of GspD$_\beta$	This study
pBT/*gspA*	TS38	Cam^R; encodes amino acids 262-489 of GspA	This study
pBT/*gspB*	TS39	Cam^R; encodes amino acids 64-139 of GspB	This study
pBT/*leoA*	TS43	Cam^R; encodes amino acids 283-577 of LeoA	This study
pBAD322C/*leoA*	TS51	Cam^R; encodes full-length LeoA	This study
pBAD322C/*leoA-myc*	TS52	Cam^R; full-length LeoA with C-terminal myc-tag	This study
pBAD322C/*myc-leoA*	TS53	Cam^R; full-length LeoA with N-terminal myc-tag	This study
pMAL-p4X/*eltB*	TS76	Cam^R; encodes LT-B subunit	This study
pTS100	TS100	Amp^R, Cam^R, Kan^R; TS12 *gspCD*$_\alpha$::FRT-*cat*-*kan*-FRT[b]	This study
TS101	TS101	Cam^R, Kan^R; *BstEII* fragment of TS100 containing FRT-*cat*-*kan*-FRT	This study
pBAD/*yghG*	TS130	Cam^R; *yghG* cloned into *NcoI/SphI* of pBAD322C	This study
pBAD/*yghG -myc*	TS131	Cam^R; 30 bp sequence encoding a myc tag at C-terminus of YghG	This study
pBAD/*yghG C25A-myc*	TS132	Cam^R; encodes *yghG* Cys 25 Ala	This study
pBAD/*yghG A26D S27D-myc*	TS133	Cam^R; encodes *yghG* Ala 26 Asp, Ser 27 Asp	This study
pBAD322C		Cam^R; *araC*, *rop*, *cat*, P$_{bad}$ promoter upstream of MCS[a]	Cronan, 2006
pRED/AMP		Amp^R; Red+, Gam+, Exo+, Rep^{ts}	Genebridges
pCP20		Amp^R Cm^R; FLP+, λ cI857+, λ p$_R$ Rep^{ts}	Cherepanov *et. al.*, 1995
pGEM-T		Amp^R; high copy number vector	Promega
pMAL-p4X		Amp^R; M13ori, creation of MalE fusion proteins	New England Biolabs
pET47B		Kan^R; encodes N-terminal His-tag	Novagen
pMMB207		Cam^R; Ptac promoter, wide host range vector	Morales *et. al.*,1991
pRJ31.1		Cam^R; 2.5kb *BstXI* fragment containing *gspAB*^{Ah} in *SmaI* of pMMB207;	Jahagirdar and Howard, 1994
pMMB/*gspAB*^{Ah}		Cam^R; 5.6 kb *EcoRI* fragment containing *gspAB*^{Ah} in *EcoRI* of pMMB207	Strozen *et. al.*, 2011
pMMB/*gspD*^{epsD}		Amp^R; *gspD*^{Vc} cloned into pMMB67;	Lybarger *et. al.*, 2009
pMMB68		Amp^r; *Ptac mob+*, *etxB* under *Ptac* control	Sandkvist *et. al.*, 1987
pMRS101		Amp^R, St^R; oriR6K, *sacBR*, pir	Sarker and Cornelis, 1997
pUC4K		Amp^R, Km^R; *lacZα*	Amersham
pBluescript SK		Amp^R; Plac, *lacZα*, f1, ColE1;	Stratagene
pBT		Cam^R; p15A ori, creation of lambda cI fusion protein	Agilent
pTRG		Tet^R; ColE1 ori, creation of RNAP-α fusion proteins	Agilent
pUT/Cam		Cam^R; R6K ori, contains a transposon	DeLorenzo, 1994

[a]MCS, multiple cloning site
[b]FRT, flippase recognition target

Table 4. Primers used in this study

Name	Sequence
US30	GTACAGCTGGCCGCTATCAT
US31	TGGTTGGGTTCAAAGCAAGT
US32	GCTGGGTTCAAGCAAAATTC
US33	TGACACATGGCGCAAAATAC
US42	ATCTCAACTACGGGTTGCAG
US43	CATCACATTGAGACGCAGCAG
UR70	GAATTTGAGGTCAGCTATCCGA
UR71	GCATAAGCGGAATTCATCGCA
US107RAN	GAATCTAGCGGCCGCATTGXXXXXXXXX
US108RAN	CTGGCACCTAGGACTTAGTTACXXXXXXXXX
US109	GAATCTAGCGGCCGCATTG
US110	CTGGCACCTAGGACTTAGTTAC
US147	TACGATCGGATCCATTGCGCTAAATCAG
US148	AAGCTCTTTACCGCGCAACAACTG
US162	TTAATCTCGAGTTACAGCACCTGTGCACGGCGAATG
US164	TTAATCTCGAGTTAATGCACCTGAGCACGGCGAATATC
US216	TCATCGGAATTCTGAACAATACGGCGCGAACTTC
US217	TACGATGAATTCGGTTGAAGAAGCCACTTTCACCGCTA
US218	ACTGGTGAATTTGAACGCCGTATC
US219	AAACCGGGAAGCGACTATGTAAGG
US243	TGACTGAATTCTGCTACGTTACCCGTTCCTG
US244	TTTCACTCGAGTCATTTCCCCTTCACTGTTG
US245	TGCTAGAATTCTGTCACTGTAGAGGCTGAAAC
US246	TGCTACTCGAGTTATCTCTCAAGTTCGTTCAG
US251	TGACTGAATTCTATTCAGCGTGTCACCCTGGCCAG
US252	TTTCACTCGAGCTATCTGGCAGTATGGTTTTCTGTG
US257	GAACAATTCAAACAGTTCAGTATTG
US258	CAGCTAGCATGCCTATCTGGCAGTATGGTTTTCTG
US260	TGATGTTCATGAAATTAAATATATTTACTAAATCTATG
US261	TACGCTTGCATGCTTACAGATCCTCTTCTGAGATGAGTTTTTGTTCTTGCTTAGTA AACGGCGCGAAG
US263	TACGTAGAATTCACCATGGAACAATTCAAACAGTTCAGTATTG
US264	TACGCTTGCATGCTTACAGATCCTCTTCTGAGATGAGTTTTTGTTCTCTGGCAGTATG GTTTTCTGTGGCAG
US265	TACGTAGAATTCACCATGGAAGAACAAAAACTCATCTCAGAAGAGGATCTGCAAT TCAAACAGTTCAGTATTGAAAAACAG
US280	TATAGCTTCATATGAATAAAGTAAAATGTTATGTTTTATTTAC
US281	TACAGCTAGGATCCCTAGTTTTCCATACTGATTGCC
US325	CTTCTCGTAGACATGGTGAATTCCTCC
US326	CAACCAGGAAAGAATAACTTCTCTTCTCGTAG
US327	ACGCCAGGTTTGACGTTTCTCACACAACCAGG
US329	GAGAATAGGAACTTCAAGAAACCAATTGTCCATA
US330	GAAAGTATAGGAACTTCAAGAAACCAATTGTCC
US332	ACGTAGTGTGGGCACGGTGAATTCC
US333	GATTCGCCAGATGGAAGGGGAAACGTAGTGTG
US334	ATTGATGGCAGCATCTTTATTATGATTCGCCAG
US337	CTTCTCGTAGACATATGCTATGGTCCTTG
US339	GAGAATAGGAACTTCGCACAATTCTCATGTTTG
US340	GAAAGTATAGGAACTTCGCACAATTCTCATG
US342	GGAAACGTAGTGTGGGCACATGCTATGGTCCTTGTTG
US355	GTTTCCGTCGGGGAACTTACAGGAATGAATGGACTGCGTCATGAAAGGACTCA ATAAAATCACCACTAAAGGGCGGCCGCGAAG

US356	CAGCGCCCAACTGGCGGGGTACGGTGAGTGAATTCTCATATGAATGCCTCACCG TGACGACTCACTATAGGGCTCGAGGA
US357	TTCCATCGCGCTGCGCTAACCGCATTTAATCCAGGAGAATCATTCATCGTGTTTT GGCGTACTAAAGGGCGGCCGCGAAG
US358	ACGGTGTTGGCGGTGGTTTCCTGTGCTACAGGCACCATTAACGCGTTCTCC CGGCATTGACTCACTATAGGGCTCGAGGA
US359	GTTTCCGTCGGGGAACTTACAGGAATGAATG
US360	CAGCGCCCAACTGGCGGGGTACGGTGAGTG
US361	TTCCATCGCGCTGCGCTAACCGCATTTAATC
US362	ACGGTGTTGGCGGTGGTTTCCTGTGCTAC
US369	CAACTGGAGGGCATGCCTTATGGAACAATTCAAACAGTTCAGTATTGAAAAACA GACTAAAGGGCGGCCGCGAAG
US370	TACGGAACTGCATTATGCATCTATCTGGCAGTATGGTTTTCTGTGGCAGGTACCCTCACT ATAGGGCTCGAGGAAGTTC
US371	CAACTGGAGGGCATGCCTTATGGAACAATTC
US372	ATAGCTATCCCGGGGAGCTCGCGAATTTCTGCCATTC
US373	ATAGCTATCCCGGGGAGCTCTTGAAATAAGATCACTACC
US381	TTTTGTCACTTGCGTTATTAATGAATAAGAAATTTAAATATAAGAAATC GACTAAAGG GCGGCCGCGAAG
US382	TTTAACATTGTTTTGTCACTTGCG
US383	ACAACAAATAACCTTTAGCCATGCTTTTTGATGTTTTTCAGCAATACCCTACT AAAGGGCGGCCGCGAAG
US384	CCAGCAAGTTACAACAAATAACCTTTAG
US385	AACCTTAGAGATTATTTACCATGTCGATAAAACAAATGCCAGGGAGGGTAA CTAAAGGGCGGCCGCGAAG
US386	TCTTATTAATAACCTTAGAGATTA
US387	GCCTTATCCGGCCTACGGGCTTACTCGGCAGACATCTTATGCTCGG TAACCTCACTATAGGGCTCGAGGAAG
US388	CGGCGCGAGCGCCTTATCCGGCCTAC
US389	TGATAATGACGTTGTTATCATTAAAACAATGCCTGTAGATAAATTGTTGCCTCA CTATAGGGCTCGAGGAAG
US390	AATGTCACTTTGATAATGACGTTGTT
US391	CAATTTATATCTACCCGACGTTATGCTTTGACTATTCCACAGGTGGTACG CTCACTATAGGGCTCGAGGAAG
US392	CAACCCGCGCCAATTTATATCTACC
US398	TACGCTTGCATGCTTATGCTTTGACTATTCCACAG
US400	TACGCTTGCATGCTTACAGATCCTCTTCTGAGATGAGTTTTTGTTC TGCTTTGACTATTCCACAG
US420	TGATGTTCATGAGCATAAAACAAATGCCAGGGAGGGTATTAATATC
US459	ACCATTCGCGAGCCTCCGGATGAC
US460	GAATGCTCATCCGGAATTTCGTATG
US461	GGATTATTAAGTGGCGCCGCCAGCCATAATG
US462	CATTATGGCTGGCGGCGCCACTTAATAATCC
US463	AAGTGGCTGTGACGACCATAATGAAAATGCCAG
US464	TTCATTATGGTCGTCACAGCCACTTAATAATCC

US164 and cloned into pBT to create plasmid TS33. The portion of the *gspA* gene that encodes amino acids 262 to 489 was amplified with primers US243 and US244 and cloned into pBT to create plasmid TS38. The portion of the *gspB* gene that encodes amino acids 64 to 139 was amplified with primers US245 and US246 and cloned into pBT to create plasmid TS39. The *leoA* gene was amplified with primers US251/US252 and cloned into pBT to create plasmid TS43.

2.4 Creation of genomic deletion mutants in ETEC and *Vibrio cholerae*

2.4.1 Suicide-vector based recombination method

Alleles $gspA^{Ah}$::kan and $gspA^{Ah}$::kan were introduced into *Aeromonas* and *Vibrio* species by marker exchange mutagenesis using the suicide vector pMRS101 (Sarker and Cornelis, 1997). A 1429 bp fragment of the *A. salmonicida gspA* gene was amplified by PCR utilizing the primers US42 and US43, blunt-end ligated into the *EcoRV* site of pBluescript II SK(+) and electroporated into *E. coli* XL-Blue. The aminoglycoside 3'-phosphotransferase gene conferring kanamycin resistance was excised from plasmid pUC4K via a *HincII* digestion and inserted into an *EcoRV* site located within *gspA* (nt position 780) of a recombinant plasmid clone. The $gspA^{Ah}$::kan gene was transferred from pSK into the suicide vector pMRS101 by *ApaI/XbaI* digestion. Recombinant plasmids were digested with *NotI* to remove the pBR322 origin of replication of the plasmid, self-ligated and electroporated into *E. coli* SM10λpir by selecting for streptomycin and kanamycin resistance. The donor strain (SM10λpir containing pMRS101/*exeA*::kan) and recipient strain AsR were grown overnight, subcultured 1:10 in brain heart infusion (BHI) broth without antibiotics and incubated at 22^{o}C for 1 hour prior to conjugation. 500 µL volumes of donor and recipient cells were mixed and collected by centrifugation. The pellet was resuspended in 100 µL BHI, applied to a pre-warmed BHI plate and incubated for 3.5 hrs at 30^{o}C. Half of the conjugation pool plate was scraped off and streaked onto LB containing kanamycin, rifampicin and sucrose (10%) and incubated at 22^{o}C to directly screen for recombinant recipient colonies. The exchange of alleles was verified by PCR.

Marker exchange mutagenesis of *Vibrio* species was conducted as described above except that 2100-3300 bp fragments containing all or part of *gspAB* were amplified from the following species using specific primer sets: UR70 and UR71 for *V. cholerae*; US30 and US31 for *V. parahaemolyticus* strain US32052027 and US32 and US33 for *V. vulnificus* strain

67181283. The aminoglycoside-phosphotransferase gene conferring kanamycin resistance was inserted into an *EcoRV* site located at positions 1106 and 1088 of the *V. vulnificus* and *V. parahaemolyticus gspA* genes respectively and into a *SphI* site located at position 1040 of *V. cholerae gspA*. *gspA*::kan fragments were transferred into pMRS101 from pSK using *BamHI/ApaI* sites for *V. vulnificus* and *V. parahaemolyticus gspA*::kan and *BamHI/SalI* for *V. cholerae gspA*::kan.

The *A. hydrophila gspAB* clone pRJ31.1 was described previously (Jahagirdar and Howard, 1994). A 5.6 kb *EcoRI* fragment containing the entire *gspAB* operon of *V. cholerae* including the presumed promoter was cloned into pMMB207 to create plasmid pMMB/*gspAB*^{Vc}. The plasmids were conjugated into *V. cholerae* and *A. hydrophila* strains from *E. coli* S17-1 as described above for complementation assays.

2.4.2 Lambda RED recombination system

Non-polar deletion mutations were constructed in the ETEC genome by use of the λRED recombination system (GeneBridges). Three clones of wild-type H10407 were transformed with the temperature-sensitive plasmids pRED ET/AMP and pRED ET/Tet (transformants selected on LB containing Amp 30 μg/mL and Tet 10 μg/mL respectively) and were utilized in triplicate to construct the various deletion mutants. Deletion mutations were created by removal of the entire open reading frame (ORF) with exception of 30 bp at the beginning and end of the gene, with insertion of a 69 bp "scar" sequence encoding a flippase recombination target (FRT) in the middle. By this method, a 53 amino acid protein was encoded that is composed of 10 amino acids of the N-terminus of the protein, 10 amino acids of the C-terminus of the protein and 33 amino acids encoded by the scar sequence.

Linear DNA encoding chloramphenicol acetyltransferase (*cat*) and aminoglycoside 3' phosphotransferase (*kan*) genes that confer chloramphenicol and kanamycin resistance respectively flanked by FRT sequences were constructed for replacement into the ETEC genome by lambda RED recombination. The *cat* gene encoded within pBAD322C was amplified with primers US372 and US373 and inserted into the *BspEI* site of FRT-*kan*-FRT encoded within plasmid TS12. The resultant plasmid TS100 was digested with *BstEII* to remove the FRT-*cat*-*kan*-FRT fragment (TS101) to be used as the template for amplification with primers outlined in Table 3. As shown in Table 4, two primers were used to amplify each

end of the FRT-*cat*-*kan*-FRT fragment and added 60 bp of H10407 genomic target sequence of the gene to be deleted. PCR reactions were performed using a 100 fold molar excess of "short" primers to "long" primers. Primers used for amplification of FRT-*cat*-*kan*-FRT and subsequently for λRed recombination are as follows: US355, US356, US359 and US360 for *gspD*$_{\alpha}$, US357, US358, US361 and US362 for *gspD*$_{\beta}$, US 381, US387, US382 and US388 for *yghJ*, US383, US389, US384 and US390 for *pppA*, US385, US391, US386 and US392 for *gspS*$_{\beta}$ and US369, US370, US371 and US372 for *leoA*. Routinely 100-200 ng of amplified product was used for electroporation into competent H10407 pRED ET/AMP made competent according to the method of Murphy and Campellone, 2003. Briefly, a 25 mL culture of H10407 pRED ET/AMP was grown in LB containing ampicillin at a concentration of 50 µg/mL to an OD_{600} of 0.1 at 30^{o}C. Arabinose was added to a final concentration of 0.3% and growth was continued at 30^{o}C until the OD reached approximately 0.6. The culture was then placed in a 42^{o}C water bath for 15 minutes and transferred to an ice water bath for 10 minutes. The cells were pelleted by centrifugation at 6,000 x g for 5 minutes and washed three times in 20% glycerol containing 1 mM MOPS. The final pellet was resuspended in 200 µL 20% glycerol containing 1 mM MOPS, using 50µL per electroporation. Following electroporation, transformation cultures were grown at 37^{o}C for 4 hours and plated onto LB containing kanamycin at a concentration of 30 µg/mL and chloramphenicol at a concentration of 2.5 µg/mL and grown at 37^{o}C. Chloramphenicol and kanamycin-resistance cassettes were removed from the genome by recombination by the FLP recombinase encoded within plasmid pCP20 (Cherepanov and Wackernagel, 1995). pCP20 was electroporated into chloramphenicol and kanamycin-resistant H10407 and transformants selected for ampicillin resistance at 30^{o}C. Ampicillin-resistant colonies were re-streaked and grown at 42^{o}C overnight, re-streaked onto LB and grown at 37^{o}C and then tested for kanamycin, chloramphenicol and ampicillin resistance. Deletion mutations that included the 69 bp scar sequence were verified by PCR and sequencing analysis.

2.4.3 Construction of H10407 strains containing alternate *gspA*$_{\alpha}$ and *gspC*$_{\alpha}$ promoters

The H10407 strain TGS59 contains a replacement of the nascent *gspA*$_{\alpha}$ and *gspC*$_{\alpha}$ promoters with an IPTG-inducible P_{tac} promoter and an arabinose-inducible P_{bad} promoter. This strain was constructed by a set of overlapping PCR reactions to generate an FRT-*cat*-*kan*-FRT

cassette flanked with P_{tac} and P_{bad} promoters on either end with 60 bp of sequence specific to $gspA_{\alpha}$ on one end and $gspC_{\alpha}$ on the other. The P_{tac} promoter from plasmid pMAL-p4X was amplified in succession by primer sets US337/US339, US326/US339 and US327/US329 to generate a 390 bp sequence that included *gspA* sequence on the upstream end and FRT-*cat*-*kan*-FRT sequence on the downstream end. Likewise, P_{bad} promoter was introduced upstream of *gspC* by a set of overlapping PCR reactions with primer sets US330/US332, US330/US333 and US330/US334. A final amplification with primers US327 and US334 generated a 3105 bp fragment that was used for lambda red recombination to construct the P_{tac} *gspA* and P_{bad} *gspC* strain.

H10407 strain TGS60 contains *gspA* and *gspC* operons inducible by arabinose and IPTG respectively. This strain was constructed by a set of overlapping PCR reactions to generate an FRT-*cat*-*kan*-FRT cassette flanked with P_{bad} and P_{tac} promoters on either end with 60 bp of sequence specific to $gspA_{\alpha}$ on one end and $gspC_{\alpha}$ on the other. The P_{bad} promoter from plasmid pMAL-p4X was amplified in succession by primer sets US325/US329, US326/US329 and US327/US329 to generate a 319 bp sequence that included *gspA* sequence on the upstream end and FRT-*cat*-*kan*-FRT sequence on the downstream end. Likewise, P_{tac} promoter was introduced upstream of *gspC* by a set of overlapping PCR reactions with primer sets US340/US332, US340/US333 and US340/US334. A final amplification with primers US327 and US334 generated a 3105 bp fragment that was used for lambda red recombination to construct the P_{tac} *gspA* and P_{bad} *gspC* strain.

2.5 Gel electrophoresis and immunodetection

Samples of whole cells were taken from broth culture and added to an equal volume of 2X sample buffer (125 mM Tris-HCl pH 6.8, 20% glycerol, 4% SDS, 0.01% bromophenol blue, 10% β-mercaptoethanol) and heated 5 min at 95^{o}C.

Samples of protein were routinely electrophoresed in 10-12% SDS-PAGE gels (Laemmli, 1970) and either soaked in Coomassie Brilliant Blue R250 stain (CBB) or transferred to PVDF membrane for immunoblotting. For CBB staining, gels were soaked in stain solution (0.125% CBB, 25% v/v isopropanol, 10% v/v acetic acid) for 1 hour, destain solution 1 (0.0025% CBB, 25% isopropanol, 10% acetic acid) for 1 hour, destain solution 2 (0.0005% CBB, 10% isopropanol, 10% acetic acid) for 1 hour (or overnight) and destain

solution 3 (10% acetic acid) until destaining was satisfactory. For immunoblotting, proteins were transferred to PVDF membranes in Tris-glycine transfer buffer. Immunoblots were visualized by blocking with 2% ECL blocking reagent (GE healthcare) for 1 hour, primary antibody in 2% blocking agent for 1 hour, goat anti-rabbit IgG-HRP conjugate (Bio-rad) secondary antibody (1/100 000 dilution) in 2% blocking agent for 1 hour and visualized by chemiluminescent detection with ECL advance (GE healthcare).

Proteins approximately 20 kDa and smaller (including YghG) were separated by 16% Tricine SDS-PAGE (Schaegger and von Jagow, 1987) and stained or transferred to PVDF membrane for immunoblotting as described above. Gels were composed of a separating gel composed of the following components: 4 mL AB3 (48% acrylamide, 1.5% bis-acrylamide), 4 mL 3 x gel buffer (37% Tris, 0.3% SDS, pH 8.45), 960 µL glycerol, 3.04 mL ddH_20, 40 µL 10% ammonium persulfate and 4 µL TEMED. The stacking gel was composed of the following components: 417 µL AB3, 1.25 mL 3 x gel buffer, 3.33 mL ddH_20, 38 µL 10% APS and 3.75 µL TEMED. Gels were electrophoresed in a 1X gradient gel running buffer composed of the following components: 6.05% Tris, 8.95 % Tricine and 1% SDS.

$GspD_\alpha$ and $GspD_\beta$ multimeric and monomeric proteins were detected by immunoblot analysis of whole cell and cell fractionated samples. Samples were electrophoresed in a 3-8% SDS gradient PAGE gel (Biorad) until an 84 kDa protein standard was approximately 1 cm from the bottom of the gel and transferred to PVDF membrane. Proteins were visualized using primary rabbit anti-$GspD_\alpha$ or anti-$GspD_\beta$ antisera (concentrations routinely used are given in Table 5), a peroxidase-conjugated anti-rabbit IgG secondary antibody (Sigma) and a chemiluminescent substrate (GE healthcare) detected with Hyper film (GE healthcare).

2.6 Isopycnic sucrose gradient fractionation

Separation of inner and outer membranes by isopycnic sucrose-density-gradient centrifugation was performed as described previously (Howard *et. al.*, 1996, Ishidate *et. al.*, 1986). Cells were grown in 100 mL CAYE media, centrifuged and resuspended in 10 mL 10 mM HEPES pH 8.0 containing 0.2 mM PMSF, 4 µg/mL DNase, 4 µg/mL RNase and 400 µL protease cocktail inhibitor (Roche) and broken by passage through a French pressure cell. Following breakage, EDTA was added to a concentration of 3 mM and whole cells were pelleted by centrifugation at 6000 x g 10 minutes at 4^oC. Supernatant was applied to the top of

the SG0 gradient composed of 0.2 mL 60% sucrose and 2.8 mL 10% sucrose in 10 mM HEPES containing 0.2 mM PMSF and centrifuged at ore 40,000 rpm 3 hours at 4^{o}C. Pelleted cell membranes located on top of the 60% sucrose bed were applied to the SG1 gradient composed of the following concentrations of sucrose in 10 mM HEPES containing 0.2 mM PMSF: 1 mL of 55%, 3 mL of 50%, 1.8 mL of 45%, 3 mL of 40%, 1.2 mL of 35% and 1.2 mL of 30% sucrose. The resultant SG1 gradient was centrifuged at 39,000 rpm 16 hrs at 4^{o}C. The SG1 was routinely separated into 14 or 15 fractions and analyzed for separation of inner and outer membrane material. Fractions containing inner membrane were identified by assay of the activity of the inner membrane protein NADH oxidase (Osborn *et. al.*, 1972). Fractions containing outer membrane material were identified by coomassie staining following SDS-PAGE to locate major outer membrane porin proteins OmpC and OmpF.

2.7 Enzyme-linked immunosorbent assay of heat-labile enterotoxin

Detection of LT in H10407 culture supernatant was performed by GM1 ganglioside enzyme-linked immunosorbent assay (ELISA) as previously described (Ristaino *et. al.*, 1983). In this protocol, 0.1 µg GM1 monosialoganglioside (Sigma G7641) in 60 mM bicarbonate buffer pH 9.6 was added to 96 well polystyrene plates and incubated overnight at 4^{o}C. Unbound GM1 was removed by three washes with 100 µL PBS-T buffer (phosphate buffered saline containing 0.5% (v/v) Tween-20). A 100 µL volume of block solution [PBS containing 5% (v/v) Fetal bovine serum (FBS)] was added to each well and incubated at 37^{o}C 1 hr. Following a set of washes, standard amounts of purified LT-B subunit (Sigma E8656) or culture supernatant was added and incubated for 2 hrs at 37^{o}C. Following another set of washes, 100 µL of primary antibody consisting of 1/5000 dilution of anti-cholera toxin antibody (Sigma C3062) in PBS-T containing 1% (v/v) FBS was added and incubated at 37^{o}C 1 hr. After washing, 100 µL of 1/1000 diluted HRP-conjugated goat anti-rabbit IgG in PBS-T 1% FBS was added and incubated at 37^{o}C 1 hour. Following another set of washes, 100 µL of developing solution (KPL) was added, the blue-green colour of the substrate was allowed to develop and was stopped prior to saturation by addition of 100 µL 1% SDS. The colour intensity was determined by the absorbance read at 410 nm.

2.8 The bacterial two-hybrid method

The Bacteriomatch II reporter strain (Table 2) (Stratagene, catalog number 200195) was transformed with the bait plasmid encoding the bait protein of interest. The resultant strain was

Table 5. **List of antibodies used in immunoblot analysis**

Specificity	**Source**	**Concentration**
Anti-GspD$_{\alpha}$	Howard laboratory	1/20 000
Anti-GspD$_{\beta}$	Howard laboratory	1/200 000
Anti-GspD^{Vc}	Sandkvist laboratory	1/100 000
Anti-GspD^{Ah}	Howard laboratory	1/50 000
Anti-myc	Cell-signalling	1/10 000
Anti-cI	Stratagene	1/10 000

made electrocompetent and electroporated with the pTRG B2H library. Following electroporation, 450 μL SOC media (2% tryptone, 0.5% yeast extract, 0.05% NaCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$) was added to cells and grown at 37°C for 2 hours at 225 rpm. Cells were spun at 2000 x g 5 min, the supernatant was removed and the pellet was resuspended in 1 mL M9+ His dropout broth, washed again in 500 μL M9+ His dropout broth and grown for 2 hours at 37°C. A volume of 100 μL of cells were routinely spread on selective screening media with and without steptomycin (according to manufacturer's instructions).

2.9 Construction of the B2H library

First-strand extension reactions consisted of 10 μg of ETEC strain H10407 genomic DNA mixed with 15 μM primer US107ran (Table 4), 1X NEB 2 buffer (50 mM NaCl, 10 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT, pH 7.9), 300 μM each dNTP and 45 Units of the Klenow fragment of *E. coli* DNA polymerase I (*exo+*) (New England Biolabs) in a 50 μL reaction volume. The genomic DNA and primer were denatured for 3 minutes at 93°C then incubated on ice for 5 minutes during which the remaining components of the reaction were added. The mixture was then incubated for 25 minutes at 25°C and 5 minutes at 50°C. Inactivation of the Klenow enzyme was achieved by the addition of 1.5 μL 0.5 M EDTA and incubation at 75°C for 10 minutes.

Second strand extension reactions were carried out in the same way as that of the first extension reaction except that the US108ran primer (Table 4) was used with 15 μL of purified first strand extension reaction (purified with QIAquick spin purification column). One-third of the volume (100 μL) of second-strand extension product was added to an equal volume of 2X urea gel loading buffer [4M urea, 10 mM EDTA, 2.5 mM Tris-HCl pH 7.5, 0.25% (w/v) bromophenol blue] and electrophoresed using two 8 M urea/4% polyacrylamide denaturing gels (20μL per well, 10 wells per gel). Three segments of the gel were excised that corresponded to regions 3-4 cm, 4-5 cm and 5-6 cm from the top of the gel. Per 100 mg of gel, 200 μL of diffusion buffer (0.5 M ammonium acetate, 10 mM magnesium acetate, 1 mM EDTA pH 8.0, 0.1 % SDS) was added and incubated at 50°C 1 hour. The supernatant was taken and purified with a QIAquick spin column. Purified second-strand extension product was PCR amplified with primers US109 and US110, digested with *NotI* and *AvrII* and ligated into plasmid pTRG digested with *NotI* and *SpeI*. Approximately 60 electroporations of purified ligation product was performed with electrocompetent *E. coli* strain DH5α to generate a library of

approximately 7 x 10^6 plasmids.

2.10 *In vivo* cross-linking

The crosslinker dithiobis(succinylpropionate) (DSP) (Pierce) was used to assess the state of the $GspD_\beta$ protein (in either monomeric or multimeric form) by crosslinking *in vivo*. Briefly, a culture volume of 1 mL (cells grown to an OD_{600} of approximately 2.0) was washed three times with phosphate-buffered saline (PBS), crosslinker was added to 100 µL of cells at a final concentration of 0.05 mM and incubated at 25°C for 30 minutes. The crosslinker was saturated by addition of Tris-HCl to a concentration of 50 mM and incubated at 25°C 15 minutes, selected samples were extracted once with phenol as previously described (Miller, 1972).

2.11 Enzyme assay of culture supernatant

Cultures were centrifuged at 31,000 x g 15 minutes and 10 mL of supernatant was taken, filter-sterilized and stored at -20°C and used to assay culture supernatant enzyme activity. When required, supernatant was concentrated approximately 10 fold by use of a Microcon-10 concentrator (10 kDa molecular weight cut-off).

Lipase activity in culture supernatants was determined by assaying the release of p-nitrophenol (pNP) from p-nitrophenol caprylate (pNPC) (Aragon *et. al.*, 2000) whereby 100 µL of culture supernatant was added to 900 µL of substrate containing 100 mM Tris pH 8.0, 0.2% Triton X-100 and 1 mM pNPC. The reaction was incubated at 25°C for 30 minutes while the absorbance at 410 nm was measured at 5 minute intervals. A unit of lipase activity equals 1 nmol pNPC hydrolyzed per minute. All enzymatic activities are expressed as units/mL supernatant per OD_{600} of culture.

The amylase activity in supernatants from cultures of *Vibrio* species was determined according to the protocol of Jiang and Howard, 1991, used to assay the T2SS-dependent amylase activity in *A. hydrophila* culture supernatants. 100 µL of supernatant was added to 250 µL 5% starch azure in 20 mM Na_2HPO_4 (pH 7), 50 mM NaCl pH 7.0 and incubated at 37°C with vigorous shaking for 16 hours. The reaction was stopped by addition of 50 µL 2.5M sodium acetate, and the supernatant was collected by 5 min centrifugation at 21000 x g and the absorbance at 595 nm was determined. Units are defined as the change in absorbance at 595 nm/hour.

Protease activity in culture supernatants was determined using azocasein as substrate. A 100 µL volume of supernatant was incubated with 300 µL of 1% azocasein in 60 mM phosphate buffer (pH 7.2). The reaction was incubated at 37°C with shaking for 2 hours and stopped by addition of 133 µL 30% trichloroacetic acid and incubation on ice for 30 minutes. Samples were then centrifuged at 21000 x g for 5 min and 400 µL of supernatant was recovered. An equal volume of 1 M NaOH was added to the supernatant and the absorbance at 450 nm was determined. Units of protease activity are expressed as the change in absorbance at 450 nm/hour.

2.12 Transposon mutagenesis of *V. cholerae*

Transposon mutagenesis of *V. cholerae* utilized a modified conjugation protocol found to generate more transconjugants than by standard methods. SM10λpir containing pUT/Cam and *V. cholerae* Bah2 $gspA^{Vc}$ were cultured in LB containing Amp (200 µg/mL), Cam (10 µg/mL) and LB containing Kan (50 µg/mL) respectively for 4 hours. A volume of 100 µL of each culture was plated on LB containing the same concentration of antibiotics and grown at 37°C overnight. A mixture of ½ of the cell mass of Bah2 $gspA^{Vc}$ was mixed with ¼ of the cell mass of the donor (SM10λpir pUT/Chl) on LB and incubated at 37°C 4 hours. The cell mixture was resuspended in 1 mL LB, 50 µL of resuspension was plated onto TCBS (Thiosulfate citrate bile sucrose) agar (EMD) containing 1 µg/mL Cam and incubated at 37°C for approximately 32 hrs. This protocol generated approximately 100 Cam-resistant *V. cholerae* Bah2 $gspA^{Vc}$ mutants per plate (2000 mutants per conjugation). Cam-resistant mutants were patched onto LB agar containing 1% skim milk to assess protease secretion competency. Those mutants that were protease-secretion negative were investigated further.

2.13 Statistical analysis

An unpaired two-sided student's *t* test was used for all statistical analysis. Values were considered significantly different at $P<0.05$.

3. Involvement of the two ETEC T2SSs in secretion of heat-labile enterotoxin

3.1 Introduction

The majority of *Escherichia coli* strains serve as integral members of the gut microbiota, although some strains have evolved or acquired genetic factors that render them capable of generating a pathogenic infection. In the non-invasive *E. coli* pathotype ETEC,

characterized by the prototypical strain H10407, the main virulence factor of the bacterium is the heat-labile enterotoxin, a toxin that is secreted by the T2SS.

In silico analysis of the H10407 genome identified that this strain encodes two T2SSs designated alpha and beta as mentioned in the introduction (Fig. 1). The $T2SS_{\alpha}$ is encoded within a divergent operon composed of a minor operon that encodes GspAB and a major operon that encodes proteins GspC-O. Both $T2SS_{\alpha}$ operons in ETEC strain H10407 are identical in nucleotide sequence to that previously described in *E. coli* K-12 (Francetic *et. al.*, 2000). The operon that encodes the $T2SS_{\beta}$ contains genes *gspC-M*, however whereas most T2SS operons encoded in other species are transcribed from a promoter upstream of *gspC* (reviewed by Filloux, 2004), this system is atypical in comparison to that of other species by including three genes (*yghJ, pppA* and *yghG*) upstream of *gspC* (Yang *et. al.*, 2007) (Fig. 7). The functions of the proteins encoded by these genes are unknown thereby suggesting that additional uncharacterized factors could be involved in the assembly and function of the $T2SS_{\beta}$ in comparison to the T2SSs encoded in other species.

Analysis of the genome of various *E. coli* pathotypes revealed that the $T2SS_{\alpha}$ and/or $T2SS_{\beta}$ operons are prevalent in pathogenic *E. coli* strains. As shown in Fig. 7, LT+ ETEC strains E24377A and B7A do not encode a $T2SS_{\alpha}$ operon but encode complete $T2SS_{\beta}$ operons. The non-pathogenic *E. coli* strain K-12 encodes the $T2SS_{\alpha}$ but has a large deletion in the operon encoding $T2SS_{\beta}$ that encompasses genes *gspD-K*. Enterohemorrhagic *E. coli* (EHEC) strain O157:H7 does not encode $T2SS_{\alpha}$ or $T2SS_{\beta}$ but does encode a T2SS within plasmid pO157. Adherent-invasive *E. coli* (AIEC) strain LF82, enteropathogenic *E. coli* (EPEC) strain O127:H6 and enteroaggregative *E. coli* (EAEC) strain 55989 only encode the *gspO* gene of the $T2SS_{\alpha}$ system and encode a complete $T2SS_{\beta}$ system. Finally, uropathogenic *E. coli* (UPEC) strain UTI89, the avian pathogenic *E. coli* (APEC) strain O1 and extraintestinal pathogenic *E. coli* (ExPEC) strain IHE3034 are similar to that of H10407 by encoding complete $T2SS_{\alpha}$ and $T2SS_{\beta}$ systems. This analysis showed that the $T2SS_{\alpha}$ and $T2SS_{\beta}$ systems are prevalent among *E. coli* pathotypes. Interestingly, the T2SS encoded on pO157 is neither an *E. coli* $T2SS_{\alpha}$ nor $T2SS_{\beta}$ and instead most closely resembles the T2SS of *Klebsiella oxytoca* (Genin and Boucher, 1994).

The $gspA_{\alpha}$ and $gspC_{\alpha}$ operons in *E. coli* K-12 MG1655 are cryptic due to repression by the histone-like nucleoid-structuring protein (H-NS) when cultured in laboratory conditions

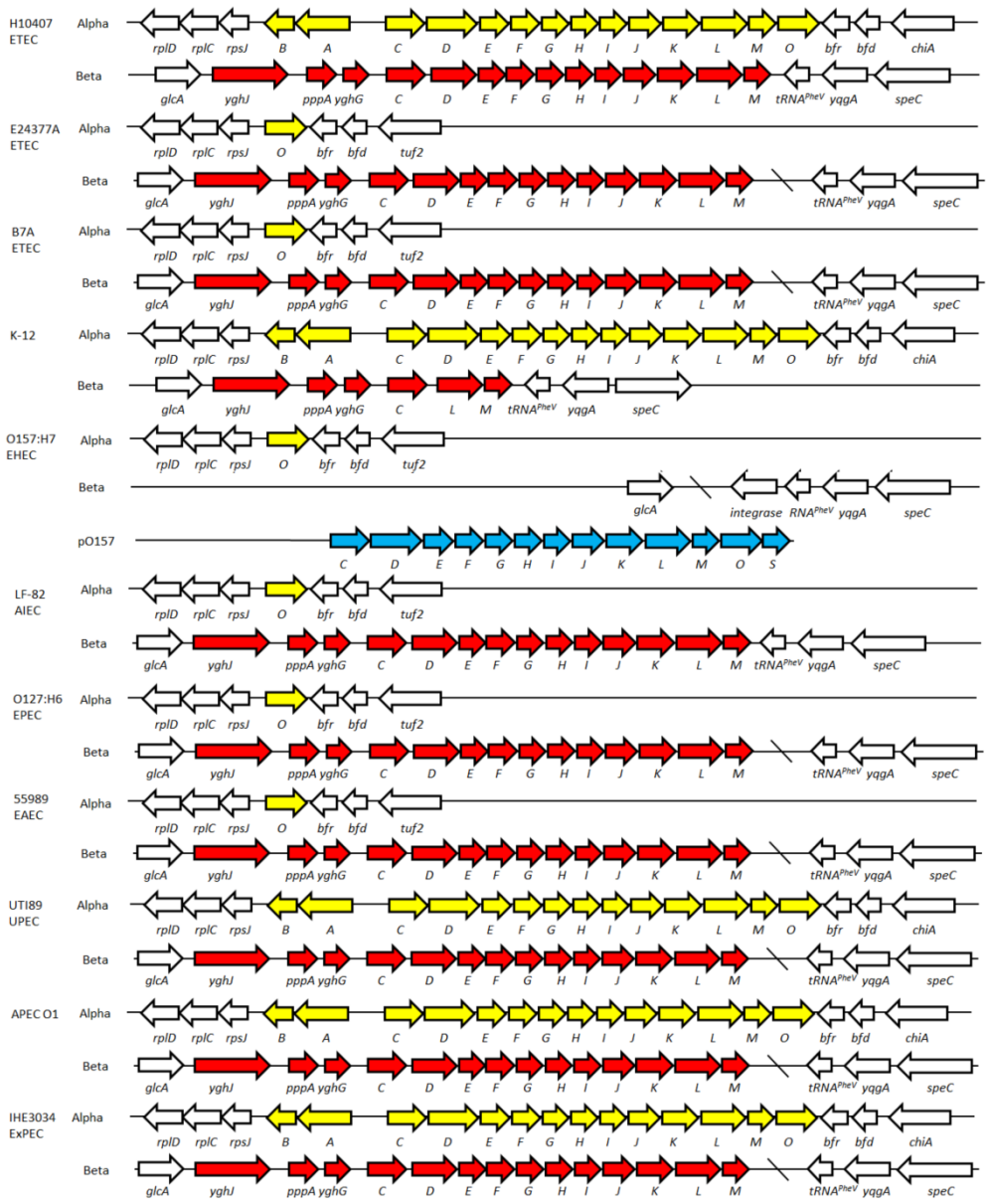


FIG. 7. Genetic organization of T2SS operons and surrounding genes in representative strains of pathogenic and non-pathogenic *Escherichia coli*. The enterotoxigenic *E. coli* (ETEC) H10407 genome (Crossman *et. al*., 2010) encodes two T2SSs termed alpha (yellow) and beta (red). The corresponding area of the genome that encodes the $T2SS_{\alpha}$ and $T2SS_{\beta}$ of H10407 is given for the following strains: LT-positive ETEC strains E24377A and B7A (Rasko *et. al*., 2008), *E. coli* strain K-12 (Blattner al., 1997), enterohaemorrhagic strain (EHEC) O157:H7 (Perna *et. al*., 2001), adherent-invasive strain (AIEC) LF82 (Peyretaillade *et al*., 2010), enteropathogenic strain (EPEC) O127:H6, enteroaggregative strain (EAEC) 55989 (Touchon *et. al.*, 2009), uropathogenic strain (UPEC) UTI89 (Chen *et. al.*, 2006), (APEC) avian pathogenic strain O1 (Johnson *et. al*., 2007) and extra-intestinal pathogenic strain (ExPEC) IHE3034. Note that the T2SS encoded on plasmid pO157 is also shown (blue) (Burland *et. al*., 1998). Genes flanking the T2SS operons are given (white).

(Francetic and Pugsley, 1996). However, when the *gspA* and *gspC-N* operons were encoded on plasmids and expressed in MG1655 *hns*, the T2SS$_\alpha$ was assembled and functional in secretion of substrates such as an endochitinase ChiA (Francetic *et. al.*, 2000). Interestingly, T2SS$_\alpha$ has also been shown to be capable of secreting LT. Expression of plasmid-encoded *gspAB*$_\alpha$ and *gspC-O*$_\alpha$ operons in an *hns* strain of K-12 supported secretion of plasmid-encoded LT, the majority of which remained associated with lipopolysaccharide on the surface of the cell, and within and on the surface of membrane vesicles (Horstman and Kuehn, 2002). The T2SS$_\beta$ has also been shown in ETEC strain H10407 to be responsible for secretion of soluble LT since deletion of *gspD*$_\beta$ renders the cell incapable of secreting LT into culture supernatant (Tauschek *et. al.*, 2002). Therefore, it is possible that T2SS$_\alpha$ could be involved in secretion of LT that remains associated with the outer membrane whereas T2SS$_\beta$ could be involved in secretion of LT into culture supernatant.

In this study, I wanted to determine the relative involvement of both the T2SS$_\alpha$ and T2SS$_\beta$ in secretion of LT. Since the studies used to identify the LT secretion capability of T2SS$_\alpha$ and the vesicle-mediated mechanism of LT secretion were performed in *E. coli* K-12, I wanted to investigate the requirement for T2SS$_\alpha$ in LT secretion in a wild-type *E. coli* strain. In addition, most reviews on the subject of LT secretion suggest that the vast majority of LT that is secreted by ETEC is vesicle associated and not soluble. Therefore I wanted to confirm if this was in fact the case and determine the relative amount of LT secreted in soluble and vesicle-associated forms.

3.2 Results

3.2.1 Involvement of T2SS$_\alpha$ and T2SS$_\beta$ in secretion of LT into culture supernatant

To determine if LT secretion into culture supernatant is entirely dependent upon the T2SS$_\beta$ as previously shown by Tauschek *et. al.*, 2002, or if the T2SS$_\alpha$ could also be involved in secretion, immunodectection of LT from culture superntatant was initially used to assess the ability of wild-type, *gspD*$_\alpha$ and *gspD*$_\beta$ strains to secrete LT. Efforts to detect the endogenous level of LT secreted in culture supernatant from H10407 when cultured in CAYE media by immunoblot with anti-cholera toxin antibody were unsuccessful since LT could not be detected in samples of 60 fold concentrated cell supernatant (data not shown). Therefore the amount of toxin made and secreted by the cell was increased by expressing a copy of the *eltB* gene (that

encodes the B subunit of LT) in ETEC. To verify that LT-B was expressed from this plasmid and could be detected by immunoblot with anti-CT antibody, whole cell samples from *E. coli* DH5α containing TS76 (pMAL-p4X/*eltB*) or empty vector pMAL-p4X were taken from cultures grown in the presence and absence of inducer molecule IPTG. As shown in Fig. 8, induction of *eltB* expression encoded on plasmid TS76 resulted in detection of two bands that likely correspond to monomers and dimers of LT-B, therefore this plasmid was used for further analysis.

Plasmid TS76 was electroporated into ETEC strains wild-type, *gspAB*$_{\alpha}$, *gspD*$_{\alpha}$ and *gspD*$_{\beta}$ and the relative amount of LT-B in concentrated samples of supernatant was detected by immunoblot. LT-B was detected in whole cell samples of wild-type, *gspD*$_{\alpha}$ and *gspD*$_{\beta}$ strains containing TS76 upon induction (Fig. 9A) and in concentrated supernatant from WT (TS76), *gspAB* (TS76) and *gspD*$_{\alpha}$ (TS76) strains but not in supernatant from *gspD*$_{\beta}$ (TS76) (Fig. 9B). This result suggested that T2SS$_{\beta}$ was required for secretion of LT. A perplexing result however was that the amount of LT-B in supernatant in strains *gspAB*$_{\alpha}$ (TS76) and *gspD*$_{\alpha}$ (TS76) was similar, and less than that observed from wild-type (TS76) supernatant. This result would suggest that the T2SS$_{\alpha}$ would be at least partially involved in secretion of LT-B, however in the *gspD*$_{\beta}$ (TS76) strain that is wild-type for the T2SS$_{\alpha}$, no secretion was observed. If T2SS$_{\alpha}$ was at least partially responsible for secretion of LT, then it would be expected that some LT-B would be observed in supernatant from the *gspD*$_{\beta}$ (TS76) strain.

Further analysis of the involvement of T2SS$_{\alpha}$ and T2SS$_{\beta}$ in LT secretion was conducted by assay of LT in supernatant directly by a GM1-ganglioside based ELISA. This technique was sensitive enough to allow assay of endogenous levels of LT in culture supernatant, thereby bypassing the requirement to use LT-B producing plasmids and immunoblot analyses. As shown in Fig. 10, supernatant from triplicate cultures of wild-type H10407, *gspD*$_{\alpha}$, *gspD*$_{\beta}$ and *eltAB* were isolated and the concentration of LT was determined by ELISA. The concentration of LT detected from wild-type H10407 supernatant is approximately 2.5 ng/mL. Absence of a functional T2SS$_{\alpha}$ by deletion of *gspD*$_{\alpha}$ had no effect on the LT secretion capability of H10407 whereas inactivation of T2SS$_{\beta}$ by deletion of *gspD*$_{\beta}$ abrogated secretion of LT into culture supernatant to levels observed in the *eltAB* strain. This result confirmed that T2SS$_{\beta}$ was required for secretion of soluble LT into culture supernatant as suggested by immunoblot analysis of LT-B in culture supernatant.

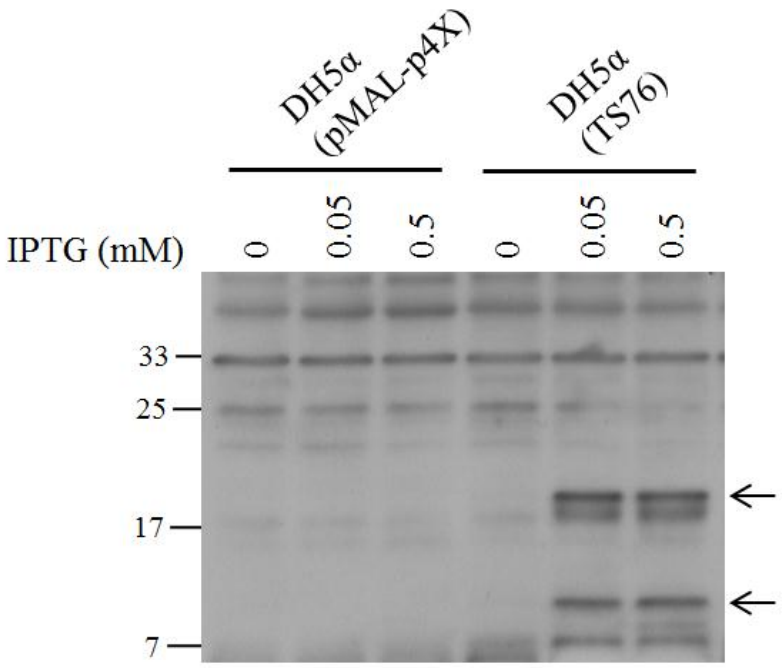


FIG. 8. Induction of *eltB* expression from plasmids TS70 and TS76 allow immunodetection of LT-B. Whole cell samples of *E. coli* DH5α containing empty vector pMAL-p4X and TS76 when cultured in the presence or absence of the inducer IPTG (as shown), were separated on 14% Tricine SDS-PAGE and immunoblotted with anti-CT antibody. The location of LT-B protomers are indicated on the right side of the panels whereas the location of protein markers is given on the left of the panel.

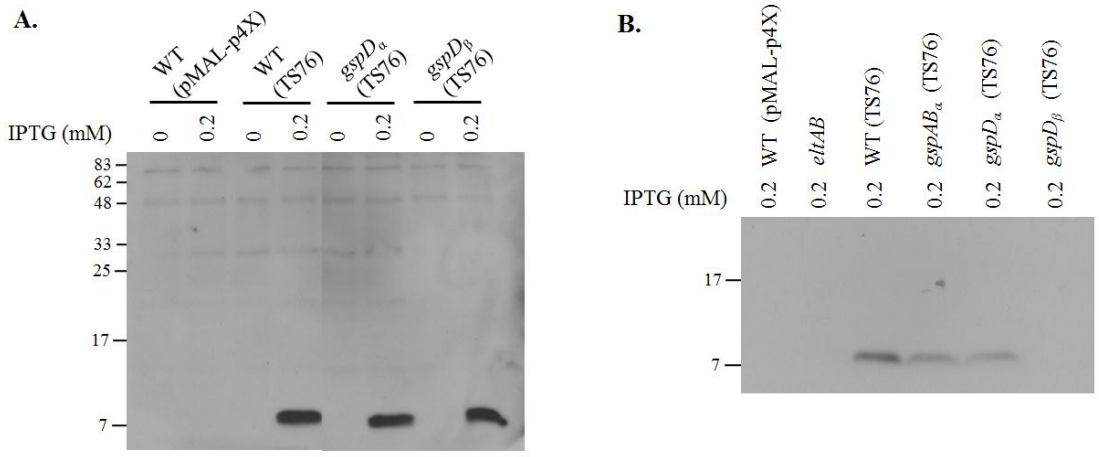


FIG. 9. **Both T2SSs may be involved in secretion of LT into culture supernatant.** (A) The relative amount of LT-B in whole cell samples of wild-type H10407 (WT), *gspD*$_{\alpha}$ and *gspD*$_{\beta}$ strains containing empty vector (pMAL-p4X) or TS76 (pMAL-p4X/eltB) cultured in the presence or absence of IPTG (as shown) was determined by immunoblot with anti-CT antibody. (B) Supernatant taken from cultures of wild-type, *eltAB*, *gspAB*, *gspD*$_{\alpha}$ and *gspD*$_{\beta}$ strains containing pMAL-p4X, TS76 or neither were concentrated 60 fold by TCA precipitation and immunoblotted with anti-CT antibody. The location of pre-stained standard protein markers is given on the left of each panel.

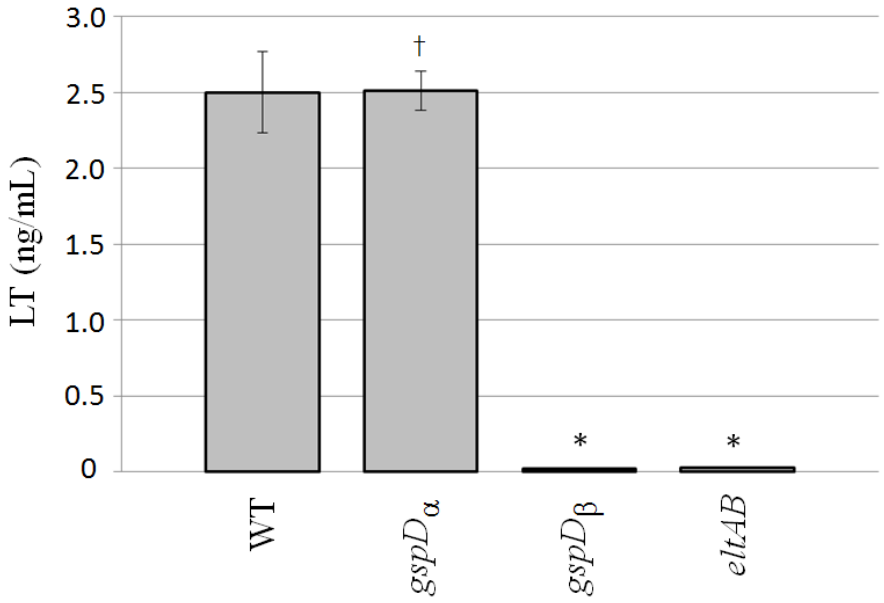


FIG. 10. The T2SS$_\beta$ is required for secretion of soluble LT into culture supernatant. The amount of LT assayed in supernatants of triplicate cultures of wild-type H10407, *gspD*$_\alpha$, *gspD*$_\beta$ and *eltAB* strains was assessed by ELISA. Significant difference († denotes non-significance and * denotes significance where $p<0.05$) was calculated based on comparison of each value against that observed in wild-type culture.

3.2.2 Immunoblot detection of $GspD_{\alpha}$ and $GspD_{\beta}$ secretin and monomer

To determine the relative level of assembly of each T2SS system in ETEC, immunoblot detection of $GspD_{\alpha}$ and $GspD_{\beta}$ secretin multimers was used. To detect $GspD_{\alpha}$ or $GspD_{\beta}$ independently by immunoblot, antibodies were raised against the periplasmic domains of $GspD_{\alpha}$ (amino acids 26-349) or $GspD_{\beta}$ (amino acids 39-364) (Fig. 11) (47% identity between $GspD_{\alpha}$ and $GspD_{\beta}$ protein fragments) since the periplasmic domain exhibits a considerable degree of heterogeneity in secretin proteins due to its possible involvement in substrate selection (reviewed by Korotkov *et. al.*, 2011). The $GspD_{\alpha}^{26\text{-}349}$ and $GspD_{\beta}^{39\text{-}364}$ proteins were encoded in plasmids TS18 and TS19 respectively and expression was induced in BL21(DE3). As shown in Fig. 12, induction of protein expression allowed detection of $GspD_{\alpha}^{26\text{-}349}$ and $GspD_{\beta}^{39\text{-}364}$ by coomassie-blue staining following SDS-PAGE (Fig. 12A). Antibodies were raised against these proteins and tested for cross-reactivity. Anti-$GspD_{\beta}$ antibody was verified to not cross-react with $GspD_{\alpha}$ (Fig. 12B, lane 3) whereas anti-$GspD_{\alpha}$ did cross-react with $GspD_{\beta}$ (Fig. 12B, lane 2), therefore requiring further purification of anti-$GspD_{\alpha}$ that rendered the purified antibody capable of detecting $GspD_{\alpha}$ (Fig. 12C) and incapable of recognizing $GspD_{\beta}$ (Fig. 12D).

To determine the relative level of assembly of each system by immunoblot detection of $GspD_{\alpha}$ and $GspD_{\beta}$ secretin multimers in ETEC, cultures were grown in conditions previously shown to induce secretion of LT (CAYE media), since presumably under these conditions expression of the T2SS(s) would be induced (Sack *et. al.*, 1980). Immunoblot analysis of whole cell samples revealed that $GspD_{\alpha}$ could not be detected in multimeric or monomeric form (Fig. 13A). Inability to detect the $GspD_{\alpha}$ protein was not due to use of an insufficient anti-$GspD_{\alpha}$ antibody since the antibody was capable of detecting the $GspD_{\alpha}$ secretin and $GspD_{\alpha}$ monomer in strains in which the *gspAB*$_{\alpha}$ and *gspC-O*$_{\alpha}$ operons were replaced by arabinose-inducible or IPTG-inducible operons (refer to section 3.2.4). These results suggested that under standard laboratory conditions the $T2SS_{\alpha}$ is not assembled. Attempts to identify a condition in which $GspD_{\alpha}$ could be detected in wild-type ETEC were not successful since variation of temperature, osmolarity and growth media did not result in detection of $GspD_{\alpha}$ (data not shown). In contrast to $T2SS_{\alpha}$, $T2SS_{\beta}$ was assembled in ETEC under standard laboratory conditions since the $GspD_{\beta}$ secretin multimer and monomer were detectable by immunoblot with anti-$GspD_{\beta}$ antibody (Fig.

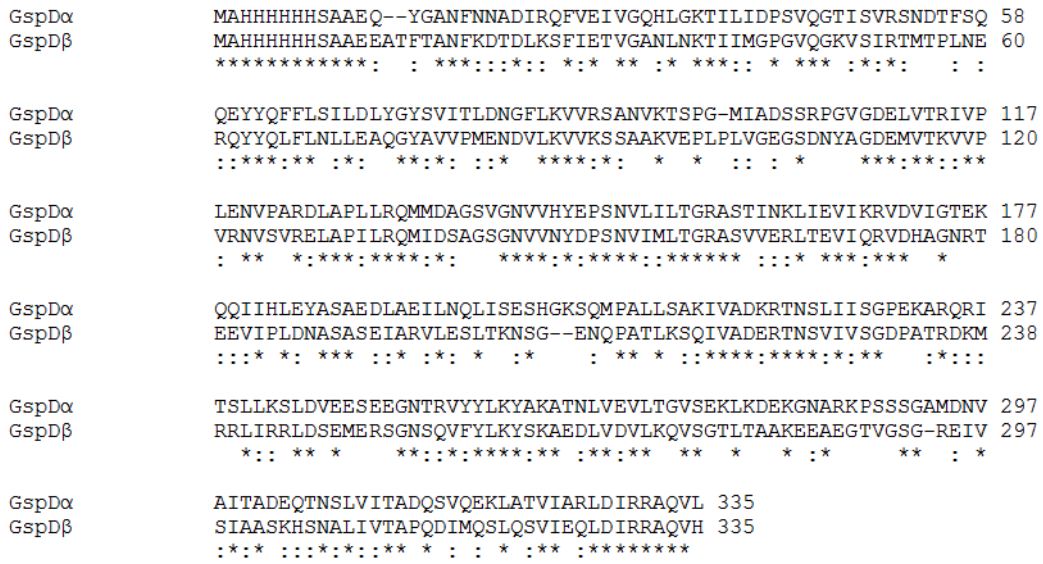


FIG. 11. Amino acid alignment of the N-terminal His-tagged putative periplasmic domain of $GspD_{\alpha}$ (amino acids 26-349, encoded by TS18) with the N-terminal His-tagged putative periplasmic domain of $GspD_{\beta}$ (amino acids 39-364, encoded by TS19). Percent identity (47%), percent similarity (69%) and score (793) of this pairwise global sequence alignment was calculated by the Needleman-Wunsch global alignment algorithm employed by the program EMBOSS (European bioinformatics institute). Identical (⋆) and similar (:) amino acids in the alignment are designated as shown.

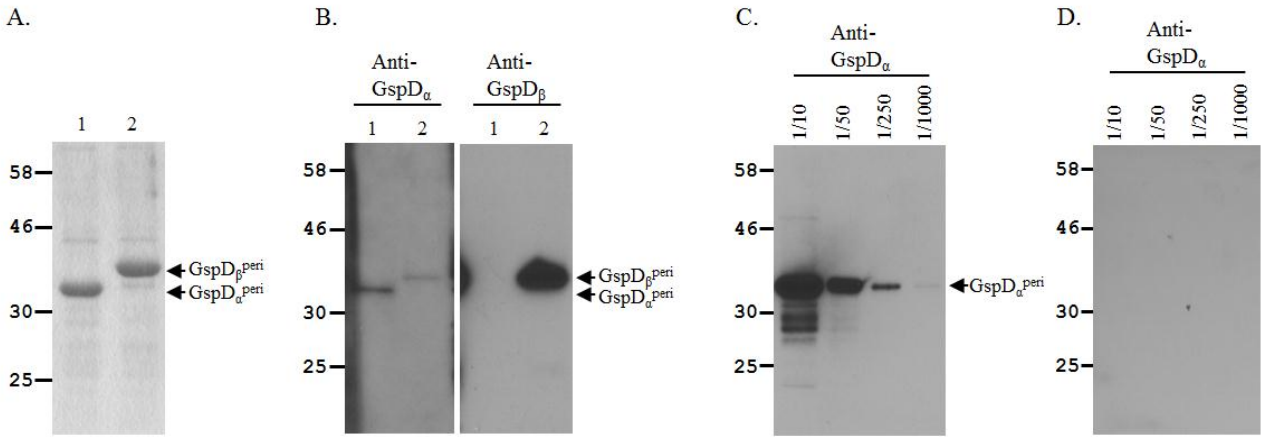


FIG. 12. Testing of anti-GspD$_\alpha$ and anti-GspD$_\beta$ antibodies for cross-reaction specificity. (A) Whole cell samples of *E. coli* BL21(DE3) containing TS18 [expressing amino acids 26-349 of GspD$_\alpha$, (GspD$_\alpha^{peri}$), sample 1] or TS19 [expressing amino acids 39-364 of GspD$_\beta$, (GspD$_\beta^{peri}$), sample 2) and induced with 0.1 mM IPTG, were separated by SDS-PAGE and stained with coomassie blue to detect induced expression of GspD$_\alpha^{peri}$ and GspD$_\beta^{peri}$. (B) The same strains as described in A were induced with 0.01 mM IPTG and immunoblotted with non-purified anti-GspD$_\alpha$ and anti-GspD$_\beta$ antibodies (1/100 000 dilution) as shown. Whole cell samples of *E. coli* BL21(DE3) containing TS18 (C) or TS19 (D) induced with 0.2 mM IPTG were diluted as shown and immunoblotted with purified anti-GspD$_\alpha$ (1/10 000 dilution). The position of GspD$_\alpha^{peri}$ and GspD$_\beta^{peri}$ is shown on the right and the location of protein markers is given on the left.

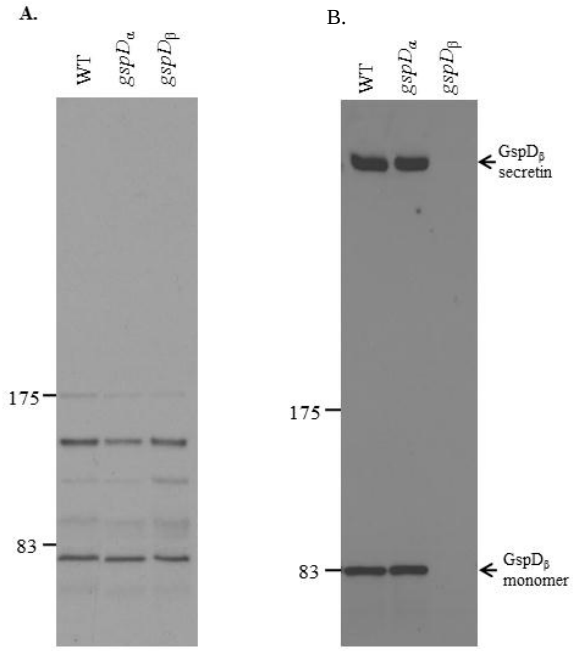


FIG. 13. In ETEC strain H10407 the T2SS$_\beta$ is assembled whereas the T2SS$_\alpha$ is not. The amount of secretin multimer and monomer in whole cell samples of H10407, *gspD*$_\alpha$ and *gspD*$_\beta$ strains was determined by immunoblot with (A) anti-GspD$_\alpha$ and (B) anti-GspD$_\beta$ antibody. The location of GspD$_\alpha$ and GspD$_\beta$ multimer and monomer is given at the right side of immunoblot and the location of standard protein markers is given on the left.

13B). Therefore according to these data, although H10407 encodes two T2SSs, under standard laboratory conditions the $T2SS_\beta$ is assembled and functional whereas the $T2SS_\alpha$ is not.

3.2.3 Assembly of the $GspD_\beta$ secretin is growth condition-dependent

The amount of LT secreted by ETEC is highly dependent upon growth conditions used for culture *in vitro*. Maximal secretion levels were observed at a temperature of 37^{o}C (Kunkel and Robertson, 1979) and in a pH range of 7.5-8.0 (Mundell *et. al.*, 1976). Addition of glucose to the growth media has also been shown to induce secretion of LT (Kunkel and Robertson, 1979; Gibert and Barbe, 1990) as well as microaerophilic conditions and increased salt concentrations (greater than 171 mM) (Trachman and Yasmin, 2004). In general, *in vitro* conditions that simulate the environment of the human small intestine have been shown to induce LT secretion, with culture in CAYE media as the standard growth condition for investigation of LT secretion by *E. coli* (Sack *et. al.*, 1980). In a study by Dorsey *et. al.*, 2006, the concentration of LT in culture supernatant was shown to be much greater when ETEC strain H10407 was cultured in CAYE media than in LB as had been previously described, but that the difference in secretion was not due to differential expression of the $T2SS_\beta$ operon or *eltAB* when cultured in these media since a constitutive level of transcription of $gspE_\beta$, $gspM_\beta$ and *eltB* was observed. The authors therefore contributed the difference in LT secretion to a post-transcriptional event that regulates LT secretion. As shown in Fig. 14, the amount of $GspD_\beta$ secretin detected in wild-type and $gspD_\alpha$ strains of H10407 when cultured in CAYE media is much greater than that observed when cultured in LB. Therefore it is likely that a post-transcriptional event that regulates expression or assembly of the $GspD_\beta$ secretin and the $T2SS_\beta$ is responsible for the difference in LT secretion observed when ETEC is cultured in LB and CAYE media.

3.2.4 Ability of $T2SS_\alpha$-inducible strains to assemble $GspD_\alpha$ secretin

The inability to detect the $GspD_\alpha$ secretin multimer or monomer in ETEC (Fig. 13) suggested that similar to *E. coli* K-12, the H10407 $T2SS_\alpha$ operon is not expressed in conditions utilized in the laboratory. In K-12, the $gspA_\alpha$ and $gspC_\alpha$ operons are cryptic due to repression by the heat-stable nucleoid-structuring protein (H-NS) when grown under laboratory conditions (Francetic and Pugsley, 1996). In a study by Francetic *et. al.*, (2000), attempts to identify a

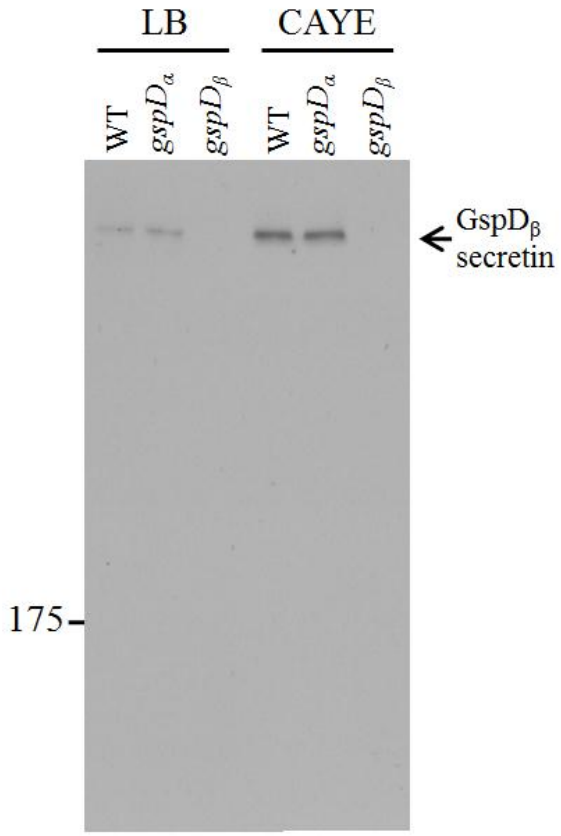


FIG. 14. In ETEC strain H10407 the amount of assembled T2SS$_\beta$ is growth condition-dependent. The amount of secretin multimer in whole cell samples of H10407, *gspD*$_\alpha$ and *gspD*$_\beta$ strains cultured in LB and CAYE media was determined by immunoblot with anti-GspD$_\beta$ antibody. The location of the GspD$_\beta$ multimer is given at the right side of immunoblot and the location of standard protein markers is given on the left.

condition that would induce expression of *gspAB*$_\alpha$ and *gspC-O*$_\alpha$ operons identifiable by expression of *gspA-lacZ* and *gspC-lacZ* reporter fusions were unsuccessful since a variety of conditions tested including anaerobiosis, carbon source variation, serum or tissue culture medium, and iron starvation did not affect transcription of the reporter fusion genes. However, when the *gspAB* and *gspC-N* operons were encoded on plasmids and expressed in an *hns* strain of MG1655, the T2SS was assembled and functional in secretion of substrates such as an endochitinase ChiA (Francetic *et. al.,* 2000) and LT that remained associated with outer membrane vesicles (OMVs) (Horstman and Kuehn, 2002).

Investigation of the role the T2SS$_\alpha$ may perform in secretion of LT would require a T2SS$_\alpha$-positive strain. Therefore to create a strain capable of assembling the T2SS$_\alpha$, the cryptic *gspA*$_\alpha$ and *gspC*$_\alpha$ promoters of the T2SS$_\alpha$ operon were replaced with inducible promoters Ptac and/or Pbad in the genome. Promoter replacement was used instead of introducing plasmids that encode the *gspAB* and *gspC-O* operons into ETEC because promoter replacement would allow more stringent control over assembly of the T2SS$_\alpha$ and because expression of the T2SS$_\alpha$ from plasmids would likely require expression in an *hns* mutant of ETEC since previous studies (as described above) have shown that T2SS$_\alpha$ assembly that is plasmid-encoded required expression in an *hns* strain of *E. coli* MG1655 (Francetic *et. al.,* 2000; Horstman and Kuehn, 2002). Since *hns* is a global regulator of the transcriptome, deletion of *hns* would introduce a non-physiologically relevant situation that would have a profound effect on the physiology of the cell. Therefore, lambda Red recombination was used to create T2SS$_\alpha$ positive strains by replacement of the nascent *gspA*$_\alpha$ and *gspC*$_\alpha$ promoters with IPTG-inducible Ptac and arabinose-inducible Pbad promoters respectively in strain TGS49. In strain TGS48, *gspA*$_\alpha$ and *gspC*$_\alpha$ promoters were replaced by Pbad and Ptac promoters respectively. To create strains that were T2SS$_\alpha$ positive and T2SS$_\beta$ negative, the *gspD*$_\beta$ gene was deleted in strains TGS49 and TGS48 to create strains TGS59 and TGS60 respectively. As shown in Fig. 15, in strains TGS49 and TGS59 detection of GspD$_\alpha$ in both monomeric and multimeric form was under stringent induction control since in the absence of arabinose that serves as the inducer for the Pbad-*gspC* promoter, GspD$_\alpha$ was not observed by immunodetection. In the presence of 0.1% Ara, a small amount of GspD$_\alpha$ monomer was detectable and upon increased induction with Ara to 1%, GspD$_\alpha$ was detected in both multimeric and monomeric forms (Fig. 15A lane 6; Fig. 15B, lane 3). In strains TGS48 and TGS60 however, the GspD$_\alpha$ monomer was observed both in the

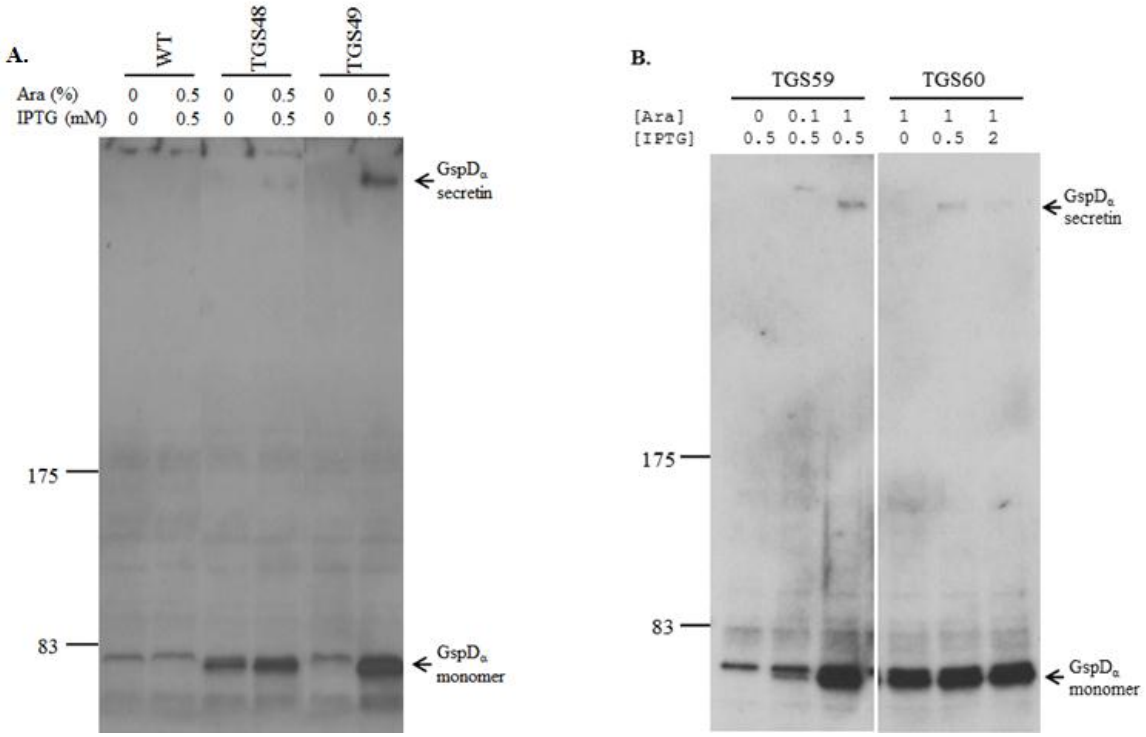


FIG. 15. Replacement of cryptic *gspA* and *gspC* T2SS$_{\alpha}$ promoters with inducible Ptac and Pbad promoters allows induction of GspD$_{\alpha}$ expression and secretin assembly. The anti-GspD$_{\alpha}$ antibody is shown to be capable of detecting the GspD$_{\alpha}$ monomer and multimer in strains in which the cryptic T2SS$_{\alpha}$ promoters have been replaced by P$_{tac}$ and P$_{bad}$ promoters (induction levels are given at top of figure). Whole cell samples were taken from strains (A) wild-type (WT), TGS48 (Pbad-*gspAB*$_{\alpha}$, Ptac-*gspC-O*$_{\alpha}$), TGS49 (Ptac-*gspAB*$_{\alpha,}$ Pbad-*gspC-O*$_{\alpha}$) and (B) TGS59 (Ptac-*gspAB*$_{\alpha,}$ Pbad-*gspC-O*$_{\alpha}$, *gspD*$_{\beta}$) and TGS60 (Pbad-*gspAB*$_{\alpha}$, Ptac-*gspC-O*$_{\alpha}$, *gspD*$_{\beta}$) grown in the presence or absence of inducer molecules arabinose and IPTG as shown. The positions of GspD$_{\alpha}$ and GspD$_{\beta}$ secretin and monomer are labeled at the right side of the panel (where applicable) and the location of standard protein markers is given at the left side of each panel.

presence or absence of the *gspC-O* inducer IPTG. Therefore expression from the Ptac-*gspC*$_\alpha$ promoter was leaky and did not allow stringent control over expression of the *gspC-O* operon. These data identified that by replacement of the endogenous *gspA*$_\alpha$ and *gspC*$_\alpha$ promoters with inducible ones, the T2SS$_\alpha$ was assembled (as detected by GspD$_\alpha$ secretin assembly) and that the antibody raised against the putative periplasmic domain of GspD$_\alpha$ was successful in detection of GspD$_\alpha$ in multimeric form.

3.2.5 The majority of LT secreted by ETEC is soluble and the T2SS$_\alpha$ does not secrete LT.

Most reviews on the subject of LT secretion state that the majority of LT remains associated with the outer membrane and outer membrane vesicles following secretion by the T2SS$_\alpha$. The results presented in Fig. 10 suggest the contrary since the T2SS$_\beta$ was absolutely required for LT secretion. Since the T2SS$_\alpha$ is not expressed in laboratory conditions, assay to determine the potential involvement of T2SS$_\alpha$ in secretion of LT would require a T2SS$_\alpha$-positive strain. As shown in Fig. 16, growth of T2SS$_\alpha$-expressing strain TGS59 in conditions shown to result in GspD$_\alpha$ secretin assembly did not result in the secretion of LT in both soluble and vesicle-associated forms. As a result, the T2SS$_\alpha$ is likely not capable of secreting LT.

To assess the amount of LT in soluble and vesicle-complexed forms, culture supernatant from wild-type, TGS59 (cultured in the presence and absence of inducer), *gspD*$_\alpha$, *gspD*$_\beta$ and *eltAB* strains were assayed for LT by GM1-ELISA before and after removal of vesicles from the supernatant. Vesicles were removed from cell supernatant both by ultracentrifugation at 100,000 x g and by passage through a 300 kDa filter. The filtration step was added to ensure that all vesicles would be removed from cell supernatant since the 300 kDa molecular weight cut-off of the filter would not allow vesicular material to pass through. As shown in Fig. 16, removal of vesicles from cell supernatant from wild-type culture did affect the level of LT detected. Ultracentrifugation to remove vesicles decreased the amount of LT assayed from supernatant from wild-type culture by 10% whereas passage through a 300 kDa filter decreased the amount of LT assayed by 19.8% in comparison to wild-type. In supernatant taken from the *gspD*$_\alpha$ strain, ultracentrifugation removed 5.9% of LT from supernatant whereas passage through a 300 kDa filter removed 19.5%. As a result, an appreciable amount of LT is present in a complexed form that is greater than 300 kDa and likely in vesicle-associated form. These data

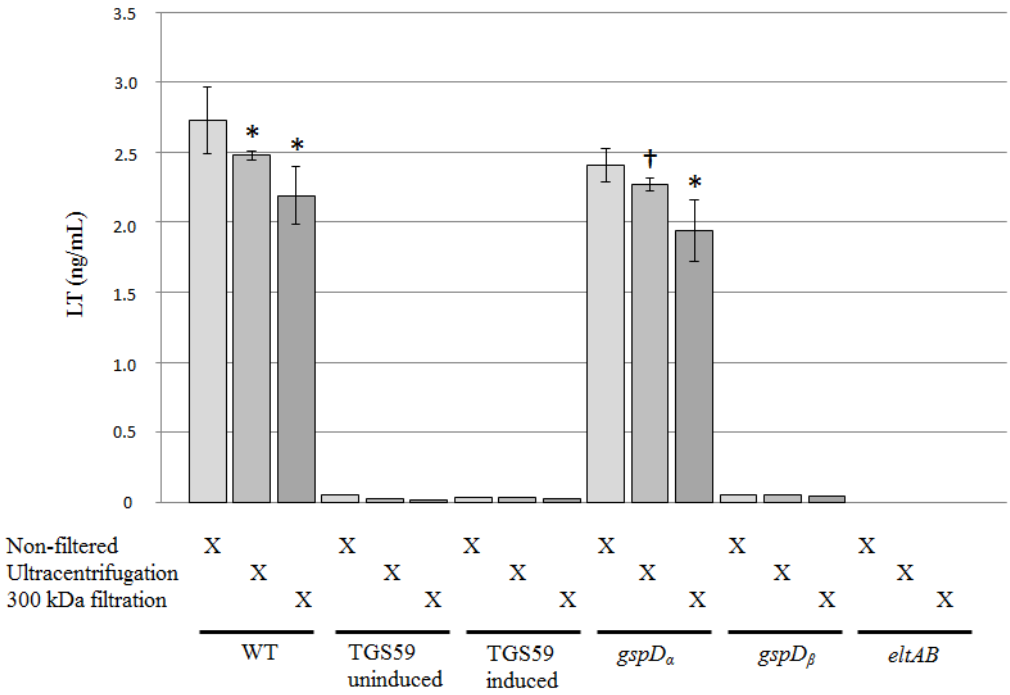


FIG. 16. The majority of LT is soluble and not vesicle-associated. The amount of LT assayed in supernatant from triplicate cultures of wild-type H10407 (WT), TGS59, TGS59 (cultured in inducing conditions; 1% Ara and 0.5 mM IPTG), *gspD*$_{\alpha}$, *gspD*$_{\beta}$ and *eltAB* strains was assessed by ELISA. Significant difference († denotes non-significance and * denotes significance where $p<0.05$) was calculated based on comparison of each value against the result from non-filtered supernatant.

also demonstrated that approximately 80% of LT secreted by ETEC was soluble and therefore soluble LT exists as the dominant form of LT secreted by the cell.

The results in Fig. 16 also identified that the T2SS$_\beta$ was absolutely required for secretion of LT in both soluble and vesicle-associated (>300 kDa) forms since only in the wild-type and *gspD*$_\alpha$ strains was LT observed in supernatant and not in T2SS$_\beta$-negative strains TGS59 and *gspD*$_\beta$. Therefore it is likely that a portion of LT is associated with vesicles as previously suggested but that the vast majority of LT remains soluble. In both cases however, the T2SS$_\beta$ is solely responsible for LT secretion.

3.3 Discussion

In numerous *E. coli* pathotypes the T2SS has been shown to be required for pathogenesis of human and animal hosts, thereby signifying this system as a major virulence factor in this species. *In silico* analysis of the H10407 genome revealed that this ETEC strain encodes two T2SSs, designated alpha and beta (Fig. 7), both of which have been identified as capable of secreting the major ETEC virulence factor LT albeit by different mechanisms (Horstman and Kuehn, 2002; Tauschek *et. al.*, 2002). The T2SS$_\alpha$ when expressed in an *hns* (histone-like nucleoid structuring protein) strain of *E. coli* K-12 containing a plasmid-encoded copy of *eltAB* was capable of secreting LT that remained associated with the outer membrane and became loaded onto outer membrane vesicles (Horstman and Kuehn, 2002). The T2SS$_\beta$ has been shown in ETEC strain H10407 to be responsible for secretion of LT into culture supernatant (Tauschek *et. al.*, 2002).

Since both systems had been shown in the literature as capable of secreting LT and conflicting reports on the dominant mode of LT secretion (soluble or vesicle-associated) exists in the literature, the subject of this study was to determine which T2SS and which mode of secretion was responsible for secretion of LT. To investigate these topics, the ETEC strain H10407 was used because the strain was a wild-type isolate, was available from ATCC and because the genome of this strain had been sequenced, thereby facilitating genetic manipulation of this strain.

Detection of the level of LT in culture supernatant was assessed initially by immunodetection and later by GM1-ELISA. Initial attempts to detect LT in culture supernatant by immunoblot were not successful (data not shown) even following concentration of

supernatant 50 to 60 fold by TCA precipitation. Therefore to increase the amount of LT to levels detectable by immunoblot, *eltB* was also expressed from a plasmid (in addition to the endogenous levels of LT expression) *in trans* in ETEC strains and the requirement for each T2SS in secretion of LT-B was determined by immunoblot (Fig. 9). The result of this analysis was perplexing since the absence of $T2SS_\beta$ abrogated detection of LT-B in supernatant yet the relative amount of LT-B in $gspAB_\alpha$ and $gspD_\alpha$ mutants were similar and less than wild-type, thereby suggesting that both systems could be involved in LT secretion. However, assay of endogenous levels of LT in culture supernatant by GM1-ELISA in $gspD_\alpha$ and $gspD_\beta$ strains confirmed that the $T2SS_\beta$ was solely required for secretion of LT when ETEC was cultured in laboratory conditions and that the $T2SS_\alpha$ was not required for this function (Fig. 10).

The $T2SS_\alpha$ operon of H10407 is identical at the nucleotide level to that of *E. coli* K-12 and similar to K-12, the $T2SS_\alpha$ is not assembled in H10407 under standard laboratory conditions (Fig. 13) presumably due to repression of the *gspA* and *gspC* promoters by HNS (Francetic and Pugsley, 1996). This scenario likely exists in ETEC since replacement of $gspA_\alpha$ and $gspC_\alpha$ promoters with IPTG and/or arabinose-inducible promoters enabled detection of $GspD_\alpha$ in both monomeric and multimeric forms (Fig. 15). It is therefore most likely that the specific *in vivo* or environmental conditions required for $T2SS_\alpha$ expression have not been duplicated in the laboratory. *In silico* analysis of representative strains of *E. coli* pathotypes identified that presence of a complete $T2SS_\alpha$ operon in the genome is not highly conserved since EHEC strain O157:H7, AIEC strain LF82, EPEC strain O127:H6 and EAEC strain 55989 do not encode this system whereas UPEC strain UTI89, APEC strain O1 strains and ExPEC strain IHE3034 do (Fig. 7).

The $T2SS_\beta$ is assembled and functional in ETEC when grown in laboratory conditions, since $GspD_\beta$ is readily detectable in both multimeric (secretin) and monomeric forms (Fig. 13) and deletion of $gspD_\beta$ rendered the strain incapable of secreting soluble LT into culture supernatant (Fig. 10). *In silico* analysis revealed that $T2SS_\beta$ is prevalent among other *E .coli* pathotypes because it is encoded in the genomes of AIEC strain LF82, EPEC strain O127:H6, EAEC strain 55898, UPEC strain UTI89, APEC strain O1 and ExPEC strain IHE3034, thereby suggesting this T2SS could function as a virulence factor in these pathotypes (Fig. 7).

The work done by Meta Kuehn's laboratory identified an ability of the $T2SS_\alpha$ when expressed in *E. coli* K-12 to secrete LT that remained bound to the cell via an interaction with

LPS on the surface of ETEC cells by binding to the outer core (Horstman and Kuehn, 2002) and the Kdo sugar (Horstman *et. al.*, 2004) of LPS. In this manner, LT would be released by the cell when vesicles bud from the outer membrane of ETEC, thereby constituting a virulence mechanism of ETEC pathogenesis that involves release of toxic packets of outer membrane material loaded with LT. Several reviews on the subject of the heat-labile enterotoxin (Mudrak and Kuehn, 2010; Patrick *et. al.*, 2010) suggest that release of LT-loaded outer membrane vesicles constitute the main virulence mechanism of ETEC. In direct contrast to the these studies and consistent with that described by Tauschek *et. al.*, 2002, the $T2SS_{\beta}$ was solely required for secretion of LT (Figs. 10 and 16), even when the $T2SS_{\alpha}$ was assembled in ETEC (Figs. 15 and 16). In addition, the majority of LT secreted was in soluble form and not vesicle-associated (Fig. 16). Therefore two possibilities exist to explain these findings. The first is that *in vivo* conditions are not recapitulated in the laboratory and in fact the $T2SS_{\alpha}$ is induced and functional in secretion of LT in vesicle-mediated form, the major mode of LT transmission *in vivo*. This possibility is unlikely however since the inability to assemble a functional $T2SS_{\beta}$ by deletion of $gspM_{\beta}$ in ETEC strain H10407 rendered the cell incapable of colonizing the mouse intestine (Dorsey *et. al.*, 2006). Therefore a second possibility is more likely, that the $T2SS_{\beta}$ is solely responsible for secretion of LT in both soluble and vesicle-associated forms *in vivo* and that the soluble form of LT is the predominant mode of LT transmission. Given that CT does not bind to the surface of *V. cholerae* due to the phosphorylation of *Vibrio* LPS (Horstman *et. al.*, 2004) and therefore a vesicle-mediated mechanism of CT transmission does not exist, it is more likely that similar to CT, soluble LT constitutes the major mode of LT transmission whereas vesicle-associated LT is a minor mode of LT transmission.

Chapter 4. Investigation of the pilotin function of YghG

4.1 Introduction

The GspD protein of the T2SS is a member of a family of outer membrane transporters termed the secretins. Other members of this family include those of the type III secretion system (T3SS), toxin co-regulated pili, type IV pili, type IV bundle-forming pili, and filamentous phage secretion systems (reviewed by Thanassi and Hultgren, 2000). In each system, the secretin functions as the outer membrane pore through which proteins or macromolecular complexes are translocated. In some systems, localization of the secretin in the

outer membrane requires the function of a small lipoprotein that serves as a pilotin to direct the secretin to the outer membrane and protect the multimer from degradation. To date, two highly similar pilotins involved in localization and protection of a T2SS secretin have been elucidated, these are $GspS^{Ko}$ of *Klebsiella oxytoca* (Hardie *et. al.*, 1996; Nouwen *et. al.*, 1999) and $GspS^{Ec}$ of *Erwinia chrysanthemi* (Shevchik and Condemine, 1998). Other pilotins involved in secretin assembly of other systems have been described and include the pilotin MxiM for assembly of the MxiD secretin of the *Shigella flexneri* type III secretion system (T3SS) (Schuch and Maurelli, 1999) and the pilotin YscW for assembly of the YscC secretin of the *Yersinia enterocolitica* T3SS (Burghout *et. al.*, 2004). The requirement for a pilotin in assembly of the T2SS secretin may not be entirely conserved in Gram-negative species since several species that encode a T2SS including *A. hydrophila* and *A. salmonicida* absolutely require the peptidoglycan-binding and ATPase functions of the inner membrane complex GspAB for secretin assembly and do not encode a recognized pilotin protein in the T2SS operon (Ast *et. al.*, 2002; Howard *et. al.*, 2006; Jahagirdar. and Howard, 1994; Li and Howard, 2010; Strozen *et. al.*, 2011).

Assembly of the $GspD_{\beta}$ secretin of the $T2SS_{\beta}$ in ETEC likely requires the function of a GspAB homologue or a pilotin protein, consistent with the requirement for these proteins in secretin assembly of other systems. *In silico* analysis of the $T2SS_{\beta}$ operon of ETEC strain H10407 did not reveal a GspS T2SS pilotin-homologue nor does it encode *gspAB*. Therefore I initially hypothesized that *gspAB* encoded in the $T2SS_{\alpha}$ operon could perform a redundant function in assembly of the secretin in both $T2SS_{\alpha}$ and $T2SS_{\beta}$ systems. Another possibility was that a gene encoded upstream of *gspC* in the atypical $T2SS_{\beta}$ operon, named *yghG*, could encode a pilotin protein since the amino acid sequence of YghG suggested that similar to pilotin proteins, YghG was an outer membrane lipoprotein.

4.2 Results

4.2.1 $GspAB_{\alpha}$ is not required for $GspD_{\beta}$ secretin assembly

I had initially hypothesized that since the $T2SS_{\beta}$ operon does not encode a protein with amino acid similarity to the GspS family of pilotins nor does it encode a GspAB homologue, that the GspAB complex encoded in the $T2SS_{\alpha}$ operon (Fig. 7) could be required for assembly of the $GspD_{\beta}$ secretin in ETEC. Immunoblot analysis of the $GspD_{\beta}$ secretin in a deletion mutant

of *gspAB*$_{\alpha}$ in ETEC however, identified that GspAB was not required for GspD$_{\beta}$ secretin assembly since the lack of GspAB had no effect on GspD$_{\beta}$ secretin assembly (Fig. 17).

4.2.2 Requirement for YghJ, PppA and YghG in secretin assembly and T2SS$_{\beta}$ function

The T2SS$_{\beta}$ operon is atypical (in comparison to T2SS operons encoded in other species) by including expression of genes *yghJ*, *pppA* and *yghG* as part of the operon encoded upstream of *gspC* (Yang *et. al.*, 2006). Since these genes are part of the T2SS$_{\beta}$ operon, they likely encode structural components or substrates of the T2SS$_{\beta}$. Therefore to more fully define the components of this system and the role these proteins perform in both assembly and function of T2SS$_{\beta}$, non-polar deletion mutants of *yghJ*, *pppA* and *yghG* were constructed by use of the lambda RED recombination system and the ability of these strains to assemble the GspD$_{\beta}$ secretin (indicative of T2SS$_{\beta}$ assembly) and secrete LT (to assess the function of T2SS$_{\beta}$) was assessed. As shown in Fig. 18, deletion of the *yghJ* gene had no effect on assembly of the GspD$_{\beta}$ secretin or secretion of LT. Deletion of *pppA* did not affect secretin assembly but did abrogate secretion of LT. This result suggested that PppA is the pre-pilin peptidase of the T2SS$_{\beta}$. Lastly, deletion of the hypothetical gene encoding the small lipoprotein YghG drastically decreased the amount of GspD$_{\beta}$ secretin assembled and amount of GspD$_{\beta}$ monomer observed, accompanied by an inability to secrete LT.

4.2.3 Complementation of the *yghG* secretin assembly-negative and LT secretion-negative phenotypes

Deletion of *yghG* resulted in nearly complete lack of GspD$_{\beta}$ in both multimeric (secretin) and monomeric forms, concomitant with abrogation of LT secretion (Fig. 18). To confirm that the secretin assembly-negative and secretion-negative phenotypes resulted from the absence of YghG, YghG and a version of YghG that contains a 10 amino acid myc sequence at the C-terminal end of the protein (to be used for immunodetection) were expressed *in trans* and the amount of secretin assembled and secretion of LT into culture supernatant was examined. Expression of YghG or YghG -myc *in trans* resulted in re-establishment of GspD$_{\beta}$ secretin and monomer levels to that of wild-type in a gradient manner such that the amount of both monomer and multimer observed increased with the amount of arabinose used to induce expression of YghG and YghG-myc (Fig. 19). Likewise, secretion of LT was re-established to

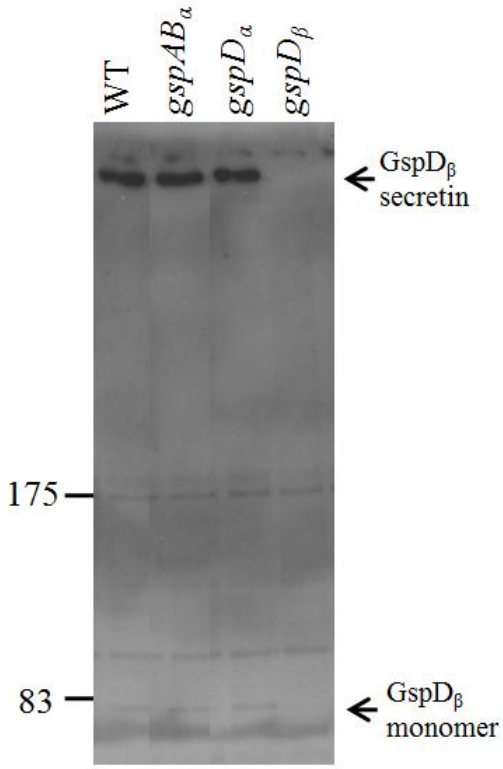


FIG. 17. Deletion of *gspAB* does not affect assembly of the GspD$_\beta$ secretin in ETEC H10407. Anti-GspD$_\beta$ immunoblot analysis of whole cell samples of H10407 wild-type (WT), *gspAB*, *gspD*α and *gspD$_\beta$* strains. The location of protein standard markers is given at left side of figure and the location of GspD$_\beta$ secretin multimer and monomer is given at right side of figure.

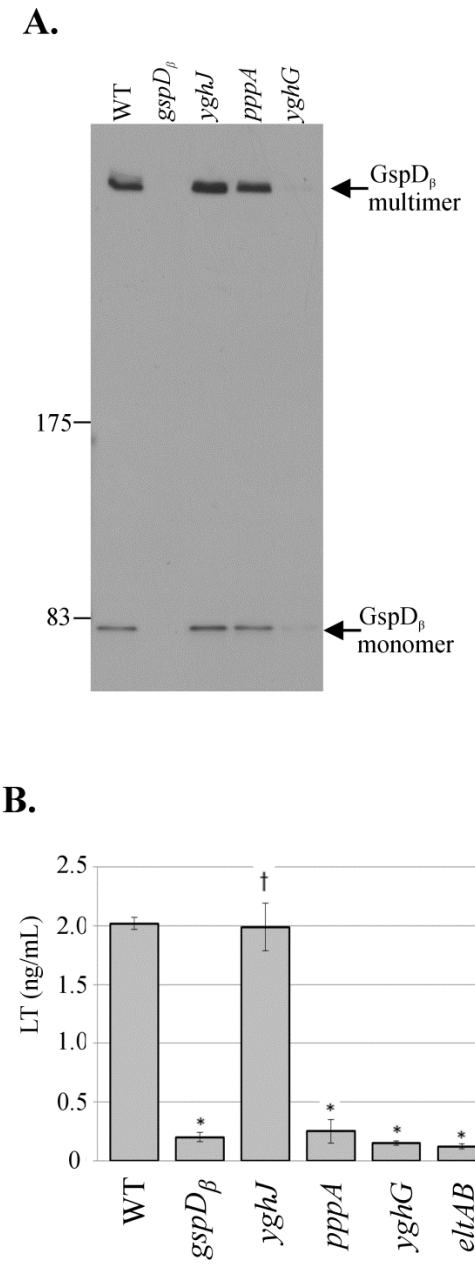


FIG. 18. Effect of mutations in hypothetical genes encoded upstream of the core *gspC-N* genes of the T2SS$_{\beta}$ operon. (A) Immunoblot analysis of whole cell samples of H10407 wild-type, *gspD*$_{\beta}$, *yghJ*, *pppA* and *yghG* strains with anti-GspD$_{\beta}$ antibody. The location of protein standard markers is given at left side of figure and the location of GspD$_{\beta}$ secretin multimer and monomer is given at right side of figure. (B) The amount of LT in supernatant taken from cultures of H10407 wild-type, *gspD*$_{\beta}$, *yghJ*, *pppA*, *yghG* and *eltAB* was assayed by ELISA. Values are shown as the average amount of LT assayed in triplicate cultures for each strain. Significant difference († denotes non-significance and * denotes significance where $p<0.05$) was calculated based on comparison of each value against that observed in wild-type culture.

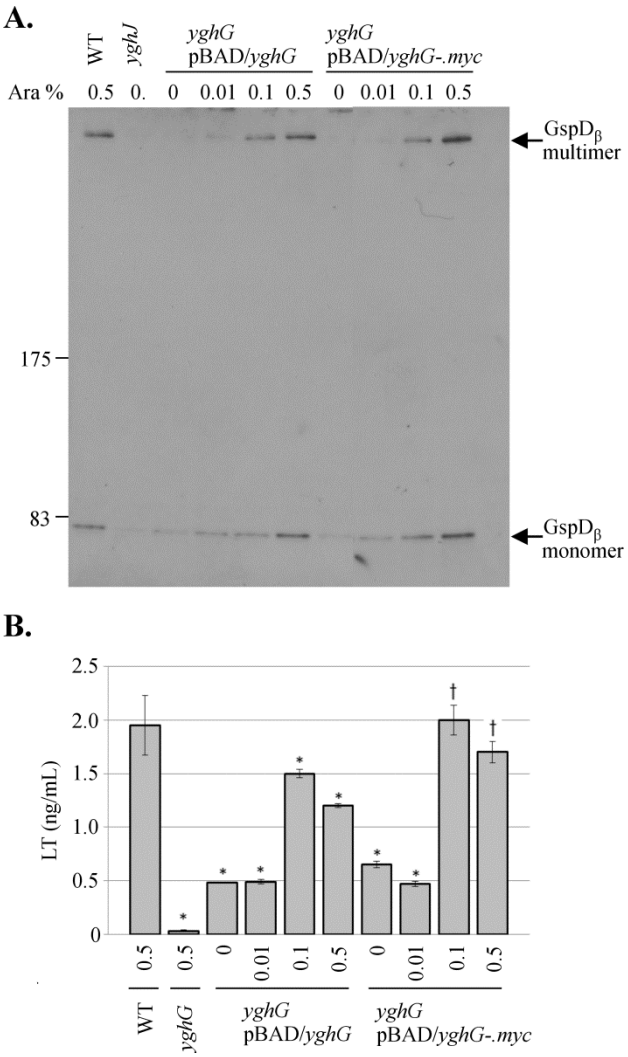


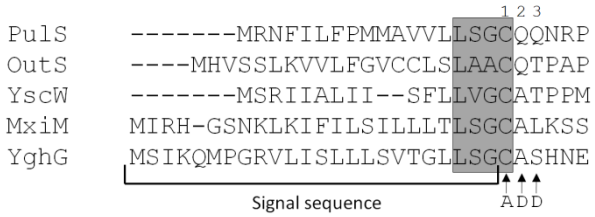
FIG. 19. The secretin assembly-negative and LT secretion-negative phenotypes of the *yghG* strain can be complemented by expression of YghG and YghG-myc *in trans*. (A) Immunoblot analysis of whole cell samples of H10407, *yghG*, *yghG* (pBAD/yghG) and *yghG* (pBAD/*yghG-myc)* with anti-GspD$_\beta$ antibody. The concentration of arabinose and standard protein markers are shown. The location of GspD$_\beta$ secretin multimer and monomer is given at right side of figure. (B) The amount of LT in supernatants from cultures of H10407, *yghG*, *yghG* (pBAD/*yghG)* and *yghG* (pBAD/*yghG*-myc) was assayed by ELISA. Values are shown as the average of triplicate assays for each strain. Significant difference († denotes non-significance and * denotes significance where $p<0.05$) was calculated based on comparison of each value against that observed in wild-type culture.

wild-type levels upon expression of YghG and YghG-myc (Fig. 19B). These results verified that the secretin assembly-negative and secretion-negative phenotypes of the *yghG* strain are due to the lack of YghG. In addition, these results also indicated that the C-terminal myc tag of YghG-myc does not interfere with the function of YghG.

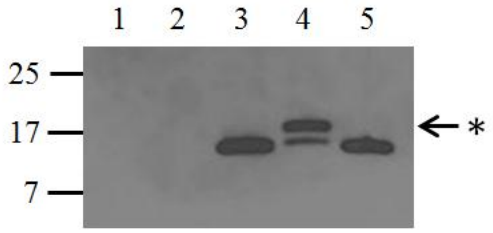
4.2.4 YghG is an outer membrane lipoprotein

The secretin assembly-negative and LT secretion-negative phenotypes of the *yghG* strain suggested that YghG could be involved in assembly of the $GspD_{\beta}$ secretin multimer in a manner similar to the GspS family of outer membrane lipoproteins that are required for assembly of the T2SS secretin in the outer membrane of *Klebsiella pneumoniae* and *Erwinia chrysanthemi* species (Hardie *et. al.*, 1996A; Nouwen *et. al.*, 1999; Shevchik and Condemine, 1998). Analysis of the N-terminal sequence of YghG revealed that like $GspS^{Ko}$ (PulS) and $GspS^{Ec}$ (OutS) of the T2SS and other outer membrane lipoproteins required for secretin assembly including YscW of the *Yersinia enterolitica* type III secretion system (T3SS) and MxiM of the *Shigella flexneri* T3SS, YghG is likely a lipoprotein because it contains a lipoprotein signal sequence comprised of a lipobox motif of the consensus sequence is L-A/S-G/A-C (Fig. 20A). As mentioned in the introduction, maturation of lipoproteins is performed at the outer leaflet of the inner membrane by addition of diacylglycerol to the Cys residue of the lipobox motif by phosphatidylglycerol/prolipoprotein diacylglyceryl transferase (Lgt), cleavage of the signal peptide by signal peptidase II (LspA) and aminoacylation of the Cys residue by phospholipid/apolipoprotein transacylase (Lnt). Therefore, to confirm that YghG was a lipoprotein, a substitution mutation was constructed to replace the cysteine located at position 25 within the lipobox motif with alanine. This substitution would be expected to prevent processing of the pre-protein by signal peptidase II due to the inability of LspA to recognize the unlipidated protein. As shown in Fig. 20B, anti-myc immunoblot of whole cell samples of cells expressing YghG C25A-myc *in trans* showed that the vast majority of YghG C25A-myc remains as a pre-protein presumably due to the inability of LspA to cleave the signal sequence. In addition, increased expression of YghG C25A-myc (generated by induction with 0.5% arabinose) conferred a growth defective phenotype possibly due to sequestration of LspA complexes with the mutant protein (data not shown).

A.



B.



C.

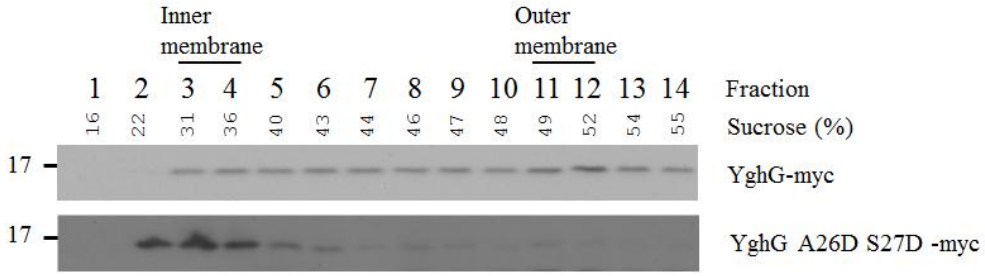


FIG. 20. **YghG is an outer membrane lipoprotein.** (A) Alignment of the N-terminal region of YghG with other lipoproteins known to be involved in secretin assembly including PulS from *Klebsiella pneumoniae* (P20440), OutS of *Erwinia chrysanthemi* (Q01567), YscW of *Yersinia enterolitica* (Q56851) and MxiM of *Shigella flexneri* (P0A1X2). The location of the signal sequence is indicated, with the lipobox (L A/S G/A C) highlighted in grey. The location of substitution mutations C25A, A26D and S27D in YghG are shown. (B) Myc-tagged proteins present in whole cell samples of WT (lane 1), *yghG* (lane 2), yghG expressing *in trans*; YghG –myc (lane 3), YghG C25A –myc (lane 4) and YghG A26D S27D –myc (lane 5) are detected by anti-myc immunoblot. The location of the unprocessed YghG C25A variant that retains the signal peptide is indicated by an asterisk. (C) YghG is localized to the outer membrane and YghG A26D S27D is localized to the inner membrane in ETEC. YghG-myc and A26D S27D YghG were detected by anti-myc immunoblot of inner and outer membranes fractions (1-14) isolated by sucrose density gradient fractionation. Fractions containing inner and outer membrane material were determined by NADH oxidase activity and the presence of outer membrane porins respectively. The concentration of sucrose in fractions isolated from the strain expressing YghG-myc is given.

The amino acid sequence of YghG suggested that the protein would be located in the inner leaflet of the outer membrane since residues +2 and +3 of the mature lipoprotein are not Asp residues as shown in Fig. 20A. During lipoprotein maturation, lipoproteins are sorted to the inner or outer membrane according to the lipoprotein sorting signal located at positions 2 and 3 of the mature protein. In general, Asp residues at position 2 (Yamaguchi *et. al.*, 1988) and 3 (Seydel *et. al.*, 1999) ensure retention of the lipoprotein in the inner membrane by evading the LolCDE complex in the inner membrane. Most other combinations of amino acids at positions 2 and 3 result in complex formation of the lipoprotein with LolCDE, transport across the periplasm by the chaperone LolA, interaction with the outer membrane lipoprotein LolB and localization to the outer membrane via the N-terminal lipid modification (reviewed by Tokuda and Matsuyama, 2004). To confirm that YghG is localized to the outer membrane, samples of inner and outer membranes from a *yghG* strain expressing YghG-myc *in trans* were separated by sucrose gradient fractionation and the location of YghG was determined by immunoblot (Fig. 20C). Although YghG-myc was present in most fractions including those of the inner membrane, the majority of YghG-myc was observed in the intermediate and outer membrane fractions. In addition, substitution of residues +2 and +3 of the mature lipoprotein with Asp residues caused YghG-myc to strongly localize within the inner membrane (Fig. 20C). These data confirmed that YghG is an outer membrane lipoprotein.

4.2.5 $GspD_{\beta}$ requires YghG for localization to the outer membrane

Pilot proteins function in localizing and protecting the secretin, therefore to determine if YghG is a pilot protein the location of the secretin when outer membrane localized (YghG-myc) and inner membrane localized (YghG A26D S27D-myc) versions of YghG were expressed in a *yghG* strain was assessed. The location of $GspD_{\beta}$ was determined by immunoblot of inner and outer membrane fractions following sucrose gradient fractionation of cell envelope preparations. As shown in Fig. 21, sucrose gradient ultracentrifugation generated 14 fractions. Fractions that contained mostly inner membrane (fractions 3 and 4) were identified by assay of the activity of the inner membrane protein NADH oxidase (Fig. 21A), whereas fractions that contained mostly outer membrane material (fractions 11 and 12) were identified by detection of outer membrane porin proteins OmpC and OmpF in samples of fractions separated by SDS-PAGE and visualized by coomassie-blue staining (Fig. 21B). It was evident from panel B that

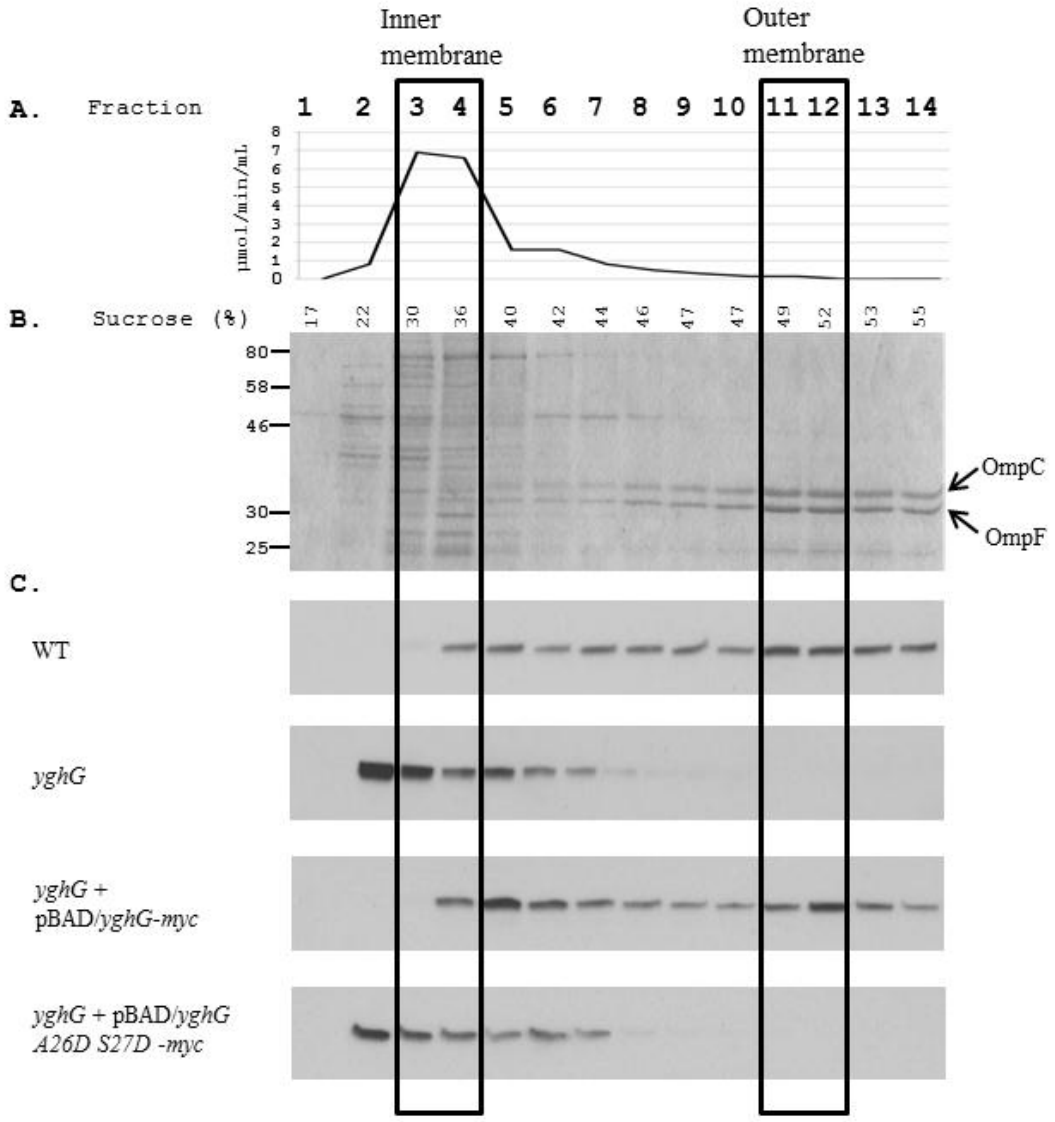


FIG. 21. GspD$_\beta$ requires outer membrane-localized YghG for localization in the outer membrane. (A) NADH oxidase activity and (B) whole protein profile of sucrose gradient ultracentrifugation fractions of inner and outer membranes of wild-type *E.coli* H10407. The concentration of sucrose observed in each fraction is given. The location of outer membrane proteins OmpC and OmpF are shown. (C) Immunoblot analysis of fractions isolated from wild-type H10407 (WT), *yghG*, *yghG* expressing *yghG*-myc *in trans* and *yghG* expressing *yghG* A26D S27D-myc *in trans* with anti GspD$_\beta$ antibody. Only the top portion of each immunoblot that corresponds to the location of secretin multimer is included due to the negligible amount of GspD$_\beta$ monomer observed after french press cell breakage (refer to results section). Fractions that contain predominantly inner (fractions 3 and 4) or outer membranes (fractions 11 and 12) are outlined.

the intermediate fractions 5 to 10 contain a mixture of both inner and outer membrane material with more inner membrane material present in fractions 5, 6 and 7 and mostly outer membrane material in fractions 8, 9 and 10. In panel C, it was observed that $GspD_{\beta}$ in wild-type H10407 mostly fractionates with the outer membrane however there was a relatively minor amount of $GspD_{\beta}$ present in the inner membrane fraction number 4. In a *yghG* mutant however, $GspD_{\beta}$ fractionated strongly with the inner membrane fractions. Expression of YghG -myc in the *yghG* strain re-establishes localization of $GspD_{\beta}$ to the outer membrane fractions, thereby modeling the situation observed in wild-type. When the inner membrane localized YghG A26D S27D was expressed in a *yghG* strain, no affect was observed in comparison to the *yghG* strain. These results showed that $GspD_{\beta}$ can only localize to the outer membrane in the presence of outer membrane-localized YghG. In the absence of YghG, $GspD_{\beta}$ was detectable and present predominantly in inner membrane fractions. These data confirm that the outer membrane localized YghG is required for localization of $GspD_{\beta}$ to the outer membrane.

The immunoblots shown in Fig. 21C are that of the $GspD_{\beta}$ secretin multimer only and do not include the area of the immunoblot that shows the $GspD_{\beta}$ monomer. This was done due to the observation that french pressure cell breakage causes spontaneous $GspD_{\beta}$ multimerization and therefore immunoblot analysis detects nearly all $GspD_{\beta}$ protein present as a multimeric complex only. As a result, this cell fractionation data could be used to determine the location of the protein as present in the inner membrane, outer membrane or neither but could not provide information regarding the relative amount of secretin multimer and monomer observed. Investigation of the relative amount of $GspD_{\beta}$ in monomeric and multimeric forms required immunoblot of whole cell samples only, as described below. Also note that although $GspD_{\beta}$ was barely detectable as monomer and multimer by immunoblot of whole cell samples in the *yghG* strain (Figs. 18 and 19), preparation of inner and outer membranes concentrates inner and outer membrane proteins considerably. As a result, relatively minor amounts of $GspD_{\beta}$ were readily detectable by immunoblot of inner and outer membranes.

4.2.6 $GspD_{\beta}$ secretin assembly requires localization to the outer membrane

Comparison of the amount of $GspD_{\beta}$ present in whole cell samples in *yghG* strains expressing inner (YghG A26D S27D-myc) or outer (YghG-myc) membrane versions of YghG revealed that the amount of $GspD_{\beta}$ observed in multimeric and monomeric form was much

greater when $GspD_{\beta}$ is localized to the outer membrane than when present in the inner membrane (Fig. 22A). In addition, expression of the inner membrane version of YghG did not alter the negligible amount of secretin multimer and monomer in the *yghG* strain. These results were not explained by a difference in amount of inner and outer membrane-localized YghG present in the cell since immunoblot detection revealed similar levels of each protein in their respective strains upon induction with arabinose (Fig. 22B). As expected, LT was not detected in culture supernatant if YghG is localized to the inner membrane, due to an absence of assembled secretin and thus functional $T2SS_{\beta}$ (Fig. 22C).

The results shown in Fig. 21 and Fig. 22 revealed that outer membrane localization of $GspD_{\beta}$ is dependent upon YghG and localization to the outer membrane is required for $GspD_{\beta}$ secretin assembly since only in the wild-type H10407 strain or in the *yghG* strain complemented with YghG-myc is $GspD_{\beta}$ present in the outer membrane fractions (Fig. 21) and detected as a multimer (Fig. 22). Therefore to confirm that localization of $GspD_{\beta}$ to the outer membrane is required for secretin assembly and to determine if the inability to detect secretin multimer when $GspD_{\beta}$ is localized to the inner membrane is not due to degradation of an inner membrane secretin that is more prone to degradation (during preparation of whole cell samples), secretin multimers of $GspD_{\beta}$ were crosslinked *in vivo* and analyzed following treatment with phenol. As an initial test, the concentration of the crosslinker DSP was varied to determine the optimal concentration required for detection of $GspD_{\beta}$ protomers. As shown in Fig. 23, addition of phenol to whole cell samples of wild-type H10407 destabilizes the $GspD_{\beta}$ secretin (lane 1) into its monomeric components (lane 2). However, addition of the crosslinker DSP to whole cells prior to treatment with phenol caused multiple bands that correspond to protomers of $GspD_{\beta}$ (likely dimers and tetramers) to be observed (lanes 3 and 4). Note that crosslinking with 0.05 mM DSP enabled the greatest amount of $GspD_{\beta}$ protomers to be observed whereas addition of 0.5 mM DSP likely saturated the cell with crosslinks since no protomers of $GspD_{\beta}$ and a minimal amount of $GspD_{\beta}$ monomer was observed in this sample. In the *gspD$_{\beta}$* strain, the absence of DSP or presence of 0.005, 0.05 and 0.5 mM DSP did not result in the detection of any bands by immunoblot. Therefore, all bands observed in the immunoblot correspond to the $GspD_{\beta}$ protein. Crosslinking of samples from *yghG* expressing *yghG-myc* or *yghG A26D S27D–myc in trans* also identified that use of 0.05 mM DSP was the optimal concentration of crosslinker for detection of $GspD_{\beta}$ protomers.

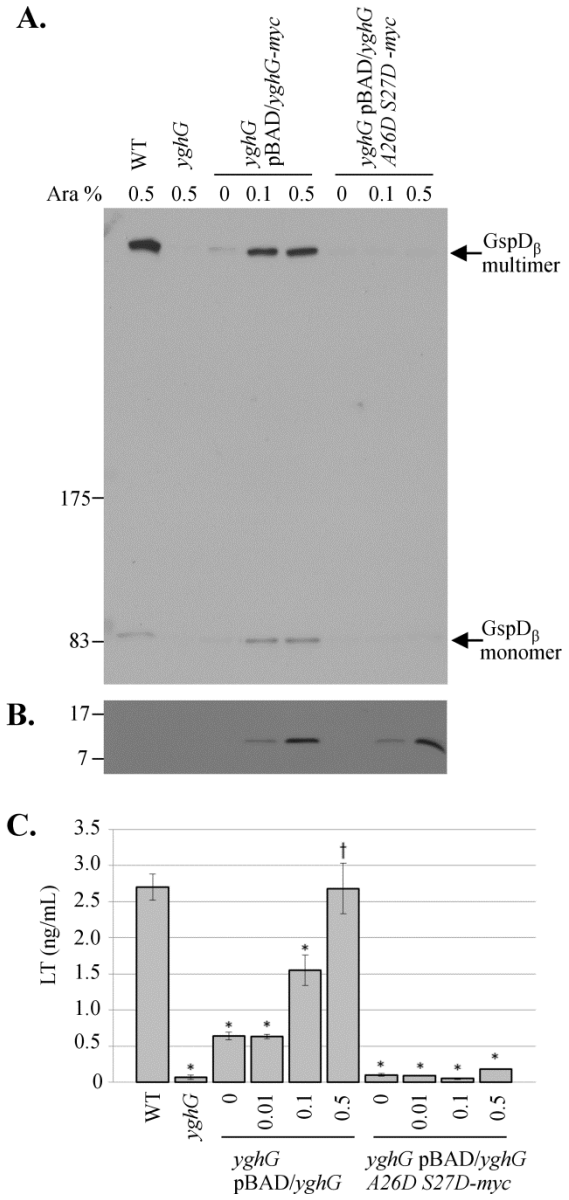


FIG. 22. GspD$_\beta$ is not detectable in the absence of outer membrane-localized YghG. Immunoblot analysis of whole cell samples of wild-type H10407 (WT), *yghG*, *yghG* expressing outer membrane localized YghG-myc or inner membrane localized YghG A26D S27D-myc with (A) anti-GspD$_\beta$ antibody or anti-myc (B) antibodies. The concentration of arabinose used for induction of plasmid-encoded proteins is given. The location of GspD$_\beta$ multimer and monomer are given. (C) Amount of LT detected in supernatant as determined by LT-ELISA. Assays of supernatant from triplicate cultures were performed with the standard deviations indicated. Significant difference († denotes non-significance and * denotes significance where $p<0.05$) was calculated based on comparison of each value against that observed in wild-type culture.

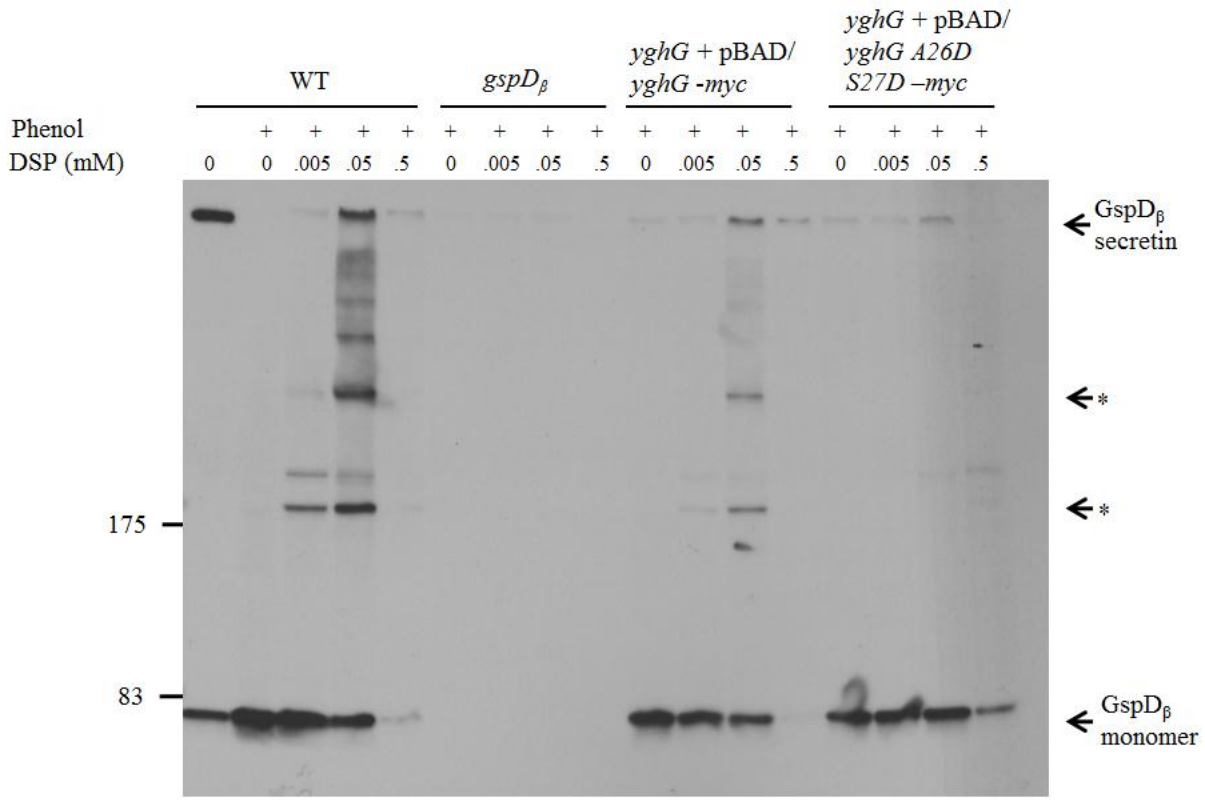


FIG. 23. Protomers indicative of GspD$_\beta$ secretin assembly are observed upon crosslinking with DSP. Immunoblot analysis with anti-GspD$_\beta$ antibody of whole cell samples of; wild-type H10407 (WT), *gspD*$_\beta$, *yghG* expressing outer membrane localized YghG-myc *in trans* or inner membrane localized YghG A26D S27D-myc *in trans*. Whole cells were treated with different concentrations of the crosslinker DSP as shown and phenol as indicated. The location of GspD$_\beta$ secretin multimer and monomer are given, asterisks indicate the positions of GspD$_\beta$ protomers.

Once the optimal concentration of the crosslinker was determined, the ability to assemble the $GspD_{\beta}$ secretin upon expression of inner membrane and outer membrane versions of YghG was determined. As shown in Fig. 24, in the absence of the crosslinker and phenol, $GspD_{\beta}$ is detectable in monomeric and multimeric forms that upon addition of phenol caused the $GspD_{\beta}$ secretin to dissociate (lane 2). Addition of the crosslinker enabled protomers of $GspD_{\beta}$ to be observed (lane 3). In the *yghG* strain, the protomers of $GspD_{\beta}$ were not detected (lane 4), however expression of YghG-myc in the *yghG* strain re-established detection of $GspD_{\beta}$ protomers in a gradient fashion whereby increased expression of YghG-myc (induced with arabinose) resulted in an increased amount of $GspD_{\beta}$ protomers observed (lanes 5-8). In the absence of outer membrane-localized YghG-myc as such is the case in the *yghG* strain (lane 4) and when the the inner membrane variant YghG A26D S27D-myc was expressed (lanes 9-12), $GspD_{\beta}$ protomers were not detected and only $GspD_{\beta}$ monomer was evident. These data therefore show that outer membrane localization of $GspD_{\beta}$ is required for secretin formation and that the $GspD_{\beta}$ that is present in the absence of outer membrane localized YghG is not present in multimeric form, that is, $GspD_{\beta}$ does not form a multimer when localized to the inner membrane.

4.2.7 YghG is a novel pilotin protein and is renamed $GspS_{\beta}$

Since YghG is encoded within the $T2SS_{\beta}$ operon, is required for assembly of the $GspD_{\beta}$ secretin and is localized to the outer membrane, YghG likely functions in a manner similar to the T2SS GspS family of pilot proteins in protection and localization of the secretin in the outer membrane. Interestingly however, amino acid alignment of YghG with the T2SS proteins $GspS^{Ko}$ (PulS), $GspS^{Ec}$ (OutS) and EtpO revealed that YghG does not share significant homology with these proteins by displaying low sequence identity (16.9), similarity (25.3) and global alignment scores (25.5) with PulS (Fig. 25A). Instead, YghG is likely part of a group of pilotins from *V. cholerae* (Q9KRD9), *Hamiltonella defensa* (C4K3F8) and *Photobacterium sp* SKA34 (Q2BX31) since these proteins exhibit significant amino acid identity, similarity and global alignment scores with YghG.

Interestingly, an alignment of the C-terminus of secretin proteins from species of the PulS/OutS and YghG families of pilotins identifies sequence conservation within PulS/OutS members and within YghG family members and not between families (Fig. 25B). Since the

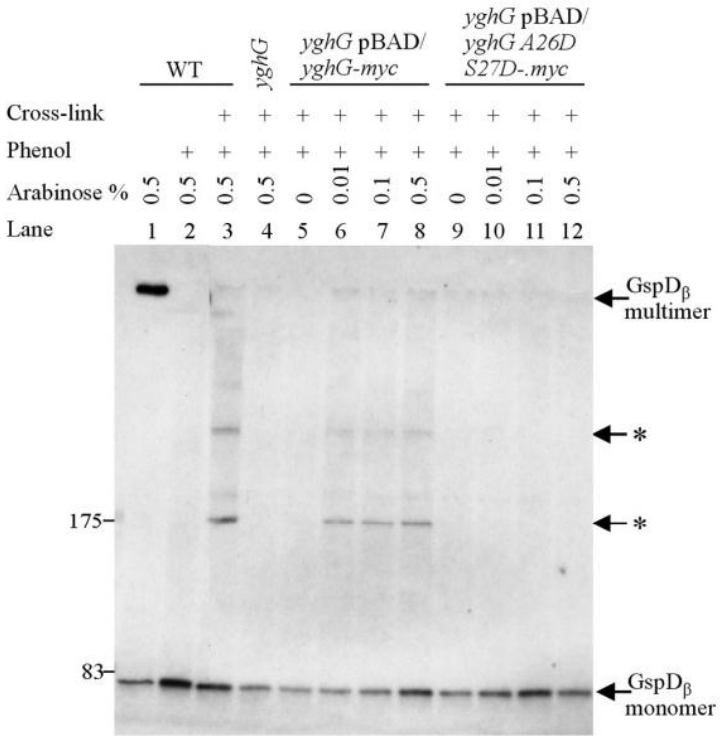
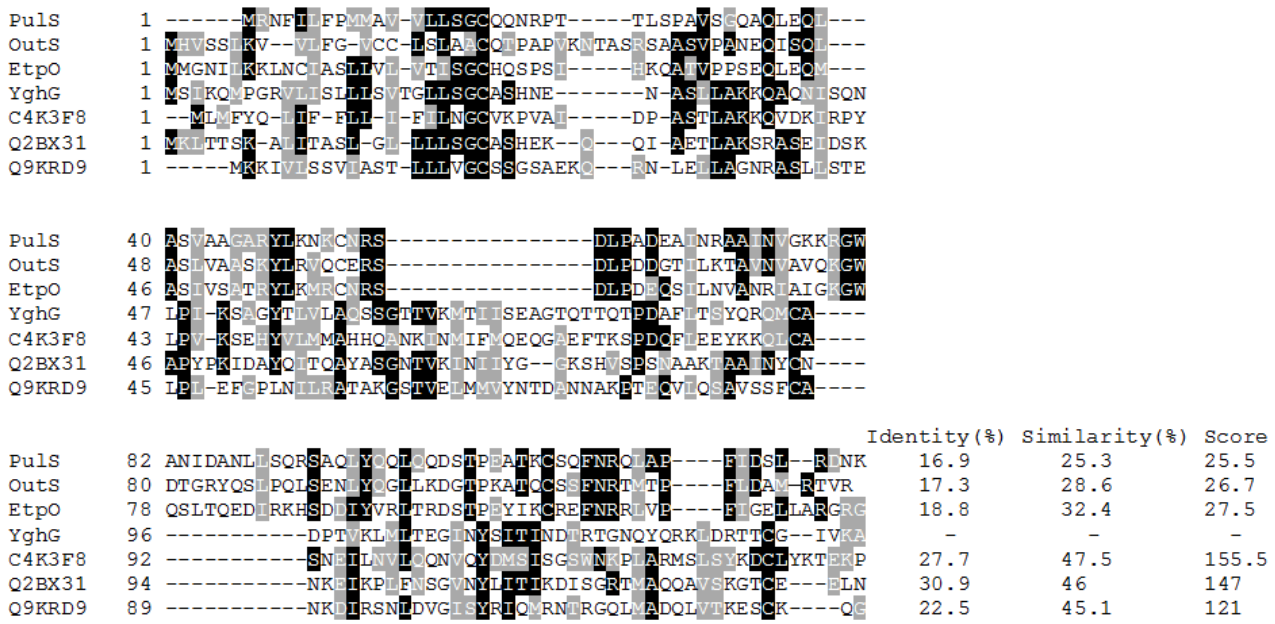


FIG. 24. Crosslinked protomers indicative of GspD$_\beta$ secretin assembly are observed only when GspD$_\beta$ is localized to the outer membrane. Immunoblot analysis with anti-GspD$_\beta$ antibody of whole cell samples of; wild-type H10407 (WT), *yghG*, *yghG* expressing outer membrane localized YghG-myc *in trans* or inner membrane localized YghG A26D S27D-myc *in trans*. Samples were treated with the crosslinker DSP and phenol as indicated. The location of GspD$_\beta$ secretin multimer and monomer are given, asterisks indicate the positions of GspD$_\beta$ protomers.

A.



B.

FIG. 25. Amino acid comparison of members of the YghG family with the PulS/OutS family of T2SS pilot proteins. (A) Full length sequence alignment of members of the PulS/OutS family of proteins including PulS from *Klebsiella oxytoca*, OutS of *Erwinia chrysanthemi* and EtpO from *E. coli* O157:H7 in addition to putative uncharacterized lipoproteins C4K3F8 from *Hamiltonella defensa*, Q2BX31 of *Photobacterium sp.* SKA34 and Q9KRD9 from *Vibrio cholerae* with YghG of H10407. (B) Amino acid alignment of the C-terminal region of secretin proteins PulD from *Klebsiella oxytoca*, OutD from *Erwinia chrysanthemi*, EtpD from *E. coli* O157:H7, EpsD from Vibrio cholerae, Q1YW23 from *Photobacterium sp.* SKA34, C4K366 from *Hamiltonella defensa* and GspD$^{\beta}$ of H10407. Percent identity, percent similarity and score were determined for each sequence in comparison to PulS by a global sequence alignment using the Needleman-Wunsch global alignment algorithm derived from the program EMBOSS (European bioinformatics institute). Identical amino acids are shown in reverse type with black background and similar amino acids are shown in reverse type with a grey background.

pilotin binding site for secretin proteins is located at the C-terminus of the secretin, this alignment identified that it is likely that the YghG homologues in *V. cholerae, Hamiltonella defensa* and *Photobacterium sp* SKA34 perform the same function in secretin assembly in their respective species as YghG performs in ETEC.

Due to the heterogeneous nature of the pilotin binding site of secretins, pilotin proteins themselves exhibit a large degree of heterogeneity, thereby making *in silico* identification of these proteins difficult. As a result, pilotin proteins are classified as such based on the similarity of the function of the protein and not solely on amino acid similarity. Therefore although YghG does not exhibit significant amino acid sequence homology with the PulS/OutS family of pilotins, the function of these proteins is similar by being required for localization of a T2SS secretin in the outer membrane, therefore a re-designation of the hypothetical protein YghG to the $T2SS_{\beta}$ pilotin protein $GspS_{\beta}$ is suggested.

4.3 Discussion

There exists an interesting conundrum in Gram-negative species regarding the requirement for accessory proteins in T2SS secretin assembly. In several species including *A. hydrophila, A. salmonicida* and *Vibrio* species *V. cholerae, V. vulnificus* and *V. parahaemolyticus* (discussed in the next section), the activity of the peptidoglycan-remodeling complex GspAB is important in T2SS secretin assembly whereby loss of GspAB causes a significant decrease in GspD secretin assembly. In secretin-containing cell envelope systems encoded within other species however, including the T2SS of *Klebsiella oxytoca* and *Erwinia chrysanthemi,* the T3SS of *Shigella flexneri* and the T4P of *Pseudomonas aeruginosa,* a GspAB homologue is not present and instead an outer membrane lipoprotein termed a pilotin is required for secretin assembly.

I was interested in the requirement for accessory proteins in assembly of the $GspD_{\beta}$ secretin in ETEC and initially thought that the *gspAB* complex encoded in the $T2SS_{\alpha}$ operon that had been previously shown as required for $GspD_{\alpha}$ secretin assembly and secretion of a chitinase in *E. coli* K-12 (Francetic and Pugsley, 1996) could perform a redundant function in ETEC by being involved in assembly of both the $GspD_{\alpha}$ and $GspD_{\beta}$beta secretins. The GspAB complex of $T2SS_{\alpha}$ however is not involved in $GspD_{\beta}$ secretin assembly since deletion of $gspAB_{\alpha}$ did not have an effect on secretin assembly (Fig. 17). Therefore another factor, possibly

a pilotin could be involved in $GspD_{\beta}$ secretin assembly. In a study by Yang *et. al.*, 2006, three hypothetical ORFs located upstream of the $gspC_{\beta}$ gene were identified as part of the $T2SS_{\beta}$ operon. These genes *yghJ*, *pppA* and *yghG* encoded a large lipoprotein, a putative pre-pilin peptidase and a small lipoprotein respectively. *In silico* analysis of the *yghG* lipoprotein revealed that it is a small lipoprotein (comparable in size to other pilotins) that is likely localized to the outer membrane. Therefore, I hypothesized that similar to other T2SSs that do not encode a GspAB complex but instead express a pilotin that is absolutely required for secretin assembly, YghG could function as a pilotin for $GspD_{\beta}$ secretin assembly in ETEC.

The *yghG* ORF encodes a putative 136 amino acid lipoprotein likely localized to the inner leaflet of the outer membrane due to the presence of an N-terminal signal sequence, a lipobox domain and amino acids that are not Asp residues at positions +2 and +3 of the hypothetical mature lipoprotein (Fig. 20A). Immunoblot detection of YghG in outer membrane fractions (Fig. 20C) following separation of inner and outer membranes of a *yghG* strain expressing YghG-myc *in trans*, confirmed that the protein is localized to the outer membrane. Substitution of the lipid-modified Cys residue of the mature lipoprotein with Ala prevented processing of the pre-protein (Fig. 20B), presumably by an inability of signal peptidase II (LspA) (Gonnet *et. al.*, 2004) to cleave the signal peptide. Lastly, YghG could be localized to the inner membrane by substitution of amino acids 2 and 3 of the mature lipoprotein with Asp residues (Fig. 20C), an alteration known to prevent detection of the protein by the LolCDE system required for transport of lipoproteins to the outer membrane. Together these results verified that YghG is an outer membrane lipoprotein.

The functions of the proteins encoded upstream of *yghG* in the $T2SS_{\beta}$ operon (*yghJ* and *pppA*) are also unknown. The *yghJ* gene encodes a large 1520 aa putative lipoprotein that is homologous to the accessory colonization factor D (AcfD) of *Vibrio cholerae* required for colonization of the mouse intestine (Peterson and Mekalanos, 1988) and resistance to the vibriocidal activity of anti-vibrio whole cell antibodies and complement (Parsot *et. al.*, 1991). Interestingly in a study by Moriel *et. al.*, 2010, anti-YghJ antibodies were found to be highly immunoprotective in mice upon challenge with ExPEC strain IHE3034. In this strain, the $T2SS_{\beta}$ was shown to be capable of secreting YghJ into culture supernatant. These data suggest that YghJ could be a substrate of the $T2SS_{\beta}$ and is not likely a structural component of $T2SS_{\beta}$. The putative designation of YghJ as a secreted protein is consistent with the results described in

this study since deletion of *yghJ* did not affect assembly of the $GspD_{\beta}$ secretin or function of $T2SS_{\beta}$ (Fig. 18).

The $T2SS_{\beta}$ *pppA* gene encodes a protein homolog of T2SS prepilin peptidases that are required for processing of T2SS pseudopilin subunits GspG, H, I, J, and K. Deletion of PppA did not affect assembly of the $T2SS_{\beta}$ (Fig. 18A) since the secretin multimer remained assembled in the *pppA* strain but loss of PppA did prevent secretion of LT (Fig. 18B). These results suggested that PppA is likely the pre-pilin peptidase of $T2SS_{\beta}$ since presumably the pseudopilus was not assembled thereby rendering the system non-functional. In fact, the vestigial $T2SS_{\beta}$ PppA of *E. coli* K-12 has been shown as capable of processing the evolutionarily similar type IV prepilin subunits of *Neisseria gonorrhoeae* and *Klebsiella oxytoca* (Francetic *et. al.,* 1998). This finding was not surprising given some species including *P. aeruginosa* and *A. hydrophila* encode one prepilin peptidase (GspO) that is required for processing of both T2SS pseudopilins and type IV pilins (Pepe *et. al.,* 1996).

Non-polar chromosomal deletion of *yghG* prevented assembly of the $GspD_{\beta}$ secretin (Fig. 18A) thereby rendering the $T2SS_{\beta}$ non-functional and resulting in the inability to secrete LT (Fig. 18B). Confirmation that the secretin assembly-negative and secretion-negative phenotypes of the *yghG* strain are attributable to a lack of YghG was verified by complementation with YghG or YghG-myc *in trans* (Fig. 19), thereby suggesting that the protein functions as a pilot protein required for $GspD_{\beta}$ secretin assembly. Consistent with the function of pilotin proteins, YghG is required for localization of $GspD_{\beta}$ to the outer membrane and in the absence of YghG, $GspD_{\beta}$ is localized to the inner membrane (Fig. 21).

Mislocalization of secretin proteins due to the absence of the pilotin protein has been previously shown to result in multimeric complexes observed in the inner membrane (Hardie *et. al.,* 1996B). In the ETEC strain H10407 however, the vast majority of the $GspD_{\beta}$ is degraded when not localized to the outer membrane and does not form a secretin in the inner membrane in the absence of its pilotin YghG. In whole cell samples of strains expressing the inner membrane or outer membrane variant of YghG (Fig. 22), it was evident that the total amount of $GspD_{\beta}$ present in secretin and monomeric forms was much greater when the outer membrane version of YghG was expressed (which localizes $GspD_{\beta}$ to the outer membrane) than when the inner membrane version of YghG was expressed. Unlike PulS that is capable of preventing PulD degradation when localized to the inner membrane (Hardie *et. al.,* 1996A), expression of

an inner membrane version of YghG (YghG A26D S27D) is incapable of preventing degradation of $GspD_{\beta}$ since the negligible amount of multimer and monomer observed in the *yghG* strain remained unchanged upon increased expression of YghG A26D S27D-myc. As a result, it is likely that localization of $GspD_{\beta}$ to the outer membrane and not YghG *per se*, is required to prevent degradation of $GspD_{\beta}$. In this manner, YghG is similar to the *Yersinia enterocolitica* T3SS pilot protein YscW that does not prevent proteolytic degradation of YscC unassembled monomers prior to secretin assembly but is solely required for localization of YscC to the outer membrane (Burghout *et. al.*, 2004). In addition, the $GspD_{\beta}$ secretin is not observed in the inner membrane in the absence of the pilotin YghG. As shown in Fig. 24, crosslinked protomers of $GspD_{\beta}$ were only observed when $GspD_{\beta}$ is localized to the outer membrane upon expression of YghG and were observed solely in monomeric form when not localized to the outer membrane. This situation is similar to that observed for the type IV pilus secretin PilQ of *Pseudomonas aeruginosa* that in the absence of its pilotin PilF, does not form secretin multimers in the inner membrane (Koo *et. al.*, 2008).

The mechanism in which $GspD_{\beta}$ is localized to the outer membrane by YghG likely involves the Lol system that is required for transfer of lipoproteins to the outer membrane. A model proposed by Okon *et. al.*, 2008 describes the putative mechanism used in localization of the T3SS secretin MxiD to the outer membrane by the pilotin protein MxiM in *Shigella flexneri* as described in the introduction (Fig. 8). This mechanism is based on structural changes that were observed upon binding of the "cracked" β-barrel type MxiM to the C-terminal 18 residues of MxiD that presumably lead to formation of a ternary complex composed of LolA-MxiM-MxiD and subsequent localization of MxiD to the outer membrane. This model is intriguing because it incorporates the known function of LolA and LolB in localization of lipoproteins to the outer membrane to explain how the secretin protein is localized by a pilotin protein. A schematic diagram that depicts the theoretical mechanism of $GspD_{\beta}$ localization via a LolA-YghG-$GspD_{\beta}$ ternary complex is given in Fig. 26. In the absence of YghG or in the presence of inner membrane localized YghG (YghG A26D S27D), $GspD_{\beta}$ is localized to the inner membrane, does not form a multimer and the majority of $GspD_{\beta}$ is degraded (scenario A). If YghG is localized to the outer membrane, $GspD_{\beta}$ is also localized to the outer membrane, the secretin is assembled, and a large amount of $GspD_{\beta}$ present as both monomer and multimer is observed. In this way, localization of $GspD_{\beta}$ to the outer membrane requires the activity of

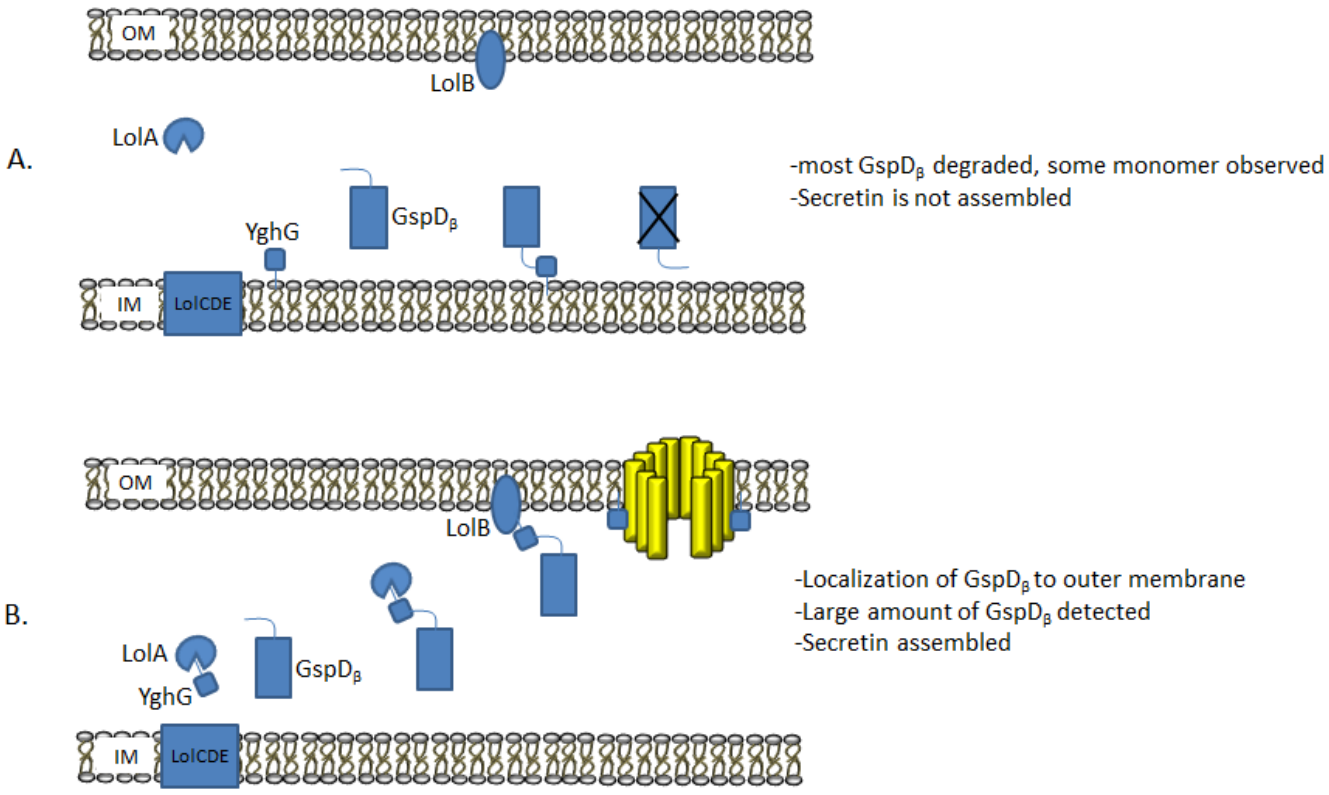


FIG. 26. Schematic diagram of the state of $GspD_{\beta}$ (in monomeric and multimeric forms) observed in the absence of YghG and in the presence of inner or outer membrane-localized YghG. (A) When YghG is absent or localized to the inner membrane (YghG A26D S27D), $GspD_{\beta}$ fractionates with the inner membrane, does not form a secretin multimer and is largely degraded. (B) If YghG is localized to the outer membrane, $GspD_{\beta}$ is also localized to the outer membrane whereupon it assembles the secretin and thereby is protected from degradation. A model based on the proposed mechanism of the pilotin MxiM in localization of the secretin protein MxiD of the *Shigella flexneri* T3SS (Okon *et al*., 2008) is given in whereby a ternary complex of LolA-YghG-$GspD_{\beta}$ would form in the periplasm. By lipoprotein transfer of the periplasmic chaperone LolA to the outer membrane lipoprotein insertion protein LolB, YghG and $GspD_{\beta}$ are localized to the outer membrane and the secretin is formed.

YghG and it is this localization rather than the presence of YghG that is required for protection of $GspD_{\beta}$ from degradation.

Pilot proteins are classified as such based on the similarity of the function of the protein and not solely on amino acid similarity. Although YghG is not similar in amino acid sequence to the PulS/OutS family of T2SS pilotin proteins (Fig. 25), YghG performs a pilotin function similar to PulS and OutS in localization of a T2SS GspD protein to the outer membrane and therefore should be re-designated as $GspS_{\beta}$. However, for the purposes of simplicity, the original designation of YghG will continue to be used in this thesis.

Instead of exhibiting similarity to the PulS/OutS family of T2SS pilotins, YghG is more similar to a group of small uncharacterized hypothetical lipoproteins including C4K3F8 from *Hamiltonella defensa,* Q2BX31 from *Photobacterium profundum* and Q9KRD9 from *Vibrio cholerae* as evident from the greater percent identity/similarity and Needleman-Wunsch global alignment score between YghG and C4K3F8, Q2BX31 and Q9KRD9 than between PulS, OutS and EtpO (Fig. 25). To explain this finding the amino acid similarity of the C-terminal ends of secretin proteins was analyzed to determine if similarity exists between GspD proteins within the S-domain along pilotin family-type (either the PulS/OutS or YghG families) since pilotins have been characterized to bind to the C-terminal S-domain of the secretin protein (Daefler *et. al.*, 1997; Burghout *et. al.*, 2004A; Okon *et. al.*, 2008). As shown in Fig. 25B, GspD proteins exhibit sequence conservation that terminates in the sequence TIIRD. The sequence of GspD proteins C-terminal to the amino acids TIIRD can be divided in two groups. The first group is comprised of PulD from *Klebsiella oxytoca*, OutD from *Erwinia chrysanthemi* and EtpD encoded on the pO157 plasmid of *E. coli* O157. The second group is composed of $GspD_{\beta}$ from H10407, EpsD from *V. cholerae*, Q1YWZ3 from *Photobacterium profundum* and C4K366 from *Hamiltonella defensa*. Therefore a correlation exists that species with similar pilotin proteins also exhibit similarity at the C-terminal end of their respective GspD target proteins. As a result, it is likely that secretin assembly in *V. cholerae, Hamiltonella defensa* and *Photobacterium profundum* require the function of a YghG-like pilotin for localization and subsequent secretin assembly in the outer membrane in a manner similar to the $T2SS_{\beta}$ of ETEC.

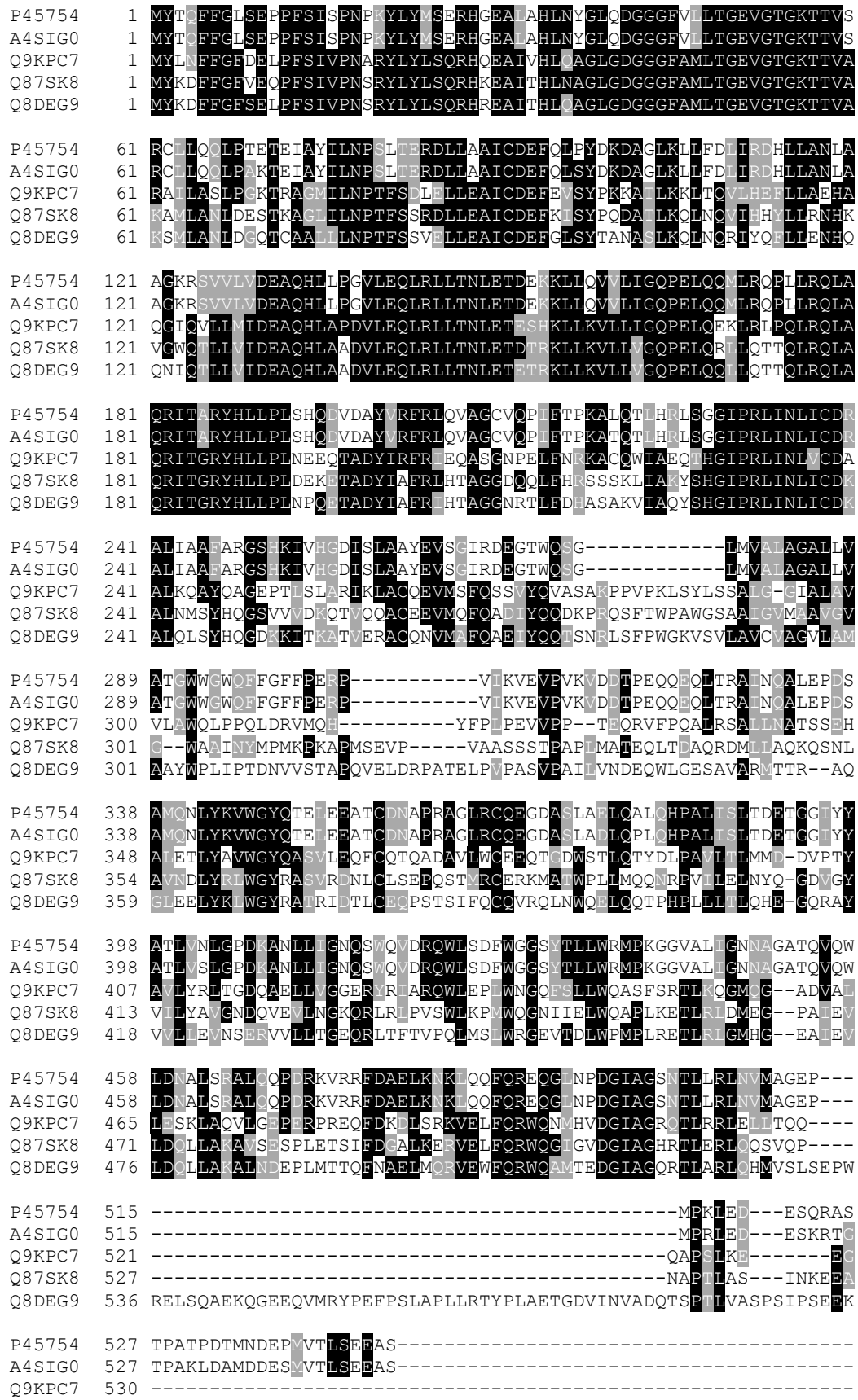
5. Requirement for GspAB in assembly of the GspD secretin in *A. salmonicida* and *Vibrio* species *cholerae, vulnificus* and *parahaemolyticus*

5.1 Introduction

Members of the *Aeromonas* and *Vibrio* genera are significant human and animal pathogens. These species, classified as part of the same clade of gammaproteobacteria (Williams *et. al.*, 2010), include *Vibrio cholerae* that is the causative agent of the gastrointestinal disease cholera, and the fish, amphibian and opportunistic human pathogen *A. hydrophila*. The T2SS encoded within these species is composed of Gsp proteins C-N, the prepilin peptidase TapD, and an additional set of proteins GspA and GspB but no identifiable GspS homologue (Jahagirdar and Howard, 1994; Jiang and Howard, 1991; Pepe *et. al.*, 1996). The GspA and GspB proteins encoded in *Aeromonas hydrophila* and *Vibrio cholerae* are alike, with GspA proteins being 40% identical and 54% similar and GspB proteins 27% identical and 44% similar. Interestingly, *V. vulnificus* encodes a 718 aa protein that contains a GspA domain within the N-terminal 530 aa of the protein and a GspB domain within the C-terminal 188 aa (Figures 27 and 28). Identification of a GspAB fusion protein is perhaps not surprising since according to the "Rosetta Stone" hypothesis for evolution of protein interactions, a propensity exists for interacting pairs of proteins to evolve into one because fusion greatly increases the affinity of each protein for the other and is therefore thermodynamically favourable (Marcotte *et. al.*, 1999).

Previous studies of $GspA^{Ah}$ and $GspB^{Ah}$ proteins in *A. hydrophila* demonstrated that these proteins span the inner membrane once and form a large heteromultimeric complex (Howard *et. al.*, 1996; Jahagirdar *et. al.*, 1994) that in some way modifies or organizes the peptidoglycan (PG) to allow assembly of the GspD secretin in the outer membrane (reviewed in section 1.8). Therefore $GspAB^{Ah}$ is required for transport and assembly of the secretin in *A. hydrophila.* In other bacteria with a functional T2SS however, the role of the GspAB complex remains unclear. In some bacteria no identifiable GspAB homologue is present, in other bacteria a GspB but no GspA is present (Condemine *et. al.*, 1992) and in others GspAB homologs have been identified *in silico* but have not been studied with respect to their role in T2SS secretion or secretin assembly.

In this study, the involvement of the GspAB complex in assembly of the T2SS secretin multimer in *A. salmonicida* and several *Vibrio* species was analyzed. Previous studies have



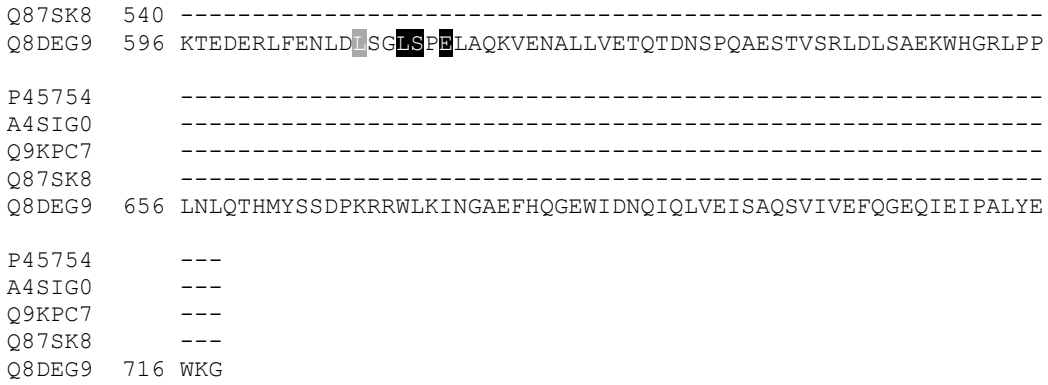


FIG. 27. **Amino acid alignment of GspA proteins.** GspA proteins from *Aeromonas hydrophila* (P45754), *Aeromonas salmonicida* (A4SIG0), *Vibrio cholerae* (Q9KPC7), *Vibrio parahaemolyticus* (Q87SK8) are aligned with the GspAB protein of *Vibrio vulnificus* (Q8DEG9).

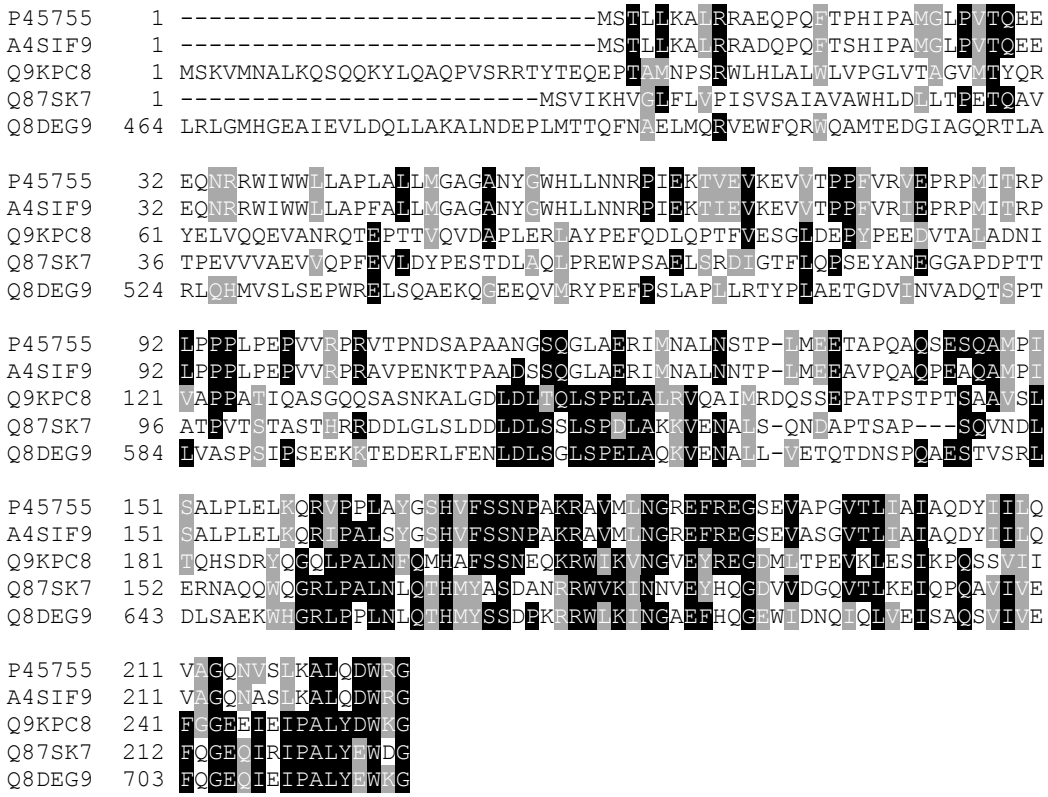


FIG. 28. Amino acid alignment of GspB proteins. GspB proteins from *Aeromonas hydrophila* (P45755), *Aeromonas salmonicida* (A4SIF9), *Vibrio cholerae* (Q9KPC8), *Vibrio parahaemolyticus* (Q87SK7) are aligned with GspAB protein (amino acids 464-718) of *Vibrio vulnificus* (Q8DEG9).

shown strong similarities between the T2SS systems of this genus, such that hybrid T2SS-component proteins between *A. hydrophila* and *V. cholerae* are often functional (Sandkvist *et. al.*, 2000), and the genomes of all sequenced *Vibrio* species contain homologous *gspAB* genes. Similar to results of previous studies in *A. hydrophila*, GspAB was found to be absolutely required in *A. salmonicida* for both assembly of the secretin multimer and secretion of T2SS substrates. Conversely, the presence of GspAB was not an absolute requirement for function of the T2SS in *V. cholerae, V. vulnificus* and *V. parahaemolyticus.* Insertional inactivation of *gspA* in these bacteria resulted in minor decreases in secretion of T2SS substrates. However a significant decrease in the amount of secretin multimer formed was observed for each of the *gspA* mutants, suggesting that the GspAB complex is involved in assembly of the secretin in *Vibrionaceae*. The minimal effect on secretion despite the clear role of GspAB in secretin assembly in *V. cholerae* was explained when expression of $gspD^{Vc}$ *in trans* in a $gspD^{Vc}$ mutant demonstrated that native levels of secretin expression in this bacterium substantially exceed the required capacity for secretion. Expression of the *V. cholerae gspAB* genes *in trans* in *A hydrophila gspAB* mutants resulted in partial complementation of secretin assembly. These results suggest that similar to the T2SS of Aeromonads, GspAB of Vibrios facilitates the assembly of the secretin, however unlike in *Aeromonas* species, other unidentified factors may exist that provide redundancy for secretin assembly in *Vibrio* species.

5.2 Results

5.2.1 Effect of *gspA* inactivation on assembly of the GspD secretin

To investigate the requirement for GspAB in the assembly and function of the T2SS in *A. salmonicida*, the relative amount of secretin assembled and glycerophospholipid:cholesterol acyl transferase (GCAT) activity in culture supernatant (as a measure of T2SS activity) were compared in wild-type and a *gspA* mutant strain of As449. GCAT has been previously shown to be secreted by the T2SS and therefore was a reliable indicator of T2SS secretion capability (Brumlik *et. al.*, 1997).

The GCAT assay results indicated that secretion was severely affected by the loss of $GspA^{As}$, as evidenced by the low lipase activity in culture supernatants from As-*gspA* in comparison to that from AsR (Table 6). As expected from previous results in *A. hydrophila*, the

Table 6: Enzymatic activities detected in culture supernatant of wild-type and *gspAB Aeromonas* and *Vibrio* species

	Lipase[a](%)	Protease[b](%)	Amylase[c](%)
A. salmonicida			
As-R	0.980+/-0.072 (100)	na	na
As-*gspA*	0.243+/-0.051 (24.8)*	na	na
V. cholerae			
Bah2-R	3.71+/-0.042 (100)	3.08+/-0.287 (100)	na
Bah2-*gspA*	3.74+/-0.378 (101)$^{\intercal}$	2.28+/-0.269 (74)*	na
V. vulnificus			
Vv-R	61.3+/-6.74 (100)	0.221+/-0.019 (100)	22.6+/-2.20 (100)
Vv-*gspA*	53.7+/-6.97 (87.6)$^{\intercal}$	0.186+/-0.011 (84.2)$^{\intercal}$	19.9+/-1.51 (88.1)$^{\intercal}$
V. parahaemolyticus			
Vp-R	0.276+/-0.031 (100)	na	19.8+/-1.10 (100)
Vp-*gspA*	0.211+/-0.001 (76.4)*	na	18.8+/-1.03 (94.9)$^{\intercal}$

[a] activity expressed in $\mu M/min/OD_{600}$
[b] activity expressed in $A_{450}/hr/OD_{600}$
[c] activity expressed in $A_{595}/hr/OD_{600} \times 10^3$
na - no detectable activity
*significantly different
$^{\intercal}$not significantly different

loss of T2SS function in the *A. salmonicida gspA* mutant was accompanied and presumably caused by the failure to assemble the GspD secretin since the amount of secretin multimer from whole cell samples was almost nonexistent in the *gspA* mutant when compared to wild-type (Fig. 29). Although a small amount of multimer was evident in concentrated samples of the *gspA* strain, the amount was negligible in comparison to wild-type (Fig. 29, lanes 3 and 4). It should be noted that the failure to assemble the secretin multimer in *A. hydrophila gspA* mutants was accompanied by the accumulation of monomer GspD in the inner membrane (Ast *et. al.*, 2002), although in *A. salmonicida* no accumulation of the monomer could be observed in the *gspA* mutant. This may result from degradation of the monomer when it cannot be assembled into the multimer in *A. salmonicida*. These results clearly demonstrated the essential requirement for GspAB in the assembly of the secretin in *A. salmonicida,* and suggested that the absolute requirement for GspAB in secretin assembly is not specific to *A. hydrophila* but is a characteristic of the *Aeromonas* species.

In order to examine the involvement of GspAB in secretin assembly in *Vibrio* species, lipase, protease and amylase activities were assayed in culture supernatant from wild-type and $gspA^{Vc}$ strains. Lipase and protease activity were detected in *V. cholerae* supernatant whereby an equivalent amount of lipase activity and 26% lower protease activity was observed in the *gspA* strain in comparison to wild-type (Table 6). Lipase, protease and amylase activities in culture supernatants of the *V. vulnificus gspA* mutant (Vv-*gspA*) were not significantly different from that observed in supernatant from wild-type culture (Table 6). Lipase and amylase activities were detected in *V. parahaemolyticus* culture supernatants, and the *gspA* mutant (Vp-*gspA*) secreted 24% less lipase activity than wild-type and an equivalent amount of amylase activity in comparison to wild-type. The mostly insignificant decrease in lipase and protease activities observed in supernatant from *gspA* strains in comparison to wild-type from *Vibrio* species could not be explained by the possibility that these proteins are not T2SS substrates since deletion of *gspD* in *V. cholerae* abrogated secretion of lipase and protease activities into culture supernatant (refer to section 5.2.2).

Although mutation of *gspA* had a minimal effect (if any) on secretion of T2SS substrates, a substantial reduction in the amount of assembled *gspD* secretin was observed in the *gspA* mutant of *V. cholerae* (Fig. 30A), *V. parahaemolyticus* (Fig. 30B) and *V. vulnificus* (Fig. 30C) when compared to wild-type. Due to the difficulty in quantifying the relative

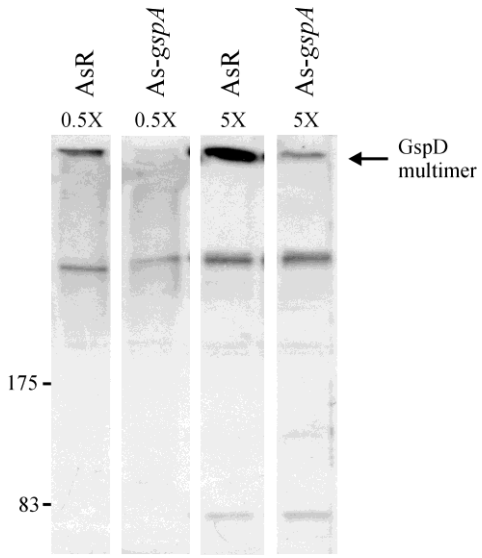


FIG. 29. In *A. salmonicida*, loss of GspA^{As} abrogates assembly of the GspD secretin multimer. Anti-GspD^{Ah} immunoblot of 0.5 and 5-fold concentrated whole cell samples taken from AsR and As-*gspA* strains grown in liquid culture to an OD_{600} of 2.0. Locations of pre-stained standard protein markers and the GspD^{As} secretin multimer are given.

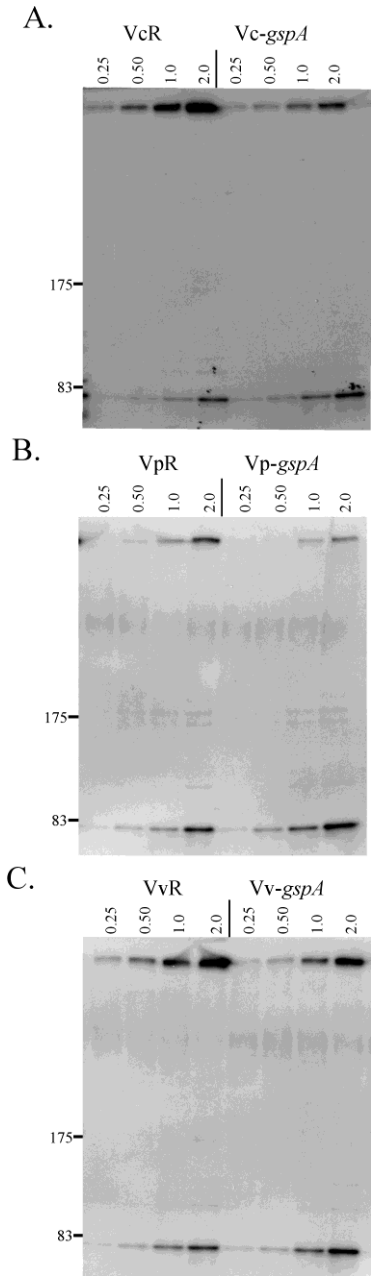


FIG. 30. Effect of *gspA* mutation on assembly of the GspD secretin in *Vibrio* species. Anti-$GspD^{Vc}$ immunoblot of whole cell samples taken from liquid culture at various stages of growth from early-exponential phase to late-exponential phase ($OD_{600\ nm}$ at the top of each lane) of (A) wild-type (Bah2-R) and *gspA V. cholerae* (Bah2-*gspA*) (B) wild-type (VpR)and *gspA V. parahaemolyticus* (Vp-*gspA*) and (C) wild-type (VvR) and *gspA V. vulnificus* (Vv-*gspA*). Locations of pre-stained standard protein markers are given

difference in amount of protein by immunoblotting techniques, multiple points were provided for comparison by repeatedly sampling cultures of the *gspA* mutant and wild type over the entire growth curve. At each point in the growth curve a lower amount of assembled secretin and increased amount of unassembled GspD monomer was observed in samples taken from the *gspA* mutant compared to the wild-type (Fig. 30). This phenotype could be complemented in *V. cholerae* since secretin levels increased in response to expression of $gspAB^{Vc}$ *in trans* and not if the complementing plasmid only encoded $gspA^{Vc}$ (Fig. 31). Apparently the *gspA* mutation also decreased the level of GspB in the cell, either through a direct polarity effect, or possibly because GspB is unstable in absence of GspA, as was previously found in *A. hydrophila* (Howard *et. al.*, 1996). In any case, this result confirmed the involvement for both GspA and GspB in secretin assembly in *Vibrio* species.

5.2.2 *V. cholerae* exhibits an excess secretion capacity

The absence of $GspAB^{Vc}$ in *V. cholerae* resulted in a substantial decrease in secretin assembly (Fig. 30A) that was not accompanied by a concomitant decrease in lipase and protease activities in culture supernatants (Table 6). This result suggested that wild-type levels of secretin are not required for efficient secretion of lipase and protease from the cell, in other words, that wild-type *V. cholerae* exhibits an excess secretion capacity. In order to assay secretion capacity, the *gspD* strain was complemented by expression of *gspD in trans* at various levels of induction. The results as shown in Fig. 32, revealed that a relatively low level of secretin in comparison to the amount expressed in wild-type cells is sufficient for full secretion, since wild-type levels of lipase were present in supernatants from cells without induction (Fig. 32B) and wild-type levels of protease were present in supernatant from cells induced with 0.002 mM IPTG (Fig. 32C). In the absence of induction and at an induction level of 0.002 mM IPTG, much less secretin was observed in the complemented *gspD* strain than in wild-type cells (Fig. 32A). These data demonstrate that at least under the growth conditions used in these experiments, there is a natural overexpression of the secretin relative to the amount of extracellular secretion, such that more secretin is assembled than is required for secretion of substrates into the media, a situation that may exist due to differential expression of genes that encode the T2SS and the substrates of the system.

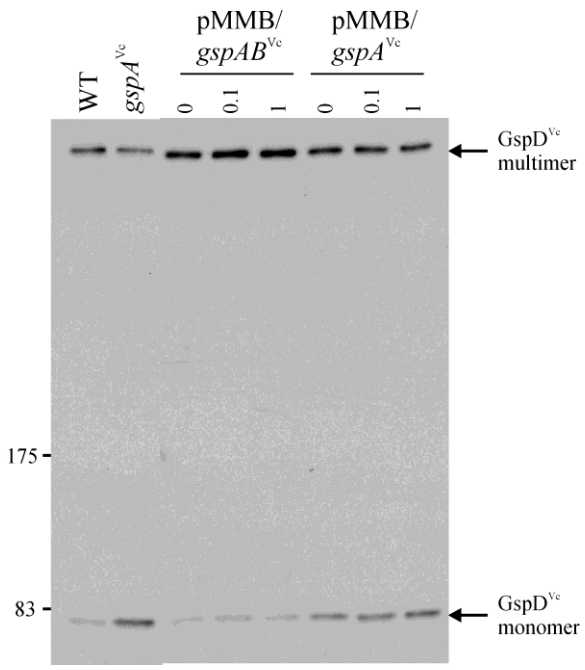


FIG. 31. **Complementation of the partial secretin-negative phenotype of *V. cholerae* Bah2-*gspA* by expression of GspAB^{Vc} or GspA^{Vc} *in trans* from plasmids pMMB207/*gspAB*^{Vc} and pMMB207/*gspA*^{Vc}.** The cultures were grown to an $OD_{600\ nm}$ of 2.0 in the absence or presence of IPTG (0.1 and 1.0 mM) and whole cell samples were electrophoresed and immunoblotted with anti-GspD^{Vc}. Locations of standard protein markers are given.

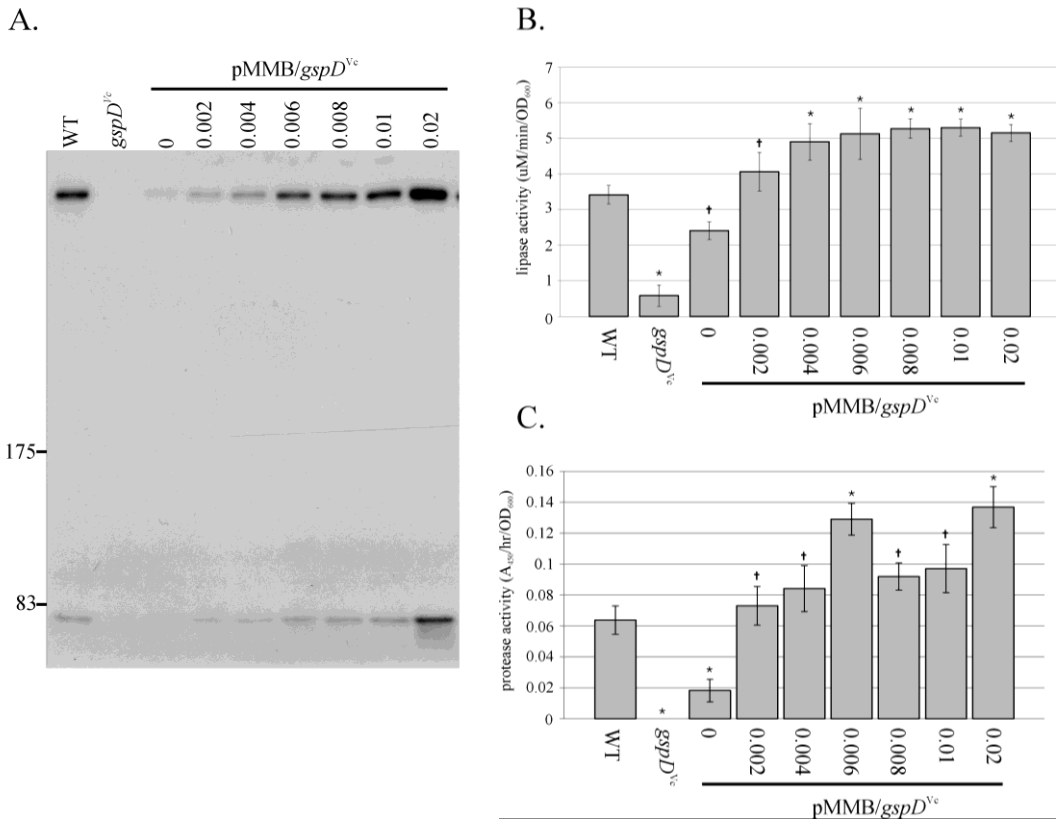


FIG. 32. **Wild-type amount of GspD^{Vc} secretin is not required for secretion of lipase and protease to wild-type levels.** (A) The amount of secretin multimer assembled in *Vibrio cholerae* TRH7000, *gspD*^{Vc} and *gspD*^{Vc} expressing *gspD*^{Vc} *in trans* (induced with 0 - 0.02mM IPTG from plasmid pMMB/*gspD*^{Vc}) was assessed by immunoblot with anti-GspD^{Vc} antibody. Triplicate assays of lipase (B) and protease (C) activity in supernatant taken from cultures of TRH7000, *gspD*^{Vc} and *gspD*^{Vc} expressing *gspD*^{Vc} *in trans* (induced with 0 - 0.02mM IPTG) are shown. Significant difference († denotes non-significance and * denotes significance where $p<0.05$) was determined based on comparison of each value against activity observed in supernatant from wild-type culture.

5.2.3 State of GspD^{Ah} secretin assembly upon expression of GspAB^{Vc} in *A. hydrophila*

The secretion and secretin assembly results suggested that although the *Vibrio* GspAB is not essential for function of the T2SS, it is nevertheless involved in assembly of the secretin. However, since the lack of GspAB in *Vibrio* species does not result in a secretin-negative phenotype such as that observed in *Aeromonas* species, the GspAB complex of Vibrios may not function in a similar manner as the GspAB of Aeromonads. Therefore to determine if GspAB^{Vc} performs the same function in Vibrios as GspAB^{Ah} in Aeromonads, GspAB^{Vc} was expressed in the *A. hydrophila gspA* mutant and the amount of secretin was analyzed to determine if complementation had taken place. The GspAB^{Vc}-expressing plasmid pMMB/*gspAB*^{Vc} was introduced into Ah65 and Ah65 *gspA* (C5.84) and analyzed for T2SS function and assembly. The recombinant plasmid contained the presumed *gspAB*^{Vc} promoter and a vector-encoded *tac* promoter to allow increased expression. Expression of *gspAB*^{Vc} *in trans* partially complemented the secretin assembly defect of C5.84 as indicated by the re-establishment of secretin multimer assembly in C5.84 (pMMB/*gspAB*^{Vc}) (Fig. 33A). The amount of secretin multimer assembled was somewhat dependent upon the level of *gspAB*^{Vc} expression since induction with 0.1 mM IPTG resulted in a slightly greater amount of ExeD multimer formed. As has been previously demonstrated, in the control experiment expression of *gspAB*^{Ah} encoded in plasmid pRJ31.1 in strain C5.84 completely restored secretin assembly (Fig. 33A) (Ast *et. al.*, 2002).

The C5.84 secretion negative phenotype was nearly fully complemented by expression of *gspAB*^{Vc} *in trans*. As shown in Fig. 34, the lipase activity in supernatants from the uninduced C5.84 (pMMB/*gspAB*^{Vc}) culture was 90% of the wild-type amount, and did not increase with further induction. In the control experiment wild type *gspAB*^{Ah} completely complemented the lipase secretion defect of the *gspA* mutant even without induction.

Interestingly, expression of *gspAB*^{Vc} in wild-type *A. hydrophila* decreased the amount of assembled secretin multimer and increased the amount of GspD monomer as visualized by immunoblot when compared to cells containing the vector only (Fig. 33B). This effect was not observed for Ah65(pRJ31.1) cells overexpressing the *gspAB*^{Ah} genes, since expression of *gspAB*^{Ah} resulted in more secretin multimer and less unassembled monomer than observed in wild type. Likewise, the amount of lipase activity in Ah65(pMMB/*gspAB*^{Vc}) culture supernatant was significantly decreased upon induction of *gspAB*^{Vc} with 0.1 mM IPTG to 45% (Fig. 34) of the lipase activity in Ah65 culture supernatant.

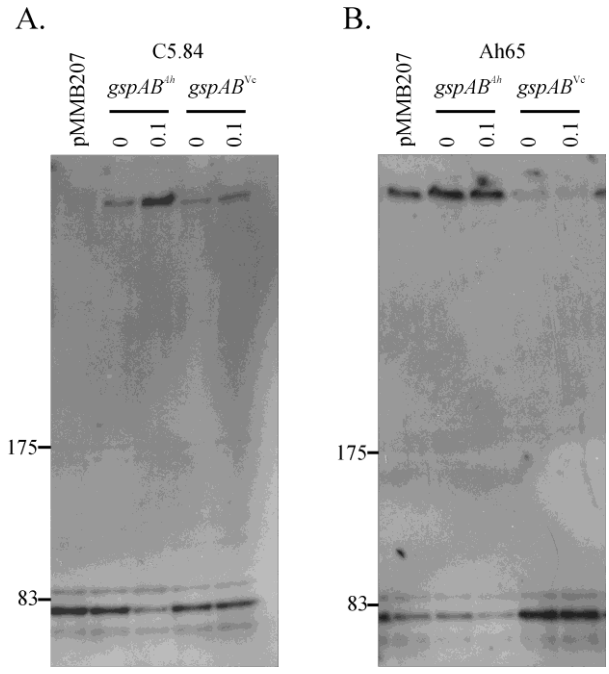


FIG. 33. Secretin assembly in *A. hydrophila gspAB* and wild-type upon expression of *V. cholerae gspAB in trans*. The amount of secretin multimer assembled upon induced expression (0.1mM IPTG) of $gspAB^{Ah}$, $gspAB^{Vc}$ or neither (pMMB207) in (A) C5.84 and (B) Ah65 was assessed by anti-$GspD^{Ah}$ immunoblot. Locations of pre-stained standard protein markers are given.

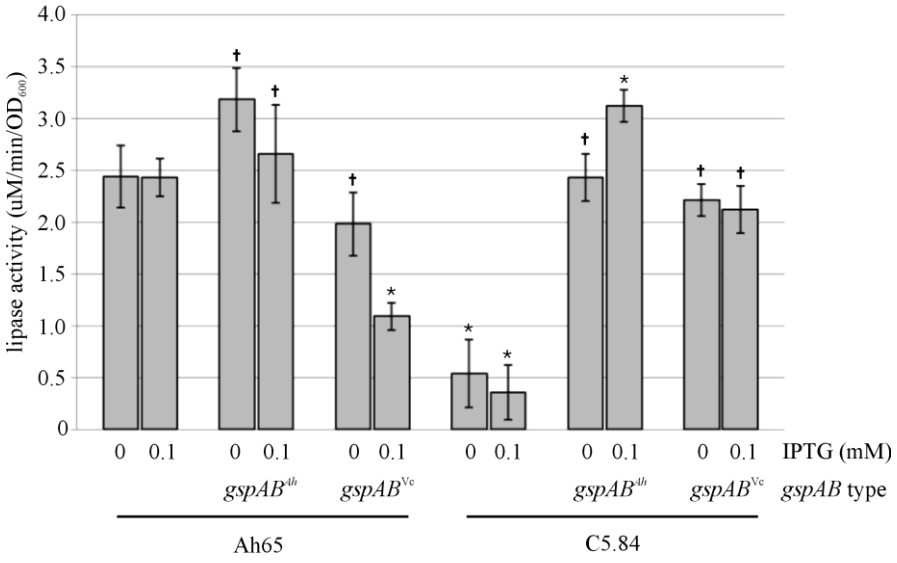


FIG. 34. Lipase activity in supernatant from cultures of *A. hydrophila* wild-type and *gspA* upon expression of $gspAB^{Ah}$ and $gspAB^{Vc}$ *in trans*. Lipase activity was assayed in supernatant taken from wild-type A. hydrophila (Ah65) and *gspA* (C5.84) cultures expressing $gspAB^{Ah}$ (from plasmid pRJ31.1) or $gspAB^{Vc}$ (from plasmid pMMB/$gspAB^{Vc}$) with or without induction with IPTG. Assays of culture supernatant were performed in triplicate with the standard deviations indicated. The results shown are representative of those obtained in multiple experiments. Significant difference († denotes non-significance and * denotes significance where $p<0.05$) was calculated based on comparison of each value against lipase activity observed in wild-type (Ah65) culture.

5.2.4 Transposon mutagenesis of *V. cholerae* to identify the GspAB redundant function

In *Vibrio* species *cholerae*, *parahaemolyticus* and *vulnificus*, the absence of GspAB only slightly affected secretion of T2SS-substrates but caused a significant deficiency in the amount of assembled secretin multimer observed (Fig. 30). However, since the secretin multimer was still formed in *gspAB*^{Vc} strains of all three *Vibrio* species, these species may contain another protein(s) involved in assembly of the secretin. Identification of other proteins involved in assembly of the secretin would be of interest because it would provide more information regarding the protein requirements for GspD multimer assembly and T2SS function in *Vibrio* and further characterize the functional overlap that likely exists between proteins of multiple systems that span the periplasm in Gram-negative bacteria.

To identify the protein(s) responsible for assembly of secretin in the absence of GspAB, transposon mutagenesis was performed employing a minitransposon Tn5 derivative that contains the chloramphenicol acetyltransferase (*chl*) gene. The minitransposon was encoded within plasmid pUT (DeLorenzo and Timmins, 1994) which contains an R6K origin of replication that allowed the plasmid to be propagated in a strain such as SM10 lambda *pir* (Miller and Mekalanos, 1988) that contains the R6K-specified replication protein π (Kolter *et. al.*, 1978) (Fig. 35). The plasmid also contains an RP4oriT origin of transfer region from the R6K conjugative promiscuous plasmid RP4 thereby allowing the plasmid to be introduced into most genera of Gram-negative bacteria.

The pUT/Chl plasmid was introduced into an *epsA* strain of *V. cholerae* Bah2 (recipient) by conjugation with SM10 lambda pir (donor) and counterselected against the donor by growth on thiosulfate citrate bile salts sucrose (TCBS) agar that allows growth of *Vibrionaceae* species and does not allow growth of *Enterobacteriaceae*.. Transconjugants of *Bah2 gspA* cells were selected on TCBS containing chloramphenicol (Cam). Ten thousand cam-resistant colonies were patched onto LB agar containing 1% skim milk to identify clones that no longer exhibited the capacity to secrete a casein-degrading T2SS substrate protease into the media (Overbye *et. al.*, 1993) (Fig. 36). Thirty-nine of ten thousand Cam^{R} colonies were identified as deficient in secretion of the protease and were further tested for the ability to assemble the secretin multimer by anti-GspD^{Vc} immunodetection of whole cell samples. The majority of clones (36) did not exhibit a secretin assembly defective phenotype whereas three mutants (numbers 15, 32 and 38) were identified as partially unable to assemble the secretin

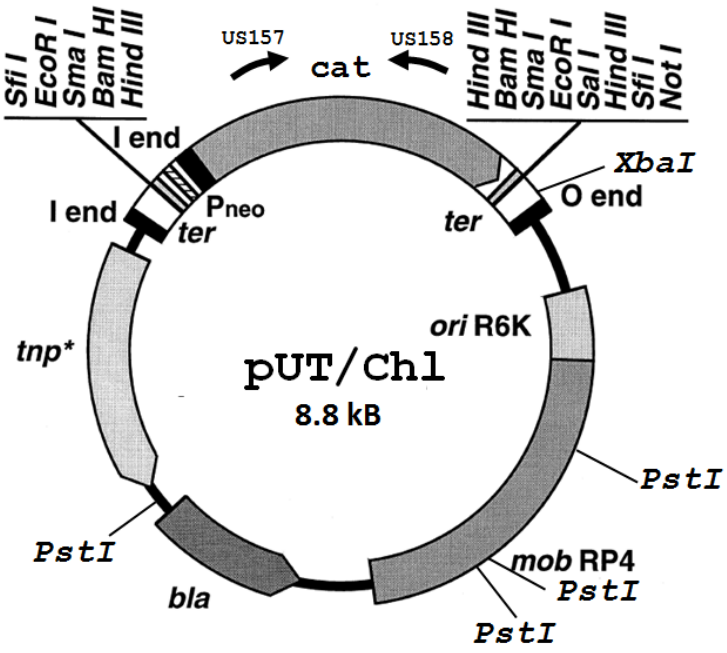


FIG. 35. Schematic diagram of the pUT-based suicide vector system utilized for transposon mutagenesis of V. cholerae Bah2 *epsA*. This figure was modified from the original figure published by DeLorenzo and Timmins, 1994. The location of BamHI, XbaI and PstI restriction endonuclease sites used in Southern blot hybridization (Fig. 38) are shown. A biotin-labeled Southern blot hybridization probe was constructed from a PCR fragment generated by amplification of pUT/Chl with primers US157 and US158 as shown.

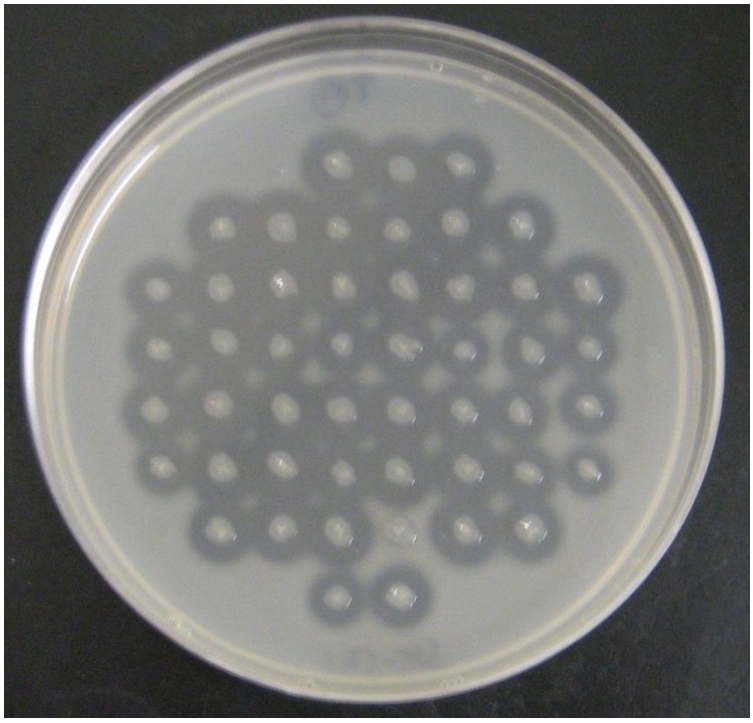


FIG. 36. Identification of potential GspD^{Vc} secretin assembly-negative mutants by patch growth of transposon mutants on skim milk agar. Ability to secrete the casein-degrading protease was assessed by a ring of clearance around the colony. Clones that did not secrete the protease were selected for further testing.

multimer and were selected for further study (Fig. 37). The ability to isolate transposon-mutated clones able to express $GspD^{Vc}$ yet deficient in secretin assembly (clones 32 and 38) suggested that $GspD^{Vc}$ secretin assembly in *V. cholerae* involved the function of another protein. Note that since the level of $GspD^{Vc}$ monomer in clone 15 was nearly undetectable, this clone was likely protease secretion and secretin-assembly deficient due to the lack of $gspD^{Vc}$ expression.

Attempts to determine the location of the transposon in the three protease-secretion negative clones were not successful. The initial attempt was to clone the Cam^R gene and surrounding genomic sequence from the genome into plasmid pSK by digestion of genomic DNA isolated from clones 15, 32 and 38 with three restriction enzymes that do not have a restriction site in the transposon. This attempt was unsuccessful except for 1 colony derived from cloning genomic DNA isolated from protease-secretion negative mutant 15 into pSK. However, in this case the entire "suicide" vector pUT/Chl was cloned into pSK. This result suggested that the plasmid was able to replicate or remain independent of the chromosome in this clone. Theoretically this result was not possible since the plasmid is a suicide vector that is dependent upon the pi protein encoded within the F' episome of the donor strain for replication. Next, genomic DNA was isolated from each clone and digested with various restriction enzymes that do or do not have sites within pUT/Chl. The resultant DNA was used in southern hybridization detection with a probe specific to the transposon. The banding pattern obtained from digested genomic DNA from each clone was compared to that of pUT/Chl and genomic DNA from the recipient strain (Bah2 *epsA*::kan). The results as shown in Fig. 38 demonstrated that clone 15 contains the suicide vector in addition to a high molecular weight band that hybridizes to the transposon-specific probe. Clones 32 and 38 contain one band that would be consistent with a random transposon insertion event into the chromosome however a transposon-specific band was not generated upon digestion of genomic DNA with *BamHI* which contains restriction sites inside of the I and O ends of the transposon and generates a 2105 bp band observed in the pUT/Chl control digest and in the *BamHI* digest of genomic DNA from clone 15. Note that non-specific hybridization of the probe to the genome was not observed since in the absence of pUT/Chl (the Bah-2 *gspA* strain) no bands were detected by southern blot.

Lastly, an arbitrary PCR method was used to identify the location of the transposon insertion. In this method, a partially degenerate oligonucleotide is used that is capable of

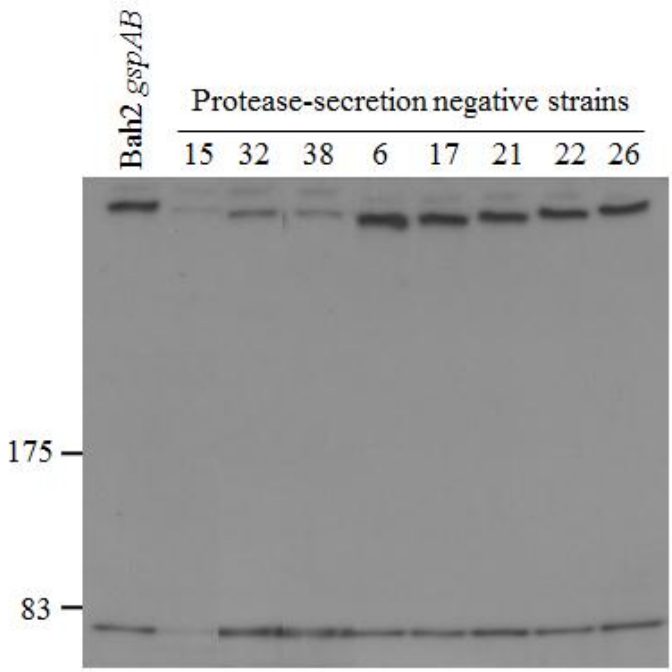


FIG. 37. Identification of GspD^{Vc} secretin-assembly deficient strains in the pool of protease-secretion negative mutants identified by transposon mutagenesis. Anti-GspD^{Vc} immunoblot of whole cell samples taken from Bah2 *gspAB*^{Vc} and eight protease-secretion negative strains as shown.

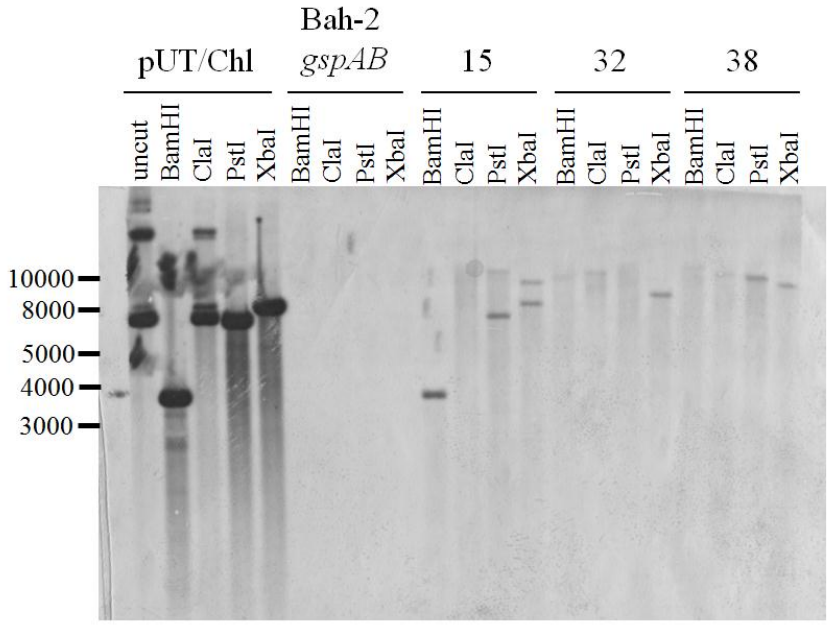


FIG. 38. Southern blot detection of the transposon. Protease secretion-negative strains 15, 32 and 38 likely contain a transposon insertion in the genome however strain 15 also contains self-replicating plasmid. Samples of undigested pUT/Chl in addition to pUT/Chl digested with *BamHI* (cuts out the transposon), *ClaI* (does not have a restriction site in pUT/Chl), *PstI* (generates a 7265 bp fragment recognized by the probe) and *XbaI* (one site that linearizes the plasmid) were electrophoresed together with genomic DNA was isolated from Bah2 *gspAB*, protease secretion-negative strains 15, 32 and 38 (digested with the enzymes as shown). The southern blot was probed with a 2105 bp fragment specific for the CamR cassette (amplified with primers US157/158) located between the I and O end of the transposon.

hybridizing randomly to the template genomic DNA and is extended by Klenow polymerase (Fig. 39). Using this single-stranded DNA (ssDNA) as template, PCR was performed with primers specific to the non-randomized portion of the degenerate oligonuleotide and the transposon. In this method, genomic DNA flanking the transposon and a portion of the transposon are amplified by PCR, is sequenced and the genomic DNA identified. The lone amplifiable product was generated from clone 38 that contained portions of the pUT/Chl suicide vector that had been interrupted by a 12 kb *V. cholerae* genomic fragment that encoded numerous open reading frames, one of which encoded a transposase. This result suggested that a transposon has been inserted into the suicide vector or that the plasmid had integrated into the genome and in some mutants had been excised from the genome, therefore excising with it a portion of the genome. Regardless of the mechanism involved, this ambiguous result did not provide information regarding the location of the transposon insertion in any of the three protease secretion-negative and secretin-deficient strains.

In characterizing the pilotin function of YghG in ETEC as described in chapter 4, a YghG-homologue in *Vibrio* species encoded by the *Vc1703* gene (protein C4K3F8) in *V. cholerae* N16961 was identified (Fig. 25). In a final attempt to determine the location of the transposon in clones 15, 32 and 38, a PCR reaction amplifying the region surrounding *yghG*^{Vc} was employed to determine if this locus could have been mutated in clones 15, 32 and 38. As shown in Fig. 40, PCR amplification of the recipient strain Bah2 *gspAB* generated a 2076 bp band indicative of the wild-type *Vc1703* gene. In clone 15, the same sized product was generated, thereby identifying that this clone is not a *yghG*^{Vc} mutant. Surprisingly, in clone 32 a much larger PCR product was amplified. This result was repeated several times and therefore suggested that the transposon (or the pUT/Chl plasmid) could be inserted into the genome within or in the vicinity of the *Vc1703* gene in this mutant. Note that since the size of the amplified product was greater than 10 kB it is likely that the entire vector had been inserted within or near the *Vc1703* gene. PCR amplification of the same region in clone 38 did not result in creation of an amplified product, therefore suggesting that the *Vc1703* locus may also not be wild-type in this mutant. These very preliminary yet interesting results suggested that the *Vc1703* gene may not be wild-type in clone 32 and therefore the lack of YghG^{Vc} could be responsible for the protease-secretion negative and secretin-assembly negative phenotypes of clone 32. Therefore, YghG^{Vc} could perform the same function in *Vibrio* species as YghG in

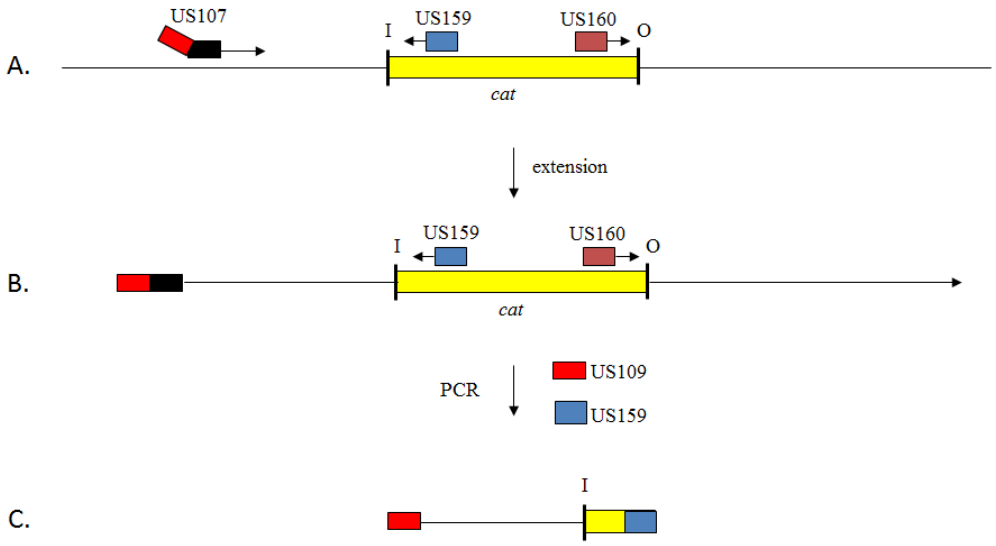


FIG. 39. Schematic diagram of the arbitrary PCR method employed in an attempt to identify the location of the transposon insertion in protease secretion negative clones 15, 32 and 38. (A) If the transposon is inserted into the genome (flanked by I and O sequences as shown) primer extension from a partially degenerate primer able to bind randomly to genomic DNA (US107) would generate a ssDNA product (B) that can serve as template for PCR with primers US109 (specific for the fixed sequence of primer US107) and US159 to generate a PCR product (C) with genomic sequence located adjacent to the insertion sequence of the transposon.

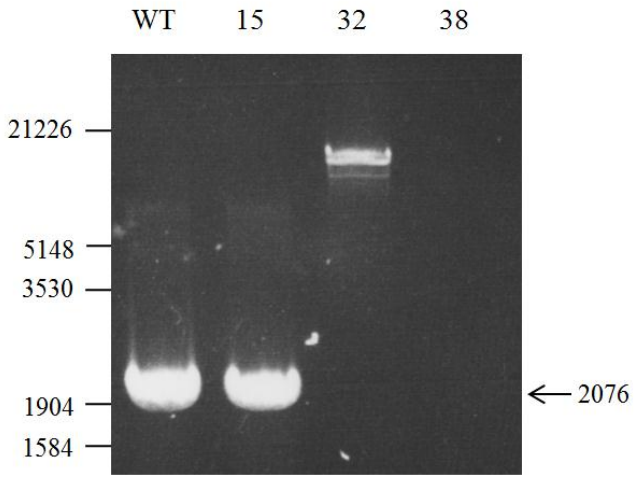


FIG. 40. PCR amplification of the region of the genome encompassing gene *Vc1703* in transposon mutagenesis clones 15, 32 and 38. The region of the *V. cholerae* genome encompassing gene *Vc1703* was amplified with primers US521 and US522 in samples from wild-type (WT) and transposon mutagenesis clones 15, 32 and 38.

ETEC. In fact, in a study performed in Dr. Howard's laboratory, expression of *yghG*^{Vc} *in trans* complemented the secretin-assembly deficient phenotype of the ETEC *yghG* strain (unpublished data). These data suggest that YghG^{Vc} also performs a pilotin function in GspD^{Vc} secretin assembly similar to YghG in assembly of the ETEC GspD$_{\beta}$ secretin and thereby potentially serves as the first example of a Gram-negative species that requires both GspAB and GspS for assembly of a T2SS secretin.

5.3 Discussion

In Gram-negative bacteria that contain a T2SS, the presence of the *gspAB* operon in addition to the set of "core" *gsp* genes is not well conserved, therefore suggesting that the role of GspAB is not integral to assembly of the T2SS. However, in some bacteria that contain *gspAB* such as *A. hydrophila*, the GspAB^{Ah} complex is absolutely required for construction of the T2SS (Ast *et. al*., 2002), characterized by assembly of the megadalton-sized GspD secretin multimer in the outer membrane. The function of the GspAB^{Ah} complex (Schoenhofen *et. al*., 2005) may involve reorganization of the peptidoglycan meshwork to allow monomeric GspD subunits or the secretin multimer itself to pass through to reach the outer membrane, or may possibly provide a scaffold for assembly of the secretin. This activity is mediated at least in part by the C-terminal peptidoglycan-binding motif of GspA^{Ah} (Howard *et. al*., 2006; Li and Howard, 2010). In this study, I analyzed if the requirement for the GspAB complex in assembly of the secretin in *A. hydrophila* also applies to another species of *Aeromonas*, *A. salmonicida*, as well as three *Vibrio* species, *V. cholerae, V. parahaemolyticus* and *V. vulnificus,* by detecting the amount of assembled GspD secretin and secretion of T2SS-substrates in culture supernatants isolated from strains with and without a functional GspAB complex.

The requirement for the GspAB^{As} complex in assembly of the secretin and function of the T2SS in *A. salmonicida* was verified since a significant reduction in lipase activity (75%) was detected in the *gspA* mutant in comparison to wild-type (Table 6) presumably caused by the lack of assembled GspD secretin multimer observed in the *gspA* strain (Fig. 29). This finding was consistent with that described by Ast *et. al*., 2002 whereby loss of the GspAB^{Ah} complex abrogated both secretion of aerolysin and assembly of the secretin multimer in *A.*

hydrophila. Together, these data suggest that GspAB is generally required for assembly of the GspD secretin in *Aeromonas* species.

Surprisingly, contrary to the essential role for GspAB in assembly of the secretin in *Aeromonas* species *hydrophila* and *salmonicida*, GspAB was not an absolute requirement for function of the T2SS in *Vibrio* species because mutation of *gspA* did not prevent secretion of T2SS substrates including lipase, amylase and protease (Table 6) in *V. cholerae, V. parahaemolyticus* and *V. vulnificus* species. However, although the secretion of T2SS-substrates remained largely unchanged in *gspA* compared to wild-type strains, a significant reduction in the amount of assembled secretin multimer in the absence of GspAB in *V. cholerae, V. parahaemolyticus* and *V. vulnificus* was observed (Fig. 30). The dichotomy that substantial decreases in secretin multimer formation had only minimal effects on secretion was explained by an apparent excess in T2SS secretion capacity at normal GspD levels, since minimal expression of *gspD* was shown to fully complement secretion in the *V. cholerae gspD* mutant (Fig. 32). In any case, the significant reduction in assembled secretin multimer in the absence of GspA suggested that the GspAB complexes of the *Vibrio* strains perform an important role in assembly of their respective secretins.

In order to more definitively determine if GspAB performs the same role in *Vibrio* and *Aeromonas* species, a plasmid-encoded copy of *gspAB*^{Vc} was expressed in the *A. hydrophila gspAB* strain C5.84 to determine if GspAB^{Vc} could complement the secretin assembly-negative and secretion-negative phenotypes of this strain. Expression of GspAB^{Vc} re-established assembly of the GspD^{Ah} secretin multimer (Fig. 33A) and secretion of lipase (Fig. 34) in strain C5.84. Although complementation was not complete, it is perhaps not surprising given that GspAB^{Ah} and GspAB^{Vc} are 36% identical and 51% similar. Therefore, although GspAB^{Vc} did not completely complement the *gspAB* phenotype of C5.84, partial complementation did show that GspAB^{Vc} and GspAB^{Ah} likely perform the same function in their respective species.

Although similar enough for partial complementation of the *gspAB*^{Ah} secretion and secretin assembly negative phenotypes of C5.84, GspAB^{Vc} and GspAB^{Ah} are apparently not sufficiently identical to function together as a heteromultimer composed of both complexes. When expressed in wild-type strain Ah65, expression of GspAB^{Vc} decreased the amount of assembled secretin (Fig. 33B) and secretion of T2SS-substrates (Fig. 34) in a dose-dependent manner. Since GspAB functions as a heteromultimeric complex (Schoenhofen *et. al.*, 2005),

this phenotype was likely caused by a dominant-negative effect of GspAB^{Vc} whereby assembly of an inactive heteromultimeric complex composed of GspAB^{Ah} and GspAB^{Vc} reduced the ability of the strain to assemble a functional secretin multimer and thus secrete T2SS-substrates into the growth media. The dominant-negative effect of GspAB^{Vc} expression was consistent with the ability to inactivate the function of a multimeric protein by expression of a mutant subunit *in trans*, a hallmark of multi-subunit proteins.

The difference in the requirement for GspAB in assembly of the T2SS secretin in *Vibrio* and *Aeromonas* species suggested that a redundant GspAB function exists in the *Vibrio* genus that is absent in the *Aeromonas* genus. The most likely candidate to possess a GspAB redundant function would be a peptidoglycan-binding protein involved in assembly and/or function of another system that spans the periplasm. Peptidoglycan-binding proteins have been shown to be involved in assembly of systems such as the flagellar apparatus (Nambu *et. al.*, 1999), type IV pili (Rambow-Larsen and Weiss, 2002) and type III secretion system (Zarhl *et. al.*, 2005) and therefore members of these systems could perform a function analogous to GspAB for assembly of the T2SS.

Conversely, the transposon mutagenesis screen of the Bah2 *gspA*^{Vc} strain identified the possibility that *yghG* encoded by the gene *Vc1703* could also be involved in secretin assembly by performing a pilotin role. If this were the case, it would be the first example of a Gram-negative species that encodes both a GspAB complex and a GspS pilotin. Since the absence of GspAB in *Vibrio* species does not result in the complete abrogation of T2SS secretin assembly and function as is observed in *Aeromonas* species, and *Aeromonas* species do not encode an obvious *yghG* homologue, does it mean that YghG performs the redundant GspAB function in *Vibrio* species? This possibility is difficult to imagine given that YghG is an outer membrane lipoprotein and GspAB are an inner membrane peptidoglycan-binding complex. However, it could be that although both proteins perform their functions in drastically different ways, the end result is the same in that both proteins are involved in localizing the GspD protein to the outer membrane for assembly of the secretin (that likely involves the BAM complex). The pilotin would accomplish this task by binding directly to the S-domain of the C-terminus of the GspD protein, forming a complex with LolA, and shuttling GspD to the outer membrane. The GspAB complex however would perform the task of shuttling GspD to the outer membrane by creating a hole through the peptidoglycan through which GspD can pass, a function that could

involve interactions directly between GspA or GspB and GspD. The possibility that GspAB and GspS do perform redundant functions would explain why most species do not encode both proteins. It is also possible however that given the heterogeneity of pilotin proteins, the lack of an identified GspS protein in species such as *Aeromonas hydrophila* that encode the GspAB complex is due to the inability to detect the presence of the pilotin protein *in silico*.

It is also entirely plausible that the function of YghG is "downstream" that of GspAB such that deletion of YghG generates a secretin-assembly negative phenotype regardless of the presence or absence of GspAB. Therefore, in future studies, a comparison of the secretion and secretin-assembly status of *yghG* and *yghG gspAB V. cholerae* strains will likely provide valuable information regarding whether YghG and GspAB perform separate or redundant functions.

Chapter 6. LeoA is not required for $GspD_{\beta}$ secretin assembly

6.1 Introduction

In a study by Fleckenstein *et. al*., 2000, inactivation of a gene encoded within a pathogenicity island encoded downstream of the adhesin Tia in the genome of ETEC strain H10407 resulted in the inability of the strain to secrete LT. This gene was correspondingly named *leoA* for labile enterotoxin output A. Since the initial study, LeoA has been characterized as an inner membrane protein with an N-terminal GTPase domain and a C-terminal domain that likely interacts with outer membrane proteins such as OmpA (Brown and Hardwidge, 2004). Given that LeoA exhibits some homology with GspA from H10407 (32% similarity and 17% identity) and is required for secretion of LT, a hypothesis was conceived that LeoA could be required for assembly of the $GspD_{\beta}$ secretin by performing a function analogous to that of $GspAB^{Ah}$ for assembly of the $GspD^{Ah}$ secretin. The theory was that since the $T2SS_{\beta}$ does not encode a GspAB complex, the LeoA protein could perform the function of GspAB in peptidoglycan remodeling required for $GspD_{\beta}$ secretin assembly in ETEC.

6.2 Results

6.2.1 The H10407 *leoA*^{UT} strain is LT secretion-negative due to the lack of $GspD_{\beta}$ secretin

The in-frame *leoA* deletion mutant of ETEC strain H10407 used in studies by Fleckenstein *et. al.*, 2000 and Brown and Hardwidge 2004, was provided by Dr. James Fleckenstein from the University of Tennessee (strain designation H10407 *leoA*^{UT}). Consistent with the hypothesis that *leoA* could perform a function similar to GspAB in $GspD_{\beta}$ secretin assembly, the *leoA*^{UT} strain was deficient in $GspD_{\beta}$ secretin assembly since the relative amount of both $GspD_{\beta}$ secretin multimer and monomer were nearly undetectable in the *leoA* strain in comparison to wild-type (Fig. 41).

6.2.2 Expression of *leoA in trans* does not complement the *leoA* secretin-assembly negative

To complement the secretin-assembly negative phenotype of the *leoA*^{UT} strain, three versions of the *leoA* gene were cloned into the medium copy number expression vector pBAD322C and expressed in the *leoA*^{UT} strain. The first plasmid expressed wild-type LeoA (pBAD/*leoA*), the second version encoded a C-terminal myc-tag (to be used for immunodetection) (pBAD/*leoA-myc*) and the third version encoded an N-terminal myc tag

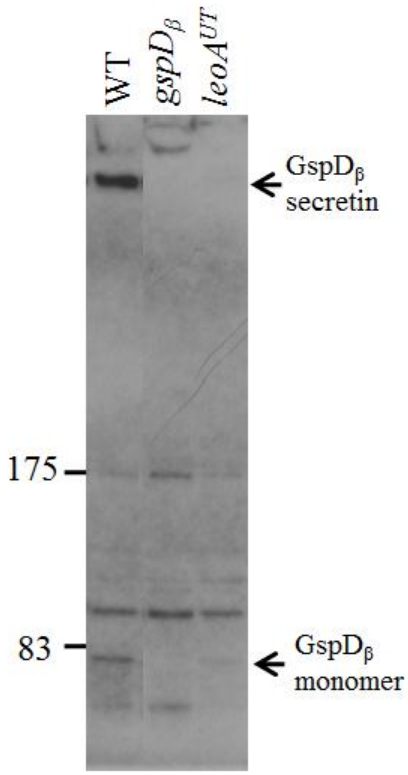


FIG. 41. GspDβ secretin-assembly deficient phenotype of a *leoA* strain. The GspDβ secretin and monomer in whole cell samples taken from cultures of wild-type H10407 (WT), *gspD*β and *leoA* were detected by immunoblot with anti-GspDβ antibody. The position of standard protein molecular weight markers is given on the left of the panel.

(pBAD/*myc-leoA*). Anti-myc immunoblot detection of whole cell samples revealed induced expression of the 65 kDa LeoA-myc and myc-LeoA proteins upon induction with 0.02% arabinose when expressed in the *leoA*^{UT} strain (Fig. 42). The wild-type protein required confirmation of expression by induction with arabinose by coomassie staining of whole cell samples. As shown in Fig. 42B, induction with 0.2% arabinose allowed detection of LeoA, LeoA-myc and Myc-LeoA proteins. Note that expression of the N-terminal tagged LeoA resulted in protein degradation (Fig. 42A) that was not observed in the C-terminal myc-tagged version of LeoA.

Although LeoA, LeoA-myc and myc-LeoA were expressed in strain *leoA*^{UT} upon induction of plasmids encoding these proteins, expression *in trans* was not successful in complementing the GspD$_\beta$ secretin-deficient phenotype of the *leoA*^{UT} strain since an identical amount of GspD$_\beta$ secretin and monomer was observed in *leoA*^{UT} strains that contained empty vector (pBAD322C) with those that contained pBAD/*leoA*, pBAD/*leoA-myc* and pBAD/*myc-leoA* plasmids (Fig. 43). To ensure that a sufficient level of LeoA protein was expressed, a relatively high induction level (0.2% arabinose) was one of the conditions utilized in the complementation experiment. Therefore, the secretin-assembly defect of the *leoA*^{UT} strain could not be complemented by expression of LeoA *in trans*.

6.2.3 *leoA* deletion mutants do not exhibit an LT secretion-negative GspD$_\beta$ secretin-negative phenotype

The inability to complement the *leoA*^{UT} secretin-assembly negative phenotype by expression of LeoA *in trans* suggested that *leoA*^{UT} contains a secondary mutation that is responsible for the GspD$_\beta$ secretin assembly-negative phenotype of this strain. To confirm this possibility, *leoA* deletion mutations in ETEC strain H10407 were constructed in duplicate and the state of GspD$_\beta$ secretin assembly was determined by anti-GspD$_\beta$ immunoblot. As shown in Fig. 44, deletion of *leoA* had no effect on GspD$_\beta$ secretin assembly since the amount of GspD$_\beta$ secretin and monomer in the *leoA* strains were identical to that observed in wild-type.

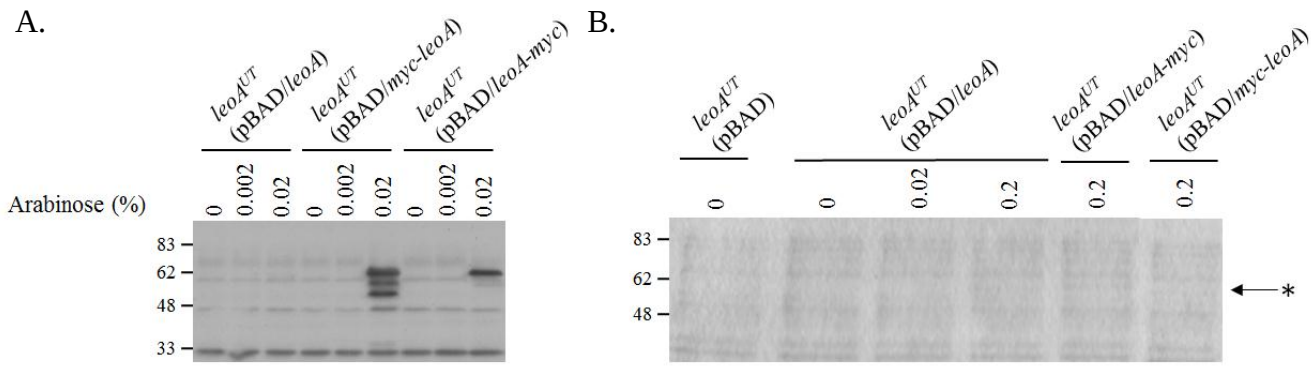


FIG. 42. Expression of LeoA variants can be detected upon expression *in vivo*. (A) Anti-myc immunoblot detection of N-terminal and C-terminal myc-tagged versions of LeoA expressed *in trans* from plasmids TS51 (pBAD/*leoA*), TS52 (pBAD/*leoA-myc*) and TS53 (pBAD/*myc-leoA*). (B) Coomassie-stain SDS-PAGE of whole cell samples from strains expressing *leoA*, *leoA-myc* and *myc-leoA*. The location of bands corresponding to LeoA and myc-tagged variants of LeoA is indicated by an asterisk. The concentration of arabinose used for induction of plasmid-encoded protein expression is given.

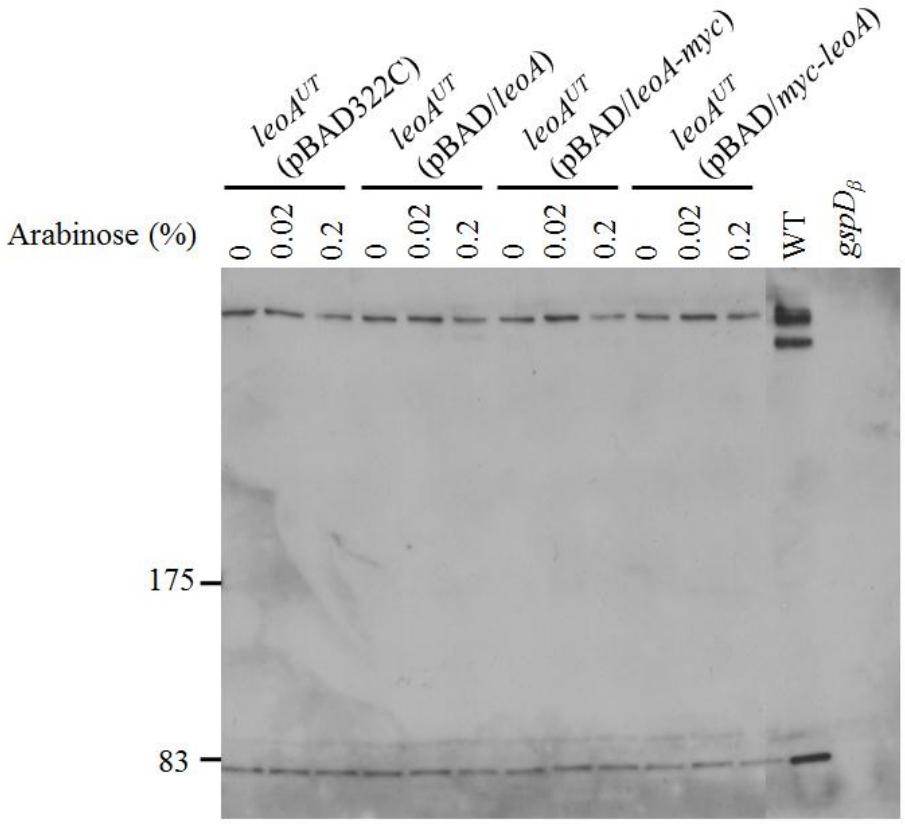


FIG. 43. Expression of LeoA *in trans* does not complement the GspD^{Vc} secretin-deficient phenotype of H10407 *leoA*^{UT}. Anti-GspD$_\beta$ immunoblot of whole cell samples taken from wild-type H10407 (WT), *gspD*$_\beta$, *leoA*^{UT} (pBAD322C), *leoA*^{UT} (pBAD/*leoA*), *leoA*^{UT} (pBAD/*leoA*-*myc*) and *leoA*^{UT} (pBAD/*myc*-*leoA*). The location of standard protein markers is given on the left.

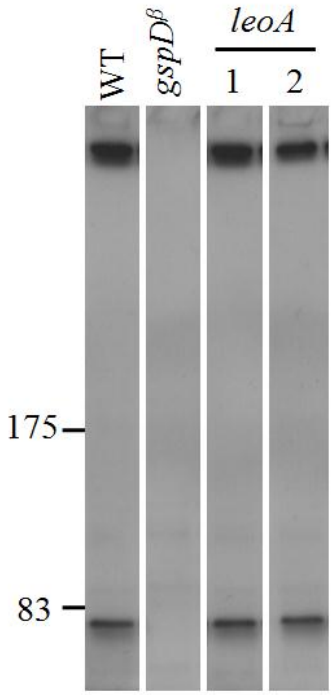


FIG. 44. **Deletion of *leoA* in H10407 did not confer a secretin-deficient phenotype.** The *leoA* deletion was constructed in H10407 in duplicate (*leoA* 1 and 2 as shown). Whole cell samples from wild-type H10407 (WT), *gspD*$^{\beta}$, *leoA* 1 and *leoA* 2 were immunoblotted with anti-GspD$^{\beta}$ antibody. The location of standard protein markers is given on the left.

6.3 Discussion

The *leoA* (labile enterotoxin output) gene is part of a pathogenicity island in the ETEC strain H10407 that also encodes the adhesin protein Tia (Fleckenstein *et. al.*, 2000). This gene was originally discovered in a mutational screen used to identify ETEC genes required for invasion of intestinal cells *in vitro*. A mutation in *leoA* generated an LT-secretion negative phenotype and a decrease in fluid accumulation in a rabbit ileal loop model of infection (Fleckenstein *et. al*., 2000), therefore suggesting that *leoA* is involved in the LT secretion mechanism of ETEC. Further analysis identified LeoA as possessing a cytoplasmic GTPase activity that was required for wild-type outer membrane vesicle production possibly due to an interaction with OmpA (Brown and Hardwidge, 2007). The finding that LeoA possesses GTPase activity suggested that LeoA could perform a function analogous to GspE in powering the T2SS for secretion of LT. I had hypothesized the LeoA could also perform the function of GspAB in peptidoglycan remodeling required for assembly of the $GspD_{\beta}$ secretin in ETEC since the $T2SS_{\beta}$ operon does not encode a GspAB homologue.

Consistent with my initial hypothesis, in the same H10407 *leoA* strain used in the studies described above (received the strain from Dr. James Fleckenstein from the University of Tennessee and have named it $leoA^{UT}$), this strain was found in fact to be $GspD_{\beta}$ secretin negative (Fig. 41). Therefore the inability to secrete LT in the $leoA^{UT}$ strain was due to the absence of assembled $GspD_{\beta}$ secretin, presumably due to the absence of LeoA. To confirm that the lack of LeoA was responsible for the $GspD_{\beta}$ secretin-negative phenotype, the *leoA* gene was cloned into the vector pBAD322C in three forms; one was wild-type, another encoded a C-terminal myc tag and the last encoded an N-terminal myc tag. Anti-myc immunodetection of whole cell samples from cultures of the $leoA^{UT}$ strain expressing the three versions of *leoA* identified myc-specific inducible bands that corresponded to expression of the N-terminal and C-terminal myc tagged versions of LeoA (Fig. 42A). Coomassie-stained whole protein profiles of the same samples confirmed induced expression of LeoA without the myc-tag (Fig. 42B). Unfortunately however, expression of any of the versions of LeoA did not complement the $GspD_{\beta}$ secretin-negative phenotype of the H10407 $leoA^{UT}$ strain (Fig. 43). This result suggested that the inability to secrete LT due to the absence of the $GspD_{\beta}$ secretin was not caused by the lack of LeoA and instead was caused by an unknown secondary mutation.

To confirm that the $GspD_{\beta}$ secretin assembly-negative phenotype of the *leoA*^{UT} strain was not actually due to the lack of *leoA*, two *leoA* mutant strains were constructed by lambda Red recombination that deleted the entire *leoA* ORF with exception of the N-terminal and C-terminal 10 amino acids encoded by the ORF. In these strains, the $GspD_{\beta}$ secretin remained assembled, therefore confirming that the $GspD_{\beta}$ secretin-negative phenotype was not caused by the lack of LeoA.

Although disappointing, the apparent non-requirement for LeoA in secretion of LT in ETEC strain H10407 was not surprising in retrospect given that in a study by Turner *et. al.*, (2006), of 116 ETEC strains isolated from humans with diarrhea (from Mexico, Guatemala and India), 34 ETEC strains isolated from cattle and several ETEC strains from the *Escherichia coli* Reference Collection, only 3% of strains contained the *leoA* gene. Therefore the vast majority of ETEC isolates that secrete LT do not encode *leoA*.

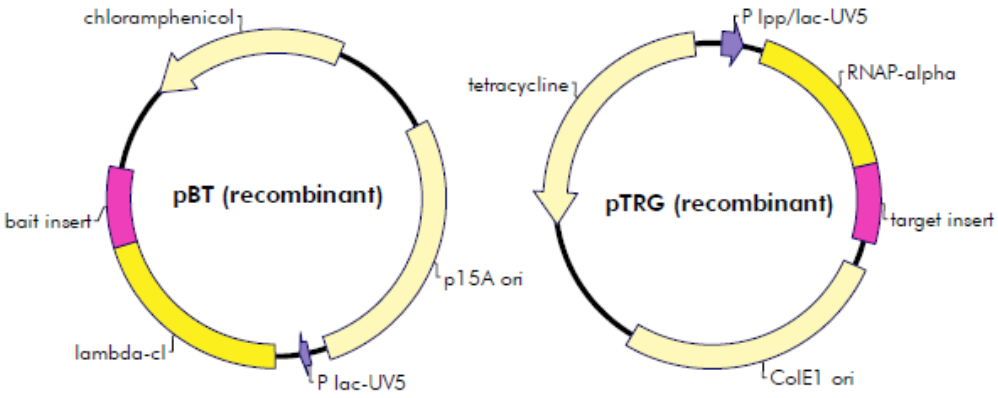
Chapter 7. Construction and use of a Bacterial II hybrid target library for identification of proteins that interact with the T2SS

7.1 Introduction

Yeast two hybrid and bacterial two hybrid methods are effective techniques used to investigate the binding characteristics between proteins. These techniques are generally used to confirm that an interaction exists between proteins and determine the relative strength of an interaction, whereby two putative interacting proteins are encoded on two plasmids and the ability of those proteins to interact is assessed. In this study, I was interested in developing a technique to utilize the capability of the bacterial II hybrid technique to screen for interactions between members of the T2SS and to identify potential interactions with members of the T2SS with all proteins in the ETEC proteome. To accomplish this, a library of overlapping ETEC strain H10407 genomic sequences was cloned into a bacterial II hybrid vector (hereafter referred to as the B2H library) and this library was used to identify potential interactions involving members of the T2SS.

The principles of the bacterial two-hybrid system as developed by Stratagene (Bacteriomatch II) are shown in Fig. 45. The gene that encodes one protein of interest (the bait) is fused to the cI lambda repressor encoded in the bait plasmid and the gene encoding the target protein is fused to the gene encoding the alpha-subunit of RNA polymerase in the target vector. Therefore, upon an interaction with the bait and the target proteins, the cI repressor and alpha-RNA polymerase are sequestered to locations adjacent to one another on the promoter of the reporter gene cassette located on the F' episome of the reporter *E. coli* strain. The promoter of the reporter gene cassette is a modified *lac* promoter whereby the CRP site has been replaced by a lambda operator (OR2). Therefore, upon recruitment of the cI repressor to OR2 and RNA polymerase to the transcriptional start via an interaction between bait and target proteins, transcription of reporter genes *his3* and *aadA* is activated. The *his3* gene encodes a protein involved in the biosynthesis of histidine that enables an *E. coli hisB* strain to grow on minimal media without histidine in the presence of the His3 competitive inhibitor 3-AT. The *aadA* gene serves as a secondary reporter gene and confers resistance to spectinomycin. By this method, clones that contain putative target-bait interactions are selected on minimal media containing 3AT and spectinomycin.

A.



B.

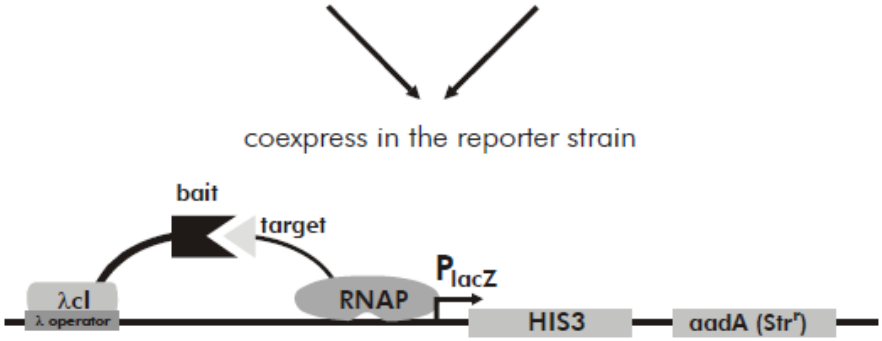


FIG. 45. The strategy used by the bacteriomatch II bacterial II hybrid system for selection of interacting proteins. (A) The bait and target proteins are cloned into plasmids pBT and pTRG respectively and co-transformed into the reporter strain. (B) The reporter strain encodes a set of reporter genes (*his3* and *aadA*) whose transcription is activated by recruitment of RNA polymerase (RNAP) mediated by the interaction between the bait and target proteins. Note that the promoter also encodes an operator sequence specific for recognition by the cI repressor protein fused to the bait protein. This figure was reproduced from the bacteriomatch II manual (Stratagene).

The advantages of using a prokaryotic two hybrid system over a yeast-two hybrid for selection of bacterial protein-protein interactions is that a bacterial cell is used to identify interactions that naturally occur in bacteria. One would expect that the likelihood of the tertiary structure of a prokaryotic protein and thus its ability to interact with other proteins would be considered to be more closely resembling wild-type when expressed in a prokaryote as opposed to a eukaryote since differences in post-translational modifications between kingdoms may seriously alter the activity of the protein of interest.

The bacterial II hybrid (B2H) technique has also been used to screen for interactions that exist involving a bait protein of interest and a library of target proteins. When attempting to identify putative target proteins, the success of the technique is dependent upon the quality of the library in terms of the representation of the proteome of the organism in the library. Traditionally, the library of DNA molecules that are cloned into the target vector for two-hybrid analysis have been constructed by random digestion of genomic DNA with frequent-cutting restriction endonucleases (Dyer *et al.*, 2010). Although this method is relatively simple, the number of translatable library molecules that can be generated in this manner is severely limited by the number of restriction sites in the genome. Therefore to increase the number of genomic sequences in the library and thus greatly enhance the representation of the proteome in the target vector, a technique previously employed in the genomic SELEX method was used to copy the ETEC genome for the B2H library.

Genomic SELEX (systematic evolution of ligands by exponential enrichment) is a method that determines the nucleic acid binding specificity of a DNA or RNA-binding protein (Singer, 1997). The process involves multiple rounds of selection whereby a library of nucleic acids molecules (DNA or RNA) derived from the genome of the organism of interest are incubated *in vitro* with the protein in question. Nucleic acids bound by the protein are separated from the protein, amplified by PCR (DNA) or RT-PCR (RNA), and incubated again with protein. With each successive round of selection, sequences recognized by the protein comprise a larger proportion of the total amplified pool of molecules. This method has been used to identify the nucleic acid binding specificity of proteins including the bacteriophage MS2 coat protein (Shtatland *et. al.*, 2000), human TAP(NXF1) (Zolotukin *et. al.*, 2001) *Drosophila* pre-mRNA splicing factor B52 (Kim *et. al.*, 2003) and the *Synechococcus* 7942 RNA-binding protein A (Strozen, 2005).

The libraries used in the genomic SELEX method are derived from the genome by copying the DNA with partially-degenerate oligonucleotides. By this method, libraries of overlapping DNA molecules (typically 2-4 x 10^6 molecules) that encode genomic sequences from the organism of interest are generated (Singer *et. al.*, 1997; Strozen, 2005). As an extrapolation of this technique, partially degenerate oligonucleotides were used to copy the H10407 genome into a set of overlapping genomic sequences that were cloned into the target vector to create the B2H library. In doing so, the sensitivity of the two-hybrid technique would be greatly enhanced due to greater representation of the proteome in the library in comparison to libraries created by restriction digestion of the genome. In addition, use of this library in the bacterial two-hybrid system could identify regions of proteins that are involved in mediating an interaction with the bait protein since the library would encode a range of overlapping DNA fragments that subsequently encode overlapping protein sequences within a specified size range. Therefore it is conceivable that the bait protein would interact with numerous protein fragments that all have in common a specific amino acid sequence. This sequence would thus serve as a putative motif involved in mediating the interaction of the target protein with the bait protein.

7.2 Results

7.2.1 pTRG library construction

The first step in construction of the B2H library was copying the H10407 genome (Fig. 46). This was accomplished by extension of the partially degenerate primer US107ran (ran<u>dom</u>) (Table 4) by the Klenow fragment of *E. coli* DNA polymerase following hybridization of the primer to template (genomic) DNA, a procedure termed "first-strand" extension. The US107ran primer consists of a 19 nucleotide fixed sequence at the 5' end and a nine nucleotide single-stranded DNA <u>ran</u>dom sequence located at the 3' end of the molecule. During first-strand extension, US107ran molecules are theoretically capable of hybridizing randomly to the genome via the degenerate portion of the primer. Following hybridization, the 3' end of the primer was extended by Klenow enzyme in the reaction mixture. Excess US107ran primer molecules that were not extended during first-strand extension were removed by spin column filtration (Qiagen). This procedure was required because in the next step (second-strand extension) a different primer, US108ran, was used instead of US107ran. The presence of excess

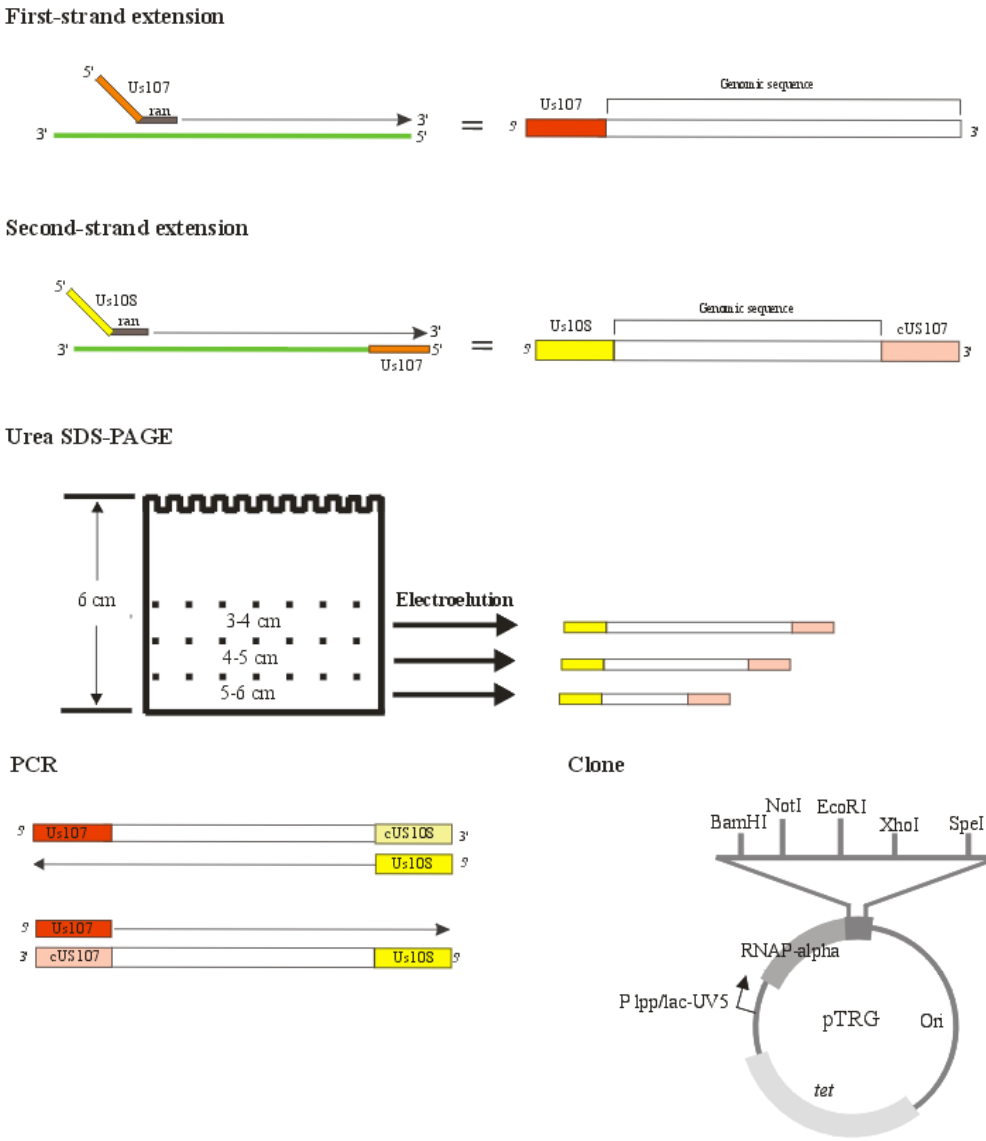


FIG. 46. Construction of the B2H H10407 library. In the first-strand extension reaction the partially degenerate primer US107ran hybridizes to the genomic template randomly via the 9 bp random nucleotide binding site (ran). The primer is extended by Klenow polymerase to generate ssDNA with the US107 sequence on the 5' end. The product of first-strand extension was used as template for the second-strand extension. Single-stranded DNA of various sizes are purified by urea SDS-PAGE. Purified ssDNA was amplified by PCR and cloned into pTRG via *NotI* and *SpeI* restriction sites.

US107ran primer molecules would interfere with US108ran extension by acting as a competitor for extension by Klenow enzyme and therefore excess first-extension primer had to be removed. The US107ran primer had encoded within it several features for construction of the library. The first was the 3' randomized end of the primer for random hybridization of the primer to the genomic template. The second was a *NotI* site encoded within the primer, this *NotI* site was introduced for cloning into pTRG such that a 6 amino acid linker would be encoded between the RNAP and the random protein fragment encoded within the H10407 genomic fragment. *NotI* was chosen for the cloning step because it recognizes an 8 bp restriction site which minimizes digestion of the genome during cloning into pTRG in comparison to that expected from a restriction enzyme with a standard 6 bp recognition site.

The second-strand extension reaction (Fig. 46) was performed using the partially degenerate primer US108ran (Table 4) that, like US107ran, consists of a fixed sequence at the 5' end of the molecule and a random sequence within the nine nucleotide positions on the 3' end of the molecule. The purpose of the second extension reaction was to introduce a second fixed sequence on the opposite 5' end of the US107ran extended products generated by first-strand synthesis, thereby creating molecules with H10407 genomic DNA sequences flanked by two different fixed sequences. The fixed sequence on the 5' end of US108ran encodes an *AvrII* site for cloning into pTRG and two stop codons encoded upstream of the *AvrII* site to ensure that upon expression of the RNAP-H10407 proteome fusion, translation will be terminated.

Once the H10407 genome was copied by the first and second strand extension reactions, ssDNA extension products of a specific size were separated from the genomic DNA template by denaturing urea polyacrylamide gel electrophoresis (PAGE) (Fig. 46). In this way the molecular size of the library and thus the length of H10407 genomic sequence encoded within each library molecule was selected.

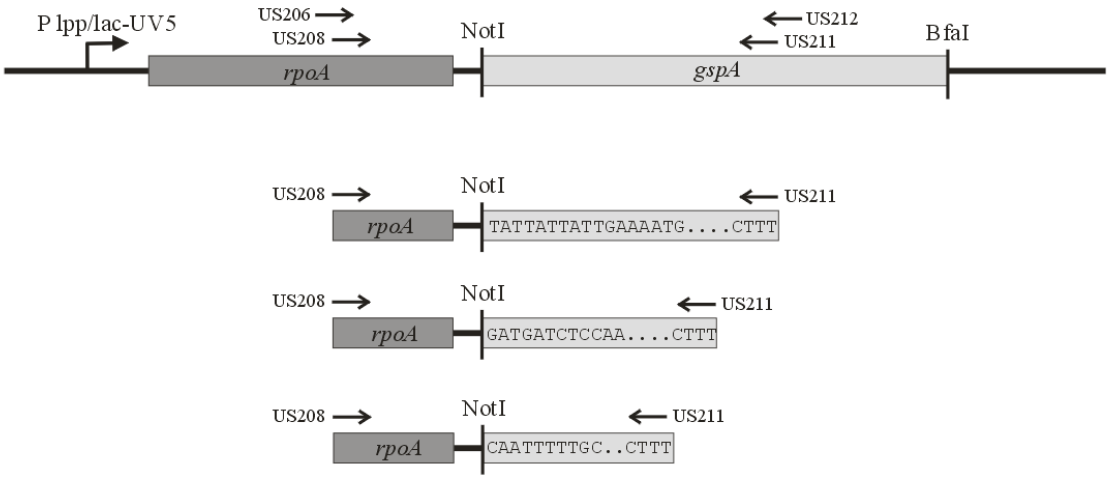
The purified 2nd strand extension product was amplified by PCR with primers US109 and US110 to generate enough DNA for cloning into B2H target vector pTRG. Amplification of the 2nd strand product revealed that the size of the library of molecules were approximately 400-500 bp on average. Therefore, once cloned, these genomic fragments would encode protein fragments approximately 150 aa in size. Following 60 electroporations of ligation mixtures of pTRG with B2H library molecules into highly competent DH5α, the library was estimated to contain approximately 7x10^6 clones.

7.2.2 Representation of the H10407 genome in the pTRG library

The B2H library contained approximately 7 x 10^6 clones of random fragments of the H10407 genome cloned into plasmid pTRG. To determine if the B2H library was sufficiently large enough to ensure that all proteins encoded by the genome are encoded in the library, the representation of a portion of the *gspA* gene in the B2H library was identified by nested-PCR. As shown in Fig. 47, primers US212 (specific for *gspA*) and US206 (specific for pTRG) were used to amplify members of the pTRG library that contained nucleotides 541 to 563 of *gspA*. To ensure that the products generated were in fact *gspA*-specific and not due to background amplification, a nested PCR reaction with primers specific for regions encoded upstream of US212 (primer US211) and US206 (US208) was used to recognize bases 473 to 497 of *gspA*. In this manner, nested PCR products were generated, cloned into plasmid pJET and sequenced. Fourteen different starting points over a span of 182 bp in the *gspA* gene were encoded adjacent to the *NotI* site in pTRG (Fig. 47). Therefore on average, a fragment of *gspA* would be encoded in the pTRG plasmid for every 13 bp in the *gspA* gene. If the two outlying clones were excluded from the analysis, there would be 12 clones over a span of 100 bp, translating into one clone per 8.3 bp of the genome. Since one clone of every 6 encoded in the pTRG library is translatable (due to three frames and two DNA strands) the library would contain approximately one translatable clone encoding a portion of the H10407 genome fused to the *rpoA* gene in pTRG per 49.8 bp. Therefore for the *gspA* gene that is 1470 bp in length, approximately 29 translatable fragments of *gspA* would be encoded in the pTRG library.

If one can assume that the representation of *gspA* in the B2H library is the same as the rest of the genome in the pTRG vector, then the size of the B2H library (assuming one genomic fragment per 8.3 bp of the genome) would be 6.18 x 10^5 given the size of the ETEC H10407 genome (5,153,435 bp) (Crossman *et. al*., 2011). Since during construction of the library, approximately 7 x 10^6 plasmids were isolated, each clone would be present 11 times in the library. In a study by Lander and Waterman (1985) that described the likelihood of obtaining a complete genomic sequence when sequencing whole genomes by shotgun approaches, a nine-fold coverage depth would be required to be 99% certain of obtaining the complete sequence assuming a poisson distribution of each base in the sequence. If that analysis can be

A.



B.

```
atgTCTACGA GAAGAGAAGT TATTCTTTCC TGGTTGTGTG AGAAACGTCA AACCTGGCGT
CTATGCTATT TGTTGGGTGA GGCTGGAAGT GGAAAAACCT GGCTGGCGCA GCAACTGCAA
AAAGATAAAC ATCGCCGTGT GATTACTTTA AGCCTCGTTG TTTCCTGGCA AGGTAAGGCC
GCATGGATCG TTACCGACGA TAACGCGGCT GAACAGGGCT GCCGTGACAG CGCCTGGACG
CGAGATGAGA TGGCGGGGCA ATTACTGCAT GCGCTTCATC GAACTGATAG TCGCTGCCCT
CTTATTATTA TTGAAAATGC TCACCTGAAT CACCGCCGCA TACTTGATGA TCTCCAACGC
GCAATATCTC TTATTCCTGA CGGTCAATTT TTGCTGATCG GCAGACCCGA TCGAAAAGTC
GAACGTGATT TTAAAAAACA GGGCATTGAA CTTGTCTCGA TAGGACGTCT GACGGAGCAC
GAACTTAAGG CAAGCATCCT TGAAGGTCAG AATATTGACC AACCCGATCT CCTTTTAACC
GCCAGAGTTC TGAAACGGAT AGCTTTATTA TGTCGGGGCG ATCGCAGAAA GCTGGCGCTT
```

FIG. 47. Identification of the representation of the H10407 genome in the B2H library. (A) Theoretical member of the B2H library that contains a portion of the *gspA* gene fused to the *rpoA* gene encoding the alpha-subunit of RNA polymerase in plasmid pTRG. (B) Examples of products generated by nested PCR amplification of the pTRG library. Nested PCR used primer sets US206/US212 and US208/US211 to amplify members of the pTRG library that encoded a portion of the *gspA* gene in plasmid pTRG. (C) Nucleotides 1 to 600 of the *gspA* gene, bases identified with a yellow background are those shown to be encoded adjacent to the *NotI* site in the pTRG library. For instance, the first underlined base T in the sequence TATTAT was encoded directly adjacent to the *NotI* site in a clone of the pTRG library and is shown in part B. Sequences recognized by primers US212 and US211 are shown in blue and red respectively.

extrapolated to this study, then coverage of each genomic fragment 11 fold in the library would ensure complete coverage of the ETEC genome (99% probability).

7.2.3 Screening of the pTRG library with T2SS proteins

The periplasmic domains of $GspD_{\alpha}$, $GspD_{\beta}$, GspA, GspB and LeoA were used as bait proteins in screens to detect members of the ETEC proteome that interact with these T2SS proteins. The most notable interactions identified include an interaction between $GspD_{\beta}$ as bait and three different overlapping clones of $GspD_{\alpha}$, as well as an interaction between $GspD_{\alpha}$ and GspA. Other proteins identified as potential interactors with $GspD_{\alpha}$ or $GspD_{\beta}$ include proteins involved in synthesis of LPS, the periplasmic protein FliY involved in cystine transport and the YjgP permease that is involved in the transport of LPS from the inner membrane to the outer membrane. Interestingly, both $GspD_{\alpha}$ and $GspD_{\beta}$ proteins have identified the highly conserved yet uncharacterized protein YdiA as a potential interactor. Although this work was preliminary, these results identify numerous interactions between components of the T2SS and other cell envelope proteins.

7.2.3.1 Screens with $GspD_{\alpha}$ and $GspD_{\beta}$

The periplasmic domains of $GspD_{\alpha}$ and $GspD_{\beta}$ encompassing subdomains N0, N1, N2 and N3 were cloned into plasmid pBT to create plasmids TS32 and TS33 respectively. TS32 encodes a 567 amino acid protein comprised of residues 26-349 of $GspD_{\alpha}$ fused to the 237 amino acid bacteriophage λ repressor (cI repressor) with a theoretical molecular weight of 62.3 kDa. The TS33 plasmid encodes a 568 amino acid protein comprised of residues 41-364 of $GspD_{\beta}$ fused to the cI repressor with a theoretical molecular weight of 62.2 kDa. The expression of cI-$GspD_{\alpha}$ and cI-$GspD_{\beta}$ fusions upon induction of TS32 and TS33 with IPTG was confirmed by immunoblot with anti-cI repressor antibody. As shown in Fig. 48, induction with 10 μM IPTG resulted in increased expression of cI-$GspD_{\alpha}$ and cI-$GspD_{\beta}$ fusion proteins.

Reporter strains containing TS32 or TS33 were electroporated with the pTRG library and the resultant transformation was spread on minimal media lacking histidine and containing 3AT with or without streptomycin. The absence or presence of streptomycin allowed selection of weak and strong interactors respectively, since the *aadA* is encoded downstream of *his3* in

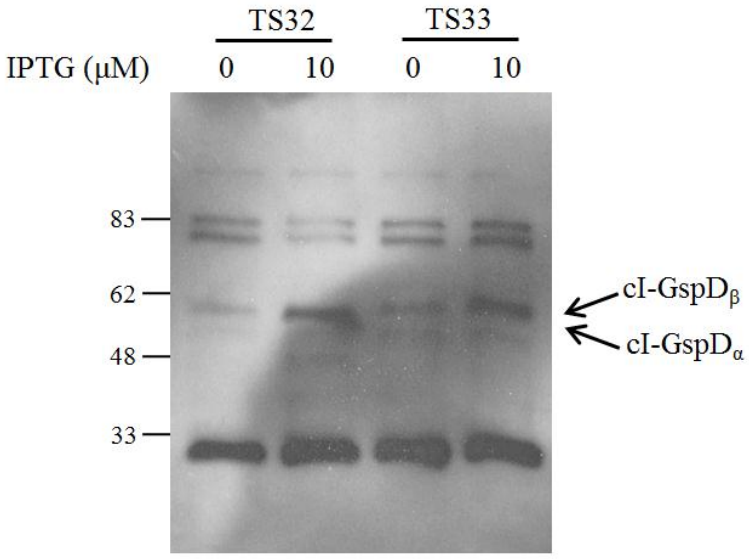


FIG. 48. **Immunoblot detection of cI fusion proteins encoded in plasmids TS32 and TS33.** Whole cell samples of bacteriomatch II reporter strain containing TS32 (pBT/cI-*gspD*$_{\alpha}$ $^{26\text{-}349}$) and TS33 (pBT/cI-*gspD*$_{\beta}$$^{41\text{-}364}$) were taken from culture grown in LB with or without the transcriptional inducer IPTG (as shown) and immunoblotted with anti-cI antibody. The location of fusion proteins cI-GspD$_{\alpha}$ and cI-GspD$_{\beta}$ are given on the right side of the panel and the location of standard protein markers are given on the left.

the reporter gene construct. Screens for interactors revealed several interesting results as described below.

In several screens for interactors with $GspD_{\beta}$ 41-364, a total of seven overlapping fragments of $GspD_{\alpha}$ were identified in a total of 30 clones (Fig. 49). Surprisingly, fragments of $GspD_{\alpha}$ were not identified in the $GspD_{\alpha}$ 26-349 screen nor were clones corresponding to $GspD_{\beta}$ observed in the $GspD_{\beta}$ 41-364 screen. As shown in Fig. 50, the portion of $GspD_{\alpha}$ identified in the screen with $GspD_{\beta}$ corresponds to the N1 domain of $GspD_{\alpha}$, therefore suggesting that the N1 domain is involved in mediating interactions between monomers in the secretin. As mentioned in the introduction, the N1 domain of the secretin is involved in an interaction with the N0 subdomain to form a globular lobe of the secretin periplasmic domain (Korotkov *et. al.*, 2009). The N0 subdomain has been shown to interact with GspC (Korotkov *et. al.*, 2009), whereas the N1 subdomain has not been designated as involved in any specific function other than interacting with N0. This result suggests that interactions between N1 subdomains are part of the interactions between monomers necessary for secretin assembly. Note that identification of overlapping fragments of $GspD_{\alpha}$ in the screen enabled refinement of the amino acids involved in putative interaction with $GspD_{\beta}$ (Fig. 49).

The amino acid similarity between $GspD_{\alpha}$ and $GspD_{\beta}$ suggests that a screen of potential interactors with each protein should reveal proteins common to both screens. Only one protein however, YdiA, was identified in a screen with both $GspD_{\alpha}$ and $GspD_{\beta}$. This protein was not identified in screens with other proteins or in the control screen (the cI repressor without a fusion protein attached to it) therefore proving that it was not selected independent of $GspD_{\alpha}$ or $GspD_{\beta}$. As shown in Fig. 51, three overlapping clones corresponding to the C-terminal end of YdiA were identified in screens with $GspD_{\alpha}$ and $GspD_{\beta}$. Although these interactions are interesting, the YdiA protein is considered a putative cytoplasmic phosphotransferase, thereby suggesting that the interaction with the GspD bait proteins may not be biologically relevant.

Interestingly, a screen with $GspD_{\alpha}$ identified one clone that encoded a portion of the GspA protein. The region of GspA identified in the screen is encoded within the periplasmic domain immediately adjacent to and within the peptidoglycan-binding site of GspA (Fig. 51). The GspA family of proteins, characterized by the $GspA^{Ah}$ protein in *A. hydrophila* and $GspA^{Vc}$ in *Vibrio* species are required for secretin assembly by functioning in peptidoglycan remodeling (Li and Howard, 2011; Li *et. al.*, 2011), a function required for assembly of the

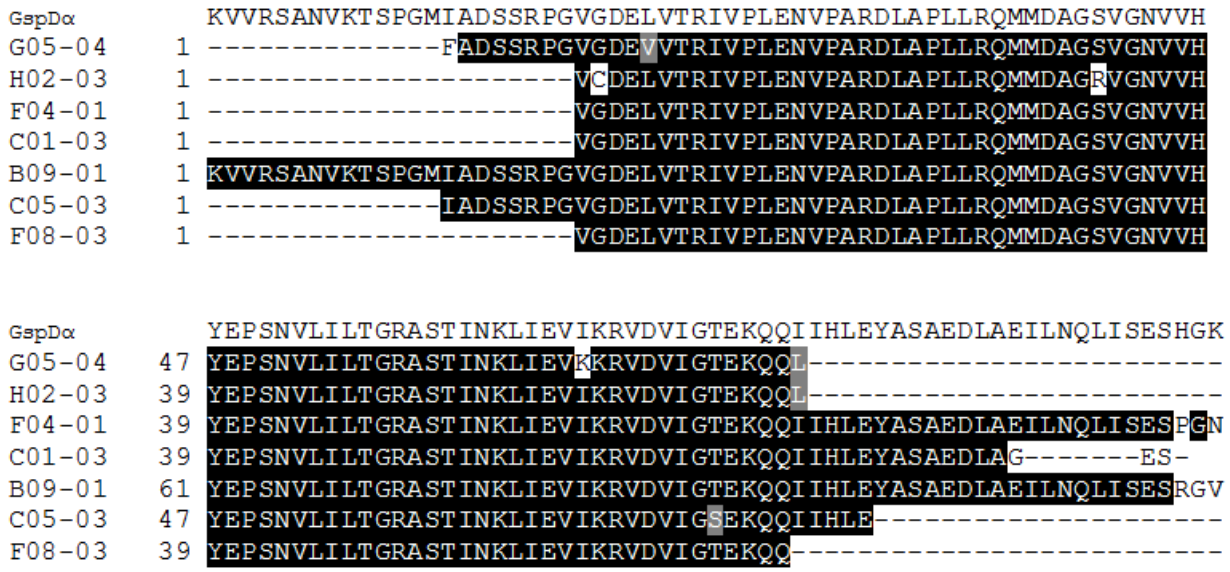


FIG. 49. Amino acid alignment of the seven different fragments of $GspD_{\alpha}$ identified in a screen with $GspD_{\beta}$ $^{41-364}$. Identical amino acids in comparison to amino acids 120-193 of $GspD_{\alpha}$ are given in reverse type.

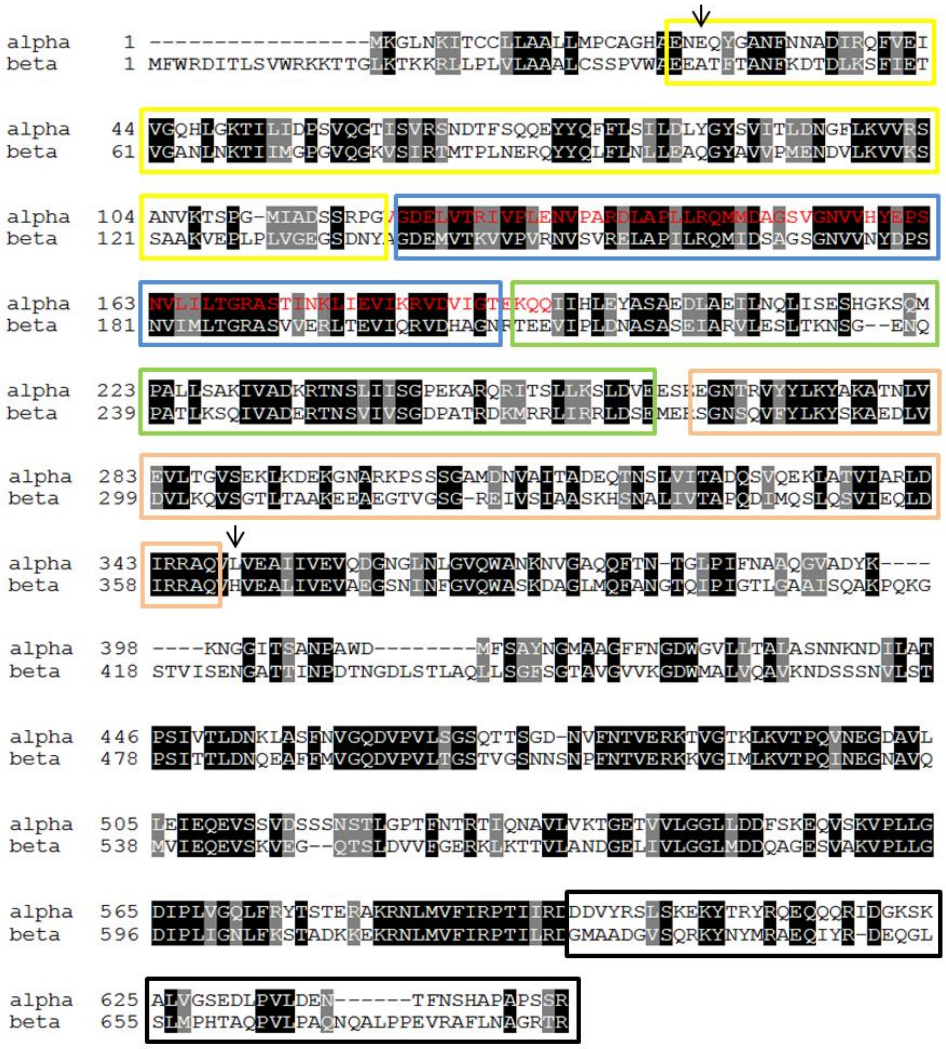


FIG. 50. The GspD$_\alpha$ fragment identified in a B2H library screen with GspD$_\beta$ corresponds to the N1 domain of secretin proteins. The N0, N1, N2 and N3 periplasmic domains of secretin proteins GspD$_\alpha$ (alpha) and GspD$_\beta$ (beta) are designated with coloured boxes yellow, blue, green and orange respectively. The S domain is identified by a black coloured box. Amino acids 121-193 of GspD$_\alpha$ identified in the GspD$_\beta$ B2H screen are designated by red lettering. The fragments of GspD$_\alpha$ and GspD$_\beta$ encoded in plasmids TS32 and TS33 are comprised of the amino acids between and including the area designated by arrows.

A.

Selected by $GspD_{\alpha}$

Clone	Sequence
E12-02	----------YRKNQIPWINSTNYSVEEIATKILDIMGLSRRMY 34
F02-02	SRKEVAEVEALYRKNQIPWINSTNYSVEEIATKILDIMGLSRRMY 45
A02-03	-QVEVAEVEALYRKNQIPWINSTNYSVEEIATKILDIMGLSRRMY 44

Selected by $GspD_{\beta}$

Clone	Sequence
A08	----------LRKNQIPWINSTNYSVEEIATKILDIMGLSRRMY 34
9	SRKEVAEVEALYRKNQIPWINSTNYSVEEIATKILDIMGLSRRMY 45
1	-QVEVAEVEALYRKNQIPWINSTNYSVEEIATKILDIMGLSRRMY 44

B.

YdiA
MDNAVDRHVF YISDGTAITA EVLGHAVMSQ FPVTISSITL PFVENESRAR
AVKDQIDAIY HQTGVRPLVF YSIVLPEIRA IILQSEGFCQ DIVQALVAPL
QQEMKLDPTP IAHRTHGLNP NNLNKYDARI AAIDYTLAHD DGISLRNLDQ
AQVILLGVSR CGKTPTSLYL AMQFGIRAAN YPFIADDMDN LVLPASLKPL
QHKLFGLTID PERLAAIREE RRENSRYASL RQCRMEVAEV EALYRKNQIP
WINSTNYSVE EIATKILDIM GLSRRMY

C.

GspA
MSTRREVILS WLCEKRQTWR LCYLLGEAGS GKTWLAQQLQ KDKHRRVITL SLVVSWQGKA
AWIVTDDNAA EQGCRDSAWT RDEMAGQLLH ALHRTDSRCP LIIIENAHLN HRRILDDLQR
AISLIPDGQF LLIGRPDRKV ERDFKKQGIE LVSIGRLTEH ELKASILEGQ NIDQPDLLLT
ARVLKRIALL CRGDRRKLAL AGETIRLLQQ AEQTSVFTAK QWRMIYRILG DNRPRKMQLA
VVMSGTIIAL TCGWLLLSSF TATLPVPAWL IPVTPVVKQD MTKDIAHVVM RDSEALSVLY
GVWGYEVPAD SAWCDQAVRA GLACKSGNAS LQTLVDQNLP WIASLKVGDK KLPVVVVRVG
EASVDVLVGQ QTWTLTHKWF ESVWTGDYLL LWKMSPEGES TITRDSSEEE ILWLETMLNR
ALHISTEPSA EWRPLLVEKI KQFQKSHHLK TDGVVGFSTL VHLWQVAGES AYLYRDEANI
SPETTVKGK

FIG. 51. A portion of the putative phosphotransferase protein YdiA was identified in a B2H screen using both $GspD_{\alpha}$ and $GspD_{\beta}$ as bait proteins. The same three overlapping clones were selected in screens with (A) $GspD_{\alpha}$ and $GspD_{\beta}$. (B) The location of the GspD-binding site in YdiA (in red lettering) is located at the C-terminus of the protein. (C) The amino acid sequence of GspA, in red is the domain recognized by $GspD_{\alpha}$ in a B2H screen. The location of the peptidoglycan-binding domain is underlined.

$GspD^{Ah}$ secretin in the outer membrane. The GspA-containing clone identified in this screen suggested that GspA could physically interact with the GspD secretin, thereby providing a tether between proteins of the inner membrane complex and the secretin.

As shown in Table 7, screening with the $GspD_{\alpha}$ and $GspD_{\beta}$ periplasmic domains revealed interactions with numerous proteins that are theoretically impossible *in vivo* due to the cellular location of target proteins in the cytoplasm. One possibility therefore is that these proteins were selected independent of an interaction with the bait protein. This possibility however is not likely because only one protein has been identified in a screen with an empty bait vector (cI repressor expressed only) and in every screen performed regardless of the bait protein attached to cI. This protein as shown in Table 7 is a putative prophage protein with no identifiable conserved domains that likely binds to the cI repressor. Therefore, the interactions between proteins shown in Table 7 are likely legitimate B2H interactions. Note that deletion of genes *fliY, wecF, kdtA, fliE, cdaR, ydiA* and *mdlA* did not affect the amount of $GspD_{\beta}$ secretin observed in ETEC or affect the ability of ETEC to secrete LT (analysis performed by Jeff Elder, unpublished), thereby providing evidence that these proteins are not involved in $GspD_{\beta}$ secretin assembly or function. The most likely scenario as to why this group of proteins were selected in these screens is that these proteins have in common a tertiary structure recognized by the KH-like domains encoded in domains N1, N2 and N3 of $GspD_{\alpha}$ and $GspD_{\beta}$.

7.2.3.2 Screens with GspA, GspB and LeoA

The periplasmic domains of GspA (residues 244-271) and GspB (residues 64-139) were cloned into pBT to create plasmids TS38 and TS39 respectively. Amino acids 283-577 of LeoA were cloned into pBT to create plasmid TS43. Screens with GspA and GspB revealed several interesting interactors located in the inner membrane, the most interesting was a putative interaction between GspB and YfhM. Screens with GspB identified four overlapping clones of the protein YfhM (Fig. 52A). YfhM is a large (1653 amino acid) lipoprotein localized to the outer leaflet of the inner membrane that likely forms a homo-oligomeric complex (Maddalo *et. al*., 2011) and an alpha2-macroglobulin (α_2M) domain (consensus sequence CxEQ) that has been characterized in eukaryotes and exists in a wide variety of species from corals to humans. Proteins with α_2M domains function to entrap attacking proteases by creating a cage-like structure whereby proteases activate the α_2M domain by degrading the "bait" region of the

Table 3: Synopsis of interactors identified by screening with the B2H library

Target	Sequence identified in screen	Location	Function
Bait Protein GspDα			
GspA	EGESTITRDSSEEEILWLETMLNRAL	Inner membrane	Peptidoglycan remodeling
FliY	GNVDLLKAVNDAIAEMQKDGTLQALSEKWFGADVTK	Periplasmic	Cysteine-binding
WecF	VAGMTVLIHVLGSDIPHHNRTVLRFFNDALAATSEHAREFMVVGKDDGLSD SCPALSVQFFPGKKSLAEAVIAKAKANRQQRFFFHGQFNPTLWLALLCGN	cytoplasmic	lipid III biosynthesis
KdtA	GERYGFYRHPLKPGGIMLHSVSVGETLAAIPLVRALRHRYPDLPITVTTMTPTGSERVQSAFGKDV QHVYLPYDLPDALNRFLNKVDPKLVLIMETELWPNLIAALHKRKIPLVIANARLSARSAAGYAKRGN	Inner membrane	KDO transferase
FliE	RNKLVAAYQEVMSMQV	Inner membrane	Flagellar basal body
HrsA	GEGLVVPHGKTAAVKEAAFAVATLSEPLQWEGVDGPEAVDLVVLLAIPPNEAGT THMQLLTALTTRLADDEIRARIQSATTPDELLSALDDKGGTQPSASFSNAPTIVCVTACPTE	Inner membrane	Phosphotransferase type permease
CdaR	TPGYYALTTPERNNLVAIVSLTEMVVLKPALNSFGRWDAEDHRKRVEQLITRMKEY GQLRFRVSLGNYFTGPGSIARSYRTAKTTMVVGNSGCQKVAAIFIRI	Cytoplasm	DNA-binding protein, activator
CysS	GVMVDREKMSKSLGNFFTVRDVLKYYDAETVRYFLMSGHYRSQLNYSEENLKQARAALERLYTAWSNIDKTVAPAGGEAFEARFY	Cytoplasm	Cysteinyl-tRNA synthetase
LeuA	SVKEKLQIALALERMGVDVMEVGFPVSSPGDFESVQTIARQVKNSRVCALARCVEKDIDVAADGLHRVNTHTLK	Cytoplasm	2-isopropylmalate synthase
gpU	HGVGQIYGMYVLTGTNTTRSEFDRYGKAKKIEFSLTLERCDEDLRERLQSSSFSDMLSGFKDKVTSSLNSAASSVKGMF		Phage protein gpU-tail protein
TrkA	QVVNFAEGKVSLAVVKAYYGGPLIGNALSTMREHMPHIDTRVAAIFRHDRPIRPQGSTIVEAGDEVFFIAASQHIRAVMSELQRLVTKS	Inner membrane	Nad-binding component of TrK potassium transporter
YdiA	SRKEVAEVEALYRKNQIPWINSTNYSVEEIATKILDIMGLSRRMY	?	Putative phosphotransferase
Bait Protein GspDβ			
MdlA	KAGQMLGICGPTGSGKSTLLSLIQRHFDVSEGDIRFHDIPLTKLQLDSWRSRLAVVSQTPFLFSDTVAN NIALGCPNATQQEIEHVARLASVHDDILRLPQGYDTEVGERGVMLSGGQKQRN	Inner membrane	Putative drug efflux pump
DUF3672	VWRMTILGAEITPEGGRYTGYGFEVRKNGVLIASRETKGAIPGSYSAGY	?	phage tail protein J
GspDα	VGDELVTRIVPLENVPARDLAPLLRQMMDAGSVGNVVHYEPSNVLILTGRASTINKLIE VIKRVDVIGTEHQQIIHLEYASAEDLAEILNQLISESPGN	outer membrane	secretin
YdiA	SRKEVAEVEALYRKNQIPWINSTNYSVEEIATKILDIMGLSRRMY	?	Putative phosphotransferase
YjgP	ARRGNEFVTVKSLNEQGQPVGVEIFHYRDDLSLESYIYARSATIVTKS	?	?
Aec62	RQHSLANRLISLFRQFSGTELRLKLVWLCWYDLMIGNSLKDWTENLKRKSEKELAEWVIDRQKQNAALTDLMDQ YVLLAYRTTVDDNRT	?	?
Mak	CSSGEWGHNPLPWMDEDELRYREEVPCYCGKQGCIETFISGTGFAMDYRRLSGHALKGSEIIRLVEE SDPVAELALRRYELRLAKSLAHVVNILDR	cytoplasm	Mannofructokinase
YcaO	ERTVTELLQGRGLKDLDVFTPPTFDDEEVAEHTNLETHFIDSSGLISWDLFKQDADYPFVDWNFSGTTEEEFAT LMAIFNKEDKEVYIADYEHLGVYACRIIVPGMSDGN	cytoplasm	Methylation of 30S ribosomal subunit S12
YedP	RGLNREGVHLHDEDPARVWRTQREGPEGWREGLDHFFSAH	cytoplasm	putative phosphatase
AtoS	LISLLLIVLFSRRLSANIDIITDGLSTLAQNIPTRLPQLPGEMGQISQSVNNLAQALRETRTLNDLIIENAAD GVIAIDRQGDVTTMNPAAKCN	Inner membrane	histidine kinase
PrpB	NEEMVDRIRAAVDAKTDPDFVIMARTDALAVEGLDAAIERAQAYVEAGAEMLFPEAITELAMYRQFADAV QVPILANITEFGATPLFTIDELRSAHVAMALYPLSAFRAMNRAAEQ	cytoplasm	2-methylisocitrate lyase
Bait Protein GspA			
RhaM	HIRPELEAVLKSHGAHNYAIYLDKARNLLFAMVEIESEERWNAVASTDVCQRWWKYMTDVMPANPDNSPVSSELQEVFYLP	cytoplasm	L-Rhamnose mutarotase
Gmd	QWMMLQQEQPEDFVMGTGVQYSVRQFVEMAAAQLGIKLRFEGTGVEEKGIVVSVPGHDAPGVKPGDVIIAVDP RYFRPAEVETLLGDATKAHEKQ	cytoplasm	GDP-mannose 4,6-dehydratase, exopolysaccharide biosynthesis
YfhW	SLGHQEILGTRPLPPLRMPFRDVIDRVEQALVSAGWQVERRGDDLQFLWVNQAVAIGDNLEADLGQVYNITA NLSVISFDDAIKIGRIVREQVQVGRVITFGGLLTDSGN	cytoplasm	Ferredoxin component of 3-phenylpropionate dioxygenase system
PcnB	LLCEYHLFQPLFPTITRYFTENGDSPMERIIEQVLKNTDTRIPNDMRVNPAFLFAAMFWYPLLETAQKIAQE SGLTYHDAFALAMNDVLDEACRSLEIRN	cytoplasm	poly A polymerase
AtoC	YQGFETHCANNGRTALHLFADIHPDVVLMDIRMAEMDGIKALKEMRSPETRTPVILMTAYAEVETAVE ALRCGAFDYVIKPFDLDEWNLIVQCGRIAIWI	cytoplasm	response regulator of AtoSC two component system
GabT	LAIAEKHPEIGDVRGLGAMIAIELFEDGDHNKPDAKLTAEIVARARDKGLILLSVT	cytoplasm	4-aminobutyrate aminotransferase
YhiD	LQGNEDHEVVAIDIQLHATTSIEDLLRLLKGMAGVKGVSIS	Inner membrane	Putative transport protein
YhdH	IGDWVVAMPQGLDARKAMIIGTAGFTAMLCVMALEDAGVRPQDGEIVVTGASGGVCSTAVALL HKLGYQVVAVSGRESTHEYLKSLGASRVLPRDEFAQ	cytoplasm	Putative oxidoreductase
YnfM	HCYMELTQSRNHDHPLWAWSSDVVFDGGYAVWFTDDLIQLAVADLCRNVTLLSRILRSPLSSQQLDRPPRKTR	Inner membrane	Putative proton-driven drug efflux system
YafK	LLIPCVSFAGLLGSSSSTTPVSKEYKQQLMGSPVYIQIFKEERTLDLYVKMGEQYQLLDSYKICN	Secreted	Required for biofilm production
Bait Protein GspB			
MurI	LSVYDEIRHLLPDLHYIYAFDNVAFPYGEKSEAFIVERVVAIVTAVQERYPLALAVVACNTASTVSIP		Glutamate racemase for peptidoglycan synthesis
YfhM	LWAGYSWQDNSDGSGAVRPDRVTLKLDKASYRPGDTIKLHIAAPTAGKGYAMVESSEGPLWWQEIDVPAQGLDLT IPVDKTWNRHDLYLSTLVVRPGDKSRSATPKR	Inner membrane	alpha-macroglobulin
ManY	GSILDEFQFHRPLVACTLVGIVLGDMKTGIIIGGTLEMIALGWMNIGAAVAPDAALASIISTILVIAGHH SIGAGIALAIPLAAAGQVLTIIVRN	Inner membrane	Part of the ManXYZ mannose permease
AcrD	MGDKIIVAAGLIVSVCGAYLSWTIMAAEVPFLAATHKAFPRIFARQNAQAAPSASLWLTNIS	Inner membrane	Component of the AcrAC-TolC drug efflux system
Bait Protein LeoA			
YejA	LFNRYALRGNPGARTEQLYDTLFTTSDDEPGSYYPLIAESARYADDYSWVEVAINPRARFHDGSPITARDVEFTFQKFM TEGVPQFRLVYKGTTVKAIAPLTVRIELAKPGKE	periplasm	Part of an ATP-dependent oligopeptide permease
NarQ	EGQPNPELPMQILPVTMQETVYGELHWQNSHVSSSEPLLNSVSSMLGRGLYFNQAQKHFQQLLLMEERATIA RELHDSLAQVLSYLRIQLTLLKRSIPEDNATEK	Inner membrane	Histidine kinase
YhiS	GVGDTIFLHKEHPTPKLYKDGQYLDKQIEGLFSTLLTINDNLLRAKAEIASTIIKFLEARITNLSYND	?	?
No Bait			
?	LCREEPYPDTFTLADRTTERLTRARTGLVHVMNEILPSVGGEQATVINSWLQKVTSLIDISLIDVESAK	?	Putative prophage protein

A.

```
H01/1-171     1 GKGWVIEMETNGTIDPEEVIRRAATILAEQLEAFVDLRDVRQTEVKEEKPEAVGSAVALTWAGYSWQDNSDGRGAVRPDR
E10/1-116     1 ----------------------------------------------------------AAALLWAGYSWQDNSDGSGAVRPDR
G09/1-114     1 ----------------------------------------------------------AAALLWAGYSWQDNSEGGGAVRPDR
F02/1-68      1 ----------------------------------------------------------AAALLWAGYSWEDNREGRGAVGPDR

H01/1-171    81 VTLKLDKASYRPGDTLKLHIAAPTAGKGYAMVESSEGPLWWQEIDVPAQGLDLTIPVDKTWNRHDVYLSTLVVRPGDKS-
E10/1-116    26 VTLKLDKASYRPGDTIKLHIAAPTAGKGYAMVESSEGPLWWQEIDVPAQGLDLTIPVDKTWNRHDIYLSTLVVRPGDKS-
G09/1-114    26 VTLKLDKASYRPGDTIKLHIAAPTAGKGYAMVESSEGPLWWQEIDVPAQGLDLTIPVDKTWNRHDIYLSTLVVRPGDKSP
F02/1-68     26 VPLKLDKASYRPGDTIKLDIAAP-----------------------------------------------------------

H01/1-171   160 -RSATPKRGVTKS
E10/1-116   105 -RSATPKRGVTKS
G09/1-114   106 LRDAQTEGN----
F02/1-68        ------------
```

B.

```
MKKLRVAACMLMLALAGCDNNDNAPTAVKKDAPSEVTKAASSENASSAKLSVPERQKLAQQSAGKVLTLLDLSEVQLDGAATLVLTFSIPLDPDQD
FSRVIHVVDKKSGKVDGAWELSDNLKELRLRHLEPKRDLIVTIGKEVKALNNATFSKDYEKTITTRDIQPSVGFASRGSLLPGKVVEGLPVMALNV
NNVDVNFFRVKPESLPAFISQWEYRNSLANWQSDKLLQMADLVYTGRFDLNPARNTREKLLLPLGDIKPLQQAGVYLAVMNQAGRYDYSNPATLFT
LSDIGVSAHRYHNRLDIFTQSLENGAAQQGIEVSLLNEKGQTLTQATSDAQGHVQLENDKNAALLLARKDGQTTLLDLKLPALDLAEFNIAGAPGY
SKQFFMFGPRDLYRPGETVILNGLLRDADGKALPNQPIKLDVIKPDGQVLRSVVSQPENGLYHFTWPLDSNAATGMWHIRANTGDNQYRMWDFHVE
DFMPERMALNLTGEKTPLTPKDEVKFSVVGYYLYGAPANGNTLQGQLFLRPLREAVSALPGFEFGDIAAENLSRTLDEVQLTLDDKGRGEVSTESQ
WKETHSPLQVIFQGSLLESGGRPVTRRAEQAIWPADALPGIRPQFASKSVYDYRTDSTVKQPIVDEGSNAAFDIVYSDAQGVKKAVSGLQVRLIRE
RRDYYWNWSEDEGWQSQFDQKDLIENEQTLDLKADETGKVSFPVEWGAYRLEVKAPNEAVSSVRFWAGYSWQDNSDGSGAVRPDRVTLKLDKASYR
PGDTIKLHIAAPTAGKGYAMVESSEGPLWWQEIDVRAQGLDLTIPVDKTWNRHDLYLSTLVVRPGDKSRSATPKRAVGVLHLPLGDENRRLDLALE
TPAKMRPNQPLTVKIKASTKNGEKPKQVNVLVSAVDSGVLNITDYVTPDPWQAFFGQKRYGADIYDIYGQVIEGQGRLAALRFGGDGDELKRGGKP
PVNHVNIVVQQALPVTLNEQGEGSVTLPIGDFNGELRVMAQAWTADDFGSNESKVIVAAPVIAELNMPRFMASGDTSRLTLDITNLTDKPQKLNVA
LTASGLLELVSDSPAAVELAPGVRTTLFIPVRALPGYGDGEIQATISGLALPGETVADQHKQWKIGVRPAFPAQTVNYGTALQPGETWAIPADGLQ
NFSPVTLEGQLLLSGKPPLNIARYIKELKAYPYGCLEQTASGLFPSLYTNAAQLQALGIKGDSDEKRRASVDIGISRLLQMQRDNGGFALWDKNGD
EEYWLTAYVMDFLVRAGEQGYSVPTDAINRGNERLLRYLQDPGMMSIPYADNLKASKFAVQSYAALVLARQQKAPLGALREIWEHRADAASGLPLL
QLGVALKTMGDATRGEEAIALALKTPRNSDERIWLGDYGSSLRDNALMLSLLEENKLLPDEQYTLLNTLSQQAFGERWLSTQESNALFLAARTIQD
LPGKWQAQTSFSAEQLTGEKAQNSNLNSDQLVTLQVSNSGDQPLWLRMDASGYPQSAPLPANNVLQIERHILGTDGKSKSLDSLRSGDLVLVWLQV
KASNSVPDALVVDLLPAGLELENQNLANGSASLEQSGGEVQNLLNQMQQASIKHIEFRDDRFVAAVAVDEYQPVTLVYLARAVTPGTYQVPQPMVE
SMYVPQWRATGAAEDLLIVRP
```

C.

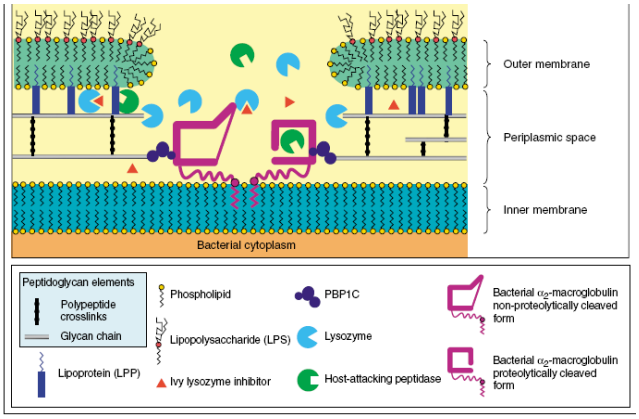


Fig. 52. YfhM was identified in a B2H screen using GspB as the bait protein. (A) Four overlapping clones were identified in screens with GspB. (B) The amino acid sequence of YfhM. YfhM is a 1653 amino acid protein with a signal peptide (red) and lipobox domain (mature N+1 Cys residue in green). The location of residues in the protein involved in an interaction with GspB are given in blue. The alpha-2 macroglobulin domain CxEQ is shown in orange. (C) The putative function of YfhM as a periplasmic defense system. Host defenses

(antimicrobial peptides, antibody or complement) have removed outer membrane material, thereby allowing lysozyme and host proteases access to the periplasmic space, leading to destruction of peptidoglycan. The bacterial alpha2-macroglobulin protein YfhM functions to entrap proteases thereby giving PBP 1C an opportunity to repair the peptidoglycan. The schematic diagram was reprinted from Budd *et. al*., (2004) with permission from the journal Genome Biology.

α_2M protein that in turn activates the thioester motif of the protein. The thioester motif in turn functions to crosslink and thereby sequester proteases in the structure. In bacteria such as *E. coli*, *yfhM* is always encoded in the same operon with the *pbp1C* gene encoding penicillin-binding protein 1C (PBP1C) (Budd *et. al.*, 2004). PBP1C is a paralog of the peptidoglycan biosynthesis proteins PBP1A and PBP1B that contain amino-terminal transglycosylase and carboxy-terminal transpeptidase domains but also includes a C-terminal 100 amino acid extension of unknown function (Budd *et. al.*, 2004) . PBP1C cannot substitute for PBP1A or PBP1B during vegetative growth since it is responsible for 3% of *de novo* peptidoglycan biosynthesis (Schiffer and Holtje, 1999). As a result, it is hypothesized that YfhM functions together with PBP1C to protect the cell envelope from degradation by sequestering invading proteases and repairing the peptidoglycan (Fig. 52). How the T2SS protein GspB would utilize the large YfhM multimer is not readily apparent, however if the GspB-YfhM interaction were valid, the GspAB complex would function in a manner similar to PBP1C by localizing YfhM to a specific region in the cell envelope given the peptidoglycan binding activity of GspA and the fact that GspA and GspB form a complex.

7.3 Discussion

The B2H library of overlapping ETEC genomic fragments represents a novel mechanism of identifying protein interactions. The usefulness of the technique lies in the fact that it employs the widely utilized technique of bacterial II hybrid to identify proteins that potentially interact with a target protein of interest. Following construction of the library itself, screening for potential interactions is simple by only requiring electroporation of the library in cells containing the bait plasmid.

Construction of the library was proven to be successful by enabling detection of gene fragments encoding multiple overlapping fragments of a protein that interacts with the target of

interest to be identified in the screen (Figs. 49, 51 and 52), thereby enabling one to define the site involved in the interaction simply by sequencing more clones grown in the screen.

The most notable interactions identified in screens using $GspD_{\alpha}$, $GspD_{\beta}$, GspA, GspB and LeoA were interactions between T2SS proteins $GspD_{\alpha}$ and $GspD_{\beta}$ in addition to GspA and $GspD_{\alpha}$. In addition to interactions between T2SS proteins, the screens identified numerous target proteins that potentially interact with the bait proteins however many of the target proteins identified are cytoplasmic proteins and therefore would not represent a physiologically relevant interaction. However, although many of these interactions would not be exhibited in the cell, it is likely that the target proteins identified in the screens represent real interactions between target and bait proteins in the B2H system since only one protein was identified in all screens including the control (Table 7). If a target protein was selected independently of the bait protein, this protein would be expected to be detected in several screens with several different bait proteins. Interestingly, only one protein was identified in the control screen and in every screen regardless of the bait protein expressed. This protein was an prophage protein of unknown function. Therefore it is likely that this putative prophage protein could interact with the cI phage repressor protein such that regardless of the bait protein fused to the cI repressor, the prophage protein of unknown function was capable of interacting with cI and therefore generate expression of the reporter genes.

The functional representation of the ETEC genome in the library, defined as the number of base-pairs in the genome required to encode an in-frame, translatable protein fused to the cI repressor of the bait vector was calculated at 1 clone per 49.8 bp of the ETEC genome (section 7.2.2). This is likely an effective representation of the genome in the pTRG library since this would translate into approximately 30 translatable fragments for a protein such as GspA that would be available for screening with the bait protein of interest.

To summarize, use of the B2H library of H10407 genomic fragments in the Bacteriomatch II bacterial II hybrid technique was successful in identifying interactions that involved members of the T2SS. As originally hypothesized, screening of the B2H library against a bait protein of interest readily resulted in the selection of overlapping fragments of target protein sequences that enabled refinement of the binding site involved in the interaction to be achieved.

8. CONCLUSION

When this study was initiated, information concerning the structure and function of the T2SS was limited to that of several species including; *Aeromonas hydrophila*, *Vibrio cholerae*, *Klebsiella oxytoca* and *Erwinia carotovora*. I had soon noticed that although *Escherichia coli* is a major human pathogen and is the most intensely studied bacterial species historically, a surprising lack of knowledge concerning the T2SS in *E. coli* species was evident in the literature. Also, the information that was present in the literature concerning the T2SS in *E. coli* was confusing since two independent studies identified separate T2SSs responsible for secretion of the major virulence factor LT by two different mechanisms. As a result, I initiated a study to determine the relative involvement of both T2SSs in the secretion of LT in both soluble and vesicle-associated means. This study identified that it is likely that the $T2SS_{\beta}$ is solely required for secretion of LT in both soluble and vesicle-mediated forms and that the $T2SS_{\alpha}$ is not required. However, a significant amount of investigation remains to be conducted to prove that a functional $T2SS_{\alpha}$ does not secrete LT in either soluble or vesicle-associated forms since although I was able to confirm that the $GspD_{\alpha}$ secretin is assembled in $T2SS_{\alpha}$-inducible strains, the function of the system must be verified. These data, in combination with *in vivo* studies using $T2SS_{\alpha}$ and/or $T2SS_{\beta}$ positive strains in a mouse model of infection would be successful in determining the requirement for $T2SS_{\alpha}$ and/or $T2SS_{\beta}$ in secretion of LT and in determining the relative involvement of soluble and vesicle-associated LT in ETEC pathogenesis.

The remaining topics of my work dealt with the function of accessory proteins in the assembly of the secretin in the T2SSs of ETEC and *Vibrio cholerae*. In ETEC, the putative hypothetical lipoprotein YghG was characterized to function as the pilotin of the $GspD_{\beta}$ secretin of the $T2SS_{\beta}$. This investigation revealed aspects of secretin assemly that are contradictory to the current dogma of secretin assembly research that suggest that the pilotin performs a role in protecting the secretin monomer from degradation and in the absence of the pilotin the secretin still forms but is mislocalized to the inner membrane. In most studies whereupon the current dogma of secretin assembly is based, T2SS proteins from species that are not *E. coli* are expressed *in trans* at levels that are not physiologically relevant in *E. coli* strains. However in this study, I made a concerted effort to conduct experiments in biologically relevant manner by using a wild-type *E. coli* strain, by constructing in-frame deletions of genes

in the chromosome and by using the endogenous levels of $T2SS_{\beta}$ and LT to investigate the system in *E. coli*. As a result, the results that I have shown in this study reveal that the secretin does not form in the inner membrane in the absence of its pilotin and that the pilotin does not perform a role in protecting the secretin but instead is solely responsible for localization of the secretin monomer to the outer membrane. This data has recently been accepted for publication in the journal Infection and Immunity (Strozen *et. al.*, 2012).

In *Vibrio cholerae* the GspAB complex was shown to only be partially required for T2SS secretin assembly unlike that in Aeromonads whereby GspAB are absolutely required for secretin assembly. This data was published in the May 2011 issue of the Journal of Bacteriology. Further investigation using a transoposon mutagenesis screen identified the possibility that both GspAB and a previously unknown pilotin could be involved in secretin assembly. This result would be the first in T2SS research to show the requirement for both the peptidoglycan-remodeling function of GspAB and the localization function of a pilotin in assembly of a T2SS secretin.

9. FUTURE STUDIES

9.1 Requirement for $T2SS_{\alpha}$ and $T2SS_{\beta}$ in secretion of LT *in vitro*

Construction of H10407 strains that contained inducible $gspA_{\alpha}$ and $gspC_{\alpha}$ promoters enabled the $GspD_{\alpha}$ to be detected in secretin form, thereby suggesting that the $T2SS_{\alpha}$ could be assembled in these strains. A critical control experiment however is to confirm that the $T2SS_{\alpha}$ system also functions normally in the cell. To accomplish this the $T2SS_{\alpha}$ substrate ChiA should be expressed *in trans* and the amount of ChiA secreted should be assayed by immunoblot. Once an assembled and functional $T2SS_{\alpha}$ is confirmed in H10407, the amount of LT secreted in soluble and vesicle-associated forms in $T2SS_{\alpha}$ and $T2SS_{\beta}$ strains can be compared.

9.2 Requirement for $T2SS_{\alpha}$ and $T2SS_{\beta}$ in ETEC pathogenesis

In vitro studies of the requirement for $T2SS_{\alpha}$ and $T2SS_{\beta}$ in secretion of LT revealed that the $T2SS_{\beta}$ is solely responsible for LT secretion and that LT is predominantly in soluble form. Although these results coupled with the finding that *gspM* mutants of H10407 are incapable of colonizing the mouse intestine (Dorsey *et. al.* 2006) suggest that the $T2SS_{\beta}$ is the major T2SS virulence factor in ETEC, the involvement of the $T2SS_{\alpha}$ in ETEC pathogenesis should also be

determined *in vivo*. It is likely that the $T2SS_{\alpha}$ is not required for secretion of LT in vesicle-associated form *in vivo* but it may be involved in colonization by secreting the known $T2SS_{\alpha}$ substrate, endochitinase ChiA. Several studies have suggested that chitinases are not only involved in the catabolism of chitin but may also be involved in colonization (reviewed by Tran *et. al.*, 2011). In *Pseudomonas aeruginosa*, the chitinase CbpD was shown to be important in the initial stages of attachment to lung epithelial cells (Folders *et. al.*, 2000; Folders *et. al.*, 2001) possibly in response to the presence of mucin in the lung (Fung *et. al.*, 2010). Expression of the *Serratia marcescens* chitinase CBP21 in *E. coli* conferred an increased rate of adhesion to colonic epithelial cells (Kawada *et. al.*, 2008). In fact, preliminary data by Tran *et. al*., 2011 suggested that in AIEC *E. coli* strain LF82, the same ChiA protein that is a substrate of the $T2SS_{\alpha}$ is important in adhesion to intestinal epithelial cells. As a result, a mouse intestinal colonization study using wild-type H10407, $gspD_{\alpha}$, $gspD_{\beta}$, *eltAB* and *chiA* strains would provide valuable information regarding the requirement for both systems in ETEC pathogenesis.

9.3 Requirement for soluble and vesicle-associated LT in ETEC pathogenesis

To identify which mode of LT transmission (soluble or vesicle-associated) is required for ETEC pathogenicity, further experiments must be conducted *in vivo* to compare the ability of wild-type strains with strains incapable of producing outer membrane vesicles loaded with LT to colonize the intestine of the animal model. Since the binding sites on LPS for LT exist in the core and KDO sugars and these sugars are essential for cell viability, creation of a LT-vesicle deficient strain by deletion of the LT-binding sites in LPS is not an option. However in a study by Mudrak *et. al.*, 2009, a T47A substitution mutation in LT-B rendered the LT holotoxin unable to bind to the bacterial surface yet did not affect expression, assembly, secretion and activity of LT. Therefore, an *in vivo* study of ETEC H10407 expressing wild-type LT in comparison to that expressing the T47A LT holotoxin that cannot bind to the surface of the ETEC cell will provide valuable information regarding the requirement for soluble and vesicle-associated LT in ETEC pathogenesis.

9.4 Characterization of the function of $YghG^{Vc}$ in assembly of the $GspD^{Vc}$ secretin in *V. cholerae*

Assembly of the GspD^{Vc} secretin in *Vibrio cholerae* may involve the function of both a GspAB complex and the lipoprotein YghG. The finding that *V. cholerae* gene *Vc1703* encodes a putative YghG homologue that may have been insertionally inactivated in a transposon mutagenesis screen to identify a protein that serves in a GspAB redundant function, coupled with the finding that expression of YghG^{Vc} can complement the GspD$_{\beta}$ secretin-deficient phenotype of an H10407 *yghG* strain suggests that YghG^{Vc} likely performs a pilotin function for localization of GspD^{Vc} to the outer membrane. To determine if YghG^{Vc} performs a GspAB redundant or downstream function will require a comparison of the GspD^{Vc} secretin assembly phenotype in *yghG*^{Vc} and *yghG*^{Vc} *gspAB*^{Vc} strains. If YghG and GspAB perform redundant functions, then deletion of YghG in a strain that is GspAB positive will allow the GspD secretin to remain assembled. However if YghG^{Vc} performs a function downstream of GspAB, deletion of YghG^{Vc} would abrogate assembly of the GspD^{Vc} secretin regardless of the presence or absence of GspAB. It is entirely plausible that this analysis will be the first to identify and characterize a T2SS that requires both GspAB and GspS for assembly and function.

9.5 Validation and use of the Bacterial II hybrid target library for assessing protein interactions *in vivo*

The bacterial II hybrid target library has the potential to be useful in identifying interactions between proteins however several important controls need to be conducted to validate the library. Due to time constraints the genomic representation of the library was not validated fully and should be investigated further by indentifying the representation of several more genes or genomic sequences in the library. Also, it would be interesting to confirm the effectiveness of the library in identifying protein interactions by using cytoplasmic proteins known to be involved in well characterized interactions in several screens to determine if their protein partner could be identified in the screen.

10. REFERENCES

Abendroth, J., Kreger, A.C., Hol, W.G.J. 2009 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., Murphy, P., Sandkvist, M., Bagdasarian, M., Hol, W.G.J. 2005 The X-ray structure of the type II secretion system formed by the N-terminal domain of EpsS and the cytoplasmic domain of EpsL of *Vibrio cholerae*. J. Mol. Biol. 348:845-855

Alami., M. 2003 Differential interactions between a twin-arginine signal peptide and its translocase in *Escherichia coli* Mol. Cell. 12:937-946

Ali, A. *et. al.* 2000 Mutations in the extracellular protein secretion pathway genes (eps) interfere with rugose polysaccharide production in and motility of *Vibrio cholerae*. Infect. Immun. 68:1967-1974

Allen, K.P., Randolph, M.M., Fleckenstein, J.M. 2006 Importance of heat-labile enterotoxin in colonization of the adult mouse small intestine by human enterotoxigenic *Escherichia coli* strains. Infect. Immun. **74**:869-875

Anderson, K. L., J. Billington, D. Pettigrew, E. Cota, P. Simpson, P. Roversi, H. A. Chen, P. Urvil, L. Du Merle, P. N. Barlow, M. E. Medof, R. A. G. Smith, B. Nowicki, C. Le Bouguenec, S. M. Lea, and S. Matthews. 2004. An atomic resolution model for assembly, architecture, and function of the Dr adhesins. Mol. Cell 15:647–657

Aragon, V., Kurtz, S., Fleiger, A., Neumeister, B., Cianciotto, N.P. 2000 Secreted enzymatic activities of wild-type and pilD-deficient *Legionella pneumophila*. Infect. Immun. **68**:1855-1863

Arts, J., van Boxtel, R., Filloux, A., Tommassen, J., Koster, M. 2007 Export of the pseudopilin XcpT of the *Pseudomonas aeruginosa* type II secretion system via the signal recognition particle-Sec pathway. *J. Bacteriol.* 189:2069-2076

Aschtgen, M.S., Bernard, C.S., Bentzmann, S.D., Lloubes, R., Cascales, E. 2008 SciN is an outer membrane lipoprotein required for type VI secretion in enteroaggregative *Escherichia coli*. J. Bacteriol. 190:7523-7531

Ast, V.M., Schoenhofen, I.C., Langen, G.R., Stratilo, C.W., Chamberlain, M.D., Howard, S.P. 2002 Expression of the ExeAB complex of *Aeromonas hydrophila* is required for the localization and assembly of the ExeD secretion port multimer. Mol. Microbiol. **44**(1):217-231

Bernadac, A., Gavioli, M., Lazzaroni, J.C., Raina, S., Lloubes, R. 1998 *Escherichia coli* tol-pal mutants form outer membrane vesicles. J. Bacteriol. 180:4872-4878

Bitter, W., Koster, M., Latijnhouwers, M., de Cock, H., Tommassen, J. 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

Black, R. 1993 Epidemiology of diarrhoeal disease; implications for control by vaccines. Vaccine 11(2):100-106

Blattner *et. al.*, 1997 The complete genome sequence of *Escherichia coli* K-12. Science **277**:1453-1474

Bolhuis, A., Mathers, J.E., Thomas, J.D., Barrett, C., Robinson, C. 2001 TatB and TatC form a functional and structural unit of the twin-arginine translocase from *Escherichia coli*. J. Biol. Chem. 276:20213-20219

Bos, M.P., Tefsen, B., Guertsen, J., Tommassen, J. 2004 Identification of an outer membrane protein required for the transport of lipopolysaccharide to the bacterial cell surface. PNAS USA 101:9417-9422

Bouley, J., Condemine, G., Shevchik, V.E. 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

Burghout, P., Beckers, F., de Wit, E., van Boxtel, R., Cornelis, G.R., Tommassen, J., Koster., M. 2004 Role of the pilot protein YscW in the biogenesis of the YscC secretin in *Yersinia enterocolitica*. J. Bacteriol. **186**:5366-5375

Brown, E.A. and Hardwidge, P.R. 2007 Biochemical characterization of the enterotoxigenic *Escherichia coli* LeoA protein. Microbiology 153(11):3776-3784

Brumlik, M.J., van der Goot, F.G., Wong, K.R., Buckley, J.T. 1997 The disulfide bond in the *Aeromonas hydrophila* lipase/acyltransferase stabilizes the structure but is not required for secretion or activity. J. Bacteriol. **179**:3116-3121

Budd, A., Blandin, S, Levashina, E.A., Gibson, T.J. 2004 Bacterial α2-macroglobulins: colonization factors acquired by horizontal gene transfer from the metazoan genome? Genome Biology 5:38

Bullock, W. O., Fernandez, J.M., Short, J.M. 1987 XL1-Blue: A high efficiency plasmid transforming *recA Escherichia coli* strain with β-galactosidase selection. BioTechniques 5:376-379

Burghout, P., Beckers, F., de Wit, E., van Boxtel, R., Cornelis, G.R., Tommassen, J., Koster., M. 2004 Role of the pilot protein YscW in the biogenesis of the YscC secretin in *Yersinia enterocolitica*. J. Bacteriol. **186**:5366-5375

Burkhardt, J., Vonck, J., Averhoff, B. 2011 Structure and function of PilQ, a secretin of the DNA transporter from the thermophilic bacterium *Thermus thermophilus* HB27. J. Biol. Chem. 286:9977-9984

Burland, V., Shao, Y., Perna, N.T., Plunkett, G., Sofia, H.J., Blattner, F.R. 1998 The complete DNA sequence and analysis of the large virulence plasmid of *Escherichia coli* O157:H7. Nucleic acid res. **26**:4196-4204

Camberg, J.L. *et. al.*, 2007 Synergistic stimulation of EpsE ATP hydrolysis by EpsL and acidic phospholipids. EMBO J. 26:19-27

Campos, M., Nilges, M., Cisneros, D.A., Francetic, O. 2010 Detailed structural and assembly model of the type II secretion pilus from sparse data. PNAS USA 107:13081-13086

Chami, M. *et. al.*, 2005 Structural insights into the secretin PulD and its trypsin-resistant core. J. Biol. Chem. 280:37732-37741

Chatterjee, S.N. and Chaudhuri, K. 2003 Lipopolysaccharides of *Vibrio cholerae*, I. Physical and chemical characterization. Biochem Biophys. Acta. 1639:65-79

Chen *et. al.*, 2006 Identification of genes subject to positive selection in uropathogenic strains of *Escherichia coli*: a comparative genomics approach. PNAS USA **103**:5977-5982

Chen, L.Y., Chen, D.Y., Miaw, J., Hu, N.T. 1996 XpsD, an outer membrane protein required for protein secretion by *Xanthomonas campestris pv. campestris*, forms a multimer. J. Biol. Chem. 271:2703-2708

Cherepanov, P. P., and W. Wackernagel. 1995. Gene disruption in *Escherichia coli*: Tc^R and Km^R cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. Gene **158**:9-14

Clavel, T., Germon, P., Vianney, A., Portalier, R., Lazzaroni, J.C. 1998 TolB protein of *Escherichia coli* K-12 interacts with the outer membrane peptidoglycan-associated proteins Pal, Lpp and OmpA. Mol. Microbiol. 29:359-367

Cline, K., and Theg, S.M. 2007 The Sec and Tat protein translocation pathways in chloroplasts. The enzymes, molecular machines involved in protein transport across cellular membranes eds Dalbey R.E., Koehler, C., Tamanoi, F. (Elsevier, San Diego), Vol XXV, pp455-485

Cline, K. and Mori, H. 2001 Thylakoid ΔpH-dependent precursor proteins bind to a cpTatC-Hcf106 complex before Tha4-dependent transport. J. Cell Biol. 154:719-729

Collin, S., Guilvout, I., Chami, M., Pugsley, A.P. 2007 YaeT-independent multimerization and outer membrane association of secretin PulD. Mol. Microbiol. 64:1350-1357

Collin, S., Guilvout, I., Nickerson, N.N. Pugsley, A.P. 2011 Sorting of an integral outer membrane protein via the lipoprotein-specific Lol pathway and a dedicated lipoprotein pilotin. Mol. Microbiol. **80**:655-665

Collins, R.F. *et. al.*, 2003 Three-dimensional structure of the *Neisseria meningitidis* secretin PilQ determined from negative-stain transmission electron microscopy. J. Bacteriol. 185:2611-2617

Condemine, G., Dorel, C., Hugovieux-Cotte-Pattat, N., Robert-Baudouy, J. 1992 Some of the *out* genes involved in the secretion of pectate lyases in *Erwinia chrysanthemi* are regulated by KdgR. **Mol. Microbiol. 6**(21):3199-3211

Coster, T.S. *et. al.*, 2007 Immune response, ciprofloxacin activity, and gender differences after human experimental challenge by two strains of enterotoxigenic *Escherichia coli*. Infect. Immun. **75**:252-259

Craig, L. *et. al.*, 2006 Type IV pilus structure by cryo-electron microscopy and crystallography: implications for pilus assembly and functions. Mol. Cell. 23:651-662

Cronan, J. E. 2006. A family of arabinose-inducible *Escherichia coli* expression vectors having pBR322 copy control. Plasmid **55**:152-157

Crossman *et. al.*, 2010 A commensal gone bad: complete genome sequence of the prototypical enterotoxigenic *Escherichia coli* strain H10407. J. Bacteriol. **192**:5822-5831

Dabney-Smith, C., Mori, H., Cline, K. 2006 Oligomers of Tha4 organize at the thylakoid Tat translocase during protein transport. J. Biol. Chem . 281:5476-5483

Daefler, S., Guilvout, I., Hardie, K.R., Pugsley, A.P., Russel, M. 1997 The C-terminal domain of the secretin PulD contains the binding site for its cognate chaperone, PulS, and confers PulS dependence on pIV^{f1} function. Mol. Microbiol. **24**:465-475

Davis, B.M., Lawson, E.H., Sandkvist, M., Ali, A., Sozhamannan, S., Waldor, M.K. 2000 Convergence of the secretory pathways for cholera toxin and the filamentous phage, CTXphi. Science 288:333-335

Deen, J.L. *et. al.*, 2008 The high burden of cholera in children: comparison of incidence from endemic areas in Asia and Africa. PLoS Negl. Trop. Dis. 2:e173

De Haan, L., and Hirst, T.R. 2004 Cholera toxin: a paradigm for multi-functional engagement of cellular mechanisms. Mol. Membr. Biol. 21:77-92

DeLorenzo, V. and Timmis, K.N. 1994 Analysis and construction of stable phenotypes in gram-negative bacteria with Tn5- and Tn10-derived minitransposons. Method Enzymol. 235:386-405

Demchick, P. and Kock, A.L. 1996 The permeability of the wall fabric of *Escherichia coli* and *Bacillus subtilis*. J. Bacteriol. **178**:768-773

Dorsey, F.C., Fischer, J.F., Fleckenstein, J.M. 2006 Directed delivery of heat-labile enterotoxin by enterotoxigenic *Escherichia coli*. Cell. Microbiol. 8:1516-1527

Douzi, B. *et. al.* 2009 The XcpV/GspI pseudopilin has a central role in the assembly of a quaternary complex within the T2SS pseudopilus. J. Biol. Chem. 284:34580-34589

Douzi, B., Ball, G., Cambillau, C., Tegoni, M., Voulhoux, R. 2011 Deciphering the Xcp Pseudomonas aeruginosa type II secretion machinery through multiple interactions with substrates. J. Biol. Chem.

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

Durand, E. *et. al.* 2005 XcpX controls biogenesis of the Pseudomonas aeruginosa XcpT-containing pseudopilus. J. Biol. Chem. 280:31378-31389

Dyer, M.D. *et al.*, 2010 The human-bacterial pathogen protein interaction networks of *Bacillus anthracis, Francisella tularensis* and *Yersinia pestis*. PLoS One 5:e12089

Ehrmann, M., Ehrle, R., Hofmann, E., Boos, W., Schlosser, A. 1998 The ABC maltose transporter. Mol. Microbiol. 29:685-694

Evans, D.G., Silver, R.P., Evans, D.J. Jr., Chase, D.G., Gorbach, S.L. 1975 Plasmid-controlled colonization factor associated with virulence in *Escherichia coli* enterotoxigenic for humans. Infect. Immun. **12**:656-667

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

Fleckenstein, J.M., Lindler, L.E., Elsinghorst, E.A., Dale, J.B. 2000 Identification of a gene within a pathogenicity island of enterotoxigenic *Escherichia coli* H10407 required for maximal secretion of the heat-labile enterotoxin. Infect. Immun. 68(5):2766-2774

Filloux, A. 2004 The underlying mechanisms of type II protein secretion. Biochem. et Biophys. Acta **1694**: 163-179

Folders, J., Tommassen, J., van Loon, L.C., Bitter, W. 2000 Identification of a chitin-binding protein secreted by *Pseudomonas aeruginosa*. J. Bacteriol. 182:1257-1263

Folders, J., Algra, J., Roelofs, M.S., van Loon, L.C., Tommassen, J., Bitter, W. 2001 Characterization of *Pseudomonas aeruginosa* chitinase, a gradually secreted protein. J. Bacteriol. 183:7044-7052

Forest, K.T. 2008 The type II secretion arrowhead: the structure of GspI-GspJ-GspK. Nat. Struct. Mol. Biol. 15:428-430

Forest, K.T., Satyshur, K.A., Worzalla, G.A., Hansen, J.K., Herdendorf, T.J. 2004 The pilus-retraction protein PilT: ultrastructure of the biological assembly. Acta Crystallogr D Biol Crystallogr. 60:978-982

Forte, L.R., Thorne, P.K., Eber, S.L., Krause, W.J., Freeman, R.H., Francis, S.H., Corbin, J.D. 1992 Stimulation of intestinal Cl- transport by heat-stable enterotoxin: activation of cAMP-dependent protein kinase by cGMP. Am. J. Physiol. 263:607-615

Francetic, O., Buddelmeijer, N., Lewenza, S., Kumamoto, C.A., Pugsley, A.P. 2007 Signal recognition particle-dependent inner membrane targeting of the PulG pseudopilin component of a type II secretion system. J. Bacteriol. 189:1783-1793

Francetic, O. and Pugsley, A.P. 1996 The cryptic general secretory pathway (gsp) operon of *Escherichia coli* K-12 encodes functional proteins. J. Bacteriol. **178**:3544-3549

Francetic, O., Lory, S., Pugsley, A.P. 1998 A second prepilin peptidase gene in *Escherichia coli* K-12. Mol. Microbiol. **27**:763-775

Francetic, O., Belin, D., Badaut, C., Pugsley, A.P. 2000 Expression of the endogenous type II secretion pathway in *Escherichia coli* leads to chitinase secretion. *EMBO J.* 19:6697-6703

Frank, D.N., Amand, A.L.S., Feldman, R.A., Boedeker, E.C., Harpaz, N., Pace, NR. 2007 Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. PNAS USA 104:13780-13785

Fung, C. *et. al.*, 2010 Gene expression of *Pseudomonas aeruginosa* in a mucin-containing synthetic growth medium mimicking cystic fibrosis lung sputum. J. Med. Microbiol. 59:1089-1100

Gaastra, W. and Svennerholm, A.M. 1996 Colonization factors of human enterotoxigenic *Escherichia coli* (ETEC). Trends Microbiol. 4:444-452

Genin, S. and Boucher, C.A. 1994 A superfamily of proteins involved in different secretion pathways in Gram-negative bacteria: modular structure and specificity of the N-terminal domain. Mol. Gen. Genet. **243**:112-118

Gerard-Vincent, M., Robert, V., Ball, G., Bleves, S., Michel, G.P.F., Lazdunski, A., Filloux, A. 2002 Identification of XcpP domains that confer functionality and specificity to the *Pseudomonas aeruginosa* type II secretion apparatus. Mol. Microbiol. **44**(6):1651-1665

Gibert, I., Barbe, J. 1990 Cyclic AMP stimulates transcription of the structural gene of the outer membrane protein OmpA of *Escherichia coli*. FEMS Microbiol. Lett. 56:307-311

Giannella, R. A., and E. A. Mann. 2003. *E. coli* heat-stable enterotoxin and guanylyl cyclase C: new functions and unsuspected actions. *Trans. Am. Clin. Climatol. Assoc.* **114:**67–85

Gill, D. M. 1976. The arrangement of subunits in cholera toxin. *Biochemistry* **15:**1242–1248

Gill, D.M., Clements, J.D., Robertson, D.C., Finkelstein, R.A. 1981 Subunit number and arrangement in *Escherichia coli* heat-labile enterotoxin. Infect. Immun. 33:677-682

Glauner, B. 1988 Separation and quantification of muropeptides with high-performance liquid chromatography. Anal. Biochem. **172**:451-464

Goins, B., and E. Freire. 1988. Thermal stability and intersubunit interaction of cholera toxin in solution and in association with its cell surface receptor ganglioside GM1. *Biochemistry* **27:**2046–2052.

Gonnet, P. Rudd, K.E., Lisacek, F. 2004 Fine-tuning the prediction of sequences cleaved by signal peptidase II: A curated set of proven and predicted lipoproteins of *Escherichia coli* K-12. Proteomics **4**:1597-1613

Gray, M.D., Bagdasarian, M., Hol, W.G., Sandkvist, M. 2011 *In vivo* cross-linking of EpsG to EpsL suggests a role for EpsL as an ATPase-pseudopilin coupling protein in the type II secretion system of *Vibrio cholerae*. Mol. Microbiol. 79:786-798

Grys, T.E., Siegel, M.B., Lathem, W.W., Welch, R.A. 2005 The StcE protease contributes to intimate adherence of enterohemorrhagic *Escherichia coli* O157:H7 to host cells. Infect. Immun. **73**:1295-1303

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

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

Hardie, K.R., Seydel, A., Guilvout, I., Pugsley, A.P. 1996B The secretin-specific, chaperone-like protein of the general secretory pathway: separation of proteolytic protection and piloting functions. Mol. Microbiol. **22**:967-976

Hlady, W.G. and Klontz, K.C. 1996 The epidemiology of *Vibrio* infections in Florida, 1981-1993. J. Infect. Dis. 173:1176-1183

Ho, T.D., Davis, B.M., Ritchie, J.M., Waldor, M.K. 2008 Type 2 secretion promotes enterohemorrhagic *Escherichia coli* adherence and intestinal colonization. Infect. Immun. **76**:1858-1865

Hobbs, M. and Mattick, J.S. 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

Honda, T. and Iida, T. 1993 The pathogenicity of *Vibrio parahaemolyticus* and the role of the thermostable direct haemolysin and related haemolysins. Rev. Med. Microbiol. 4:106-113

Horstman, A.L. and Kuehn, M. 2002 Bacterial surface association of heat-labile enterotoxin through lipopolysaccharide after secretion via the general secretory pathway. J. Biol. Chem. **277**:32538-32545

Horstman, A.L., Bauman, S.J., Kuehn, M.J. 2004 Lipopolysaccharide 3-deoxy-D-manno-octulosonic acid (Kdo) core determines bacterial association of secreted toxins. J. Biol. Chem. 279:8070-8075

Howard, S.P. and Buckley, J.T. 1985 Protein export by a gram-negative bacterium: production of aerolysin by *Aeromonas hydrophila.* J. Bacteriol. **161**:1118-1124

Howard, S.P., Gebhart, C., Langen, G.R., Li, G., Strozen, T.G. 2006 Interactions between peptidoglycan and the ExeAB complex during assembly of the type II secretin of *Aeromonas hydrophila*. Mol. Micro. **59**(3):1062-1072

Howard, S.P., Meiklejohn, H.G., Shivak, D., Jahagirdar, R. 1996 A TonB-like protein and a novel membrane protein containing an ATP-binding cassette function together in exotoxin secretion. Mol. Microbiol. **22**:595-604

Hu, N.T., Leu, W.M., Lee, M.S., Chen, A., Chen, S.C., Song, Y.L., Chen, Y.L. 2002 XpsG, the major pseudopilin in *Xanthomonas campestris pv. campestris*, forms a pilus-like structure between cytoplasmic and outer membranes. Biochem. J. **365**:205-211

Hudault, S., Guignot, J., Servin, A.L. 2001 *Escherichia coli* strains colonizing the gastrointestinal tract protect germfree mice against *Salmonella typhimurium* infection. Gut 49:47-55

Huq, Q. et. al. 1986 Colonization of the gut of the blue crab by *Vibrio cholerae*. Applied and Environ. Microbiol. **52**:586-588

Ishidate, K., Creeger, E.S., Zrike, J., Deb, S., Glauner, B., McAlister, T.J., Rothfield, L. 1986 Isolation of differentiated membrane domains from *Escherichia coli* and *Salmonella typhimurium*, including a fraction containing attachment sites between the inner and outer membranes and the murein skeleton of the cell envelope. J. Biol. Chem. **261**:428-443

Islam, M.S., Drasar, B.S., Bradley, D.J. 1989 Attachment of toxigenic Vibrio cholerae O1 to various freshwater plants and survival with a filamentous green algae, Rhizoclonium fontanum. J. Trop Med Hyg. **92**:396-401

Izore, T. *et. al.*, 2011 Structural characterization and membrane localization of ExsB from the type III secretion system (T3SS) of Pseudomonas aeruginosa. J. Mol. Biol.

Jahagirdar, R. and Howard, S.P. 1994 Isolation and characterization of a second *exe* operon required for extracellular protein secretion in *Aeromonas hydrophila.* J. Bacteriol. **176**:6819-6826

Jain, S. *et. al.*, 2011 Structural characterization of outer membrane components of the type IV pili system in pathogenic *Neisseria*. PLoS One 6:e16624

Jiang, B. and Howard, S.P. 1991 Mutagenesis and isolation of *Aeromonas hydrophila* genes which are required for extracellular secretion. J. Bacteriol. **173**(3):1241-1249

Johnson *et. al.*, 2007 The genome sequence of avian pathogenic *Escherichia coli* strain O1:K7:H7 shares strong similarities with human extraintestinal pathogenic *E. coli* genomes. J. Bacteriol. **189**:3228-3236

Kay, W.W., Buckley, J.T., Ishiguro, E.E., Phipps, B.M., Monette, J.P.L., Trust, T.J. 1981 Purification and disposition of a surface protein associated with virulence of *Aeromonas salmonicida.* J. Bacteriol. **147**:1077-1084

Kawada, M., *et. al.*, 2008 Chitinase 3-like-1 enhances bacterial adhesion to colonic epithelial cells through the interaction with bacterial chitin-binding protein. Lab invest. 88:883-895

Kennedy, D.J., Greenberg, R.N., Dunn, J.A., Abernathy, R., Ryerse, J.S., Guerrant, R.L. 1984 Effects of *Escherichia coli* heat-stable enterotoxin STb on intestines of mice, rats, rabbits and piglets. Infect. Immun. 46:639-643

Kesty, N.C., Mason, K.M., Reedy, M., Miller, S.E., Kuehn, M.J. 2004 Enterotoxigenic *Escherichia coli* vesicles target toxin delivery into mammalian cells. EMBO J. 23:4538-4549

Kim, K. *et. al.*, 2006 Crystal structure of PilF: functional implication in th type 4 pilus biogenesis in *Pseudomonas aeruginosa*. Biochem Biophys. Res. Commun. 340:1028-1038

Kirn, T.J., Lafferty, M.J., Sandoe, C.M.P., Taylor, R.K. Deliniation of pilin domains required for bacterial association into microcolonies and intestinal colonization by *Vibrio cholerae*. Mol. Microbiol. **35**:896-910

Knowles, T.J., Scott-Tucker, A., Overduin, M., Henderson, I.R. 2009 Membrane protein architects: the role of the BAM complex in outer membrane protein assembly. Nat. Rev. Microbiol. 7:206-214

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

Kolter, R., Inuzuke, M., Heliniki, D. 1978 Trans-complementation-dependent replication of a low molecular weight origin fragment from plasmid R6K. Cell 15(4):1199-1208

Koo, J., Tammam, S., Ku, S.Y., Sampaleanu, L.M., Burrows, L.L., Howell, P.L. 2008 PilF is an outer membrane lipoprotein required for multimerization and localization of the *Pseudomonas aeruginosa* Type IV pilus secretin. J. Bacteriol. **190**:6961-6969

Korotkov, K.V., Gonen, T., Hol, W.G.J. 2011A Secretins: dynamic channels for protein transport across membranes. Trend Biochem. Sci. 36:433-443

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

Korotkov, K.V. *et. al.* 2009 Calcium is essential for the major pseudopilin in the type II secretion system. J. Biol. Chem. 284:25466-25470

Korotkov, K.V., Pardon, E., Steyaert, J., Hol, W.G.J. 2009 Crystal structure of the N-terminal domain of the secretin GspD from ETEC determined with the assistance of a nanobody. Structure 17(2):255-265

Korotkov, K.V. *et. al.*, 2011B Structural and functional studies on the interaction of GspC and GspD in the type II secretion system. PLoS Pathog. 7:e1002228

Kreger, A. and Lockwood, D. 1981 Detection of extracellular toxins produced by V*ibrio vulnificus*. Infect. Immun. 33:583-590

Kunkel, S.L. and Robertson, D.C. 1979 Factors affecting release of heat-labile enterotoxin by enterotoxigenic *Escherichia coli*. Infect. Immun. 23:652-659

Kulkarni, R., Dhakal, B.K., Slechta, E.S., Kurtz, Z., Mulvey, M.A., Thanassi, D.G. 2009 Roles of putative type II secretion and type IV pilus systems in the virulence of uropathogenic *Escherichia coli*. PLOS ONE **4**(3):e4752

Lander, E.S. and Waterman, M.S. 1988 Genomic mapping by fingerprinting random clones: a mathematical analysis. Genomics 2:231-239

Lapointe, C.F. and Taylor, R.K. 2000 The type IV prepilin peptidases comprise a novel family of aspartic acid proteases. J. Biol. Chem. 275:1502-1510

Lario, P.I. *et. al.* 2005 Structure and biochemical analysis of a secretin pilot protein. EMBO J 24:1111-1121

Le Bouguenec, C., L. Lalioui, L. Du Merle, M. Jouve, P. Courcoux, S. Bouzari, R. Selvarangan, B. J. Nowicki, Y. Germani, A. Andremont, P. Gounon, and M. I. Garcia. 2001. Characterization of AfaE adhesins produced by extraintestinal and intestinal human *Escherichia coli* isolates: PCR assays for detection of afa adhesins that do or do not recognize Dr blood group antigens. J. Clin. Microbiol. 39:1738–1745.

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

Letellier, L., Howard, S.P., Buckley, J.T. 1997 Studies on the energetics of proaerolysin secretion across the outer membrane of *Aeromonas* species. Evidence for a requirement for both the protonmotive force and ATP. J. Biol. Chem. 11109-11113

Li, G. and Howard, S.P. 2010 ExeA binds to peptidoglycan and forms a multimer for assembly of the type II secretion apparatus in *Aeromonas hydrophila*. Mol. Microbiol. **76**:772-781.

Li, G., Miller, A., Bull, H., Howard, S.P. 2011 Assembly of the type II secretion system: identification of ExeA residues critical for peptidoglycan binding and secretin multimerization. J. Bacteriol. 193:194-204

Lybarger, S.R., Johnson, T.L., Gray, M.D., Sikora, A.E., Sandkvist, M. 2009 Docking and assembly of the type II secretion complex of *Vibrio cholerae*. J. Bacteriol. **191**(9):3149-3161

Marcotte, E. M., Pellegrini, M., Ng, H.L., Rice, D.W., Yeates, T.O., Eisenberg, D. 1999 Detecting protein function and protein-protein interactions from genome sequences. Science **285**:751-753

Martin, Y.P. and Bianchi, M.A. 1980 Structure, diversity and catabolic potentialities of aerobic heterotropic bacterial population associated with continuous cultures of natural marine phytoplankton. Microbial. Ecology **5**:265

Matsuyaman, S., Tajima, T., Tokuda, H. 1995 A novel periplasmic carrier protein involved in the sorting and transport of *Escherichia coli* lipoprotein destined for the outer membrane. EMBO J. 14:3365-3372

Matsuyama, S., Yokota, N., Tokuda, H 1997 A novel outer membrane lipoprotein, LolB (HemM), involved in the LolA (p20)-dependent localization of lipoproteins to the outer membrane of *Escherichia coli*. EMBO J. 16:6947-6955

Michel, G., Bleves, S., Ball, G., Lazdunski, A., Filloux, A. 1998 Mutual stabilization of the XcpZ and XcpY components fo the secretory apparatus in *Pseudomonas aeruginosa*. Microbiology 144:3379-3386

Miller, J.H. 1972 Experiments in molecular genetics. Cold spring harbor, NY: Cold Spring Harbor Laboratory

Miller, V.L. and Mekalanos, J.J. 1988 Use of a novel suicide vector in the construction of site directed mutations: Evidence that osmoregulation of virulence gene expression requires toxR. J. Bacteriol. 170:2575-2583

Mizoguchi, E. 2006 Chitinase 3-like-1 exacerbates intestinal inflammation by enhancing bacterial adhesion and invasion in colonic epithelial cells. Gastroenterology 130:398-411

Morales, V.M., Backman, A., Bagdasarian, M. 1991 A series of wide-host range low-copy-number vectors that allow direct screening of recombinants. Gene **7**:39-47

Moriel, D.G. *et. al.*, 2010 Identification of protective and broadly conserved vaccine antigens from the genome of extraintestinal pathogenic *Escherichia coli.* PNAS **107**:9072-9077

Mudrak, B. and Kuehn, M.J. 2010 Heat-labile enterotoxin: beyond GM1 binding. Toxins 2:1445-1470;

Mudrak, B., Rodriguez, D.L., Kuehn, M. 2009 Residues of heat-labile enterotoxin involved in bacterial cell surface binding. J. Bacteriol. 191:2917-2925

Mundell, D.H., Anselmo, C.R., Wishnow, R.M. 1976 Factors influencing heat-labile *Escherichia coli* enterotoxin activity. Infect. Immun. 14:383-388

Murphy, K. C., and K. G. Campellone. 2003. Lambda Red-mediated recombinogenic engineering of enterohemorrhagic and enteropathogenic *E. coli*. BMC Mol. Biol. **4:**11

Nambu, T., Minamino, T., Macnab, R.M., Kutdukake, K. 1999 Peptidoglycan-hydrolyzing activity of the FlgJ protein, essential for flagellar rod formation in *Salmonella typhimurium.* J. Bacteriol. **181**:1555-1561

Nataro, J.P. and Kaper, J.B. 1998 Diarrheagenic *Escherichia coli*. Clin. Microbiol. Rev. **11**:142-201

Nickerson, N.N. *et. al.*, 2011 Outer membrane targeting of secretin PulD relies on disordered domain recognition by a dedicated chaperone. J. Biol. Chem.

Nikaido, H. 1998 Multiple antibiotic resistance and efflux. Curr. Opin. Microbiol. 1:516-523

Nouwen, N., Ranson, N., Saibil, H., Wolpensinger, B., Engel, A., Ghazi, A., Pugsley, A.P. 1999 Secretin PulD: Association with pilot PulS, structure, and ion-conducting channel formation. PNAS **96**:8173-8177

Nunn, D.N., and Lory, S. 1992 Components of the protein excretion apparatus of *Pseudomonas aeruginosa* are processed by the type IV prepilin peptidase. PNAS USA 89:47-51

Nunn, D.N., and Lory, S. 1993 Cleavage, methylation, and localization of the *Pseudomonas aeruginosa* export proteins XcpT, -U, -V and –W. J. Bacteriol. 175:4375-4382

Ohno, A., Marui, A., Castro, A. 1997 Enteropathogenic bacteria in the La Paz river of Bolivia. Am. J. Trop. Med. Hyg. **57**:438-444

Okon, M., Moraes, T.F., Lario, P.I., Creagh, A.L., Haynes, C.A., Strynadka, N.C.J., McIntosh, L.P. 2008 Structural characterization of the Type III pilot-secretin complex from *Shigella flexneri*. Cell Struct. **16**:1544-1554

Okuda, S., Tokuda, H. 2011 Lipoprotein sorting in bacteria. Annu. Rev. Microbiol. 65:239-259

Okuda, S., Tokuda, H. 2009 Model of mouth-to-mouth transfer of bacterial lipoproteins through inner membrane LolC, periplasmic LolA, and outer membrane LolB. PNAS USA 106:5877-5882

Opalka, N. *et. al.*, 2003 Structure of the filamentous phage pIV multimer by cryo-electron microscopy. J. Mol. Biol. 325:461-470

Osborn. M.J., Gander, J.E., Parisi, E., Carson, J. 1972 Mechanism of assembly of the outer membrane of *Salmonella typhimurium*. Isolation and characterization of cytoplasmic and outer membrane. J. Biol. Chem. **247**:3962-3972

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

Paranjpye, R.N., Lara, J.C., Pepe, J.C., Pepe, C.M., Strom, M.S. 1998 The type IV leader peptidase/N-methyltransferase of *Vibrio vulnificus* controls factors required for adherence to HEp-2 cells and virulence in iron-overloaded mice. Infect. Immun. 66:5659-5668

Parsot, C., Taxman, E., Mekalanos, J.J. 1991 ToxR regulates the production of lipoproteins and the expression of serum resistance in *Vibrio cholerae*. Proc. Natl. Acad. Sci. **88**:1641-1645

Patrick, M., Gray, M.D., Sandkvist, M., Johnson, T.L. Type II secretion in *Escherichia coli*. Ecosal. Module 4.3.4 American Society for Microbiology

Patrick, M. Korotkov, K.V., Hol, W.G.J., Sandkvist, M. 2011 Oligomerization of EpsE coordinates residues from multiple subunits to facilitate ATPase activity. J. Biol. Chem. 286:10378-10386

Pearson, G.D., Woods, A., Chiang, S.L., Mekalanos, J.J. 1993 CTX genetic element encodes a site-specific recombination system and an intestinal colonization factor. Proc. Natl. Acad. Sci. USA **90**:3750-3754

Pepe, C.M., Eklund, M.W., Strom, M.S. 1996 Cloning of an *Aeromonas hydrophila* type IV pilus biogenesis gene cluster: complementation of pilus assembly functions and characterization of a type IV leader peptidase/N-methyltransferase required for extracellular protein secretion. Mol. Microbiol. **19**:857-869

Perna *et. al.*, 2001 Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. Nature **409**:529-33

Peterson, K.M. and Mekalanos, J.J. 1988 Characterization of the *Vibrio cholerae* ToxR regulon: identification of novel genes involved in intestinal colonization. Infect. Immun. **56**:2822-2829

Peyretaillade, E. *et. al.* 2010 Complete genome sequence of adherent invasive *Escherichia coli* strain LF82 isolated from a patient with Crohn's disease. PLoS ONE e12714

Possot, O.M., Gerard-Vincent, M., Pugsley, A.P. 1999 Membrane association and multimerization of secreton component PulC. J. Bacteriol. 181:4004-4011

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

Pugsley, A.P. 1993 Processing and methylation of PulG, a pilin-like component of the general secretory pathway of *Klebsiella oxytoca*. Mol. Microbiol. 9:295-308

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

Qadri, F., Svennerholm, A.M., Faruque, A.S., Sack, R.B.. 2005 Enterotoxigenic *Escherichia coli* in developing countries: epidemiology, microbiology, clinical features, treatment and prevention. *Clin. Microbiol. Rev*. **18**:465-483

Rasko, D.A., Rosovitz, M.J., Myers, G.S., Mongodin, E.F., Fricke, W.F., Gajer, P., Crabtree, J., Sebaihia, M., Thomson, N.R., Chaudhuri, R., Henderson, I.R., Sperandio, J., Ravel, J. 2008. The pangenome structure of *Escherichia coli*: comparative genomic analysis of *E. coli* commensal and pathogenic isolates. J. Bacteriol. **190:**6881-6893.

Rath, H.C. *et. al.*, 1996 Normal luminal bacteria, especially baceriodes species, mediate chronic colitis, gastritis, and arthritis in HLA-B27/human beta(2) microglobulin transgenic rats. J. Clin. Invest. 98:945-953

Reichow, S.L., Korotkov, K.V., Hol, W.G.J., Gonen, T. 2010 Structure of the cholera toxin secretion channel in its closed state. Nat. Struct. Mol. Biol. 17:1226-1232

Reichow *et. al.*, 2011 The binding of cholera toxin to the periplasmic vestibule of the type II secretion channel. Channels 5:215-218

Ristaino, P.A., Levine, M.M., Young, C.R. 1983 Improved GM1-enzyme-linked immunosorbent assay for detection of *Escherichia coli* heat-labile enterotoxin. J. Clin. Microbiol. **18**:808-815

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

Sack, D.A., Sack, R.B., Nair, G.B., Siddique, A.K. 2004 Cholera. Lancet 363:223-233

Sack, D.A., Huda, S., Neogi, P.K.B., Daniel, R.R., Spira, W.M. 1980 Microtiter ganglioside enzyme-linked immunosorbent assay for *Vibrio* and *Escherichia coli* heat-labile enterotoxins and antitoxin. J. Clin. Microbiol. 11:35-40

Sadlack, B., Merz, H., Schorle, H., Schimpl, A., Feller, A.C., Horak, I. 1993 Ulcerative colitis-like disease in mice with a disrupted interleukin-2 gene. Cell 75:253-261

Sambrook, J., Fritch, E.F., Maniatis, T. 1989 Molecular cloning: a Laboratory Manual, 2nd edn. Cold Spring Harbour, New York: Cold Spring Harbour Laboratory Press

Sanchez, J. and Holmgren, J. 2005 Virulence factors, pathogenesis and vaccine protection in cholera and ETEC diarrhea. Curr. Opin. Immunol. 17:388-398

Sandkvist, M., Bagdasarian, M., Howard, S.P., DiRita, V.J. 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., Hirst, T.R., Bagdasarian, M. 1987 Alterations at the carboxy terminus change assembly and secretion properties of the B subunit of *Escherichia coli* heat-labile enterotoxin. J. Bacteriol. **169**:4570-4576

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

Sandkvist, M., Keith, J.M., Bagdasarian, M., Howard, S.P. 2000 Two regions of EpsL involved in species-specific protein-protein interactions with EpsE and EpsM of the general secretion pathway in *Vibrio cholerae.* J. Bacteriol. **182**:742-748

Sandkvist, M., *et. al.*, 1997 General secretion pathway (eps) genes required for toxin secretion and outer membrane biogenesis in *Vibrio cholerae*. J. Bacteriol. 179:6994-7003

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

Sarker, M.R. and Cornelis, G.R. 1997 An improved version of suicide vector pKNG101 for gene replacement in gram-negative bacteria. Mol. Microbiol. **23**:410-411

Sauvonnet, N., Vignon, G., Pugsley, A.P., Gounon, P. Pilus formation and protein secretion by the same machinery in *Escherichia coli*. EMBO J. 19:2221-2228

Schaegger, H. and von Jagow, G. 1987 Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Anal. Biochem. **166**:368-379

Schamberger, G.P., Phillips, R.L., Jacobs, J.L., Diez-Gonzalez, F. 2004 Reduction of *Escherichia coli* O157:H7 populations in cattle by addition of colicin E7-producing *E. coli* to feed. Appl. Environ. Microbiol. 70:6053-6060

Schiffer, G. and Holtje, J.V. 1999 Cloning and characterization of PBP 1C, a third member of the multimodular class A penicillin-binding proteins of *Escherichia coli*. J. Biol. Chem. 274:32031-32039

Schleifer, K. H. and Kandler, O. 1972 Peptidoglycan types of bacterial cell walls and their taxonomic implications. Bacteriol. Rev. **36**:407-471

Schoenhofen, I.C., Li, G., Strozen, T.G., Howard, S.P. 2005 Purification and characterization of the N-terminal domain of ExeA: a novel ATPase involved in the type II secretion pathway of *Aeromonas hydrophila*. J. Bacteriol. **187**(18):6370-6378

Schraidt, O. and Marlovits, T.C. 2011 Three-dimensional model of *Salmonella*'s needle complex at subnanometer resolution. Science 331:1192-1195

Schuch, R. and Maurelli, A.T. 1999 The Mxi-Spa type III secretory pathway of *Shigella flexneri* requires an outer membrane lipoprotein, MxiM, for invasin translocation. Infect. Immun. **67**:1982-1991

Seydel., A., Gounon, P., Pugsley, A.P. 1999 Testing the "+2 rule" for lipoprotein sorting in the *Escherichia coli* cell envelope with a new genetic selection. Mol. Microbiol. **34**:810-821

Shevchik, V.E. and Condemine, G. 1998 Functional characterization of the *Erwinia chrysanthemi* OutS protein, an element of a type II secretion system. Microbiology **144**:3219-3228

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

Shirai, H. *et. al.*, 1990 Molecular epidemiologic evidence for association of thermostable direct hemolysin (TDH) and TDH-related hemolysin of *Vibrio parahaemolyticus* with gastroenteritis. Infect. Immun. 58:3568-3573

Simon, R., Prifer, U., Puhler, A. 1983 A broad range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in Gram-negative bacteria. Biotechnology **1**:784-791

Singer, B.S., Shtatland, T., Brown, D., Gold, L. 1997 Libraries for genomic SELEX. Nucleic Acid Res 25:781-786

Smith, G.C. and Merkel, J.R. 1982 Collagenolytic activity of *Vibrio vulnificus*: potential contribution to its invasiveness. Infect. Immun. 35:1155-1156

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

Strom, M. S. and Paranjpye, R.N. 2000 Epidemiology and pathogenesis of *Vibrio vulnificus*. Microbes Infect. 2:177-188

Strozen, T.G., Stanley, H., Gu, Y., Boyd, J., Bagdasarian, M., Sandkvist, M., Howard, S.P. 2011 Involvement of the GspAB complex in assembly of the type II secretion system secretin of *Aeromonas* and *Vibrio* species. J. Bacteriol. **193**:2322-2331

Strozen, T.G., Li, G., Howard, S.P. 2012 YghG ($GspS_{\beta}$) is a novel pilot protein required for localization of the $GspD_{\beta}$ Type II secretion system of enterotoxigenic *Escherichia coli*. Accepted for publication in the journal Infection and Immunity.

Svennerholm, A.M. 2011 From cholera to enterotoxigenic *Escherichia coli* (ETEC) vaccine development. Ind. J. Med. Res. 133:188-194

Svennerholm, A.M., Vidal, Y.L., Holmgren, J., McConnell, M.M., Rowe, B. Role of PCF8775 antign and its coli surface subcomponents for colonization, disease and protective immunogenicity of enterotoxigenic *Escherichia coli* in rabbits. *Infect. Immun*. **56**:523-528

Tauschek, M., Gorrell, R.J., Strugnell, R.A., Robins-Browne, R.M. 2002 Identification of a protein secretory pathway for the secretion of heat-labile enterotoxin by an enterotoxigenic strain of *Escherichia coli. Proc. Natl. Acad. Sci. USA* **99**:7066-7071

Taylor, R.K., Miller, V.L., Furlong, D.B., Mekalanos, J.J. 1987 Use of *phoA* fusions to identify a pilus colonization factor coordinately regulated with cholera toxin. PNAS USA 84:2833-2837

Tenaillon, O., Skurnik, D., Picard, B., Denamur, E. 2010 The population genetics of commensal *Escherichia coli*. Nat. Rev. Microbiol. **8**:207-217

Thanassi, D.G. and Hultgren, S.J. 2000 Multiple pathways allow protein secretion across the bacterial outer membrane. Curr. Opin. Cell. Biol. **12**:420-430

Tobias, J., Lebens, M., Kallgard, S., Nicklasson, M., Svennerholm, A.M. 2008 Role of different genes in the CS6 operon for surface expression of enterotoxigenic *Escherichia coli* colonization factor CS6. Vaccine **26**:5373-5380

Tomkiewicz, D., Nouwen, N., Driessen, A.J.M. 2008 Kinetics and energetics of the translocation of maltose binding protein folding mutants. J. Mol. Biol. 377:83-90

Tokuda, H. and Matsuyama, S. 2004 Sorting of lipoproteins to the outer membrane in *E. coli*. Biochem. Biophys. Acta **1693**:5-13

Tokuda, Matsuyama, S., Tanaka-Masuda, K. 2007 Structure, function and transport of lipoproteins in *Escherichia coli*. In The periplasm, ed. M Ehrmann, pp. 67-79 Washington, DC: ASM press

Touchon *et. al.*, 2009 Organized genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. PLoS Genet. **5**:E1000344-E1000344

Trachman, J.D. and Yasmin, M. 2004 Thermo-osmoregulation of heat-labile enterotoxin expression by *Escherichia coli*. Curr. Microbiol. 49:353-360

Tran, H.T., Barnich, N., Mizoguchi, E. 2011 Potential role of chitinases and chitin-binding proteins in host-microbial interactions during the development of intestinal inflammation. Histol. Histopathol. 26:1453-1464

Trindade, M.B., Job, V., Contreras-Martel, C., Pelicic, V., Dessen, A. 2008 Structure of a widely conserved type IV pilus biogenesis factor that affects the stability of secretin multimers. J. Mol. Biol. 378:1031-1039

Turner, S.M. *et. al.*, 2006 Phylogenetic comparisons reveal multiple acquisitions of the toxin genes by enterotoxigenic *Escherichia coli* strains of different evolutionary lineages. J. Clin. Microbiol. 44:4528-4536

Turner, S.M., Scott-Tucker, A., Cooper, L.M., Henderson, I.R. 2006 Weapons of mass destruction: virulence factors of the global killer Enterotoxigenic *Escherichia coli*. FEMS Microbiol. Lett. 263:10-20

Vignon, G., Kohler, R., Larquet, E., Giroux, S., Prevost, M.C., Roux, P., Pugsley, A.P. 2003 Type IV-like pili formed by the type II secreton: specificity, composition, bundling, polar localization, and surface presentation of peptides. J. Bacteriol. 185:3416-3428

Voulhoux, R., Ball, G., Ize, B., Vasil, M.L., Lazdunski, A., Wu, L.F., Filloux, A. 2001 Involvement of the twin-arginine translocation system in protein secretion via the type II pathway. *EMBO J.* **20**:6735-6741

Voulhoux, R., Bos, M.P., Guertsen, J., Mols, M., Tommassen, J. 2003 Role of a highly conserved bacterial protein in outer membrane protein assembly. Science 299:262-265

Waldor, M.K. and Mekalanos, J.J. 1996 Lysogenic conversion by a filamentous phage encoding cholera toxin. Science, 272:1910-1914

Walker, R.I., Steele, D., Aguado, T., 2007 Analysis of strategies to successfully vaccinate infants in developing countries against enterotoxigenic *E. coli* (ETEC) disease. Vaccine 25:2545-2566

Wenneras, C., Erling, V. 2004 Prevalence of enterotoxigenic *Escherichia coli*-associated diarrhoea and carrier state in the developing world. J. Healt Popul. Nutr. 22(4):370-382

Williams, K.P., Gillespie, J.J., Sobral, B.W.S., Nordberg, E.K., Snyder, E.E., Shallom, J.M., Dickerman, A.W. 2010 Phylogeny of Gammaproteobacteria. J. Bacteriol. **192**(9): 2305-2314

Wu, T., Malinverni, J., Ruiz, N., Kim, S., Silhavy, T.J., Kahne, D. 2005 Identification of a multicomponent complex required for outer membrane biogenesis in *Escherichia coli*. Cell 121:235-245

Xu, J. *et. al.*, 2007 Evolution of symbiotic bacteria in the distal human intestine. PLoS Biol. 5:1574-1586

Yahr, T.L. and Wickner, W.T. 2001 Functional reconstitution of bacterial Tat translocation *in vitro*. EMBO J. 20:2472-2479

Yakushi, T., Masuda, K., Narita, S., Matsuyama, S., Tokuda, H. 2000 A new ABC transporter mediating the detachment of lipid-modified proteins from membranes. Nat. Cell Biol. 2:212-218

Yang, J., Baldi, D.L., Tauschek, M., Strugnell, R.A., Robins-Browne, R.M. 2007 Transcriptional regulation of the *yghJ-pppA-yghG-gspSDEFGHIJKLM* cluster, encoding the type II secretion pathway in Enterotoxigenic *Escherichia coli*. J. Bacteriol. **189**:142-150

Yamaguchi., K., Yu, F., Inouye, M. 1988 A single amino acid determinant of the membrane localization of lipoproteins in *E. coli*. Cell **53**:423-432

Yamanaka, H., Nomura, T., Fuji, Y., Okamoto, K. 1998 Need for TolC, an *Escherichia coli* outer membrane protein, in the secretion of heat-stable enterotoxin I across the outer membrane. Microb. Pathog. 25:111-120

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

Yu, J., Webb, H., Hirst, T.R 1992 A homologue of the *Escherichia coli* DsbA protein involved in disulphide bond formation is required for enterotoxin biogenesis in *Vibrio cholerae*. Mol. Microbiol. 6:1949-1958

Zalewska-Piatek, B., K. Bury, R. Piatek, P. Bruzdziak, and J. Kur. 2008. Type II secretory pathway for surface secretion of DraD invasin from the uropathogenic *Escherichia coli* Dr(+) strain. J. Bacteriol. 190:5044–5056

Zahrl, D., Wagner, M., Bischof, K., Bayer, M., Zavecz, B., Beranek, A., Ruckenstuhl, C., Zarfel, G.E., Koraimann, G. 2005 Peptidoglycan degradation by specialized lytic transglycosylases associated with type III an type IV secretion systems. Microbiology, **151**:3455-3467

Zhang, H. Z., Lory, S., Donnenberg, M.S. 1994 A plasmid-encoded prepilin peptidase gene from enteropathogenic *Escherichia coli*. J. Bacteriol. **176**:6885-6891

Zhang, R.G., Scott, D.L., Westbrook, M.L., Nance, S., Spangler, B.D., Shipley, G.G., Westbrook, E.M. 1995 The three-dimensional crystal structure of cholera toxin. J. Mol. Biol. 251:563-573