Identification of human proteins interacting with the protein IcsB of *Shigella flexneri*

Ashwag Alzahrani

Thesis submitted to the

Faculty of Graduate and Postdoctoral Studies for the

Master degree in Chemistry

Department of Chemistry and Biomolecular Sciences

Faculty of Science

University of Ottawa

Candidate	**Supervisor**
Ashwag Alzahrani	Professor Francois-Xavier Campbell-Valois

Acknowledgements

I thank **God** for giving me this opportunity to be a student in Canada and making my dream of graduation a reality.

Then first and foremost, I thank my supervisor **Francois-Xavier Campbell-Valois**. It has been an honor to be his first MSc student. Since I began working on the thesis, Professor Campbell-Valois has offered valuable guidance, scholarly advice, and consistent encouragement throughout my research work. I will not forget all the moments we shared that talking about different topics such as science, politics, and religions. He always considers us as a family and understood the struggle that we, international students, have to deal with being away from home in a completely different environment and culture. This success was possible only because of his unconditional support.

Second, I am grateful for **The Ministry of Education in Saudi Arabia** and **Saudi cultural bureau in Canada** for supporting my M.sc. studies. I would like to thank my committee members, **Dr. John Pezacki** and **Dr. Chris Boddy** for their constant guidance and support. Also, I would like to thank all the members and professors of the Marion lab for their assistance.

To my nephew **Ayman,** I want to thank you for everything you have done for me. Without your support and your great patience, this dream would not have been possible.

To my family, I express my deepest gratitude to my awesome **parents** and **siblings**, even though you are far away, you have encouraged and helped me at every stage of my personal and academic life. Most importantly, I thank my loving and supportive husband, **Rashed**, who has supported me in every possible way to see the completion of this work. Thank you for your understanding and for holding my hand from the beginning of my study abroad until I reached this stage of prosperity; and especial thanks to my lovely kids, **Anas, Rital,** and **Ritaj** who have provided unending inspiration.

And last, but not least warm thanks to all my friends, especially**, Salimah** , **Jessica,** and **Nwara.**

Dedicated to

My Mom and Dad

Thanks for everything; particularly your patience.

Abstract

Problem: *Shigella* is a gram-negative enteropathogen that, when passed through fecal particles from one host to the oral cavity of another host, causes an infectious disease known as shigellosis. One of the distinctive features of the infection by *Shigella* is its ability to bypass its host's autophagic defenses. It does this through the use of a Type III secretion system, found in gram-negative pathogens like *Shigella*, which injects virulent proteins into the host cell. One of these proteins is IcsB; however, its exact function is not well understood. This study aims to better understand the role of this protein in the infection.

Methods: A yeast two-hybrid screening test is used in this case to examine the interactions between variations of the protein IcsB, and a library of host proteins. Given IcsB's high yeast toxicity and that resulted in the total absence of yeast colony formation, the first aim was to identify IcsB variants which expression would not prevent yeast growth. The second aim was to use the mutant with reduced cytotoxicity to perform a Y2H screen that will allow for the identification of candidate host proteins interacting with IcsB.

Results: Two mutations of the IcsB protein grew in the Y2HG yeast strain, indicating a significant reduction in the protein's toxicity. Of the cultures that reacted, high stringency and strong interaction was observed between four genes and IcsB proteins. Among the four identified clones that grew, three corresponded to the gene RNF2, while the last one corresponds to a non-coding sequence. Key control experiments revealed that the interaction of IcsB with RNF2 is likely false-positive. Thus, when screened full-length IcsB using new epithelial cells cDNA

libraries, strong interaction was observed between three genes and our IcsB proteins. All the three genes DDX3X, FANCL, and SGT1 passed the false-positive interaction tests. It is interesting to notice that DDX3X and SGT1 interacted with catalytically active and inactive IcsB, suggesting that the interactions established between IcsB and prey proteins does not require the catalytic - C306A mutation and that IcsB most likely does not function as a protease against these two proteins. By contrast, FANCL bound catalytically inactive, but not catalytically active IcsB, suggesting it could be a substrate of IcsB. The literature provides some support for the putative role of DDX3X, FANCL, and SGT1 in regulating the vacuole escape of *Shigella* through IcsB action.

Conclusion: The aim of this study was to determine the functional of IcsB in the vacuole escape of Shigella. This study successfully identified three candidates interacting partner proteins for IcsB. Key control experiments confirmed the interaction of IcsB with DDX3X, FANCL and SGT1. This study provides a basis for further research, with further study aimed at confirming these results during *Shigella* infection.

Table of Contents

List of figures and tables

List of abbreviations

T3SS	Type three secretion system
AbA	Aureobasidin
Y2HG	Yeast-two-hybrid Gold
DNA-BD	Binding domain
DNA-AD	Activation domain
WB	Western blot
DDO	Double Dropout synthetic medium
QDO	Quadruple Dropout synthetic medium
EV	Empty vector
DM	*pGBKT7 IcsB C306A Δ80* or double mutant
SM	*pGBKT7 IcsB Δ80* or single mutant
DM1	*pGBKT7 IcsB domain1*(residues 81-319)
DM2	*pGBKT7 IcsB domain 2* (residues 320-494)

Chapter 1: Introduction and Literature Review

In a recent publication, the World Health Organization (WHO) has reported that infectious diseases such as lower respiratory infections (rank 4th), tuberculosis (rank 10th) and diarrheal diseases (rank 9th) are each among the ten most common causes of death in the world (WHO, 2017). These numbers are based on the complete data from higher-income countries as well as lower-income countries. Death toll of infectious diseases is higher in the latter where statistical data is, however, less accurate due to lack of specific organization in charge of documenting causes of death. This implies that annual deaths due to infectious diseases may well be higher. The specific causes of diarrheal diseases are feces-contaminated water, which as a result contains bacterial, viral and parasitic organisms. Furthermore, it is the lower respiratory infections that cause the most fatalities with about 3.2 million deaths/year according to 2015 statistics (WHO, 2017). Many diarrheal and respiratory tract infections are caused by bacterial pathogens. Hence, these infections can often be treated with antibiotics and/or better sanitation. Antibiotic therapies have proven to be effective in the treatment of infectious diseases and significantly reduced the number of fatalities due to infectious diseases.

The rise in antibiotic resistance makes the treatment of bacterial infection more challenging. Excessive use of antibiotics in humans and farm animals has a disadvantage owing to the development of resistance. For example, *Salmonella* spp. and *Campylobacter jejuni* have developed antibiotic resistance primarily due to use of antibiotics in food animals (CDC, 2017). The impact of antibiotic resistance according to the data from Center for Disease Control is implicated in over 2 million infections per year in the USA alone (CDC, 2017). This effect is due

to the natural tendency of bacteria to develop mechanisms to resisting the antibiotics (CDC, 2016). One potential solution to this serious problem will be based on expedient and practical detection and identification of the key microbial agents implicated in the infection. This alternative strategy to treating infectious diseases will focus on targeting host-pathogen processes specific to microbes detected in the patient. Specifically better and more effective treatment for diarrheal diseases can therefore be developed if specific host-pathogen processes are well understood. Currently, such treatment methods based on direct target of pathogens intracellular lifestyles are underdeveloped. Hence, if such methods of treatment are to be developed they must focus on targeting intracellular pathogens directly. As it will become apparent below, this element constitutes a potential application that could stem from my thesis work.

Gram-negative bacteria are characterized by their unique cell envelopes that are made up of a thin peptidoglycan cell wall located between the outer membrane and the cytoplasmic membrane (Baron, 1996). In terms of their abundance in nature, Gram-negative bacteria are commonly found in virtually every ecosystem . Some of the prototypical examples of gram-negative bacteria include *Escherichia coli (including Shigella spp.), Pseudomonas aeruginosa, Chlamydia trachomatis* and *Yersinia pestis*. Due to their unique double membrane structure, gram-negative bacteria are well shielded from antibiotics, detergents, lysozyme and antimicrobial enzymes and peptides. In addition, the membrane of the gram-negative bacteria most often also contains lipopolysaccharide (LPS). Lipid A is one component of LPS, which unfortunately cause toxicity known as a septic shock when released in the host. Thus gram-negative bacteria are well equipped for bypassing the host cell immune system and thereby cause serious illness.

Shigella spp. are gram-negative enteroinvasive bacteria that infect the large intestine and have a highly contagious nature (Kothary, 2007). They are rod-shaped nonspore-forming bacteria that are virtually identical to *Escherichia coli*. The most critical difference between them is the presence of a virulence plasmid that encodes most of the virulence factors of *Shigella spp*. They are responsible for a condition called shigellosis, one of the forms taken by bacterial dysentery. This condition is a diarrheal disease characterized by the presence of mucus and sometime blood in liquid stools. *Shigella* is a genus that fall into the *Enterobacteriaceae* class and is broken up into four pathovars named *S. dysenteriae, S. flexneri, S. boydii,* and *S. sonnei* (Kang et al., 2018). One of the ways to diagnose shigellosis involves direct analysis of stool samples from the infected individual. However, a more accurate diagnostic might be achieved by cultivation of the etiologic agent from the stools and their identification using specific growth and metabolic requirements (Keusch, 1996).

The key mode of transmission of *Shigella* bacteria is through the fecal-oral route, which makes it highly contagious through person-to-person contact (such as in childcare), consumption of contaminated food and/or fluids. Another factor that makes these bacteria so highly contagious is their overall low infection doses; an inoculum of 200 bacteria is generally sufficient to cause symptoms of the disease in the majority of individuals (Kothary, 2007). This is in stark contrast to many other pathogens that mostly require a larger inoculum for causing disease. For example, about 10^3 salmonella bacteria are needed to cause infection (Kothary, 2007). One of the potential reasons invoked to explain this phenomenon is that *Shigella spp.* can survive in extreme environments within the body such as the highly acidic pH encountered in the stomach (Gorden & Small, 1993). The characteristic mode of action of *Shigella* bacteria is the secretion of proteins

called effectors and, less commonly, exotoxins (limited to *S. dyssenteriae*) that allow the colonization of the mucosae of the large intestine. The characteristic symptoms develop within 3 days and include significant ulcers and swelling of the infected mucosal surface. These inflammatory damages are followed by diarrhea containing mucus and blood, nausea and fever.

Analysis of the statistical data shows that yearly shigellosis represents 10 to 20% of all the reported diarrheal disease cases and one half of all reported bloody diarrhea cases. The numbers are even more alarming when considering that the death toll due to shigellosis amounts to over 500,000 deaths/year, mostly children under the age of five (*Shigella* spp, 2016). It appears that not all geographic regions are equally affected by this highly contagious disease as both cases and death cases in Canada, for example, are rare. By contrast, the higher number of cases and fatality rate in lower-income countries is due to poor sanitation conditions and the overall health and access to care, respectively (CDC, 2017).

Although a one-week treatment of shigellosis can be achieved with antibiotics, they are becoming less effective in the treatment of this disease due to the emergence of antimicrobial resistance in certain *Shigella* strains. Thus, the efficiency of treatment based on antibiotics is in jeopardy and directly stimulate research toward identifying alternative treatment methods (CDC, 2017). The World Health Organization ranks *Shigella* between the 10^{th} to 12^{th} positions in the list of 12 bacteria for which novel antibiotics are most urgently needed (WHO, 2017). The development of alternative therapeutic strategies necessitates that we further our understanding of *Shigella* pathogenesis.

Effectors are highly diverse in both their structure and their function. Pathogenic bacteria secrete a large variety of effectors with various functions. For example, the type III secretes effectors, which have been shown to catalyze reactions such as ADP ribosylation, proteolysis, phosphorylation and ubiquitin ligation (Banfield et al., 2017). T3SS effectors can be found in the bacterial cytosol or at the tip of the needle of the T3SS prior to secretion, and in the host cell when the secretion process is completed.

Specifically, type three secretion system (T3SS) is the basis for virulence of *Shigella* (T3SS) and was shown to be a key part of pathogenicity for many other gram-negative pathogenic bacteria such as *Burkhoderia spp.*, *Enteropathogenic E. coli*, *Yersinia enterocolitica*, and *Salmonella* (F. X. Campbell-Valois, Sachse, Sansonetti, & Parsot, 2015). The T3SS is a syringe shaped multimolecular proteinaceous complex embedded in the cell envelope of the bacterium. The T3SS and translocator protein complexes direct the secretion mechanism. Regulation of the expression of the T3SS components and effectors occurs mostly at the transcriptional level. Secretion is also regulated at the post-translational level by T3SS-specific chaperons, which modulate the abundance and availability of effectors that are expressed before the secretion is activated (F. X. Campbell-Valois et al., 2014). Critical for the mechanism of action are protein molecules called effector. Bacterial effectors often function via selective binding to host proteins and thereby regulating their specific function or activity through enzymatic or non-enzymatic processes. The specific mechanism of action involves T3SS contact-mediated detection of the host cell surface, which consequently promotes the host epithelial cell plasma membrane to uptake the bacterium (F. X. Campbell-Valois et al., 2015). During this entry process, the T3SS secretes tens of effectors directly into the host cell. Following entry, *Shigella* lyses the entry vacuole. Once

liberated in the cytosol, *Shigella* use an interesting mechanism to achieve cell-to-cell spread, the mechanism through which it is capable of invading neighboring epithelial cells from the initially infected cell (Ireton, 2013). The cell-to-cell spread is commenced with pathogen activation of host filamentous (F)-actin, which forms F-actin tail or comets structure that enable its movement in the cytoplasm. This general mechanism for cell-to-cell spread is shared with other human bacterial pathogens such as *Listeria monocytogenes*, *Burkholderia pseudomallei* and *Ricckettsia* spp. (Sitthidet et al., 2008). To move inside the cytosol, these bacteria remodel the F-actin network. This process required the outer membrane protein IcsA in *Shigella* spp. (Bernardini, Mounier, d'Hauteville, Coquis-Rondon, & Sansonetti, 1989). The cytosolic movement favors the encounter of bacteria with the basolateral side of the plasma membrane. The invasion of the neighboring cell occurs via the engulfment of the protrusion formed by the bacteria through the deformation of the plasma membrane of the initially infected cell and its neighbor. During that step, *Shigella* requires the function of the effector IcsB to escape the vacuole resulting from the protrusion engulfment (Ogawa et al., 2005). Interestingly, the function of IcsB in vacuole escape has also been linked to autophagy since the autophagosome marker LC3 was recruited in high amount around IcsB-deficient *Shigella* (F. X. Campbell-Valois et al., 2015).

The classical role of autophagy is to carry out the regulated degradation and recycling of cellular components. The defense against microbes that make it into host cells is also relying principally on the autophagy pathway, which is also often referred to as xenoautophagy in the case where foreign entities are directly targeted (Mizushima, 2007). The bacterial xenoautophagy process commences with double-membrane autophagosome formation around intracellular bacteria. The nascent bacterial autophagosome then merges with lysosomal compartments that

contain molecules (e.g. proteases, etc.) that are detrimental to bacterial fitness (Mizushima, 2007). ATG8/MAP1LC3 (LC3) is the class of proteins that promote the fusion of autophagosome with lysosome yielding the degradative compartment dubbed the autophagolysosome (F.-X. Campbell-Valois & Pontier, 2016). These effectors alter the physiology of the host cell through various mechanisms thereby allowing the pathogen to colonize the large intestine (Figure 1.1). During their genesis, autophagosome membranes entrap cytoplasmic content including organelles, cytosolic proteins, viruses or bacteria just to name a few and destroy them allowing the recycling of amino acids and other small degradative products of biomolecules. The hallmark of autophagosomes is the double-membrane protective outside structure. Interestingly, there exists a membrane compartment called PAS, which precedes in the genesis of the autophagosome. Hence, autophagosome genesis is initiated at the pre-autophagosome structure (PAS), such as phosphatidylinositol-3-phosphate (PtdIns3P) rich domain on the cytosolic side of the ER membrane. Overall, infection signal induction, which promotes phagophore membrane nucleation that precedes formation of autophagosome. Fusion with lysosome then initiates the degradation process of the infection (Huang & Brumell, 2014).

A somewhat more specific autophagy called selective autophagy has cargo selection as an initial step in autophagosome formation. It is suggested that this extra step is activated by the detection of certain proteins by their cognate cargo receptors and adaptors. These specific receptors most commonly contain LC3-interacting region that triggers the recruitment of LC3-bound autophagosomes to the cargo domains (Huang & Brumell, 2014). Some of these adaptor proteins have been shown to contain ubiquitin-binding domain such as NBR1, NDP52 and optineurin. Selective autophagy is often associated with xenoautophagy. For example, NBR1 was shown to

be specific for protein aggregates, peroxisome and the bacterial pathogen *Francisella tularensis* (Huang & Brumell, 2014). In a separate study focusing on *Salmonella enteri*ca, NDP52 was found to be important for autophagy (Randow et all, 2009). *S. enterica* belongs to the enterobacteriace family just as *Shigella* spp. However, *Shigella* uses at least two effectors known as VirA and IcsB to avoid LC3 recruitment in their vicinity. VirA has been shown to act as a GTPase activating protein (GAP) for Rab1, a process that could potentially affect LC3 recruitment to *Shigella* (Dong et al., 2012). The mechanism of IcsB anti-autophagy activity has been more disputed.

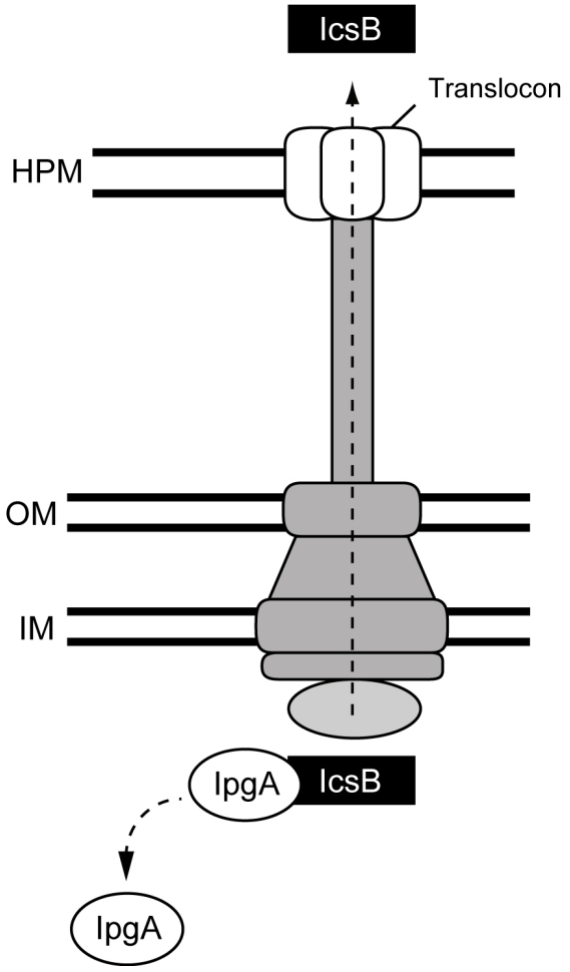

Figure 1.1. Scheme of the Type III Secretion System(T3SS).
The T3SS is a multiprotein nanomachine spanning the bacterial inner and outer membranes that allow the transfer of effectors such as IcsB.

Now, let me describe in more details the properties of IcsB. IcsB is an effector protein secreted by *Shigella* via the T3SS during infection (Ogawa, Suzuki, Tatsuno, Abe, & Sasakawa, 2003). It is composed of 494 amino acid residues, which confer it a molecular weight of approximately 56 kDa. It is a relatively basic protein since its pI is around 9.7. A sequence alignment showed that IcsB and BopA are homologous proteins. This sequence alignment was used to characterize mutations $IcsB_{81\text{-}494}$ (Δ80) and $IcsB_{81\text{-}494}$ (Δ80) C306A. Also, it was used to identify $IcsB_{81\text{-}319}$ (domain 1 or DM1) C306A and $IcsB_{320\text{-}494}$ (domain 2 or DM2) (Figure 1.2).

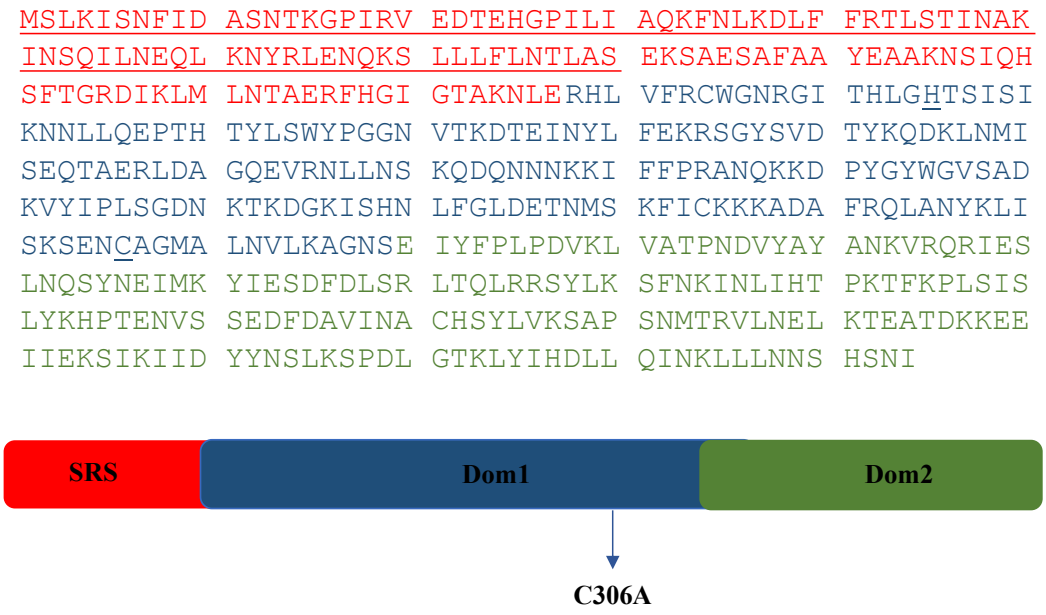

Figure 1.2. The domain architecture of IcsB.

A. Alignments of IcsB primary structure with remote sequence homologs help defined domain 1 (DM1) containing a putative catalytic site (C306 underlined in the sequence) and a domain 2 (DM2) with unknown function. The Sequence Required for Secretion (SRS) extending from residues 1 to 127 is the domain normally involved in recognition of IcsB by the T3SS, thus it is necessary only for IcsB secretion. B. Schematic representation of the arrangement of IcsB in three regions: (red: SRS, blue: Domain 1 and green: Domain 2).

As depicted in Figure 1.1, the chaperone of IcsB is IpgA; the role of IpgA is to stabilize IcsB in the bacterial cytosol and promotes its T3SS-dependant secretion (Ogawa et al., 2003). It was previously shown that during the course of infection in the absence of IcsB, an additional *Shigella* protein named IcsA interacts with the Atg5 protein, an essential component of the host autophagy machinery (2). The model presented in this study suggests that the recognition of IcsA by ATG5 enable bacterial capture by the autophagosome. In contrast, Campbell-Valois et al. have proposed a model where LC3 is not recruited around cytosolic bacteria but rather around bacteria found in vacuoles formed during cell-to-cell spread (F. X. Campbell-Valois et al., 2015). They observed that LC3 is directed within the vicinity of secreting bacteria found in these vacuoles but not around non-secreting bacteria located in the cytosol. Since LC3 was also observed around both *ΔicsB* and WT *Shigella* albeit to a lesser extent, these data suggest that IcsB might play a role in vacuole escape. In the absence of IcsB and VirA function, LC3 continues accumulating on the vacuole hence stimulating fusion with lysosome, which ultimately kill the trapped bacteria (Figure 1.3).

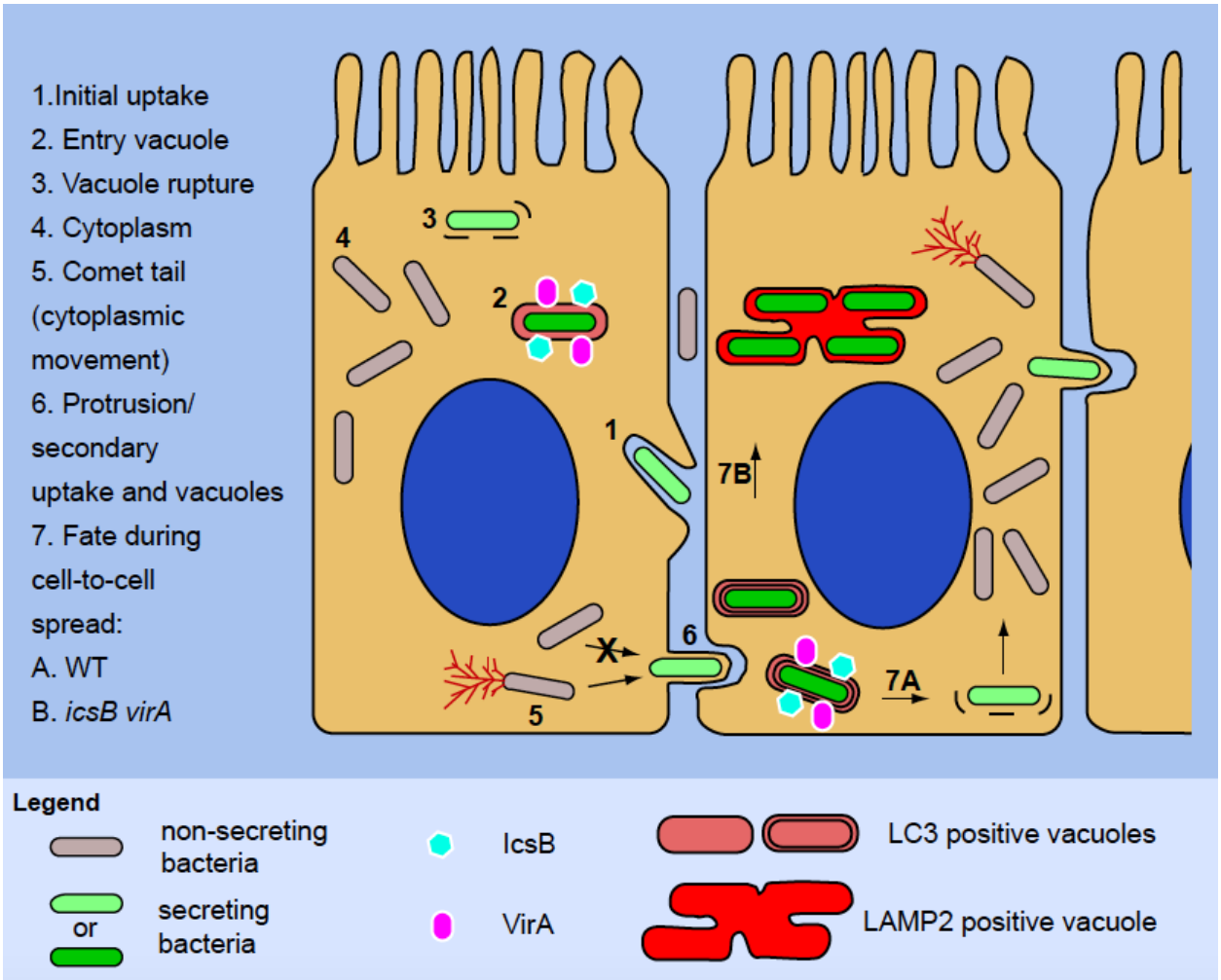

Figure 1.3. Model describing the role of the effectors IcsB and VirA.
IcsB and VirA in the escape of actively secreting *S. flexneri* from LC3- positive vacuoles formed in epithelial cells (F. X. Campbell-Valois et al., 2015).

These observations put into question the targeting of IcsA by ATG5 since it seems counterintuitive that IcsA would be targeted by autophagy in the lumen of the dissemination vacuole but not in the cytosol where it is totally exposed and IcsB is not secreted. Hence, the working hypothesis of my thesis work is: IcsB is a classical effector that is targeting one or several host proteins to facilitate vacuole escape (Figure 1.4).

It is clear that if we are to gain valuable insight into the mechanism of escape of *Shigella* from the host autophagy defense system and the key role played by IcsB, the host targets of this effector must be clearly identified. At the moment where this thesis work started, there were no known host proteins that interacted with IcsB. Nevertheless, valuable information was available in the scientific literature. For example, a cholesterol-binding domain (CBD) was found in IcsB, which enabled interaction with cholesterol found in the inner face of the host plasma membrane (Kayath et al., 2010). A more critical finding reported in the same study was that the mutant Y297F Y340F that is impaired in cholesterol-binding was not able to block LC3 recruitment, which suggests a role of the CBD in vacuole escape.

Furthermore, a theoretical study conducted by Pei and Grishin have shown that the N-terminus domain of IcsB and of its homolog BopA, and the Rho GTPase inactivation domain (RID) of *Vibrio cholera* MARTX toxin have homology with the inverted papain like fold (Pei & Grishin, 2009) (Figure 1.2). The sequence alignment they generated allowed them to identify a conserved catalytic dyad (H145 and C306 in IcsB), which is the hallmark of serine proteases or acyltransferase. This led the authors to speculate that IcsB may use this protease or acyltransferases activity to escape the autophagy defense system. This study supported the notion that IcsB may be targeting a host protein to enable this phenomenon. Lastly, another study demonstrated the toxic nature of IcsB expression in yeast cells (Slagowski, Kramer, Morrison, LaBaer, & Lesser, 2008). This phenotype has precluded the use of the yeast-two hybrid to identify human proteins interacting with IcsB.

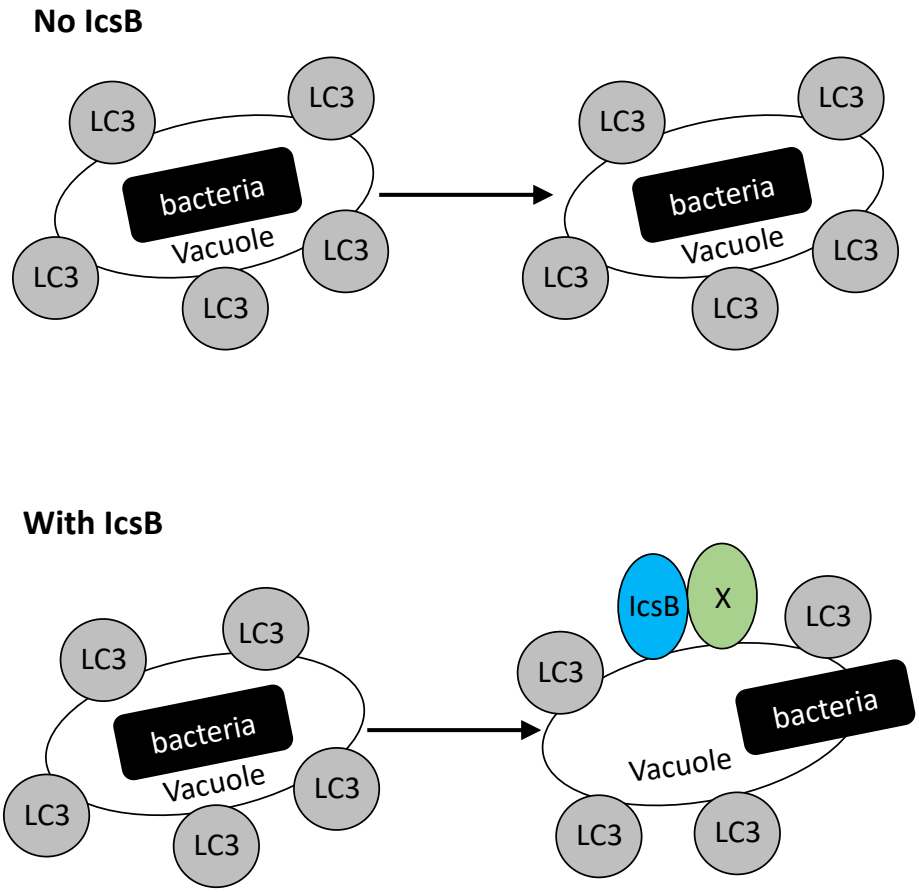

Figure 1.4. IcsB is interacting with host protein X to enable vacuole escape.
The presence of IcsB enables the escape of positive vacuoles through its localization around the bacterial surface.

During my thesis work, significant efforts were put toward the identification of mutations in IcsB that might rescue the yeast cytotoxicity phenotype. This was a precondition to use the yeast two-hybrid (Y2H) system to identify human proteins that might interact with IcsB, thereby providing new insights about how IcsB contributes to vacuole escape. The basis for the Y2H system used in this study is based on the fusion of the bait and prey proteins with the Gal4 DNA binding (DBD) and activating domain (AD), which upon interaction of the bait and prey activates the transcription of 4 reporters genes: AUR1-C, ADE2, HIS3, and MEL 1 (Van Criekinge &

Beyaert, 1999). The promoters of these 4 reporter genes contain different promoters that all contain Gal4 responsive elements, thus, activating the process of transcription in a Gal4-binding dependent manner. Each reporter gene was subjected to various conditions to monitor its activation. For instance, to test the *AUR1-C* gene, it was exposed to Aureobasidin A (AbA) that is an antibiotic responsible for inhibiting the yeast enzyme inositol phosphoryl ceramide synthase, critical to the synthesis of the ceramide class of phospholipids. The activation of *AUR1-C* gene allows expression of a mutant inositol phosphoryl ceramide synthase that is resistant to AbA. Hence, yeast colonies would only be formed if *AUR1-C* was sufficiently expressed. Additionally, to test the reporter gene *MEL1*, yeasts were incubated in the presence of *X-α-Gal*. If *MEL1* was activated, leads to the expression of an *α*-galactosidase (a glycoside hydrolase enzyme), and in the presence of *X-α-Gal* it is responsible for the blue color of the yeast colonies. The intensity of the color depends on the level of the expression of *MEL1* and is therefore an indirect measure of the strength of the interaction between the bait and the prey protein. The Y2HGold or diploid cells obtained by their mating with Y187 cannot grow in synthetic medium lacking histidine and adenine. However, if reporter genes *HIS3* and *ADE2* are activated by the Y2H, they recover their ability to grow these conditions. The combination of the four reporter genes allows to minimizing the number of false positives by increasing the stringency of the selection. One of the key steps when using this method is to ensure that the expression of the bait protein is not toxic. A direct indication that one protein is toxic is measured by slower growth of the yeast culture and smaller colonies on the agar. Using this method it is also important to confirm the positive interactions by performing false positive tests and by validating the interaction between the bait and the prey proteins by other methods in their natural context.

During my thesis, I have been able to identify that a truncation of 80 amino acids in the N-terminus of IcsB and the mutation of the catalytic Cysteine 306 in alanine rescued the cytotoxicity phenotype in yeast. Therefore, we used *IcsB Δ80 C306A* as a bait in two yeast-two-hybrid screens. This allowed us to identify three host proteins that are putative targets of IcsB during infection. I will discuss the possible role that these interactions could play in the vacuole escape of *Shigella*.

Chapter 2: Materials & Methods

2.1 IcsB mutagenesis

pGBKT7-IcsB was obtained from Ashley Weeks a former student in the lab. It is a plasmid that allows the expression of the Gal4-DNA binding domain-IcsB fusion and the GAL4 DBD-IcsB Δ80, respectively. PCR mutagenesis and ligation were used for creating two types of IcsB mutants known as *IcsB-delta80NT* (also known as *IcsB Δ80*) and *IscB-C306A-delta80NT),* as depicted by the graphical representation of *pGBKT7-IcsB* WT in Figure A1 of the appendix. Moreover, ThermoFisher Scientific (Waltham, MA, USA) and Bio-Rad (Hercules, CA, USA) provided the materials required for conducting these assays.

2.1.1 Polymerase Chain Reaction amplification

IcsB-*Δ80*: To perform this mutagenesis PCR we used primers 80N_ter_IscB_S and 80N_ter_IscB_R using pGBKT7 IcsB WT as a template (primers are listed in Table A1 of the appendix) and a reaction mix composed of distilled water, buffer (5X Phusion HF), 10mM dNTPs, 50 ng of template DNA (*pGBKT7 IcsB*) and 0.02 U/µL of reaction of the Phusion High-Fidelity DNA polymerase (Thermofisher). The PCR was run using this cycle: 98˚C for denaturation, 62˚C for annealing and 72˚C for the extension. This program was set for only 20 cycles to prevent the accumulation of undesirable mutations. Finally, a 5 minutes 72˚C final extension time was programed. The resulting product was 8560 bp in length yielding a linear version of *pGBKT7 IscB Δ80*. The product was analyzed by agarose electrophoresis: 5uL of 1% ethidium bromide were mixed with 1.5% of agarose in the 1X TAE buffer system. After that, the wells were filled with 10 µL of 1X GeneRuler 1 kb Plus DNA Ladder and 15 µL of the sample, which was run for a period

30 minutes set a 100 V. Lastly, a photograph of the resulting gel under UV light was taken to document the result.

IcsB C306A* Δ80*: To perform this mutagenesis PCR, we used 10 µM of IcsB_C306A_S (forward primer) and IcsB_C306A_R (reverse primer) using *pGBKT7 IcsB Δ80* as a template (primers are listed in Table A1 of the Appendix) and the PCR mix, program and electrophoresis protocol described above. However, 0.5 mM of additional $MgCl_2$ was added to the PCR mix.

2.1.2 Digestion, phosphorylation and Ligation of PCR product

DnpI digestion*:* 25 to 50 µl of the PCR reaction was mixed with an equal volume of 1x Fast Digest Buffer and 1 µl of the DpnI enzyme that degrades specifically the methylated template DNA but not the unmethylated PCR product. The PCR products were digested at 37˚C for a period of 2 hours.

Purification: The GeneJET Gel Extraction Kit was used to purify the PCR product that had undergone digestion following the procedure suggested by the manufacturer.

Phosphorylation of 5' ends*:* 15 uL of the purified DNA (e.g. 345 to 750 ng) was incubated at 37˚C with 1 µL of T4 polynucleotide kinase (PNK enzyme), 2 µL of 10 nM ATP, and 2 µL of PNK Buffer A (forward reaction) for 30 minutes. Afterwards, the enzyme was heat inactivated at 75˚C for 10 minutes

***Intramolecular ligation**:* Finally, ligation was initiated by combining 5 to 10 microl of the PNK reaction with 4-4.5 µl of 10X T4 DNA ligase buffer, 0.5 µl T4 DNA ligase and water to obtain a final reaction volume of 50 microl.

2.1.3 Bacterial Transformation

The Escherichia coli MC1061 bacterial strain was used for the transformation of IcsB mutants and they were spread onto LB agar plates containing 50 µg/mL kanamycin. The GeneJET Plasmid Miniprep Kit was used for plasmid DNA purification following the procedure suggested by the manufacturer.

2.1.4 DNA Sequencing

The isolation of the relevant mutants was confirmed through Sanger sequencing using either SeqO8 or SeqO7 primer (Table A1). The pGBKT7_MCS3'_seq_R (SeqO7) primer that is a reverse primer because it allows for sequencing from the 3' end, helped in characterizing *pGBKT7-IcsB-C306A*, as it consisted of the desirable point C306A, a mutation located within the second half of the coding sequence. Additionally, the SeqO7 primer was also used to sequence *IcsB C306A Δ80* because *pGBKT7 IcsB Δ80* was used at the DNA template for the mutagenesis PCR and its 3' end contained the mutated. SeqO8 was used to sequence the 5' end of *pGBKT7 IcsB* Δ80 and derivatives. Sanger sequencing was performed at the Genome Québec Innovation Centre at McGill University (Montreal, QC). Lastly, the standard nucleotide BLAST from NCBI database (National Centre for Biotechnology Information) was helpful for conducting the analysis of the sequencing results. For validating these sequences, Blind BLAST and double BLAST against

wild-type IcsB were executed. Furthermore, the glycerol stocks of positive clones were formulated by combining 900µL of the relevant bacterial culture and 600µl of 50% glycerol.

2.2 Yeast cytotoxicity assay

2.2.1 Yeast transformation

After the process of sequencing, the Yeast-two-hybrid Gold (Y2HG) *Saccharomyces cerevisae* strain (Clontech, Mountain View, CA) was transformed with the various *pGBKT7-IcsB* variants according to the PEG/LiCl protocol. Consequently, the control group consisted of cells transformed with the vector *pGBKT7* (negative control) and the wild-type IcsB (positive control). Transformed cells were plated on SD-Trp containing 50 µg/mL kanamycin to inhibit the growth of putative bacterial contaminants. Plates were finally incubated at 30˚C for 3-5 days. All drop outs were for yeast media were from Sigma-Aldrich. Yeast nitrogen base and agar were from Bioshop. Glucose was from Fisher Scientific.

2.2.2. Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Yeast Protein Extraction: The transformed products of yeast were extracted, diluted to OD_{600} 2.5 and were subjected to centrifugation. This phase was run at room temperature for a period of 5 minutes, and consisted of suspending the pellet again in 100 µL of H_20 and 100 µL of 0.2 M NaOH. This process was repeated in the presence of 50 µL of 1X loading buffer and 0.5 µL of 100x PMSF and was boiled for 3 minutes. Samples were centrifuged at 11,000xg to recover the soluble fraction (Kushnirov, 2000).

SDS-PAGE: According to the classical protocol a 12% resolving gel was layered with a 5% stacking gel. After that, the wells were filled with the preheated samples and the Precision Plus Protein Standards marker (Cat. No. 161-0374) set at a temperature of 95°C for 5 minutes. The electrophoresis was performed at 100 V for the first 20 minutes, after which the migration was completed at 180 V in 1X SDS running buffer.

Western blot: According to the protocol of the manufacturer, the proteins were moved using the Protein Trans-Blot Turbo Blotting System onto the membrane of polyvinylidene difluoride (PVDF). Lastly, the membrane was transferred after adding the buffer and keeping it under the standard conditions (e.g. 2-3 mA at 110 v) for a duration of 45 minutes. Moreover, the membrane was exposed to 5% TBST-Milk and was subjected to a primary and secondary antibody, 1/2000 dilution of mouse anti-human c-Myc monoclonal antibody (Cedarlane, Burlington, ON. Reference: CLAT211) and 1/20,000 dilution of peroxidase-conjugated AffiniPure goat anti-mouse IgG (H+L) antibody (Jackson ImmunoResearch Labs, West Grove, Pa. Reference: 115-035-003). The c-Myc antibody recognizes the c-Myc epitope tag present in *pGBKT7* and its derivative. Throughout this process, the membrane was washed prior to the antibody treatment and afterwards with 1X TBST and was left to sit in the 1X TBS solution.

Imaging: The WB was visualized using the Chemidocs imaging system (Bio-Rad, Hercules, CA) and Clarity Western ECL solution (Bio-Rad). The results were revealed using either Chemi or High Chemi Sensitivity settings depending of the variation in the level of signals.

2.2.3 Dot Assay

Large and healthy colonies of yeast strains carrying the relevant *pGBKT7* derivatives were taken from SD-Trp plates and incubated overnight at 30 ˚C in 5 mL of SD-Trp under vigorous shaking (250 rpm). The value of OD_{600} was assessed the following morning and it was adjusted to OD_{600}= 0.10 in 5mL of SD-Trp. Then, it was incubated under the same conditions as above for 4 hours and the new OD_{600} was recorded. Generally, strains harboring *pGBKT7 IcsB C306A* or *pGBKT7 IcsB Δ80* were compared with those harbouring the empty vector *pGBKT7*. Serial 10X dilution between 0.1 and 0.0001 were obtained and 4.5 µL of each dilution for each strain were deposited on the surface of pre-dried SD-Trp agar plates. Observation were typically visualized after three days of incubation at 30°C

2.3 Yeast-Two-Hybrid **(**Additional details about the yeast-two-hybrid are presented in the appendix*)*

2.3.1 Auto-activation

Y2H gold strains containing the empty vector *pGBKT7*, *pGBKT7 IcsB Δ80* and *pGBKT7 IcsB C306A Δ80* were plated on different media to test for their capacity to autoactivate the *MEL1* and *AUR1-C* reporter genes: i) SD-Trp; ii) SD-Trp/AbA; iii) SD-Trp/X-α-Gal; iv) SD-Trp/AbA/X-α-Gal. Each plate consisted of 8µL of ABA (150 ng/mL), 40µL of X-α-Gal (40 µg/mL) and 20mL of SD-Trp. Cells were either plated when freshly transformed or using the dot assay protocol described above. Aba was purchased from Clontech (Cat Nos. 630466 & 630499). X-α-Gal was purchased from Sigma-Aldrich.

2.3.2 Matchmaker Gold Yeast-Two Hybrid Mating and Screening

The Clontech Laboratories (Mountain View, CA, USA) (Clontech, 2010) provided the materials and kits essential for conducting the Yeast Two-Hybrid assay including media. A screen was conducted to identify prey proteins derived from cDNA libraries cloned into plasmid *pGADT7-RecAB* that interacted with the bait protein *IcsB C306A Δ80*. The bait plasmid *pGBKT7* harbouring the GAL4 DNA-binding-domain (DNA-BD) coding sequence was used as the entry vector for cloning the *IcsB C306A Δ80* gene as described above. This plasmid was initially transformed into Y2H Gold. Similarly, the prey plasmid *pGADT7-RecAB* consisted of the GAL4 activation domain (AD) coding sequence fused with the cDNA library were obtained in a ready to use format directly transformed into the Y187 strain (Clontech). Both of these haploid cells population Diploid cells were mated to obtain the diploid cells employ in the screen. Hence, the resulting diploid strain harboured the four types of reporter genes (e.g. *AUR1-C*, *ADE2*, *HIS3*, and *MEL1*) controlled by the Gal4-responsive promoter elements described above and the bait and prey plasmids required to perform the screen.

In the first screening, we used the Normalized Mate & Plate Universal Human cDNA library (Reference number: 630480). This library contains cDNA from different human tissues.
In the second screening, we used the Normalized Mate & Plate HeLa S3 cDNA library (Reference number: 630479). Both libraries were inserted in *pGADT7 RecAB* (Figure A2). A concentrated overnight culture of the Y2H Gold bait strain containing *pGBKT7 IcsB C306A Δ80* was prepared by inoculating one fresh and medium-sized colony into 50 mL of SD/–Trp liquid medium as indicated per manufacturer's protocol. It was incubated with shaking at 260 rpm at 30°C after 16 hr the OD_{600} reached approximately 0.8 as expected. The overnight culture was centrifuged at

1,000 x g for 5 minutes to pellet the cells. The supernatant was removed and 5 mL of SD/–Trp medium was used to resuspend the pellet to a cell density of >1x10^8. Moreover, a hemocytometer was used to count the cells by making 1/50 dilution of the resuspension (2 µL cell x 98 µL SD/–Trp medium). Usually, the concentration of cell was such that the 5 mL resuspension at to be used in full for the mating. Prior to the mating, 10 µl of the library Y187 strain was removed to measure the library titer. Then the rest of the library Y187 cells suspension (e.g. 1 mL) was mated with the 5 mL cell suspension of the Y2H Gold bait strain *pGBKT7 IcsB C306A Δ80* in a sterile 2 L flask. 45 mL of 2xYPDA liquid medium (with 50 µg/mL kanamycin) was add to the 2 L flask. The mated culture was incubated at 30°C for 20 hr with slow shaking 40 rpm. After 20 hr, a drop of the culture was checked under a phase contrast microscope (40X) to make sure zygotes were present. Then the mated culture was split in two 50 mL conical tubes and centrifuged at 1,000 x g for 10 minutes to pellet the cells. The 2L flask was then rinsed twice with 45 mL 0.5xYPDA (with 50 µg/mL kanamycin) and used to resuspend the pelleted cells. All pelleted cells were resuspended in 10 mL of 0.5X YPDA/Kan liquid medium. Typically, the total volume of the cell suspension obtained at that step was 12.5. Finally, serial 10X dilution between 0.1 and 0.00001 of the mated cultures were plated onto variety of synthetically defined medium (SD). Consequently, after 5 days the library titer was measured and their viability was tested using the single dropout media. The bait plasmid's pGBKT7 viability was measured using SD/-Trp. The prey plasmid's *pGADT7-RecAB* viability was tested using the SD/-Leu media, since it harbored LEU2 a gene essential for leucine synthesis. Therefore, the yeast cells without plasmids did not form colonies, as they did not have the capacity to synthesize these essential amino acids. The diploid cells containing both bait (*pGBKT7*) and the prey plasmid (*pGADT7 RecAB*) were selected using a Double Dropout synthetic medium (DDO

or SD/-Leu/-Trp). Typically, during the Y2H screen 55-60 150 mm DDO plates containing AbA and X-α-Gal at the concentration indicated above were used to plate the resulting mated culture.

Another dropout technique known as the Quadruple Dropout (QDO or SD/-Leu/-Trp/-Ade/-His) or SD/-Leu/-Trp/-Ade/-His was used on diploid cells that consisted of the prey and bait plasmids; the purpose of the QDO was not only to select for diploid cells containing both plasmids but also test for the activation of *HIS3* and *ADE2* reporter genes. The QDO was also supplemented with *X-α-Gal* and *AbA* (X/A) at the concentration indicated above to test for the activation of *AUR1-C* and *MEL1* reporter gene. The QDO X/A is thus more stringent than the DDO X/A and allowed to validate further a binary interaction between the bait and a prey. Library plasmid responsible for selective of each positive clones were isolated by three successive streaking on the selective medium. After the first streaking, a mixture of blue and white colonies was observed, indicating segregation of positive interactors (blue) from non-interactors (white). After streaking two more times, only blue colonies were observed. The library plasmids were then rescued from one of these resulting colonies.

2.3.3 Zymoprep yeast plasmid DNA

The modified Zymoprep Yeast Plasmid Miniprep I (Zymo Research, Catalog No. D2001) was used to isolate positive clones obtained from the screening of the Y2H. Colonies picked from the plate obtained at the previous step, were used to seed 4 mL of DDO. The culture was incubated overnight at 30°C under vigorous shaking (e.g. 250 rpm). 2mL of the resulting yeast culture was centrifuged at 600 x g for 2 minutes. The supernatant was removed and 150 µl of Solution 1 and 2 µl of Zymolyase (previously buffered with 200 µl of the supplied Storage Buffer) of

concentration 5 unites/µL was poured onto the cell pellet and was incubated for 60 minutes. Moreover, the centrifugation process was conducted two minutes after the addition of 150 µl of Solution 2 and 150 µl of Solution 3. After this step, the GeneJET Plasmid Miniprep Kit was used for transferring supernatant to the spin column of GeneJet, which was washed by 500 µl of wash solution and it was centrifuged twice. The DNA was eluted with 35 µl of elution buffer (Tris 5 mM pH= 8.0). The typical yield was equal to or smaller than 2 ng.

2.3.4 Rescue of prey plasmids by transformation in DH10B

DNA isolated from yeast was transformed in the bacterial strain *Escherichia coli* DH10ß by electroporation. These bacterial transformed variants were plated onto LB agar plates supplemented with ampicillin (100 µg/l mL) in order to select for transformation of the *pGADT7 RecAB* that carried the prey cDNA. Lastly, the GeneJET Plasmid Miniprep Kit was used to isolate the plasmid DNA of the four clones. DNA concentration and purity was quantified with a NanoDrop spectrophotometer.

2.3.5 Restriction Enzyme Digestion

In the cDNA library cloning process, an EcoRI and a XhoI restriction site were respectively maintained in the 5’ end and 3’ end of each clone. To assess whether the clones appeared to be different from one another, we therefore digested each clone with EcoRI and XhoI and analyzed the result using 1.5% agarose electrophoresis gel.

2.3.6 Sanger Sequencing

The prey clones that were positive underwent Sanger sequencing. The prey plasmid *pGADT7 RecAB* contained a T7 promoter in proximity of the restriction site MCS A, which we therefore used for sequencing. Again, all sequencing runs were performed at the McGill and Genome Quebec Innovation Centre (Montreal, QC). Lastly, the standard nucleotide BLAST from NCBI database (National Centre for Biotechnology Information) was used to identify the nature of the gene encoded by each unique clone identified in the screens.

2.3.7 Dot assay against pGBKT7 IcsB Δ80 C306A, pGBKT7 IcsB Δ80, IcsB domain 1, and IcsB domain 2 .

Using a small-scale Matchmaker Yeast Two-Hybrid procedure, the empty vector *pGBKT7*, *pGBKT7 IcsB C306A Δ80, pGBKT7 IcsB Δ80, pGBKT7 IcsB domain1 (DM1) C306A*, and *pGBKT7 IcsB domain 2 (DM2)* in Y2H Gold were mated with each of the three prey DNA plasmids previously re-transformed in Y187. 4.5 µl of the undiluted , 0.1 and 0.01 dilutions of the mated suspensions were plated on DDO/X and QDO/X/A medium.

Chapter 3: Results

Hypothesis:

IcsB is targeting one or several host proteins to facilitate the vacuole escape of *Shigella*.

Objectives :

•Create IcsB mutants to solve the yeast cytotoxicity issue

•Screen for protein-protein interaction with the yeast two-hybrid system

•Analyze and characterize potential prey proteins

3.1 IcsB Mutagenesis

This project focused on assessing the yeast cytotoxicity of different IcsB mutants via *S. cerevisiae* yeast transformations, in light of the fact that the wild type IcsB was characterized as a highly toxic protein. The first objective of this project was to resolve the toxicity issue attributed to wildtype IcsB. The most ideal IcsB mutant with the greatest attenuation in toxicity may be used in order to identify the protein-protein interaction through yeast two-hybrid screening.

Using classical molecular biology methods, including PCR site-directed mutagenesis and ligation protocols, three different IcsB mutants were initially constructed – *pGBKT7 IcsB C306A, pGBKT7-IcsB Δ80,* and *pGBKT7 IcsB C306A Δ80* (Figure A3). The C306A mutation was believed to be involved in the putative catalytic site of IcsB protein, while the Δ80 mutation was proposed in playing a role in the facilitation of chaperon binding and secretion. Finally, *pGBKT7 IcsB C306A Δ80* combined both of the above mutations. This double mutant was generated in two steps by using the single mutant C306A as a template. Ultimately, these mutants were carefully crafted to decrease IcsB's cytotoxicity in yeast. Upon the successful transformation of competent *E. coli* MC1061 cells with the DNA resulting from IcsB mutagenesis, GeneJET Plasmid Miniprep Kit was used for plasmid purification of a few colonies that were selected from each mutant. The purified plasmid DNA was subsequently sequenced to confirm the presence of the desired mutation.

3.1.1 Sequencing Results of Three IcsB Mutants

The sequence of *pGBKT7 IcsB C306A* and *pGBKT7 IcsB Δ80* was obtained using Sanger sequencing, also known as the chain termination method. Standard nucleotide BLAST from the NCBI database was used to analyze the results of the sequencing. The blind BLAST determined that all three IcsB mutants shared 100% sequence identity with IcsB (Accession number: WP_01092 1665) except, of course, at the mutated site in the case of C306A; we also confirmed the deletion of the sequence coding for the 80 amino-terminal residues. Table 3.1 indicates that the wild-type IcsB plasmid correctly showcased the site-specific mutations of the individual mutants, which was revealed by the PAIRWISE BLAST. That means we used the theoretical sequence as the subject for this blast instead of searching a database. These tests show that all three IcsB mutagenesis procedures described in section 3.1 were successful.

Sanger DNA sequencing was applied to *pGBKT7 IcsB C306A Δ80*. Once again, standard nucleotide BLAST from the NCBI database was used to analyze the sequencing results. Serial Cloner detected the mutagenesis primers IcsB_C306A_S and IcsB_C305_R within the sequences; however, the mutagenesis primers D_80N-ter_IcsB_S and D_80N-ter_IcsB_R were not found within the sequences, for they were further up the sequence and sequencing reactions performed with the *SeqO7* sequencing primer did not allow to confirm the sequence at the 3' end of the gene (N-terminus of the resulting protein). This was not problematic as this part of the plasmid was sequenced initially when the N-terminal truncation mutant was generated (see above). According to the blind BLAST results, *pGBKT7 IcsB C306A Δ80* shared 100% sequence identity with IcsB (accession: WP_01092 1665).

Table 3.1: Summary of Sanger sequencing analysis results of three IcsB mutants.
Sequencing results were analyzed via standard nucleotide BLAST from NCBI database. Mutagenesis primers were found within the sequences using Serial Cloner.

IcsB Mutants	Sequencing Primer	Mutagenesis Primers*	Mutation	Sequence identity (BLAST)
pGBKT7-IcsB-C306A	SeqO7	IcsB_C306A_S IcsB_C305_R	Cysteine replaced by alanine at residue 306	99% identity (260/261 amino acid)
pGBKT7-IscB-D80	SeqO8	D_80N-ter_IcsB_S D_80N-ter_IcsB_R	80 amino acids deleted from the N-terminus	100% identity (288/288 amino acid)
pGBKT7-IcsB-C306A-delta80NT	SeqO7	IcsB_C306A_S IcsB_C305_R	80 amino acids deleted from the N-terminal and cysteine replaced by alanine at residue 306	99% identity (221/222 amino acid)

*See Table A1 for a detailed description of the primers.

Table 3.1 illustrates that the wild-type IcsB plasmid correctly showcased the –C306A mutation, as revealed by the PAIRWISE BLAST. Taken together, the *pGBKT7-IcsB-C306A- D80* clone was designed successfully.

3.1.2 Transformation and toxicity phenotype in the yeast strain Y2HG

In order to assess the cytotoxicity in yeast, plasmids derived from pGBKT7 that harbored the IcsB mutants described above were transformed into the Yeast-Two-Hybrid Gold (Y2HG) *Saccharomyces cerevisae* cells (Figure 3.1). The empty vector *pGBKT7* and the *pGBKT7 IcsB* (WT) were also transformed and acted as controls. The positive control *pGBKT7* yielded regular sized pink colonies because no toxic protein was expressed. With regards to the negative control *pGBKT7 IcsB* , no colonies were observed, which was expected given its cytotoxicity in yeast (Slagowski. et al.2008). Both *pGBKT7 IscB Δ80* and *pGBKT7 IcsB C306A* resulted in relatively equal numbers of pink colonies (Refer to Figure 3.1). However, *pGBKT7 IscB Δ80* yield larger colonies than *pGBKT7 IcsB C306A*, suggesting that the rescue of the phenotype was less complete in the latter. *pGBKT7 IcsB C306A Δ80* formed colonies of similar size than *pGBKT7 IscB Δ80* (data not shown). It is noteworthy that the C306A mutation should disrupt the putative proteolytic or acetylation catalytic activity predicted by the sequence alignment (Pei and Grishin, 2009). C306 is in a functionally conserved region of IcsB that we have simply called domain 1 (*DM*1). The fact that C306A rescues the cytotoxicity phenotype indicates that the catalytic activity is implicated in the cytotoxicity phenotype. On the other hand, the rescue of the phenotype observed upon the deletion of the 80 N-terminal amino acids indicates that this domain implicated in secretion and chaperon binding is also imparting toxicity to IcsB. Taken together these data, suggested that these mutants were able to rescue the cytotoxicity phenotype albeit to various levels. We hypothesized that the truncation of the 80 N-terminal amino acids will not have any impact on the effectors

function in yeast since secretion is not necessary in this context. Also, the toxicity was not fully rescued by IcsB C306A, so we were not comfortable to perform the Y2H screen with this mutant.

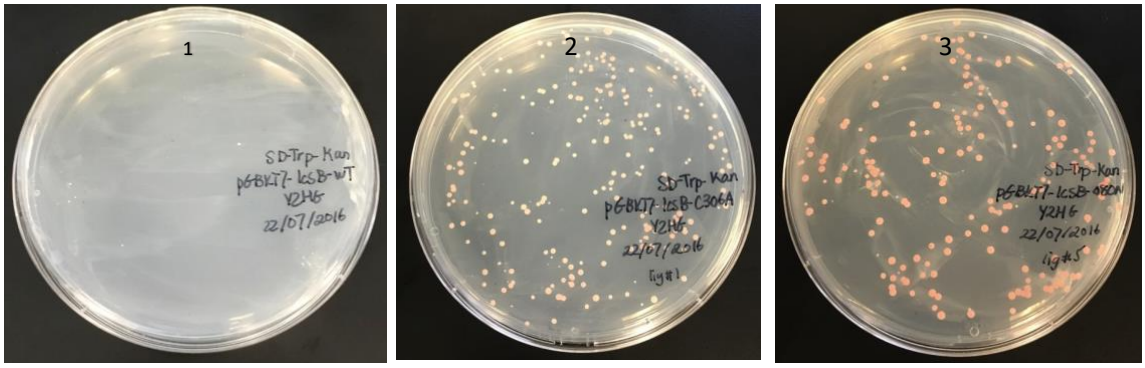

Figure 3.1: Transformation of pGBKT7 IcsB mutants to test for IcsB yeast cytotoxicity.
This test was done in the Y2HG strain on SD-Trp. 1: *pGBKT7 IscB*; 2: *pGBKT7 IcsB C306A*; 3: *pGBKT7 IscB Δ80*.

3.1.3 Western Blot

We next wondered whether the expression level of these mutants by Y2HG cells were similar. To do this, protein extraction was performed on the relevant Y2HG transformants described in the previous section. The samples were then used to perform a western blot. Figure 3.2 illustrates that the western blot of *pGBTK7 IcsB C306A* and *pGBTK7 IcsB Δ80* were very informative, despite the presence of background signals. pGBKT7 allowed the expression of the Gal4 DBD fused with the Myc-tag (DBD-Myc). All *IcsB* harboring plasmids allowed the expression of the DBD linked to IcsB variants through a Myc tag and other plasmids derived sequences that were constant across all constructs. DBD-Myc should have a molecular weight

(MW) of approximately 21.7 kDa in size. DBD-Myc-IcsB Δ80 should have a MW of 68.2 kDa, while DBD-Myc-IcsB C306A (well 3) should have a MW of 77 kDa.

The WB revealed that the expression of DBD-Myc-IcsB C306A was undetectable, while both DBD-Myc-IscB Δ80 and DBD-Myc-IscB Δ80 C306A were highly expressed to similar levels (Figure 3.2 and 3.3).

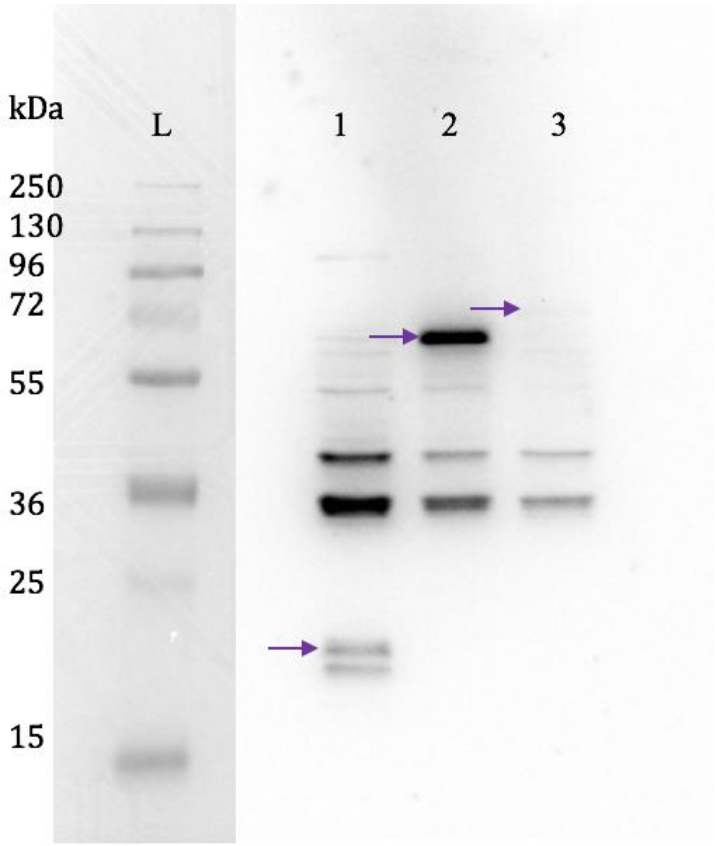

Figure 3.2: Assessment of the expression of IcsB C306A *and* IcsB Δ80 by Western Blot.
Lysates of colonies obtained from yeast transformations described in section 4.1.2 were used to assess the expression of IcsB protein fusion by Western blot. Western blot of lysate of cells transformed with the empty vector pGBTK7 (DBD-Myc: well 1), *pGBKT7 IcsB Δ80* (DBD-Myc-IscB Δ80: well 2), and *pGBKT7 IcsB C306A* (DBD-Myc-IscB C306A: well 3) was treated with 1/2000 of Myc-tag (primary antibody) and 1/20,000 anti-mouse IgG (secondary antibody). 3µL of protein molecular weight ladder was loaded in lane L. 5µL of each sample were loaded in numbered wells. The gels were visualized with ECL substrate using a colorimetric imaging system. Signal was accumulated for 30 seconds with the regular chemiluminescence sensitivity settings.

The location of the observed bands accurately matched the predicted MW of the proteins. It was previously expected that both IcsB mutants would successfully translate into proteins, given the decrease in yeast cytotoxicity; however, the western blot revealed that DBD-Myc-IcsB Δ80 (well 2) was better expressed in comparison to DBD-Myc-IcsB C306A (well 3), which was undetected. This correlated with the smaller colonies formed by cells expressing the latter. It was thus concluded that *pGBTK7 IcsB Δ80* was the most ideal because it possesses the greatest reduction in cytotoxicity and higher expression level. We then decided to compare the expression level of DBD-Myc-IcsB Δ80 to DBD-Myc-IcsB Δ80 C306A (Figure 3.3). The latter provided it behaves correctly might be important for future screening purposes because it should bind putative substrate but would not be able to transform them into a product, thereby stabilizing their interaction with IcsB. Both DBD-Myc-IcsB Δ80 and DBD-Myc-IcsB Δ80 C306A have a predicted MW of 68.2 kDa, thus migrated right below the 70 kDa band. The migration of the bands accurately corresponded to the weight of the proteins. DBD-Myc-IcsB Δ80 C306A (well 3) demonstrated a comparable level of protein expression than DBD-Myc-IcsB Δ80 (well 2). It was thus concluded that DBD-Myc-IcsB Δ80 C306A was well expressed in yeast, and furthermore, may constitute a good candidate for Yeast-Two Hybrid screening because it combines a strong expression with complete rescue of the cytotoxicity, and no or low catalytic activity.

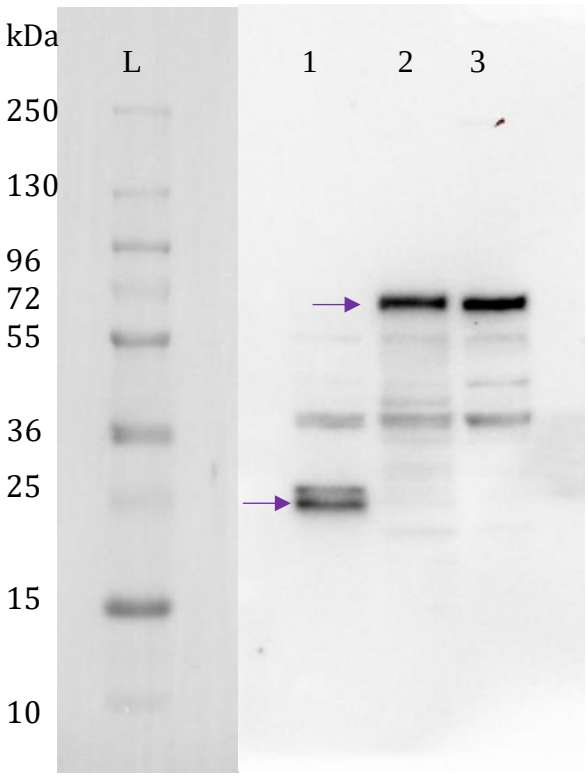

Figure 3.3: Assessment of the expression of IcsB Δ80 and IcsB Δ80 C306A by Western Blot.
Lysates of colonies obtained from yeast transformations described in section 4.1.2 were used to assess the expression of IcsB protein fusion by Western blot. Western blot of cell lysates transformed with the empty vector pGBTK7 (DBD-Myc: well 1), *pGBKT7-IcsB Δ80* (DBD-Myc-IscB Δ80: well 2), and *pGBKT7-IcsB Δ80 C306A* (DBD-Myc-IscB Δ80 C306A: well 3) was treated with 1/2000 of Myc-tag (primary antibody) and 1/20,000 anti-mouse IgG (secondary antibody). 3µL of protein molecular weight ladder was loaded in lane L. 5µL of each samples was loaded in the numbered well. The gels were visualized with ECL substrate using a colorimetric imaging system. Signal was accumulated for 30 seconds with the regular chemiluminescence sensitivity settings.

3.2 Yeast-Two Hybrid

3.2.1 Matchmaker Gold Yeast-Two Hybrid Mating and Screening

The yeast two-hybrid system was used to identity novel protein interaction between the bait protein IcsB C306A Δ80 (*pGBKT7 IcsB C306A Δ80*) and the library of prey proteins cloned into pGADT7. Bait protein and prey proteins are cloned as genetic fusion with the Gal4 DNA-binding domain (DBD) and Gal4 activation domain (AD), respectively. The DBD-Myc-IcsB and a given AD-Prey protein fusion are expressed in each diploid cell obtained through mating. If IcsB interacts with the prey protein, the AD is recruited to DBD binding sites in the promoters of four independent reporter genes (*AUR1-C*, *ADE2*, *HIS3*, and *MEL1*), thereby activating their transcription. This allows growth of those diploid cells in the presence of the antibiotic aerobasidin A, and in the absence of adenine and histidine and render the resulting colony blue when the medium contains X-α-Gal. For this first screen, we chose to use a universal human library that included cDNA obtained from a diverse collection of human tissues. This library was also normalized to preferentially remove abundant cDNAs derived from high-copy-number mRNAs. That reduces the number of clones that must be screened to identify positive interactions, and facilitates the identification and characterization of novel protein-protein interactions.

Prior to screening of the library, a control mating was done in order to familiarize myself with the mating procedure. The mating strains consisted of the positive control bait plasmid *pGBKT7-53* and the positive control prey plasmid *pGADT7-T* harbored by the Y2H gold and Y187 cells, respectively. Based on the viability of the prey of 1.2×10^6 CFU/mL and the viability of the diploids of 4.7×10^5 CFU/mL, the mating efficiency was calculated to be 39%.

When we performed the mating between the universal library harbored by Y187 cells and the bait plasmid *pGBKT7 IcsB C306A Δ80* harbored by the Y2HG cells, we obtained a 17.89% mating efficiency (Refer to Table 3.2). This percentage of diploid comfortably exceeded the minimal recommended efficiency of 2-5%. From the 55 low stringency DDO/X/A plates, only 6 out of the 13 blue colonies that were re-patched on DDO/X/A grew. Among the six clones, four clones yielded blue colonies while two clones yielded pink colonies. These observations suggested that these four clones expressed human proteins that might bind to IcsB. Segregation of the library plasmid in yeast was done for these positive clones to increase the chance of rescuing the prey plasmid responsible for activating the reporter genes via *E. coli* transformation. This was done by streaking out the four positive clones that grew on DDO/X/A onto DDO/X and on higher stringency QDO/X/A agar plates three successive times, each time isolating only one blue colony. This step was highly critical, as during the library preparing phase, the yeast cells that have been transformed have the tendency to uptake numerous plasmids. Thus, this could result in yeast cells containing several plasmids, out of which one may be able to activate the reporter genes while other(s) may be unable to express a protein or may be expressing a protein unable to interact with the bait. Because of the natural tendency that plasmids have to segregate during yeast division, streaking clones multiple times ensured that a blue colony isolated after the third passage should contain a single prey plasmid generating a positive signal. The plasmid DNA transformed into *E. coli* selected on ampicillin plate was then purified and submitted to Sanger sequencing. Sequence analysis identified the gene present in the four clones as RNF2 (data not shown). However, along with every two-hybrid screen comes the chance of detecting false positives; it is thus important to confirm that interactions are genuine. One needs to ensure that both bait and prey are required to activate the Gal4-responsive reporters to conclude that the interaction is genuinely

positive. By contrast, in the case of false-positive interactions, the prey can activate the Gal4-responsive reporters in the absence of the bait. To perform a key experiment that addressed this, two colonies from the prey RNF2 were mixed either with the empty vector *pGBKT7* or with *pGBKT7 IcsB C306A Δ80*. Plated on DDO/X/A and on QDO/X/A. After 5 days at 30°C similar numbers of blue colonies were observed on all plates. This suggested that RNF2 was prone to generate false positive because it did not require the bait to activate the reporters (Figure 3.4).

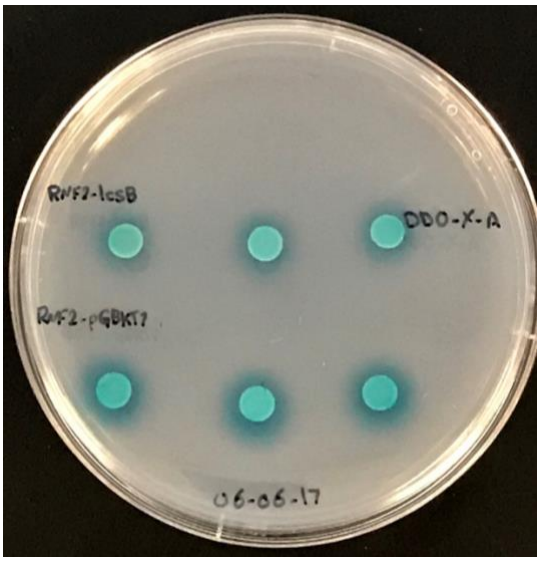

Figure 3.4: Dot Assay confirmed false positive interactions between IcsB and RNF2.
RNF2 that was identified in the yeast two-hybrid screen #1 allows growth on the Aerobasidin A plate and generates positive X-α-Gal signal (blue color) with pGBKT7 (DBD-Myc), hence IcsB is not necessary for activation of the reporter genes.

Table 3.2: IcsB yeast two-hybrid Screening.
Data obtained with yeast two-hybrid screen of bait *pGBKT7 IcsB Δ80 C306A* and *pGBKT7-IcsB Δ80* against a universal human library.

	pGBKT7 IcsB Δ80 C306A	***pGBKT7 IcsB Δ80***
Number of Clones screened (diploids)	$6.24x10^7$ diploid cells	$3\ x10^7$ diploid cells
Experimental Library titer	$2.64x10^8$ CFU/mL	$1.32x10^9$ CFU/mL
Theoretical Library clones	$> 5x10^7$ CFU/mL	$> 5x10^7$ CFU/mL
Viability of the Prey Library	$2.2x10^7$ CFU/mL	$1.56x10^7$ CFU/mL
Viability of diploids	$3.9x10^6$ CFU/mL	$2.4x10^6$ CFU/mL
Mating efficiency	17.89%	15%
Number of diploid clones on 55 DDO/X/A (150 mm)	13	33

Hence, the next experiments focused on using as bait IcsB C306A Δ80 and independent domains 1 and 2 of IcsB to screen a new epithelial cells cDNA library. This normalized HeLa S3 cDNA library transformed into the yeast strain Y187 was mated to the bait Y2HGold reporter strains as described above.

Data about all those three screens are shown in Table 3.3. In the interest of time, I was able to analyze only the screen with bait IcsB C306A Δ80. Therefore, only the latter is described in detail below. Based on the viability of the prey of $5x10^7$ CFU/mL and the viability of the diploids

of 7.6 x10^6 CFU/mL, the mating efficiency of the screen with *pGBKT7 IcsB C306A Δ80* was calculated to be 15.2%. This percentage of diploid comfortably exceeded the minimal recommended efficiency of 2-5%. According to the manufacturer of the cDNA library and Y2H kit, it is imperative that at least 1 million diploids are screened. Otherwise, there would be less chance of detecting genuine interaction. Based on the suspension volume and the number of colonies detected on DDO plates, the total number of diploid clones screened was determined to be 9.5 x10^7 diploid cells. The theoretical number of independent clones in the library was suggested to be >5x10^7 CFU/mL. Theoretically speaking, the library was therefore barely saturated. Given that the mating was randomized and that the number of diploid clone screened was only roughly 1.2 times greater than the library titer, it is probable that some low-abundance genes were not screened. From the 55 low stringency DDO/X/A plates obtained to screen the library, we isolated 4 blue colonies. They all grew back when patched on DDO/X/A plates, but only 3 clones yielded blue colonies (Table 3.3). The prey plasmids of these three clones were rescued and activated the reporter genes upon re-transformation and mating with the bait plasmid but not with the control *pGBKT7* (Table 3.4).

Table 3.3: IcsB Yeast Two-Hybrid Screening.

Data obtained with Yeast Two- Hybrid screen of bait *pGBKT7 IcsB C306A Δ80, pGBKT7-IcsB domain 1 (DM1) C306A*, and *pGBKT7-IcsB domain 2 (DM2)* against the Hela S3 cell cDNA library.

	pGBKT7 IcsB Δ80 C306A	***pGBKT7 IcsB Δ80***
Number of Clones screened (diploids)	6.24×10^7 diploid cells	3×10^7 diploid cells
Experimental Library titer	2.64×10^8 CFU/mL	1.32×10^9 CFU/mL
Theoretical Library clones	$> 5 \times 10^7$ CFU/mL	$> 5 \times 10^7$ CFU/mL
Viability of the Prey Library	2.2×10^7 CFU/mL	1.56×10^7 CFU/mL
Viability of diploids	3.9×10^6 CFU/mL	2.4×10^6 CFU/mL
Mating efficiency	17.89%	15%
Number of diploid clones on 55 DDO/X/A (150 mm)	13	33

Table 3.4: High stringency screening of potential interactors and retransformation of mated diploids. Four diploid clones isolated from the yeast two-hybrid screen #2 were re-patched onto low stringency DDO/X/A plates. The four clones were patched onto high stringency QDO/X/A plates. The prey protein of the three clones that grew on the QDO/X/A plates were re-mated with IcsB bait protein. Mated cultures were re-transformed in yeast.

Plates	Number of Clones Regrowth	Color of Clones			
		Clone 1	Clone 2	Clone 3	Clone 4
Patch on DDO/X/A	4	Blue	Blue	Blue	Blue
Patch on QDO/X/A	3	Blue	Blue	Blue	Pink
Patch on DDO/X/A of mating performed with the re-transformed plasmid	3	Blue	Blue	Blue	-----

3.2.2 Sequencing and Dot Assay

The extracted prey DNA plasmids were then digested using the restriction enzymes EcoRI and XhoI to determine whether the three prey plasmids shared the same insert (gene)or not. Agarose gel electrophoresis indicated that the three clones contained insert of different sizes, suggesting that the gene candidates in the three prey plasmids were unique. These plasmids were sent for sequencing to validate these preliminary observations. Indeed, sequencing results supported this assertion as the clone 1 contained a fragment of the Homo sapiens DDX3X, the clone 2 contained a fragment of the Homo sapiens FANCL, and the clone 3 contained a fragment of the Homo sapiens SGT1 (Table 3.5). I will describe below which fragment of these genes were covered in the sequence contained in these prey plasmids.

As described above, it is important to confirm that putative interactions with IcsB identified by the Y2H are not false positive. To do so, Y2HG cells harboring *pGBKT7* and *pGBKT7 IcsB Δ80 C306A* were re-mated with the three prey plasmids encoding the DDX3X, FANCL and SGT1 gene. To map the interaction of the prey proteins into IcsB primary structure, Y2HG harboring *pGBKT7 IcsB Δ80* and *pGBKT7 IcsB DM1 C306A* and *pGBKT7 IcsB DM2* were also mated with strains containing each of the prey plasmids.

The full-length DDX3X protein is composed of 662 amino acids (Figure 3.5), which corresponds to approximately 1986 bp. The sequence obtained was incomplete at the 5' end of the gene, but also at the 3' of the gene, which is less common. We tried to obtain the complete sequence using several primers but we always obtained a sequence that was of poor quality at the 3' end of the gene. The reason for this phenomenon is unclear, but is possibly due to the extremely low sequence diversity of the 3' end of the gene, which contains 29.5% glycine, 27.9% serine and 8.2% tyrosine codons. The DDX3X that we could confirm by sequencing corresponded to residues 381 to 601 (221 amino acids). This corresponded roughly to the helicase domain of DDX3X. The dot assay results for DDX3X indicated activation of reporter genes with *pGBKT7 IcsB Δ80 C306A*, but not with *pGBKT7*. This confirmed that this protein-protein interaction was genuine and not false positive. The expression of MEL1 by cell harboring the DDX3X prey plasmid as judge from the blue color on the DDO/X plate was similar for *pGBKT7 IcsB Δ80 C306A* and *pGBKT7 IcsB Δ80* diploid cells. It was lower for *pGBKT7-IcsB DM2*, but the latter appeared to be stronger than for *pGBKT7-IcsB DM1 C306A*. This conclusion is reinforced by the observation that cells

harboring *pGBKT7 IcsB DM1 C306A* did not grow on QDO/X/A plate while *pGBKT7 IcsB DM2* did.

The full-length SGT1 protein (isoform A) is composed of 333 amino acids (Figure 3.6), which corresponds to 999 bp. The sequence encoded by the prey plasmid corresponded to the 154 amino acids in the N-terminus. Downstream of that region, there is a frame shift due to the absence of two nucleotides. The quality of the sequencing in that region was excellent upon inspection of the chromatogram, suggesting that these two nucleotides are truly absent from the sequence contained in the plasmid. Downstream of the missing nucleotides, we noticed that the sequence corresponds to the rest of the coding sequence of SGT1. It is therefore likely that only the first 154 amino acids of the protein are correctly expressed and are sufficient to binding IcsB. This region corresponded to the 3 tetraticopeptide repeats and about 40% of the CHORD-containing proteins and SGT1 (CS) domain. The dot assay results for SGT1 indicated activation of reporter genes with *pGBKT7 IcsB Δ80 C306A*, but not with *pGBKT7*. This confirmed that this protein-protein interaction was genuine and not false positive. SGT1 appeared to bind IcsB similarly to DDX3X. The expression of MEL1 by cell harboring the SGT1 prey plasmid as judge from the blue color on DDO/X plate was similar for *pGBKT7 IcsB D80* and *pGBKT7 IcsB Δ80* diploid cells. The signal was lower for *pGBKT7 IcsB DM2*, but the latter appeared to be stronger than that for *pGBKT7 IcsB DM1 C306A*. This conclusion is reinforced by the observation that cells harboring *pGBKT7-IcsB DM1 C306A* did not grow on QDO/X/A plate while those with *pGBKT7-IcsB DM2* did.

The FANCL protein is composed of 375 amino acids (Figure 3.7), which corresponds to approximately 1125 bp. The sequence of FANCL contained in the isolated prey plasmid was almost complete as it encompassed residues 24-375 (351 of 375 residues). Therefore, the sequence

harbored by this prey plasmid contained the full-length RING domain. The dot assay results for FANCL indicated activation of reporter genes with *pGBKT7 IcsB C306A Δ80*, but not with *pGBKT7*. This confirmed that this protein-protein interaction was genuine and not false positive. FANCL appeared to bind IcsB differently than DDX3X and SGT1. The expression of MEL1 by cell harboring the FANCL prey plasmid as judge from the blue color on DDO/X plate was much higher for *pGBKT7 IcsB C306A Δ80* than for *pGBKT7 IcsB Δ80* diploid cells. This was confirmed by the poor growth of *pGBKT7 IcsB Δ80* compared to *pGBKT7 IcsB C306A Δ80* on the QDO/X/A plate. In addition, *pGBKT7 IcsB DM1 C306A.* Furthermore, the signal was higher for *pGBKT7-IcsB DM1 C306A*, while it was very low for *pGBKT7 IcsB DM2*. This conclusion is reinforced by the observation that cells harboring *pGBKT7 IcsB DM2* did not grow on QDO/X/A plate while those with *pGBKT7-IcsB DM1 C306A* did.

Table 3.5: Sequencing results of the 3 prey plasmids isolated from the Y2H screening of the HeLa S3 cDNA library with *pGBKT7 IcsB Δ80 C306A*.

Sequencing results were analyzed via standard nucleotide BLAST from NCBI database

pGADT7-prey gene	**Blind Blast**		
	Identity	**Gene/Protein**	**Uniprot Accession Number**
Clone 1	99%	Homo sapiens DEAD-box helicase 3 X-linked (DDX3X)	O00571 (isoform 1)
Clone 2	99%	E3 ubiquitin-protein ligase FANCL	Q9NW38 (isoform 1)
Clone 3	99%	Protein SGT1 homolog	Q9Y2Z0 (isoform A or 2)

A

MSHVAVENAL GLDQQFAGLD LNSSDNQSGG STASKGRYIP PHLRNREATK
GFYDKDSSGW SSSKDKDAYS SFGSRSDSRG KSSFFSDRGS GSRGRFDDRG
RSDYDGIGSR GDRSGFGKFE RGGNSRWCDK SDEDDWSKPL PPSERLEQEL
FSGGNTGINF EKYDDIPVEA TGNNCPPHIE SFSDVEMGEI IMGNIELTRY
TRPTPVQKHA IPIIKEKRDL MACAQTGSGK TAAFLLPILS QIYSDGPGEA
LRAMKENGRY GRRKQYPISL VLAPTRELAV QIYEEARKFS YRSRVRPCVV
YGGADIGQQI RDLERGCHLL VATPGRLVDM MERGKIGLDF CKYLVLDEAD
RMLDMGFEPQ IRRIVEQDTM PPKGVRHTMM FSATFPKEIQ MLARDFLDEY
IFLAVGRVGS TSENITQKVV WVEESDKRSF LLDLLNATGK DSLTLVFVET
KKGADSLEDF LYHEGYACTS IHGDRSQRDR EEALHQFRSG KSPILVATAV
AARGLDISNV KHVINFDLPS DIEEYVHRIG RTGRVGNLGL ATSFFNERNI
NITKDLLDLL VEAKQEVPSW LENMAYEHHY KGSSRGRSKS SRFSGGFGAR
DYRQSSGASS SSFSSSRASS SRSGGGGHGS SRGFGGGGYG GFYNSDGYGG
NYNSQGVDWW GN

B

Feature Key	Positions	Description	Length
Domain	211 - 403	Helicase ATP-binding	193
Domain	414 -575	Helicase C-termina	162

C

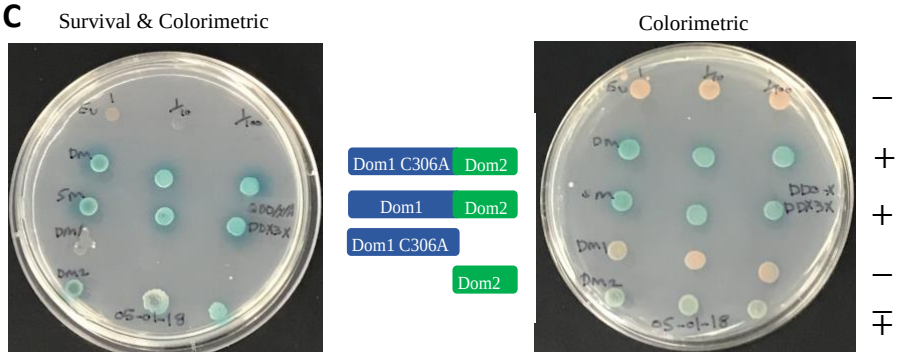

Figure 3.5. Prey DDX3X.

(A) This is DDX3X primary structure. The underlined sequence represents the coding sequence present in the prey plasmid that was isolated. Colored segments highlight the various domains listed in the next panel. (B) List of known domains found in DDX3X. (C) Dot assay of reporter genes activation test of DDX3X with various IcsB constructs. A Y187 strain harboring *pGADT7-DDX3X* was mated with Y2HG strains harboring either: the empty vector control *pGBKT7* (EV), *pGBKT7 IcsB Δ80 C306A* (DM), *pGBKT7 IcsB Δ80* (SM), *pGBKT7-IcsB DM1 C306A*, and *pGBKT7-IcsB DM2*. Not diluted, 1/10 and 1/100 dilution of each diploid strain culture were patched into the plates.

A

MAAAAAGTAT SQRFFQSFSD ALIDEDPQAA LEELTKALEQ KPDDAQYYCQ
RAYCHILLGN YCVAVADAKK SLELNPNNST AMLRKGICEY HEKNYAAALE
TFTEGQKLDS ADANFSVWIK RCQEAQNGSE SEVWTHQSKI KYDWYQTESQ
VVITLMIKNV QKNDVNVEFS EKELSALVKL PSGEDYNLKL ELLHPIIPEQ
STFKVLSTKI EIKLKKPEAV RWEKLEGQGD VPTPKQFVAD VKNLYPSSSP
YTRNWDKLVG EIKEEEKNEK LEGDAALNRL FQQIYSDGSD EVKRAMNKSF
MESGGTVLST NWSDVGKRKV EINPPDDMEW KKY

B

Feature Key	Positions	Description	Length
Repeat	11 - 44	TPR 1	34
Repeat	45 - 78	TPR 2	34
Repeat	79 - 112	TPR 3	34
Domain	169 - 258	CS	90
Domain	276 - 365	SGS	90

C

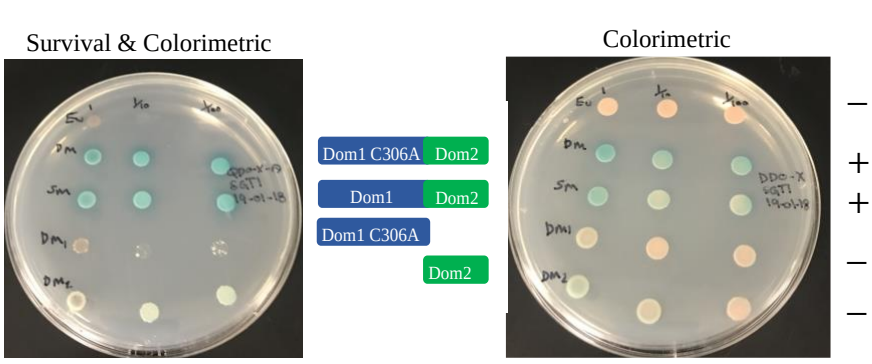

Figure 36: Prey SGT1.

(A) This is SGT1 (isoform A) primary structure. The underlined sequence represents the coding sequence present in the prey plasmid that was isolated. Colored segments highlight the various domains listed in the next panel. (B) List of known domains found in SGT1 (isoform A). (C) Dot assay of reporter genes activation test of FANCL with various IcsB constructs. An Y187 strain harboring *pGADT7-SGT1* (isoform A) was mated with Y2HG strains harboring either: the empty vector control *pGBKT7* (EV), *pGBKT7 IcsB Δ80 C306A* (DM), *pGBKT7 IcsB Δ80* (SM), *pGBKT7 IcsB DM1 C306A*, and *pGBKT7 IcsB DM2*. Not diluted, 1/10 and 1/100 dilution of each diploid strain culture were patched into the plates.

A

MAVTEASLLR QCPLLLPQNR SKTVYEGFIS AQGRDFHLRI VLPEDLQLKN
ARLLCSWQLR TILSGYHRIV QQRMQHSPDL MSFMMELKML LEVALKNRQE
LYALPPPPQF YSSLIEEIGT LGWDKLVYAD TCFSTIKLKA EDASGREHLI
TLKLKAKYPA ESPDYFVDFP VPFCASWTPQ SSLISIYSQF LAAIESLKAF
WDVMDEIDEK TWVLEPEKPP RSATARRIAL GNNVSINIEV DPRHPTMLPE
CFFLGADHVV KPLGIKLSRN IHLWDPENSV LQNLKDVLEI DFPARAILEK
SDFTMDCGIC YAYQLDGTIP DQVCDNSQCG QPFHQICLYE WLRGLLTSRQ
SFNIIFGECP YCSKPITLKM SGRKH

B

Feature Key	Positions	Description	Length
Region	104 - 294	UBC-RWD region (URD)	191
Zinc finger	307 - 365	RING-type	59

C

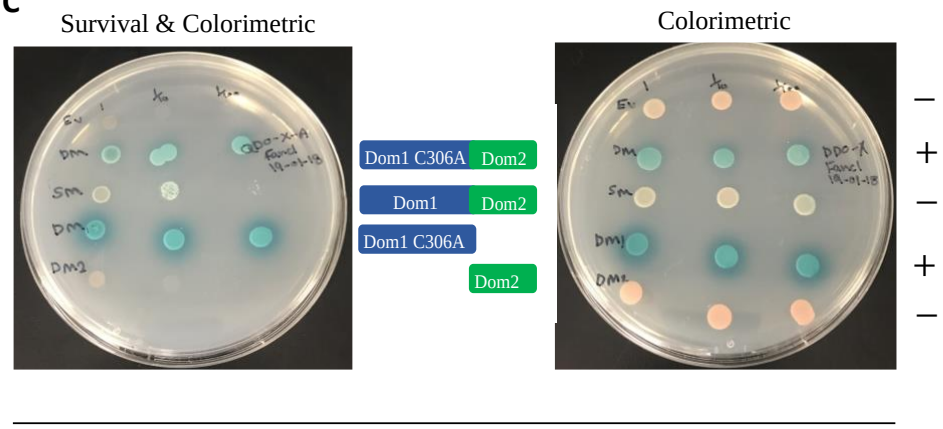

Figure 3.7: Prey FANCL.

(A) This is FANCL primary structure. The underlined sequence represents the coding sequence present in the prey plasmid that was isolated. Colored segments identify the various domains listed in the next panel. (B) List of known domains found in FANCL. (C) Dot assay of reporter genes activation test of FANCL with various IcsB constructs. A Y187 strain harboring *pGADT7-FANCL* was mated with Y2HG strains harboring either: the empty vector control *pGBKT7* (EV), *pGBKT7 IcsB C306A Δ80* (*DM*), *pGBKT7 IcsB Δ80*(*SM*), *pGBKT7 IcsB DM1 C306A*, and *pGBKT7 IcsB DM2*. Not diluted, 1/10 and 1/100 dilution of each diploid strain culture were patched into the plates.

Chapter 4: Discussion of Results

IcsB belongs to a unique class of virulence proteins that is also encountered in other pathogenic or symbiontic bacteria (Waad Bajunaid, personal communication). IcsB is highly critical for the vacuole escape that is triggered by the T3SS during cell-to-cell spread. Despite the extensive research that has already been conducted on IcsB, the specific molecular interactions it may form during the vacuole escape are unknown. My work has identified specific host proteins that might be binding IcsB during the infection. Below, I discuss my findings.

The first finding of my work was the identification of two mutations in IcsB that relieved its cytotoxicity in yeast. The first mutation C306A targeted a predicted catalytic residue in IcsB (Pei & Grishin, 2009). This result supports the notion that C306 is truly a catalytic residue and that this putative catalytic activity is responsible for the yeast cytotoxicity phenotype. This catalytic activity may therefore play a role in the function of IcsB during infection. Using *pGBKT7 IcsB Δ80 C306A* for the screening was important because if IcsB binds to its substrate when it is catalytically inactive, it will not be able to transform it to a product and releasing it. Therefore, its substrate may be found by using this mutant. The second mutation that rescued the cytotoxicity phenotype corresponded to the deletion of the first 80 amino-terminal residues of IcsB, which comprised the T3SS signal peptide and the chaperon-binding domain. The 80 amino acids in the amino-terminus seem to have the capacity to kill the yeast independently of the catalytic activity. If that was not the case, the mutation of the catalytic residue would have been enough to fully rescue the toxicity. In other words, the catalytically residue is toxic and the 80 amino-terminal residues are toxic as well. The reason for the rescue of toxicity of Δ80 is unclear.

I can speculate that this domain might be toxic on its own because it modifies the localization of IcsB. An alternative may be that in the absence of the chaperone the N-terminus leads to cellular stress because of the aggregation of IcsB. Indeed, chaperon binding domain adopts a random coil structure that may make IcsB prone to aggregation. During these experiments, we have expressed IcsB without its chaperone. It might thus be a good idea to express IcsB variants with the chaperon IpgA and assess if the toxicity of IcsB C306A is decreased and its expression level is increased to the same level as IcsB Δ80 C306A.

Interestingly, although RNF2 was initially thought to interact with IcsB based on the yeast two-hybrid screen using IcsB C306A Δ80 as bait, further analysis demonstrated it was a false positive. Following this observation, identification of other genes that showed promising interactions with IcsB specifically DDX3X, FANCL and SGT1 were found in a second yeast two-hybrid screen also using IcsB C306A Δ80 as bait. In contrast to RNF2 those three genes passed the false-positive test. The library contained at least 50 million clones (Refer to Table 3.3). Given that there are an estimated 20,000 human protein-coding genes, each gene should be present multiple times among the library clones. The library was normalized to selectively remove highly abundant transcripts, thereby enhancing the representations of low-abundance and rare cDNAs. However, despite the normalization of the libraries, there was only one colony for each of DDX3X, FANCL and SGT1 that we identified in our screen. Therefore, it can be assumed that DDX3X, FANCL and SGT1 are probably present in low copies in the library. We observed that FANCL bound catalytically inactive IcsB but not WT IcsB. Furthermore, FANCL bound the catalytic domain 1, but not domain 2. These observations suggested that FANCL could be a

substrate of the putative acyl transferase or protease activity of IcsB. In principle, this could be verified by SDS-page because both of these post-translational modifications should alter the electrophoretic mobility of FANCL. By contrast DDX3X and SGT1 binding to IcsB was independent of the catalytic activity as it bound equally well to IcsB Δ80 C306A and IcsB Δ80; its binding appeared weaker with DM2 than with IcsB Δ80, while it did not seem to bind to the catalytic DM1 at all. Taken together, interaction of DDX3X and SGT1 with IcsB seems to depend mostly on *DM2* or in the region of *DM2* at the interface with the *DM1*. Below, I discuss these three genes and their potential connections to my initial hypothesis stating that IcsB target human proteins to facilitate vacuole escape.

In one study, SGT1 was found to strongly interact with the Skp1-Cullin-F-box (SCF) ubiquitin ligase complex, which main function is to carry protein degradation particularly but not limited to those involved in the cell cycle (Kitagawa, Skowyra, Elledge, Harper, & Hieter, 1999). It was shown that SGT1 directly interacts with Skp1-4 and as the result of that causes defects in yeast kinetochore function. Another important key finding of this study is that SGT1 was shown to also bind with the SCF protein domain. Taken together these results suggest that SGT1 is directly involved in multiple protein domain interactions. A significant step forward was made when the crystal structure of SGT1-Skp1 interaction complex was reported (Willhoft et al., 2017). SGT1 activates Hsp90 chaperone, which in turn is involved in the interaction with SCF E3 ubiquitine ligases and kinetochore. During this mechanism Sgt1 directly interacts with Skp1, which is known to be a small size protein, which also interacts with F-box domain. Although Skp1-Cullin-F-box ubiquitin ligase protein complexes may differ in structure from IcsB protein domains, they are both present in high enough concentrations to significantly interact with the same gene

domain SGT1. The study also found that SCF protein domain is directly involved in the G1–S and G2–M transition of the cell cycle in eukaryotes (yeast). As the study also suggested, SGT1 also functions to provide defense against various diseases and most importantly direct resistance against bacterial pathogens. In a study published by Mysore and coworkers, SGT1 is essential to induce plant cell death during bacterial infections (Wang, Uppalapati, Zhu, Dinesh-Kumar, & Mysore, 2010). Interestingly, their study demonstrated that plant SGT1 was activated in response to infection by *Pseudomonas syringae* and *Erwinia carotovora*. SGT1 was doing so through eliciting the hypersensitive response. This study provides precedent for the implication of SGT1 in the host response to bacterial pathogen possessing a T3SS.Another study conducted by Finlay and coworkers, implicated the *Salmonella* Typhimurium effector SspH2 in dampening the immune response in mammals and plants by acting on SGT1. The function of SGT1 that SspH2 seemed to interfere with was the co-chaperon activity of NLR proteins such as Nucleotide-binding oligomerization domain-containing protein 1 (NOD1) in humans. In addition, the effector SspH2 was shown to interact with SGT1 protein domains using a tight binding motifs without any noticeable alteration of SGT1 cell cycle functions. Furthermore, SspH2 in the presence of SGT1 was able to monobubiquitinate NOD1. The ubiquitination specificity for the leucine abundant region of NOD1 was controlled by the E3 ubiquitin ligase activity of SspH2. This post-translational modification rendered NOD1-activity independent of ligand binding, which resulted in an increase of NOD1-mediated IL-8 secretion. Hence it can be concluded that no significant protein shape changes occur upon binding of the effector to SGT1 but that its main role during infection was to activate SspH2 and facilitate the interaction of the latter with NOD1. In a separate study, the induction of autophagy was linked to interaction between NOD1 or Nucleotide-binding oligomerization domain-containing protein 2 (NOD2) and the ATG16L1 protein (Travassos et al.,

2010). In this study NOD1 and NOD2 were shown to be critical to the mechanism of defense against bacteria. During the defense action both NOD2 and NOD1 activated protein called ATG16L1 that serves as the defense against *Shigella* during its residence in the entry vacuole. However, when naturally occurring mutations of NOD2 were introduced, this defense mechanism was ineffective. Therefore, it is tantalizing to speculate that IcsB could also depend on interaction with SGT1 to target NOD1 or NOD2. Degradation of inhibition of NOD proteins by IcsB would potentially allow down regulating this pathogen-specific autophagy-inducing signal, thereby facilitating the vacuole escape mechanism described in the Introduction of this thesis. A detailed study should be conducted to link IcsB function to SGT1 during cellular invasion.

Another putative target of IcsB is DEAD-box protein 3, X-chromosomal (DDX3X). Specifically, DDX3X belongs to a group of specific RNA-binding proteins called DEAD-box which contain a characteristic amino acid sequence composed of the amino acids D-E-A-D (Asp-Glu-Ala-Asp). A special characteristic of DDX3X is the presence of an RNA-independent ATPase activity, while most DEAD-box proteins have RNA-dependent ATPase activity. Furthermore, it can probably bind not only RNA, but also DNA. Structurally this protein contains multiple domains including a helicase and ATPase, which serve as binding domains during translation, cell-to-cell signaling and viral replication (Weizmann, 2014). Several studies showed that cell growth and development were significantly affected upon silencing of the DDX3X gene, which highlight its importance in cell immune response mechanism (M.-C. Lai, Chang, Shieh, & Tarn, 2010) (M. C. Lai, Sun, Wang, & Tarn, 2016)(Grallert et al., 2000). It was found that DDX3X functions as a critical component for promoting and controlling G1 to S transition in the cell cycle. In another study, DDX3X was shown to function in concert with TBK1 to stimulate the IFN promoter (Soulat

et al., 2008). In addition, TBK1 was implicated in xenoautophagy of *Salmonella* Typhimurium (Thurston, 2009). Although it is not confirmed, it is reasonable to hypothesize that IcsB might be able to interfere with host autophagy response by perturbing TBK1 function through binding to DDX3X. In this case as well, further experiments will be necessary to test this hypothesis.

The last putative target of IcsB identified during my studies is the Fanconi Anemia Complementation Group L (FANCL). Of the three interacting partners identified herein, it is the only one that seemed to behave as a substrate for IcsB. Fanconi anemia (FA) is a genetically inherited disorder characterized y genetic instability and DNA repair defects. FA patients experience bone marrow failure and cancer in the majority of cases before 40 years old. There are eight identified FA proteins: FANCA, FANCB, FANCC, FANCE, FANCF, FANCG, FANCL and FANCM that form the core complex (Meetei et al., 2003). Specifically, it was shown that FANCL protein contains a PHD/RING-finger domain and possesses E3 ubiquitin ligase activity (Meetei et al., 2003). The RING finger domain was shown to be important for its structural integrity FANCL monoubiquitinate FANCD2, which is a key step in the DNA damage pathway, as well as FANCI (Longerich, San Filippo, Liu, & Sung, 2009). More recently, FA complementation group proteins, including FANCC and FANCL as well as others were implicated in mitophagy and virophagy (Orvedahl et al., 2011) (Sumpter et al., 2016). It is tantalizing to suggest that IcsB could facilitate the vacuole escape by inhibiting the autophagy-inducing activity of FANCL. As for the previous examples, this hypothesis will be challenged by future research.

Chapter 5: Conclusion and Future Perspectives

During my thesis I have identified three human proteins with which IcsB could be interacting during the infection. SGT1, DDX3X and FANCL have been implicated as indirect positive regulators of autophagy. It is therefore possible that IcsB could facilitate vacuole escape by either one of these channels. None of them seems to be more likely to be involved in that process than the others with the experimental and theoretical evidence at hand at the moment of the submission of my thesis. Since I have obtained some experimental evidence that FANCL might be a substrate of IcsB, I would certainly tackle it in priority if I was pursuing the work.

At the end of July 2018, a study reported that IcsB is capable of acylating the primary amine group of the side chain of lysine residues. Similarly to my thesis findings, the authors found that mutation of the predicted catalytic residue C306A, as well as other residues rescued the cytotoxicity phenotype in yeast. Indeed, they also reported that H145A as well as a previously unpredicted D195A rescued the cytotoxicity. Their conclusion was that H145, D195 and C306 constituted the catalytic triad of IcsB acyltransferase activity. Interestingly, they expressed the chaperon IpgA in all their yeast phenotypic assays. This might have prevented the toxicity associated with the N-terminus of the Sequence Required for Secretion (SRS), which I rescued in my own experiments by deleting the first 80 amino acids of IcsB (Δ80). Since our objective was to perform yeast two-hybrid screening, it did not seem reasonable to co-express IpgA because it could have competed for IcsB binding with our prey proteins. However, one might be able to test

whether the partial rescue observed with IcsB C306A can be improved to "full" rescue by co-expressing IpgA in yeast.

This study identified the charged multivesicular body protein 5 (CHMP5) as the substrate of IcsB that is implicated in the vacuole escape in human cells (Liu et al., 2018). It is not possible to discard the possibility that another unknown substrate of IcsB might be involved in this process. Furthermore, this study did not identify neither test FANCL has a substrate of IcsB. Likewise, these findings did not discard as well the possibility that IcsB might bind DDX3X and SGT1 in a none catalytic manner. Furthermore, the identification of the catalytic activity of IcsB will help in numerous ways to design future studies that will address whether FANCL may be yet another IcsB substrate implicated in vacuole escape. For example, one could perform an experiment to verify whether FANCL expressed in the presence of WT IcsB but not of the C306A mutant is acylated. Acylation can be detected simply by performing a WB with the relevant diploid strains; if acylation happens in this system, the apparent molecular weight of FANCL should be increased specifically when co-expressed with WT IcsB. In fact, these experiments could be initially performed in yeast with the constructs already described in my thesis. Of course, this would have to be confirmed in transfected and infected human cells. Another important control is to verify whether FANCL, SGT1 and DDX3X are expressed in the cultured mammalian cells that we would use for the infection experiment. We could then also confirm where the three candidates are located inside infected human cells. A critical point would be to verify whether they are recruited to the dissemination vacuole formed by *Shigella*. The article of Liu et al. also provided protocols to detect acylation of proteins by IcsB using a fluorescently or biotin labeled alkyn-analogue of stearic acid, allowing to perform in-gel or affinity purification of acylated proteins, respectively.

The affinity purification would allow us to directly test whether FANCL is enriched in the acylated fraction in the presence of WT IcsB. As SGT1 and DDX3X are probably not substrates as indicated by my data, they should not be acylated in an IcsB dependent manner. Since Liu et al. reported that CHMP5 is the only IcsB substrate that they identified to be critical for the escape of the dissemination vacuole, it represents a valid positive control for these experiments. Hence, it should be systematically included in experiments alongside with FANCL. Furthermore, using mass spectrometry and mutagenesis, we should be able to determine which lysine of FANCL are acylated. What could be the functional consequences of the acylation of FANCL? First, it could modify the cellular partitioning of FANCL by allowing its association with membrane, most likely with the vacuole membrane. This should be measurable by an increase of FANCL in the detergent resistant fraction of cells transfected with IcsB. Secondly, the acylation of the side-chain amine of one or several lysine residues of FANCL may inhibit FANCL ubiquitylating activity by preventing the formation of important protein-protein interactions. Alternatively, it might prevent the ubiquitylation of FANCL by preventing lysine residues to react with ubiquitin; this might prevent the ubiquitin-proteasome degradation of FANCL in the vicinity of secreting *Shigella*. Both events would likely lead to increase FANCL ubiquitylation activity in the vicinity of the vacuole by increasing its local concentration. Concerning SGT1 and DDX3X, we will perform IcsB immunoprecipitation in human tissue culture cell line to validate their binding to IcsB. Each validated interaction will be mapped onto the structure of the candidates and IcsB to identify point mutations thereof that would disrupt formation of these protein complexes. These mutant will constitute key controls to further validate our findings. If FANCL, SGT1 and DDX3X are involved in any way, a functional consequence of their absence in a cell line should be the complementation of IcsB deficiency in *ΔicsB* strains. Therefore, one may use siRNA or CRISP/Cas9 to generate cell

lines devoid of FANCL, SGT1, DDX3X or CHMP5 to verify this hypothesis. As apparent from this brief discussion of the future perspectives of my thesis, there are obviously many more experiments that need to be performed to validate my findings. I leave this to the group of researchers that will continue the project from then on. I am looking forward to those developments.

References

Baron, S. (1996). *Medical Microbiology*. (S. Baron, Ed.) (4th editio). Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK7627/

Bernardini, M. L., Mounier, J., d'Hauteville, H., Coquis-Rondon, M., & Sansonetti, P. J. (1989). Identification of icsA, a plasmid locus of Shigella flexneri that governs bacterial intra- and intercellular spread through interaction with F-actin. *Proceedings of the National Academy of Sciences*, *86*(10), 3867–3871. https://doi.org/10.1073/pnas.86.10.3867

Campbell-Valois, F.-X., & Pontier, S. M. (2016). Implications of Spatiotemporal Regulation of Shigella flexneri Type Three Secretion Activity on Effector Functions: Think Globally, Act Locally. *Frontiers in Cellular and Infection Microbiology*, *6*(March), 1–13. https://doi.org/10.3389/fcimb.2016.00028

Campbell-Valois, F. X., Sachse, M., Sansonetti, P. J., & Parsot, C. (2015). Escape of actively secreting Shigella flexneri from ATG8/LC3-Positive vacuoles formed during cell-to-cell spread is facilitated by IcsB and VirA. *MBio*, *6*(3), 1–11. https://doi.org/10.1128/mBio.02567-14

Campbell-Valois, F. X., Schnupf, P., Nigro, G., Sachse, M., Sansonetti, P. J., & Parsot, C. (2014). A fluorescent reporter reveals on/off regulation of the Shigella type III secretion apparatus during entry and cell-to-cell spread. *Cell Host and Microbe*, *15*(2), 177–189.

https://doi.org/10.1016/j.chom.2014.01.005

CDC. (2016). Antimicrobial Resistance. Retrieved from http://www.cdc.gov/drugresistance/about.html

CDC. (2017). Shigella - Shigellosis. Retrieved from https://www.cdc.gov/Shigella/general-information.html

Clontech. (2010). Matchmaker ™ Gold Yeast Two-Hybrid System User Manual. *Pt40841 Pr033493*, *1*(630489), 1–41. Retrieved from http://www.clontech.com/

de Lichtenberg, U., Jensen, L. J., Fausbøll, A., Jensen, T. S., Bork, P., & Brunak, S. (2005). Comparison of computational methods for the identification of cell cycle-regulated genes. *Bioinformatics*, *21*(7), 1164–1171. https://doi.org/10.1093/bioinformatics/bti093

Dong, N., Zhu, Y., Lu, Q., Hu, L., Zheng, Y., & Shao, F. (2012). Structurally distinct bacterial TBC-like GAPs link Arf GTPase to Rab1 inactivation to counteract host defenses. *Cell*, *150*(5), 1029–1041. https://doi.org/10.1016/j.cell.2012.06.050

Franceschetti, M., Maqbool, A., Jiménez-Dalmaroni, M. J., Pennington, H. G., Kamoun, S., & Banfield, M. J. (2017). Effectors of Filamentous Plant Pathogens: Commonalities amid Diversity. *Microbiology and Molecular Biology Reviews*. https://doi.org/10.1128/MMBR.00066-16

Gorden, J., & Small, P. L. C. (1993). Acid Resistance in Enteric Bacteria. *Infection and Immunity*, *61*(1), 364–367.

Grallert, B., Kearsey, S. E., Lenhard, M., Carlson, C. R., Nurse, P., Boye, E., & Labib, K. (2000). A fission yeast general translation factor reveals links between protein synthesis and cell cycle controls. *J Cell Sci*, *113 (Pt 8*, 1447–1458. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10725227

Huang, J., & Brumell, J. H. (2014). Bacteria-autophagy interplay: A battle for survival. *Nature Reviews Microbiology*, *12*(2), 101–114. https://doi.org/10.1038/nrmicro3160

Ireton, O. B. K., & Ireton, K. (1300). bacterial pathogens cell spread of intracellular − Molecular mechanisms of cell Subject collections Molecular mechanisms of cell – cell spread of intracellular bacterial pathogens, (step 5). Retrieved from http://rsob.royalsocietypublishing.org/content/3/7/130079.full.html#ref-list-1%5Cnhttp://dx.doi.org/10.1098/rsob.130079

Kang, E., Crouse, A., Chevallier, L., Pontier, S. M., Alzahrani, A., Silué, N., … Malo, D. (2018). Enterobacteria and host resistance to infection. *Mammalian Genome*, *0*(0), 1–19. https://doi.org/10.1007/s00335-018-9749-4

Kayath, C. A., Hussey, S., El hajjami, N., Nagra, K., Philpott, D., & Allaoui, A. (2010). Escape

of intracellular Shigella from autophagy requires binding to cholesterol through the type III effector, IcsB. *Microbes and Infection*, *12*(12–13), 956–966. https://doi.org/10.1016/j.micinf.2010.06.006

Keusch, T. L. H. and G. T. (1996). Medical Microbiology (4th editio). Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK8038/

Kitagawa, K., Skowyra, D., Elledge, S. J., Harper, J. W., & Hieter, P. (1999). SGT1 encodes an essential component of the yeast kinetochore assembly pathway and a novel subunit of the SCF ubiquitin ligase complex. *Molecular Cell*, *4*(1), 21–33. https://doi.org/10.1016/S1097-2765(00)80184-7

Kothary, M. h. (2007). Infective dose of Foodborne pathogens in volunteers. *Food Safety*. Retrieved from https://www.genecards.org/cgi-bin/carddisp.pl?gene=DDX3X

Kushnirov, V. V, & Kushnirov, V. V. (2000). Yeast Functional Analysis Report Rapid and reliable protein extraction from yeast, 857–860. https://doi.org/10.1002/1097-0061(20000630)16

Lai, M.-C., Chang, W.-C., Shieh, S.-Y., & Tarn, W.-Y. (2010). DDX3 Regulates Cell Growth through Translational Control of Cyclin E1. *Molecular and Cellular Biology*, *30*(22), 5444–5453. https://doi.org/10.1128/MCB.00560-10

Lai, M. C., Sun, H. S., Wang, S. W., & Tarn, W. Y. (2016). DDX3 functions in antiviral innate immunity through translational control of PACT. *FEBS Journal*, *283*(1), 88–101. https://doi.org/10.1111/febs.13553

Liu, W., Zhou, Y., Peng, T., Zhou, P., Ding, X., Li, Z., … Shao, F. (2018). Nε-fatty acylation of multiple membrane-associated proteins by Shigella IcsB effector to modulate host function. *Nature Microbiology*. https://doi.org/10.1038/s41564-018-0215-6

Longerich, S., San Filippo, J., Liu, D., & Sung, P. (2009). FANCI binds branched DNA and is monoubiquitinated by UBE2T-FANCL. *Journal of Biological Chemistry*, *284*(35), 23182–23186. https://doi.org/10.1074/jbc.C109.038075

Meetei, A. R., De Winter, J. P., Medhurst, A. L., Wallisch, M., Waisfisz, Q., Van de Vrugt, H. J., … Wang, W. (2003). A novel ubiquitin ligase is deficient in Fanconi anemia. *Nature Genetics*, *35*(2), 165–170. https://doi.org/10.1038/ng1241

Mizushima, N. (2007). Autophagy: Process and function. *Genes and Development*. https://doi.org/10.1101/gad.1599207

Ogawa, M., Suzuki, T., Tatsuno, I., Abe, H., & Sasakawa, C. (2003). IcsB, secreted via the type III secretion system, is chaperoned by IpgA and required at the post-invasion stage of Shigella pathogenicity. *Molecular Microbiology*, *48*(4), 913–931. https://doi.org/10.1046/j.1365-2958.2003.03489.x

Orvedahl, A., Sumpter, R., Xiao, G., Ng, A., Zou, Z., Tang, Y., … Levine, B. (2011). Image-based genome-wide siRNA screen identifies selective autophagy factors. *Nature*. https://doi.org/10.1038/nature10546

Pei, J., & Grishin, N. V. (2009). The Rho GTPase inactivation domain in Vibrio cholerae MARTX toxin has a circularly permuted papain-like thiol protease fold. *Proteins: Structure, Function and Bioinformatics*, *77*(2), 413–419. https://doi.org/10.1002/prot.22447

Shigella spp. (2016). Micrococcus Pathogen Safety Data Sheet - Infectious Substances Section I - Infectious Agent, (12), 3–6.

Sitthidet, C., Stevens, J. M., Chantratita, N., Currie, B. J., Peacock, S. J., Korbsrisate, S., & Stevens, M. P. (2008). Prevalence and sequence diversity of a factor required for actin-based motility in natural populations of Burkholderia species. *Journal of Clinical Microbiology*, *46*(7), 2418–2422. https://doi.org/10.1128/JCM.00368-08

Slagowski, N. L., Kramer, R. W., Morrison, M. F., LaBaer, J., & Lesser, C. F. (2008). A functional genomic yeast screen to identify pathogenic bacterial proteins. *PLoS Pathogens*, *4*(1), 0096-0107. https://doi.org/10.1371/journal.ppat.0040009

Soulat, D., Bürckstümmer, T., Westermayer, S., Goncalves, A., Bauch, A., Stefanovic, A., … Superti-Furga, G. (2008). The DEAD-box helicase DDX3X is a critical component of the TANK-binding kinase 1-dependent innate immune response. *EMBO Journal*, *27*(15), 2135–

2146. https://doi.org/10.1038/emboj.2008.126

Sumpter, R., Sirasanagandla, S., Fernández, Á. F., Wei, Y., Dong, X., Franco, L., … Levine, B. (2016). Fanconi Anemia Proteins Function in Mitophagy and Immunity. *Cell*. https://doi.org/10.1016/j.cell.2016.04.006

Thurston, T. L. M. (2009). The tbk1 adaptor and autophagy receptor ndp52 restricts the proliferation of ubiquitin-coated bacteria. *Nature Immunology*, *10*(11), 1215–1222. https://doi.org/10.1038/ni.1800

Travassos, L. H., Carneiro, L. A. M., Ramjeet, M., Hussey, S., Kim, Y. G., Magalhes, J. G., … Philpott, D. J. (2010). Nod1 and Nod2 direct autophagy by recruiting ATG16L1 to the plasma membrane at the site of bacterial entry. *Nature Immunology*, *11*(1), 55–62. https://doi.org/10.1038/ni.1823

Van Criekinge, W., & Beyaert, R. (1999). *Yeast Two-Hybrid: State of the Art*. *Biological Procedures Online* • (Vol. 2). Retrieved from www.biologicalprocedures.com

Wang, K., Uppalapati, S. R., Zhu, X., Dinesh-Kumar, S. P., & Mysore, K. S. (2010). SGT1 positively regulates the process of plant cell death during both compatible and incompatible plant-pathogen interactions. *Molecular Plant Pathology*. https://doi.org/10.1111/j.1364-3703.2010.00631.x

Weizmann. (2014). DDX3X Gene (Protein Coding). Retrieved from https://www.genecards.org/cgi-bin/carddisp.pl?gene=DDX3X

WHO. (2017). The top 10 causes of death. Retrieved from http://www.who.int/mediacentre/factsheets/fs310/en/.

Willhoft, O., Kerr, R., Patel, D., Zhang, W., Al-Jassar, C., Daviter, T., … Vaughan, C. K. (2017). The crystal structure of the Sgt1-Skp1 complex: The link between Hsp90 and both SCF E3 ubiquitin ligases and kinetochores. *Scientific Reports*, *7*(December 2016), 1–13. https://doi.org/10.1038/srep41626

Appendix

Table A1: Mutagenesis and sequencing primer. ***PCR mutagenesis was performed as per protocol to design IcsB mutants.***

Primer	Sequence (5'-3')
IcsB_EcorI_S	AGAGAGGAATTCATGAGCCTCAAAATTAGCAATTTCA
IcsB_Bam_R	AGAGAGGGATCCTTAGCTATATATTAGAATGAGAGTTATTC
IcsB_C306A_S	GCCGCTGGTATGGCACTTAATGTTC
IcsB_305_R	GTTTTCAGATTTACTGATTAATTTATA
D_80N-ter_IcsB_S	GAAAAATCTGCAGAAAGTGCGTT
D_80N-ter_IcsB_R	GAATTCGGCCTCCATGGC
IcsB-dDom 1-3`-F	GAA ATA TAC TTT CCA CTC CCT GAC G
pGBKT7 5`-MCS-R	GAA TTC GGC CTC CAT GGC C
IcsB-dDom 2-5`-R	AGA ATT TCC TGC TTT AAG AAC ATT AAG
pGBKT7 3`-MCS-F	GGA TCC GTC GAC CTG CA
SeqO7- pGBKT7_MCS3'_seq_R	GAGTCACTTTAAAATTTGTATACAC
SeqO8- pGBKT7_seq_S	GTCAAAGACAGTTGACTGTATCG
T7	TAATACGACTCACTATAGGG

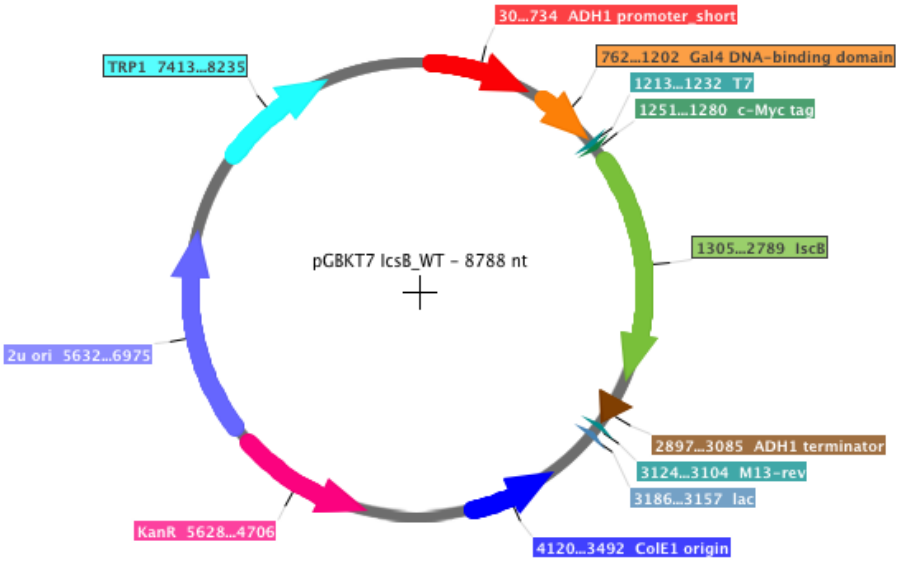

Figure A1: Graphic map of pGBKT7-IcsB.

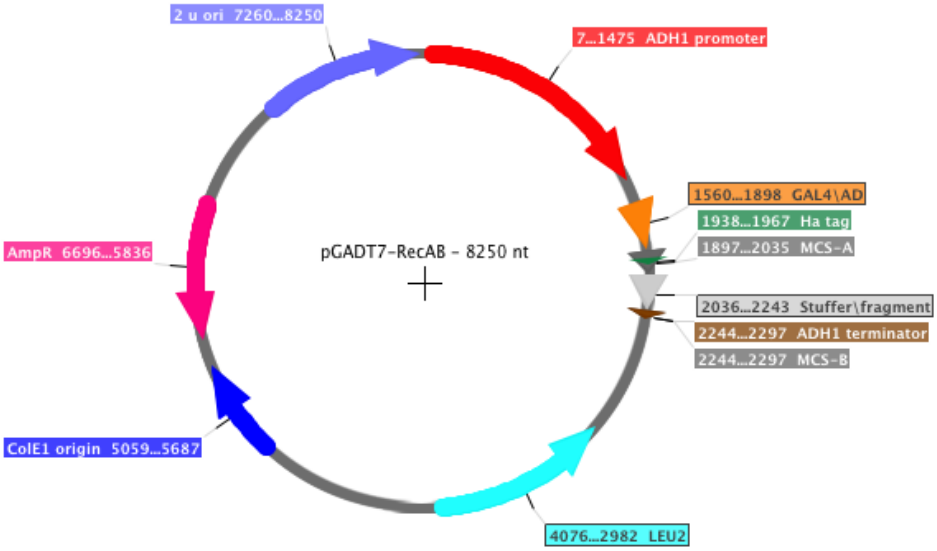

Figure A2: Graphic map of pGADT7-RecAB.

Figure A3: Secondary structure prediction with different algorithms of IcsB threaded over the experimentally determined secondary structure of distant *Salmonella* Typhimurium homolog SptP. Note that residues 80 is located at the C-terminus of an α-helix.