Investigation of a Putative Type I Secretion System and Potential Substrates in *Treponema pallidum*, the Causative Agent of Syphilis

by

Claudia Gaither
Bachelor of Science, University of Victoria, 2013

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Biochemistry and Microbiology

Supervisory Committee

Investigation of a Putative Type I Secretion System and Potential Substrates in *Treponema pallidum*, the Causative Agent of Syphilis

by

Claudia Gaither
Bachelor of Science, University of Victoria, 2013

Supervisory Committee

Dr. Caroline Cameron, Department of Biochemistry and Microbiology
Supervisor

Dr. Caren Helbing, Department of Biochemistry and Microbiology
Departmental Member

Dr. Steve Perlman, Department of Biology
Outside Member

Abstract

Supervisory Committee
Dr. Caroline Cameron, Department of Biochemistry and Microbiology
Supervisor
Dr. Caren Helbing, Department of Biochemistry and Microbiology
Departmental Member
Dr. Steve Perlman, Department of Biology
Outside Member

Recent bioinformatic analyses identified an operon encoding a potential Type I Secretion System (T1SS) in *Treponema pallidum* that we hypothesize functions to export key treponemal virulence factors that may contribute to the unique invasiveness and pathogenesis of this spirochete. The membrane fusion protein component (MFP) of T1SSs in other organisms has been shown to play a role in substrate recognition. Hence, the objective of this project is to use the putative MFP, Tp0965, of the potential *T. pallidum* T1SS to investigate protein-protein interactions with the *T. pallidum* virulence factor pallilysin (Tp0751) and assess the possibility of the latter being a T1SS substrate. Moreover, protein-protein interactions between Tp0965 and a *Treponema phagedenis* lysate are investigated with the goal of identifying putative T1SS substrates in this spirochete that could result in the discovery of novel *T. pallidum* virulence factors via amino acid sequence similarity.

Plate-based binding studies and pull-down assays showed a low level of interaction between recombinant Tp0965 and the previously characterized host-component-binding protease, pallilysin, suggesting that the export of this virulence factor could occur via the putative T1SS.

Additionally, bioinformatic analyses of the related but cultivable model spirochete *T. phagedenis* predicted the presence of a potential T1SS homologous to the putative T1SS

in *T. pallidum*. Thus, a more global and unbiased pull-down assay using "bait" Tp0965 and a "prey" *T. phagedenis* lysate was carried out, followed by mass spectrometric analysis to identify putative novel T1SS substrates with potential homologs in *T. pallidum*. We successfully identified a *T. phagedenis* protein, TphBIg, that showed evidence of an interaction with Tp0965. TphBIg seems to possess characteristics of a T1SS substrate suggesting it may be secreted via this system in *T. phagedenis*. Upon bioinformatic analysis, it was found that TphBIg showed weak amino acid sequence similarity as well as some structural similarity to the *T. pallidum* protein, Tp0854.

Tp0854 is predicted to contain a sialidase and a phosphatase domain with an RTX motif, which is characteristic of some T1SS substrates. Thus, it was hypothesized that if Tp0854 had characteristics of a T1SS, it may interact with Tp0965. Therefore, the phosphatase domain containing the RTX motif was produced recombinantly and plate-based binding studies indeed suggested an interaction with Tp0965, confirming the *in silico*-predicted interaction.

Future experiments to characterize the potential T1SS and substrates in *T. pallidum* could comprise the functional and structural characterization of the novel putative T1SS substrate, Tp0854. This would include assays to investigate the putative sialidase and phosphatase activities of Tp0854, as well as the identification of Tp0854-Tp0965 interacting sites. Moreover, as a more definite test for T1SS substrate secretion, *T. pallidum* pallilysin and/or Tp0854 could be expressed heterologously in an *E. coli* strain harbouring an endogenous T1SS and test for secretion. Similarly, the reconstitution of the *T. pallidum* putative T1SS in liposomes could be used to further investigate the secretion of pallilysin and/or Tp0854 via this system.

Additionally, the optimized unbiased pull-down technique could be further applied to detect more protein-protein interactions within *T. pallidum* and potentially lead to the identification of more virulence factors that may be secreted via the T1SS.

These studies constitute the first investigation of a putative T1SS and substrates within *T. pallidum*. Thus, insight gained will lead to a better understanding of the mechanisms facilitating *T. pallidum* host invasion and may reveal new potential vaccine targets to prevent bacterial dissemination and chronic infection.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Abbreviation	Description
ABC	ATP-Binding Cassette
ACN	Acetonitrile
ATP	Adenosine triphosphate
BAM	β-barrel assembly machinery
BAP	Biofilm-Associated Protein
BCA	Bicinchoninic acid assay
BIg	Bacterial immunoglobulin-like
BLAST	Basic Local Alignment Search Tool
BpfA	Bap/RTX cell surface protein
BSA	Bovine serum albumin
cAMP	Cyclic adenosine monophosphate
CBDPS	Cyanurbiotindipropionylsuccinimide
CD	Conserved domain
CDD	Conserved domains database
CID	Collision-induced dissociation
CLD	C39-like domain
CNS	Central nervous system
DDA	Data-dependent acquisition
dH_2O	Deionized water
ECM	Extracellular matrix
FPLC	Fast protein liquid chromatography
HAMP	Histidine kinase, Adenylyl cyclase, Methyl-accepting protein, and Phosphatase
HCCA	α-Cyano-4-hydroxycinnamic acid
HIV	Human immunodeficiency virus
HlyA	Hemolysin A
ICC-CLASS	Isotopically-coded cleavable cross-linking analysis software suite
IM	Inner membrane
IMAC	Immobilized metal ion affinity chromatography
IMF	Inner membrane fraction
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LOL	Lipoprotein sorting system
MALDI	Matrix-assisted laser desorption/ionization
MFP	Membrane fusion protein
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
NBCI	National Center for Biotechnology Information
NBD	Nucleotide binding domain
Ni-HRP	Nickel-conjugated horseradish peroxidase
Ni-NTA	Nickel-conjugated nitrilotriacetic acid
OM	Outer membrane
OMF	Outer membrane fraction
OMP	Outer membrane proteins

OMV	Outer membrane vesicles
ORF	Open reading frame
PDD	Papillomatous digital dermatitis
PBS	Phosphate buffered saline
POTRA	Polypeptide translocation associated
PTM	Post-translational modification
RT	Room temperature
RTX	Repeats-in-toxin
SBP	Substrate binding protein
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
Sec	General secretory system
T. pallidum	*Treponema pallidum* spp. *pallidum*
T. phagedenis	*Treponema phagedenis*
T1SS, T2SS, … T6SS	Type 1, 2, … 6 secretion systems
TAT	Twin-arginine translocation
TBS	Tris buffered saline
TBS-T	TBS-0.1% Tween-20
TCEP	Tris (2-carboxyethyl) phosphine
TFA	Trifluoroacetic acid
TMB	Tetramethylbenzidine
TMD	Transmembrane domain
TOF	Time-of-flight
TPR	Tetratricopeptide repeat
WHO	World Health Organization
WT	Wildtype

Acknowledgments

So many people have been a part of this project and I am so thankful! First of all, I would like to thank Dr. Caroline Cameron for the opportunity to work on such an amazing project and with such a wonderful team, for all the guidance and for all the support throughout the past three years. I would also like to thank all the Cameron lab members: Dr. Simon Houston, Charmaine Wetherell, Brigette Church and Karen Lithgow – I have learned so much from all of you and will miss you!

My committee members, Dr. Steve Perlman and Dr. Caren Helbing for all their advice throughout this project, for all the positive (and even the negative!) feedback.

Dr. Martin J. Boulanger for all the insight in structural Biology and for providing us with the cloning vector used for the expression of some of our *T. pallidum* proteins.

Finally, to everyone at the UVic – Genome BC Proteomics Centre, especially Dr. Christoph H. Borchers, Dr. Evgeniy Petrotchenko, Dr. Jason Serpa, and Derek S. Smith for playing a key role in the proteomics aspect of my project.

Dedication

I will dedicate my work, as always, to my family. My family back home who I miss dearly: my parents and my grandma. My family in Victoria who make me smile everyday, my fiancé and my two pups, Canela and Lyla. Robert, thanks for keeping the house together while I wrote my thesis; pups, thanks for keeping those tails wagging!

I love you all, my accomplishments are yours.

Chapter 1: Introduction

1.1 Syphilis

1.1.1 Syphilis; a modern "old" infectious disease

Syphilis was first recognized in the 15th century in Europe as a devastating, painful and repulsive disease that burdened society; it was known that its causative agent was able to invade the whole human body, was resistant to medical treatment, easily spread and extremely torturous to the patient (Frith, 2012). Today, this ancestral disease remains a public health threat, with estimates of 12 million new cases per year globally, despite the availability of effective antibiotic treatment since the last years of World War II (World Health Organization, 2015).

Syphilis can be transmitted sexually, via direct contact with an infectious lesion resulting from either the primary or secondary stages of syphilis, or vertically from a pregnant woman to her fetus *in utero,* which leads to fetal loss or congenital syphilis in the newborn (LaFond and Lukehart, 2006). Indeed, syphilis is currently the most significant disease affecting fetuses worldwide, with 305,000 fetal and neonatal deaths every year, as well as 215,000 infants left at an increased risk of death from prematurity, low-birth-weight or congenital disease (World Health Organization, 2015).

Furthermore, there is evidence of complex interactions between syphilis and human immunodeficiency virus (HIV) infections, indicating a greater risk of transmission and acquisition of HIV infection after syphilis infection (Hook, 1989; Nusbaum *et al.*, 2004). Peterman *et al.* showed that the risk of subsequent HIV infection was 3.6% higher in the first year after syphilis diagnosis, and reached 17.5% ten years after syphilis

diagnosis in Florida (Peterman *et al.*, 2014). Additionally, previous studies have suggested that individuals with syphilis are three to five times more likely to acquire HIV if exposed to the virus via sexual contact (Buchacz *et al.*, 2004; Wasserheit, 1992).

The global incidence of syphilis and its close relation to HIV, in addition to the current epidemic of syphilis, not only in developing countries, but also in the United States of America, Canada and Europe, strongly suggest the need for the development of an effective vaccine as a preventative measure to reduce both syphilis, congenital syphilis and HIV rates (Brown and Frank, 2003; Cameron and Lukehart, 2014; Ho and Lukehart, 2011). However, successful vaccine development depends on the identification of target virulence factors and vaccine candidates, hence highlighting the need to better understand the molecular mechanisms underlining syphilis infection.

1.1.2 Syphilis is a multi-stage disease

Syphilis is a multi-stage disease with characteristic sequelae, as shown in Figure 1A, resulting from infection with the spirochete *Treponema pallidum* (depicted in Figure 1B). Initial localized *T. pallidum* infection leads to bacterial replication *in situ*, along with bacterial dissemination via the circulatory system and throughout host body organs, tissues and even the central nervous system (CNS). Indeed, it has been shown that *T. pallidum* is so invasive and infectious that only ten organisms or less are sufficient for infection in rabbits and humans, when inoculated intradermally (Magnuson *et al.*, 1948). Furthermore, studies in the early 20th century showed that organisms could be detected in the dermis of rabbits only 2-3 hours following *T. pallidum* exposure, indicating a fast traversion of the genital mucosa (Mahoney and Bryant, 1933).

Approximately 3-6 weeks after infection, the patient enters primary syphilis, during which regional lymphadenopathy occurs and a painless chancre appears at the site of infection. This stage can be easily missed since the characteristic chancre, shown in Figure 1C, can remain unnoticed and local clearance of *T. pallidum* results in spontaneous healing and resolution of the lesion within 3-8 weeks, even in the absence of treatment (Radolf and Lukehart, 2006).

The following stage, secondary syphilis, in which *T. pallidum* reaches systemic levels, occurs within three months of infection and is characterized by generalized lymphadenopathy and a disseminated rash, most commonly on the trunk and extremities, including the patient's palms of their hands and soles of their feet, as shown in Figure 1D. Upon healing of the disseminated rash, patients enter a latent syphilis stage and up to 25% of individuals show a single or multiple recurrences of secondary syphilis symptoms.

Following the latent stage of syphilis, up to 28% of untreated patients enter tertiary syphilis, whereas approximately 72% show no further complications. During tertiary syphilis, several organs are affected leading to clinical manifestations that include gumma, shown in Figure 1E, cardiovascular syphilis, as well as late neurological complications that characterize neurosyphilis, and can occur years or even decades after initial infection with *T. pallidum* (LaFond and Lukehart, 2006). It is important to note that even though neurological complications are generally associated with tertiary syphilis, *T. pallidum* is able to disseminate to the CNS shortly after infection and therefore patients with early syphilis can develop such symptoms as well (Lukehart *et al.*, 1988).

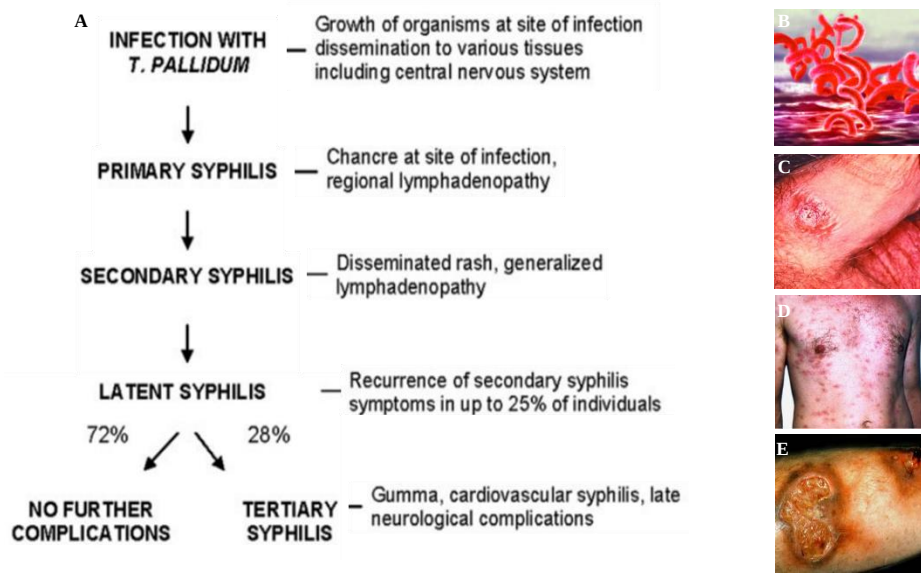


Figure 1. Diagram of the syphilis sequelae.

(A) The multiple stages of untreated syphilis (LaFond and Lukehart, 2006); (B) picture of *T. pallidum*, the spirochete that causes syphilis (French, 2007); (C) penile chancre, characteristic of primary syphilis (French, 2007); (D) disseminated lesions characteristic of secondary syphilis (Baughn and Musher, 2005); (E) gumma on the leg of a patient with tertiary syphilis (Carlson *et al.*, 2011). Printed with permission.

1.2 *Treponema pallidum* spp. *pallidum*

1.2.1 Morphology of *Treponema pallidum* spp. *pallidum*

The causative agent of syphilis, *Treponema pallidum* spp. *pallidum,* is a highly invasive Gram-negative-like spirochete of approximately 0.2 μm in diameter and between 6 and 15 μm in length, with unique characteristics and an unusual architecture that make it challenging to study its pathogenesis (Izard *et al.*, 2009; LaFond and Lukehart, 2006). The invasiveness of *T. pallidum* is well shown in Figure 2A, where a single *T. pallidum* bacterium can be seen moving in between rabbit testicular tissue after infection. Indeed, this spirochete is one of the most invasive pathogens known, being able to cross both the blood-brain and placental barriers (LaFond and Lukehart, 2006; Lukehart *et al.*, 1988; Norris *et al.*, 2001). Like all spirochetes, *T. pallidum* possesses corkscrew motility due to the presence of endoflagella, which are flagellar structures located within the periplasmic space, and like typical Gram-negative bacteria, it has both a cytoplasmic (inner) membrane (IM) and an outer membrane (OM), as depicted in Figure 2B (LaFond and Lukehart, 2006; Limberger, 2004). However, unlike conventional Gram-negative bacteria, the OM of *T. pallidum* has very few integral outer membrane proteins (OMPs).The freeze-fracture electron microscopy pictures in Figure 2C and 2D, show that the *T. pallidum* IM has a rougher surface, indicative of a greater number of integral membrane proteins, in comparison with the OM surface which appears smoother and thus devoid of OMPs (Walker *et al.*, 1989). It is important to note however, that lipidated integral membrane proteins cannot be detected via freeze-fracture techniques and thus, no inferences could be made regarding the presence of such type of proteins on *T. pallidum* membranes via these studies. As shown in Figure 2E and 2F, *T. pallidum* also differs from typical Gram-negative

bacteria in that its peptidoglycan layer is found closer to the IM rather than underlying the OM (Cox *et al.*, 1992; Ruiz *et al.*, 2009). Moreover, *T. pallidum*'s OM is devoid of lipopolysaccharide (LPS), a strong pro-inflammatory glycolipid (Radolf and Norgard, 1988), and thus, the dominant immunogens are lipoproteins and cytoplasmic membrane-associated proteins (Blanco *et al.*, 1997).

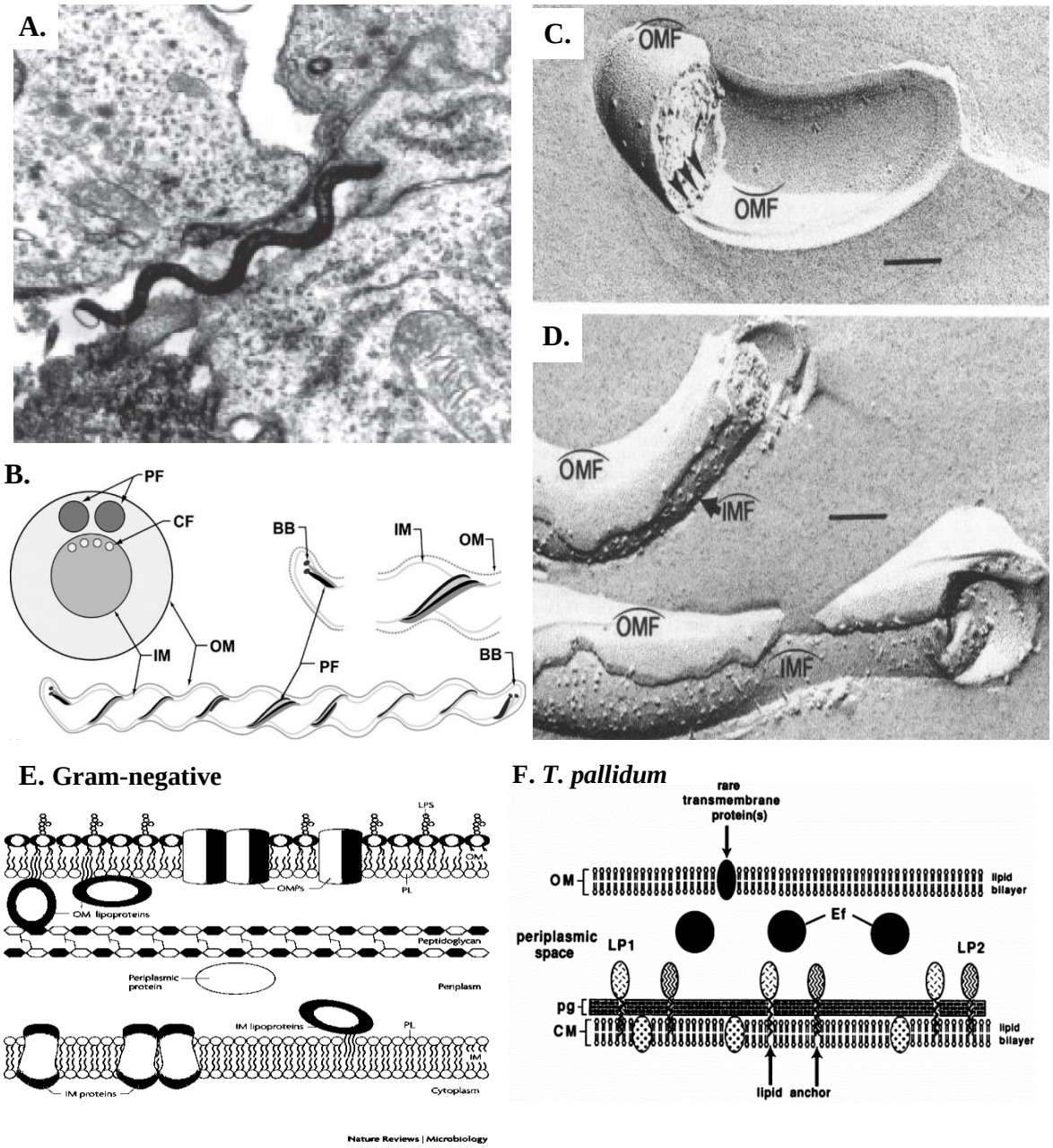


Figure 2. Invasiveness and morphology of *Treponema pallidum*.

(A) Transmission electron microscopy image showing the invasive nature of *T. pallidum* on rabbit testicular tissue post-infection (Norris *et al.*, 2001); (B) spirochetal morphology of *T. pallidum*, PF: Periplasmic Flagella, CF: Cytoplasmic Filaments, IM: Inner Membrane, OM: Outer Membrane, BB: Basal Body (Limberger, 2004); (C) and (D) freeze-fracture electron microscopy images of the outer and inner membranes (OMF and IMF, respectively) of *T. pallidum* (Walker *et al.*, 1989); (E) membrane structure of a typical Gram-negative bacterium (Ruiz *et al.*, 2009); (F) membrane structure of *T. pallidum* (Cox *et al.*, 1992). Printed with permission.

1.2.2 *Treponema pallidum*, the "stealth" pathogen

It is widely accepted in the syphilis field that the minimalistic features of *T. pallidum* result in the poor antigenicity that characterizes this pathogen. Studies have shown that serum against *T. pallidum* is only reactive on treponemal cells that have been compromised in some way, either through storage or handling, resulting in OM disruption (Cox *et al.*, 1992; Hardy Jr and Nell, 1957). Moreover, *in vitro* antibody and complement killing of *T. pallidum* only occurs after long exposure times of a minimum of 4 hours to achieve any degree of bacterial killing, and 16 hours for complete killing (Bishop and Miller, 1976; Nelson and Mayer, 1949). Finally, the *in vitro* recognition and uptake of *T. pallidum* by macrophages has been reported to be only 65% in a period of 24 hours (Alder *et al.*, 1990). *T. pallidum* is thus recognized as the "stealth" pathogen and the few OMPs present on its surface contribute to the low reactivity characteristic of this spirochete. Furthermore, this poor antigenicity explains, at least to some extent, the persistence and invasive properties unique to *T. pallidum* (LaFond and Lukehart, 2006; Radolf and Lukehart, 2006). Hence it remains of great interest to identify and decipher the nature of the rare OMPs that may be important not only as antigens for immunodetection, but also as *T. pallidum* virulence determinants that could potentially make good vaccine candidates.

1.2.3 Virulence of *Treponema pallidum*

Virulence of most pathogens is strongly related to their ability to attach and successfully invade their host in order to achieve tissue colonization (Ribet and Cossart, 2015) and it is thought that *T. pallidum* is no exception to this mechanism. In addition to being a "stealth" pathogen due to the reduced protein content on its OM, *T. pallidum*

virulence depends upon adherence to epithelial cells and to the host's extracellular matrix (ECM) components as an important first step to establishing infection (LaFond and Lukehart, 2006; Radolf and Lukehart, 2006). *T. pallidum* is able to adhere to and invade epithelial cell surfaces, traverse the tissue barrier, and undergo widespread dissemination by gaining access into the bloodstream through disruption of the tight junctions between endothelial cells, and including the blood-brain and placental barriers (Fitzgerald, Cleveland, *et al.*, 1977; Fitzgerald, Johnson, *et al.*, 1977; LaFond and Lukehart, 2006; Lukehart *et al.*, 1988; Thomas *et al.*, 1988).

Moreover, throughout the three stages of syphilis, it is evident that the fast and extensive dissemination of *T. pallidum* results in a high degree of damage to the host, from the chancre in primary syphilis, to the rash in secondary syphilis, to gumma in tertiary syphilis. However, the exact molecular mechanism responsible for the invasiveness and overall pathogenesis of this spirochete has not been fully elucidated.

1.2.3.1 Virulence factors contribute to *T. pallidum* pathogenesis

A wide number of bacterial pathogens interfere or alter host processes by secreting effectors (usually proteinaceous in nature, and also known as virulence factors) to the bacterial cell surface, into the host environment, or in some cases, directly into host cells (Gauthier and Finlay, 2001), resulting in a bacterial advantage that favours pathogen survival. Different secreted virulence factors possess diverse functionalities that are pathogen-dependent, such as nutrient acquisition, adhesion to host cells or ECM, biofilm formation, host cell lysis, serum resistance, and protein, lipid and carbohydrate degradation (Henderson and Nataro, 2001).

In the case of *T. pallidum*, previous literature has stated that although this pathogen possesses several virulent characteristics (briefly outlined above), it lacks a specific molecule or constituent identified as a virulence factor in the classical sense (Radolf and Lukehart, 2006). Nonetheless efforts remain to reveal genes and proteins that enable this spirochete to invade and colonize its host. Indeed, previous investigations tailored to detect *T. pallidum* OMPs identified what are now believed to be virulence factors that may be key for invasion and dissemination (Cameron, 2003), and their specific roles in *T. pallidum* pathogenesis are underway to being uncovered (Cameron *et al.*, 2008; Houston *et al.*, 2011; Houston *et al.*, 2014).

Previous binding studies have shown that both fibronectin and laminin, the most abundant component of the basement membrane, are key for treponemal cytadherence (Baughn, 1987; Fitzgerald *et al.*, 1984), but it wasn't until 2003 that a key laminin-binding adhesin, Tp0751 (now also known as pallilysin), was identified (Cameron, 2003). Cameron's binding studies showed the specific attachment of pallilysin to laminin in a dose-response manner. Moreover, pallilysin-specific antibodies were detected in serum samples from syphilis infections, suggesting that this virulence factor is expressed *in vivo* during infection conditions and might be involved in the attachment of *T. pallidum* to host cells and tissues, thus playing an important role in bacterial dissemination (Cameron, 2003). Follow-up work by the Cameron lab further showed that soluble recombinant pallilysin is also able to bind human fibrinogen (Houston *et al.*, 2011), that antibodies against pallilysin prevent *T. pallidum* attachment to laminin-coated surfaces (Cameron *et al.*, 2005) and that pallilysin is a target of opsonic antibodies (Houston *et al.*, 2012).

Furthermore, it has been shown that *Treponema phagedenis*, a model treponeme discussed in Section 1.3, could not only heterologously express pallilysin on its surface, but, as shown in Figure 3, pallilysin confers upon this spirochete the ability to bind laminin (Cameron *et al.*, 2008). Thus, WT *T. phagedenis* and *T. phagedenis* transformed with an empty plasmid, showed no ability to bind to laminin, whereas *T. phagedenis* transformed with a *pallilysin* (*tp0751*)-encoding plasmid could bind laminin.

In addition, another putative virulence factor, Tp0750, was recently shown to be co-expressed with pallilysin. Tp0750 is able to bind host components, and degrade the major coagulation proteins fibrinogen and fibronectin by means of its serine protease activity (Houston *et al.*, 2014).

The recent advances in the discovery of potential novel virulence factors represents a major step toward understanding the mechanisms whereby *T. pallidum* is able to invade every host tissue whilst evading the immune system, however, much research remains to be carried out in order to fully understand this minimalistic, yet complex pathogen.

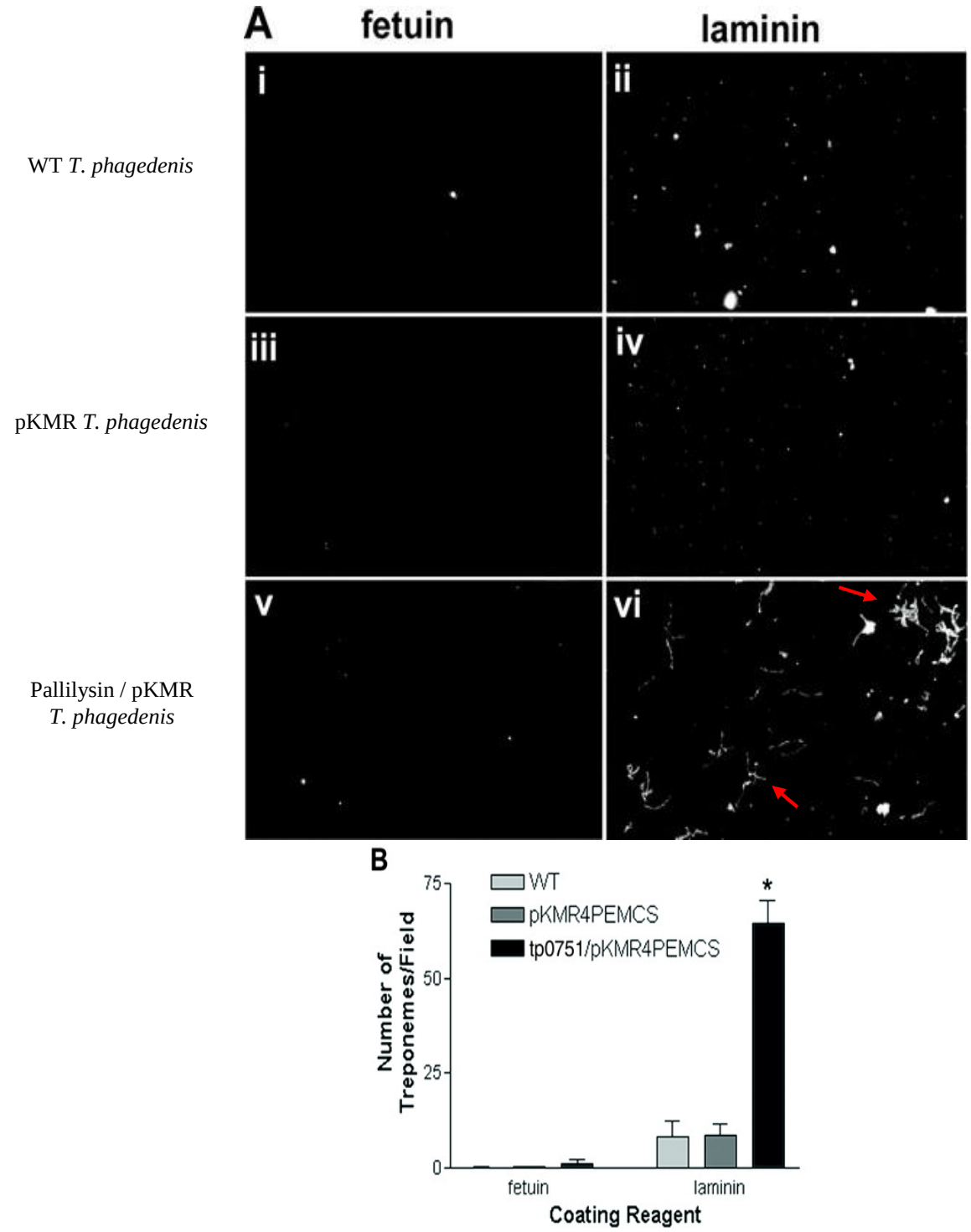


Figure 3. Attachment of WT and transformed *T. phagedenis* to laminin and fetuin.

(A) Slides were covered with either fetuin (negative control) or laminin and spirochetes were visualized via dark-field microscopy with a Nikon Eclipse E600 microscope. (B) Quantitation of *T. phagedenis* attachment to laminin. Student t-test was used to compare the level of attachment of WT *T. phagedenis*, pKMR *T. phagedenis*, and pallilysin / pKMR *T. phagedenis* to laminin with the level of attachment to fetuin (*, $P < 0.0001$). Adapted from (Cameron *et al.*, 2008). Printed with permission.

1.2.4 Challenges of *Treponema pallidum* research

The genome of the syphilis spirochete was fully sequenced by the end of the 20th century (Fraser *et al.*, 1998), yet *T. pallidum* research remains associated with many technical difficulties. Hence, elucidating the virulence of this spirochete and the molecular mechanisms behind its pathogenesis poses significant challenges that have contributed to a slow progress in the syphilis field.

T. pallidum is a slow-growing, persistent and fastidious obligate human pathogen with a minimalistic genome of just over 1000 open reading frames (ORF) (Fraser *et al.*, 1998), which is approximately four times less than that of *E. coli*, and is genetically intractable. Due to its reduced genome, this bacterium relies solely on glycolysis for energy production, since it lacks the components required for the Krebs cycle and to carry out oxidative phosphorylation. It is also incapable of synthesizing most amino acids, fatty acid and enzyme cofactors (LaFond and Lukehart, 2006). Thus, to account for a lack of metabolic and anabolic machineries, it possesses a repertoire of proteins dedicated to import resources from the host environment (Fraser *et al.*, 1998). Indeed, a significant portion, approximately 5%, of the genome of this pathogen encodes for channels and transport systems predicted to have a wide specificity for a variety of nutrients from host origin (Fraser *et al.*, 1998).

The inability of *T. pallidum* to self-sustain and adapt makes this pathogen unable to survive under *in vitro* growth conditions and makes the direct investigations on virulence determinants challenging and almost impossible. *T. pallidum* must be passaged via intradermal or intratesticullar inoculation of rabbits, which is expensive and poses ethical issues. Moreover, even after harvesting bacteria from rabbit models, the spirochetes are so

delicate and susceptible to environmental stress, that they are only viable and adequate for experimentation for a few hours following harvest (LaFond and Lukehart, 2006). The fragile nature of *T. pallidum* can be attributed in great part to the delicate nature of its protein-devoid OM, which can be easily disrupted during centrifugation and other experimental procedures if not handled with care (Cox *et al.*, 1992; Radolf and Norgard, 1988).

Although the delicate nature of *T. pallidum* does hamper syphilis research, it is the inability to successfully culture, and thus genetically manipulate this pathogen that has hindered the field the most. Despite the challenges faced by *T. pallidum* researchers, significant advances in more recent years, such as the identification of novel putative virulence factors, have been achieved. The development of powerful bioinformatic tools, recombinant expression systems and model organisms adapted to heterologously express *T. pallidum* proteins, have made the indirect, yet significant, investigation of *T. pallidum* virulence determinants possible. One such model is the non-pathogenic spirochete *Treponema phagedenis*.

1.3 *Treponema phagedenis*; a model treponeme

Since it is impossible to carry out genetic manipulations on *T. pallidum*, direct investigation of the function of individual gene products is challenging and heterologous expression of specific candidate genes thought to be involved in pathogenesis is sometimes required. Fortunately, the strict anaerobe, *T. phagedenis* is a good candidate for the heterologous expression of *T. pallidum* putative virulence factors since it is an easily cultivable spirochete that is non-adhesive, non-invasive and non-pathogenic to humans (Blanco *et al.*, 1997; Moskophidis and Muller, 1984). Originally isolated from a human

lesion, *T. phagedenis* provides an alternative means to study *T. pallidum* at the molecular level, not only for the above reasons, but also because there are a number of *T. phagedenis* polypeptides that are cross-reactive with *T. pallidum* proteins (Radolf *et al.*, 1986).

Indeed, the shuttle vector pKMR4PEMCS (or for simplicity, pKMR) has been successfully used for the heterologous expression of the *T. pallidum* virulence factor adhesin, pallilysin, in *T. phagedenis.* Thus, the immunofluorescence images depicted in Figure 4, show that pallilysin is appropriately localized on the surface of *T. phagedenis* transformed with *pallilysin/pKMR* (Figure 4 E and F), but not on *T. phagedenis* transformed with the empty vector *pKMR* (Figure 4 G and H). The flagellar protein FlaA, used as a positive control, could only be detected upon treatment with the detergent Triton X-100, ensuring the cellular integrity of the organisms used (Figure 4A-D) (Cameron *et al.*, 2008). Related investigations further showed that unlike wild type (WT) *T. phagedenis*, *pallilysin/pKMR/T. phagedenis* is able to bind to laminin, suggesting that pallilysin confers upon this bacterium the ability to bind to host component-coated surfaces, an interaction that is inhibited by pallilysin-specific serum (Cameron *et al.*, 2008; Houston *et al.*, 2011). These indirect studies involving *T. phagedenis* further support previous experimental observations that suggested surface exposure / secretion of pallilysin, and its interaction with host components (Cameron *et al.*, 2005; Houston *et al.*, 2011; Houston *et al.*, 2012). However, the means whereby this virulence factor is secreted/exported remains to be elucidated.

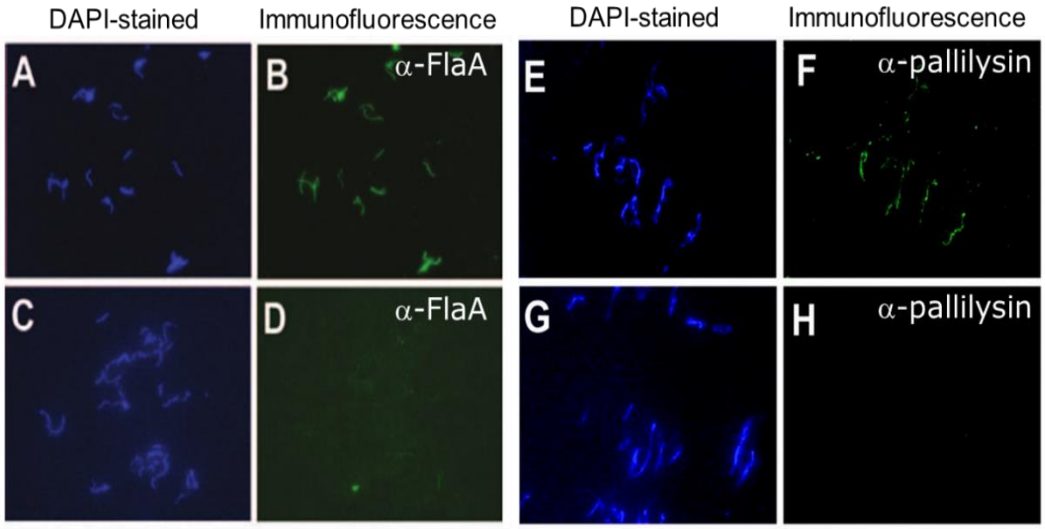


Figure 4. Surface expression of pallilysin on *T. phagedenis*.
(A), (C), (E), (G) show DAPI-stained images of the *T. phagedenis* constructs. (B), (D), (F), (H) show the corresponding immunofluorescence images. (A) and (B): WT *T. phagedenis*, treated with Triton X-100 (positive control). (C) and (D): WT *T. phagedenis*, no Triton X-100 treatment. (E) and (F): *pallilysin/pKMR/T. phagedenis*. (G) and (H): *pKMR/ T. phagedenis*. Adapted from (Cameron *et al.*, 2008). Printed with permission.

1.4 Bacterial secretion systems

Gram-negative bacteria have developed a number of secretion systems employed for the transport of effectors. These macromolecular systems can transport not only virulence factors, but also other substrates like small molecules and DNA (Costa *et al.*, 2015). Secretion machineries are widely present both in disease-causing bacteria, such as those responsible for skin, oral cavity, gastrointestinal tract and sexually transmitted infections, as well as in commensal bacteria, found throughout the body, to fight off competing microbes.

The secretion systems found in Gram-negative bacteria include the general secretory system (Sec), the Twin-arginine translocation (TAT) pathway, the β-barrel assembly machinery (BAM) complex, the localization of lipoproteins (LOL) system and

type I, II, III, IV, V, VI secretion systems (T1SS, T2SS, T3SS, T4SS, T5SS, T6SS). The Sec and TAT pathways are responsible for the transport of proteins across bacterial plasma membranes in the unfolded and folded state, respectively (Natale *et al.*, 2008), and the rest of the secretory machineries are either Sec-dependent or Sec-independent.

1.4.1 Sec and TAT-dependent pathways

Sec and TAT-dependent bacterial secretion systems comprise a two-step secretion mechanism involving a periplasmic intermediate. Thus, the substrate is first recognized, transported across the IM, and delivered to the periplasm via the Sec system or the TAT system (Costa *et al.*, 2015), followed by recognition via another secretory machinery for further transport when required. Periplasmic intermediates that are not meant to remain in the periplasm can then either be positioned in the OM or secreted into the extracellular space, each pathway carried out by different specialized bacterial machineries (Costa *et al.*, 2015).

1.4.1.1 The β-barrel assembly machinery (BAM)

The β-barrel assembly machinery (BAM) is responsible for the appropriate folding and positioning of β-barrel proteins into bacterial OMs (Kim *et al.*, 2012). Upon secretion of substrates into the periplasmic space via the Sec or TAT systems, β-barrel-containing proteins destined to the OM are recognized by the BAM complex (Silhavy *et al.*, 2010). In *E. coli*, this complex is composed of five proteins, as shown in Figure S 1: the OMP BamA, and the four lipoproteins BamB through BamE, which are anchored to the periplasmic side of the OM (Wu *et al.*, 2005). BamA is predicted to possess a C-terminal β-barrel that crosses the OM, and has up to five polypeptide translocation

associated (POTRA) domains at the N-terminus that extend into the periplasm (Kim *et al.*, 2012). These domains are thought to function as interacting sites for BamB, BamC, BamD and BamE, and may have substrate chaperone capabilities (Kim *et al.*, 2007). BamB is thought to be involved in delivering β-barrel precursors to BamA, and BamC seems to play a regulatory role via interactions with BamD (Kim, Aulakh, *et al.*, 2011). Additionally, BamD may be involved in initial substrate recognition via interactions between the C-terminus sequence of substrate OMPs and Tetratricopeptide repeat (TPR) motifs found on BamD (Albrecht and Zeth, 2011; Sandoval *et al.*, 2011). Finally, BamE is thought to play a structural role on the BAM complex as a whole (Sklar *et al.*, 2007), and to improve the insertion of OMP into bacterial OMs (Endo *et al.*, 2011). Hence, all the five components play an important role in proper folding and localization of β-barrel OMPs, and are required at the same ratio to achieve a fully functional BAM complex (Hagan *et al.*, 2010). However, the molecular mechanism whereby the BAM complex is able to properly fold and insert OMPs into the OM remains to be fully elucidated (Kim *et al.*, 2012).

1.4.1.2 Lipoproteins and the localization of lipoproteins (LOL) system

Lipoproteins in Gram-negative bacteria can have a variety of functions, including roles in biogenesis, maintenance of cell surface structures, transport of substrates and pathogenesis (Bernadac *et al.*, 1998; Clavel *et al.*, 1998; Ehrmann *et al.*, 1998; Nikaido, 1998; Okuda and Tokuda, 2011). Thus, the appropriate positioning of lipoproteins is of great importance and different signals dictate their final destination, which can be either the IM or the OM (Schulze and Zuckert, 2006; Tokuda, 2009). Lipoproteins are synthesized as prolipoproteins and contain at the N-terminus a signal peptide with a consensus sequence, Leu-Ala/Ser-Gly/Ala-Cys, called lipobox (Hayashi and Wu, 1990;

Tokuda and Matsuyama, 2004). Upon transport across the IM, most commonly via the Sec system, prolipoproteins become anchored in the outer leaflet of the IM through a reaction carried out by an enzyme called Phosphatidylglycerol/ prolipoprotein diacylglyceryl transferase (Lgt). This enzyme catalyzes the formation of a thioester linkage between a diacylglycerol and a conserved Cys residue at the N-terminus of the prolipoprotein, anchoring the protein to the IM. The signal peptide is then cleaved by a lipoprotein signal peptidase (Lsp, also called signal peptidase II), and the N-terminal Cys residue is acylated by a phospholipid: apolipoprotein transacylase (Lnt) (Sankaran and Wu, 1994). If the mature lipoprotein contains a LOL avoidance signal, it remains anchored to the outer leaflet of the IM (Narita and Tokuda, 2007; Tanaka *et al.*, 2007; Terada *et al.*, 2001). Lipoproteins lacking the LOL avoidance signal are recognized by the localization of lipoproteins (LOL) system and delivered to their final destination in the OM (Okuda and Tokuda, 2011; Zuckert, 2014). The *E. coli* LOL system, depicted in Figure S 2, is composed of the five proteins LolA, LolB, LolC, LolD and LolE. The LolCDE complex is found embedded in the IM and four domains form a functional unit. Hence, a cytoplasmic LolD homodimer forms the ABC transporter, and a LolCE heterodimer spans the IM, extending into the periplasm (Yakushi *et al.*, 2000; Yasuda *et al.*, 2009). Once acylated, the lipoprotein is recognized by LolCE and is transferred to the periplasmic chaperone LolA. Thus, ATP hydrolysis carried out by LolD provides the energy required for the transfer (Okuda and Tokuda, 2011). LolA then interacts with LolB, transferring the lipoprotein to the latter. The exact molecular mechanism whereby LolB delivers lipoproteins to the periplasmic side of the OM remains unknown (Matsuyama *et al.*, 1997; Okuda and Tokuda, 2011). Moreover, the delivery of lipoproteins to the outer leaflet of the OM requires transversal of the surface

through the membrane, and although it is not known how bacteria carry out such a transport, a model involving a flippase complex located within the OM has been proposed. Thus, the lipoprotein becomes anchored in the periplasmic side of the OM, and interacts with a flippase complex that enables translocation through the OM, delivering the protein to its final destination, the bacterial surface (Chen and Zückert, 2011; Schulze *et al.*, 2010).

1.4.1.3 The Type II Secretion System

The T2SS acts in concert with the Sec or TAT system to secrete folded proteins in a two-step manner (Costa *et al.*, 2015). This secretory machinery consists of 12-15 components divided into four parts, an OM complex, a periplasmic pseudopilus, an IM platform and a cytoplasmic ATPase (Costa *et al.*, 2015). During transport across the IM the Sec or TAT system, the N-terminal signal peptide of the substrates is cleaved and the mature protein released into the periplasm where it adopts a folded conformation (Filloux, 2004). Export from the periplasm and into to the extracellular space (or directly into target host cells) occurs upon interaction of the substrate with periplasmic domains of the T2SS, such as the pseudopilus tip, which results in ATPase activity that powers substrate secretion via a piston-like activity of the pseudopilus that is thought to push the substrate across the OM complex. Substrates of the T2SS include lipases, proteases, carbohydrate-degrading enzymes and toxins (Korotkov *et al.*, 2012). Although it seems like a folded conformation is necessary for secretion, the specific interactions between the T2SS and its substrates remain to be elucidated.

1.4.1.4 The Type V Secretion System

The T5SS, also called autotransporter, depends on the Sec or TAT and BAM complexes to cross the inner and outer membranes, respectively. Thus, the multi-domain autotransporters secrete substrates in a two-step manner. A signal peptide found at the N-terminus of autotransporters mediates secretion across the IM via the Sec system. Upon cleavage of the signal peptide and release into the periplasm, the C-terminal β-barrel domain of the autotransporter gets inserted into the OM via the BAM complex (van Ulsen *et al.*, 2014). It is this C-terminal β-barrel (or translocator domain), that acts as a pore whereby the N-terminal passenger domain of the autotransporter is able to cross the OM. The passenger domain adopts a hairpin conformation allowing for transport across the OM in a carboxy- to amino- direction. Once across the OM, the autotransporter undergoes proteolytic cleavage whereby the passenger domain is released from the surface into the extracellular space, or remains attached to the bacterial cell via non-covalent interactions (van Ulsen *et al.*, 2014). Due to the multiple domains that characterize autotransporters, these proteins are generally large, however, the size constraints of the β-barrel pore suggest that the passenger domain remains mostly unfolded during secretion and fold upon contact with the extracellular milieu (Leyton *et al.*, 2011; van Ulsen *et al.*, 2014). The passenger domain of different autotransporters possesses different functionalities including proteolysis, toxicity, adhesion and biofilm formation. Hence, autotransporters are important virulence determinants since they secrete factors that contribute to invasion, colonization and immune evasion of Gram-negative pathogens (van Ulsen *et al.*, 2014).

1.4.2 Sec or TAT-independent pathways

The other group of secretion systems, the T1SS, T3SS, T4SS and T6SS, are independent of the Sec or TAT pathways and thus transport their substrates across both the IM and OM of Gram-negative bacteria, without a periplasmic intermediate (Costa *et al.*, 2015). These machineries are able to localize their respective substrates either on the bacterial OM, into the extracellular milieu, or directly into a target cell, which can be either bacterial or eukaryotic (Gerlach and Hensel, 2007).

1.4.2.1 The Type I Secretion System

The first protein secretion system discovered was the *E. coli* hemolysin A (HlyA) T1SS (Mackman and Holland, 1984), which is one of the few well-characterized T1SSs, along with the Hemophore HasA T1SS from *Serratia marcescens*.

The general T1SS spans the entire Gram-negative bacterial cell envelope, crossing both the IM and OM, and as shown in Figure 5, this system is an oligomeric channel with three multimeric components: an ATP-Binding Cassette (ABC) transporter, a membrane fusion protein (MFP) and a TolC, which is an OMP (Kanonenberg *et al.*, 2013).

The ABC transporter is found embedded within the IM of Gram-negative bacteria and provides energy via ATP hydrolysis required for substrate transport. The structure of ABC transporters consists of four modules from two types of domains, which are the cytoplasmic nucleotide binding domain (NBD), and the IM transmembrane domain (TMD) (Balakrishnan *et al.*, 2001; Kerr, 2002; Zolnerciks *et al.*, 2011). Bacterial ABC transporters consist of two polypeptides, each with a NBD and a TMD, that interact to form a functional transporter (Davidson *et al.*, 2008; Kerr, 2002; Zolnerciks *et al.*, 2011). For instance, the hemolysin ABC transporter, HlyB, in *E. coli*, contains a C-terminal NBD and an N-

terminal TMD, expressed as a single polypeptide which then dimerizes upon ATP binding to form the active conformation of HlyB (Schmitt *et al.*, 2003; Zaitseva *et al.*, 2005; Zaitseva *et al.*, 2006). While the NBD is generally highly conserved amongst transporters, the TMD shows little sequence conservation between TMDs of ABC transporters in different bacteria. Furthermore, the specific domain responsible for ATP binding is the NBD, whereas the TMD is in charge of forming a channel toward the periplasmic space, via six to eight predicted transmembrane segments that span the IM toward contact with the MFP (Dawson and Locher, 2006).

The second component of the T1SS, the MFP, is located within the periplasm and is involved in substrate recognition (Balakrishnan *et al.*, 2001; Nicaud *et al.*, 1985; Thomas *et al.*, 1988; Zhang, Yin, *et al.*, 1995). In the *E. coli* hemolysin transport system, the MFP HlyD has an N-terminal cytoplasmic region with a single transmembrane α-helix that spans the IM and has been shown to interact with T1SS substrates (Balakrishnan *et al.*, 2001; Jorgensen *et al.*, 1980; Moayeri and Welch, 1994). Indeed, *in vivo* cross-linking experiments carried out with *E. coli* harboring HlyD and the substrate HlyA, but not HlyB, showed a HlyD-HlyA interaction (Balakrishnan *et al.*, 2001). Moreover, HlyD seems to affect HlyA folding upon contact, either after or during substrate transition through the system. These studies showed that mutations within the HlyD periplasmic domain affected HlyA translocation and/or final folding (Pimenta *et al.*, 2005). In addition, the MFP component interacts with the TMDs of the ABC transporter and extends toward the OM. There is evidence showing that the MFP forms trimers and hexamers to constitute a functional unit. Indeed, the structure of HlyD was recently solved by X-ray crystallography and suggested a hexameric complex with an α-helical periplasmic domain (Kim *et al.*,

2016), as shown in the model in Figure 5. Finally, the MFP of the T1SS is the component that links the ABC transporter to the TolC.

The third component of the T1SS is the OMP TolC, which forms a channel within the OM through which the substrate gets secreted (Balakrishnan *et al.*, 2001). This OMP can either be part of the T1SS operon, as is the case of the *B. pertussis* CyaA T1SS, which encodes for all the components within a single operon, or can be found elsewhere in the bacterial chromosome, as is the case of the hemolysin T1SS from *E. coli* (Angelos *et al.*, 2003; Linhartová *et al.*, 2010). The solved structure of the *E. coli* TolC showed that this OMP is composed of a short β-barrel embedded in the OM, and long α-helices that extend into the periplasm. Moreover, TolC proteins appear to form homotrimers resulting in a channel of 140Å in length (Koronakis *et al.*, 2000) that forms a pore of maximum 40Å in external diameter, or 20Å in internal diameter, and narrowing to 3.5Å at the periplasmic end (Koronakis *et al.*, 1997; Thanabalu *et al.*, 1998). Although the internal diameter appears to be too small for transport of appropriately folded proteins and even ions, (Delepelaire, 2004) there is evidence showing an iris-like movement that results in the α-helices rearranging into a larger opening of 30Å, which is large enough for substrate transport, including the transport of unfolded proteins that retain some degree of secondary structure (Andersen *et al.*, 2002; Eswaran *et al.*, 2003; Sharff *et al.*, 2001).

The assembly mechanism of the three components of the T1SS is not fully understood. However, two of the *E. coli* hemolysin T1SS components, the ABC transporter and the MFP, have been shown to be essential for the secretion of HlyA, since they provide substrate specificity prior to secretion (Nicaud *et al.*, 1985; Zhang, Yin, *et al.*, 1995). This results in the recruitment of the TolC, leading to the finalized assembly of the T1SS

(Balakrishnan *et al.*, 2001; Thomas *et al.*, 1988; Zhang, Yin, *et al.*, 1995). Moreover, early studies using the substrates PrtC and HasA suggested that the ABC and MFP are able to interact with each other in the absence of TolC. Additionally, the ABC transporter and MFP associate with TolC only upon substrate binding. During substrate secretion however, the TolC is found bound to the complex (Letoffe *et al.*, 1996; Thanabalu *et al.*, 1998).

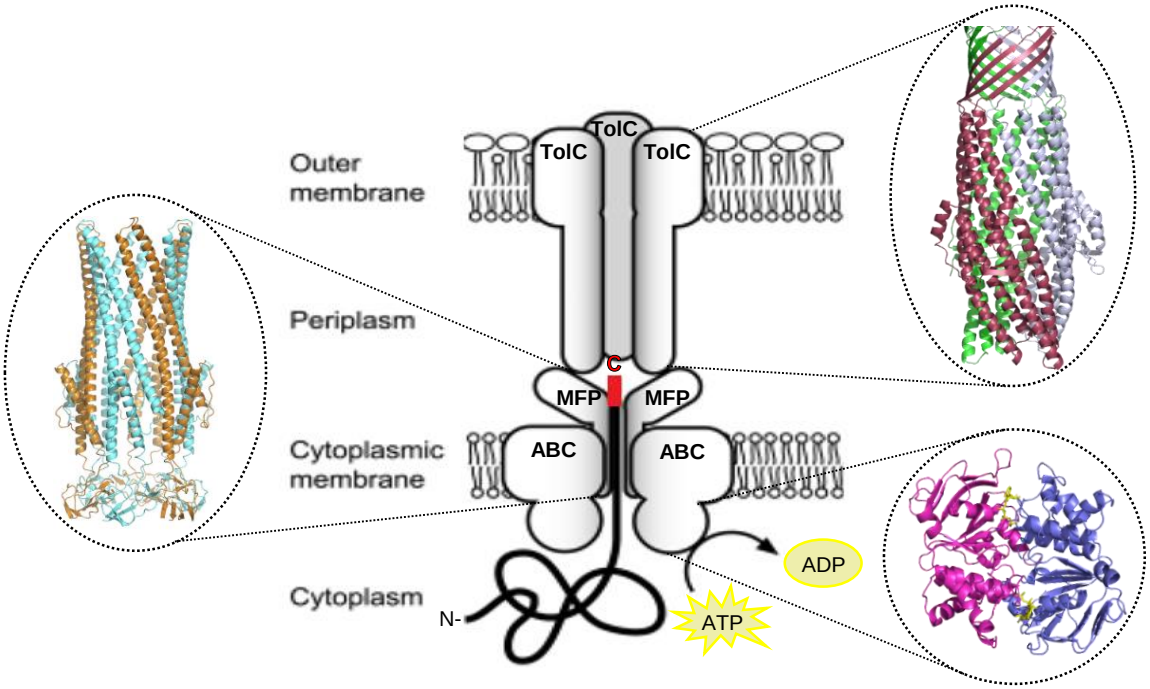


Figure 5. Model of the general structure of a Type 1 Secretion System (T1SS). Adapted from (Linhartová *et al.*, 2010). The solved crystal structures of all the components of the T1SS are shown, from top to bottom: the trimeric conformation of TolC (Koronakis *et al.*, 2000), hexameric MFP, adapted from (Kim *et al.*, 2016), and dimeric NBD of the ABC transporter (Jumpertz *et al.*, 2005). Printed with permission.

1.4.2.1.1 Mechanism of substrate secretion by the T1SS

Translocation of polypeptides via the T1SS is achieved in a single-step across both the inner and outer membranes, from the cytoplasm into the extracellular environment without a periplasmic intermediate (Bakás Laura, 2012; Mackman *et al.,* 1985; Oropeza-Wekerle *et al.,* 1989). Most known T1SS substrates are released from within the bacterial cell and into the host extracellular environment, however there are exceptions where the substrate remains loosely associated with the bacterial cell surface. There is evidence, for instance, that the adhesin SiiE from *Salmonella enterica* can exist as an exoprotein and remain loosely associated to the bacterial OM until host cell contact, which triggers bacterial cell surface retention of SiiE (Gerlach *et al.,* 2007). Similarly, BapA (a protein also from *Salmonella,* that is required for biofilm formation), and the adhesin LapA from *P. fluorescens* are both secreted extracellularly and remain in a loose association with the bacterial cell surface (Hinsa *et al.*, 2003; Latasa *et al.*, 2005). Finally, HlyA has been shown to associate with the OM of *E. coli* and form outer membrane vesicles (OMV), as shown in Figure 6, that are released into the host environment carrying more than one HlyA per vesicle (Balsalobre *et al.,* 2006).

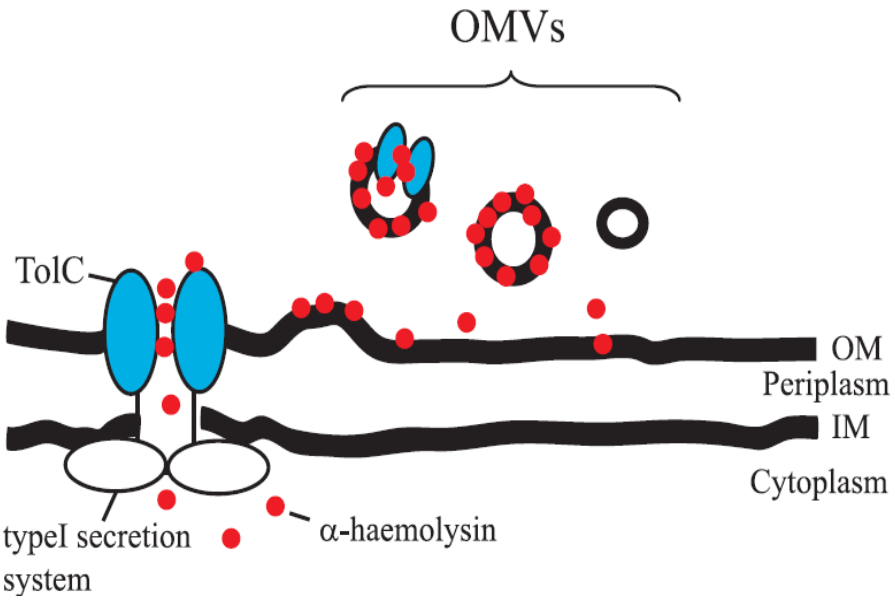


Figure 6. Model of the secretion of HlyA in Outer Membrane Vesicles (OMVs) via the *E. coli* T1SS.
Adapted from (Balsalobre *et al.*, 2006) and printed with permission.

1.4.2.1.2 The Type 1 Secretion System substrates

In pathogenic Gram-negative bacteria, T1SSs are known to secrete virulence factors with a wide range of functions. Some of the known T1SS substrates are mentioned in Table 1 and include toxins, adenylate cyclases, adhesins, leukotoxins, nodulation proteins, lipases, nutrient acquisition factors and proteases, among others (Lenders *et al.*, 2015). These T1SS substrates vary greatly in size, from small proteins such as *E. coli* 5.8 kDa bacteriocin Colicin V or the 19kDa HasA, to the 900kDa adhesion factor LapA from *P. fluorescens* (Hinsa *et al.*, 2003; Letoffe *et al.*, 1994; Satchell, 2011). The known T1SS substrates are secreted in an unfolded conformation (or partially folded), and adopt a stable structure once in contact with the extracellular milieu (Holland *et al.*, 2005). It is not completely understood how the substrates are able to remain unfolded within the bacterial cell and then spontaneously fold appropriately once translocated. However, there are a few theories thought to be involved in misfolding prevention, such as the presence of

a substrate binding protein (SBP), a C39-like domain (CLD) on the ABC transporter, or ion concentration differences between the intra and extra-cellular environments (Delepelaire and Wandersman, 1998; Lecher *et al.*, 2012; Linhartová *et al.*, 2010).

In *S. marcescens,* chaperone SecB functions as a SBP, which is able to interact with the T1SS substrate HasA and is essential to maintain it in the unfolded conformation (Delepelaire and Wandersman, 1998). Indeed, studies have shown that if HasA is allowed to fold in the cytoplasm, secretion is inhibited (Delepelaire and Wandersman, 1998). A CLD on the ABC transporter of a T1SS, on the other hand, is found in the hemolysin system, where it tethers HlyA and prevents its aggregation and/or degradation during secretion (Lecher *et al.*, 2012). Finally, it has been shown that certain types of T1SS substrates interact with divalent cations, which are present in higher concentrations in the extracellular environment (in comparison to the lower concentrations found within bacterial cells), and lead to a stable, active substrate conformation (Linhartová *et al.*, 2010).

Table 1. Examples of T1SS substrates in Gram-negative bacteria.

Organism	Virulence Factor	Function	Reference
Escherichia coli	HlyA	Toxin	(Bailey *et al.*, 1992; Bakkes *et al.*, 2010)
Escherichia coli	Colicin V	Bacteriocin	(Gérard *et al.*, 2005)
Bordetella pertussis	CyaA	Adenylate cyclase-toxin	(Basler *et al.*, 2007)
Salmonella enterica	SiiE	Adhesin	(Morgan *et al.*, 2007)
Pseudomonas fluorescens	LapA	Adhesin	(Hinsa *et al.*, 2003)
Mannheimia haemolytica	LktA	Leukotoxin	Davies et al., 2002)
Rhizobium leguminosarum	NodO	Nodulation protein	(Scheu *et al.*, 1992),
Serratia marcescens	LipA	Lipase	(Akatsuka *et al.*, 1997),
Serratia marcescens	HasA	Nutrient (iron) acquisition factor	(Arnoux *et al.*, 1999)
Erwinia chrysanthemi	PrtB, PrtC	Protease	(Hege and Baumann, 2001)

1.4.2.1.2.1 Repeats in toxin (RTX) proteins

Repeats in toxin (RTX) proteins are a highly diverse family of proteins secreted by the T1SS. These T1SS substrates possess a variable number of tandem repeats with the typically nonapeptide glycine-rich sequence consensus Gly-Gly-X-Gly-X-Asp-X-U-X (where X is any amino acid and U is a nonpolar amino acid). The number of tandem repeats within these proteins varies greatly, from less than 5 to more than 40 repeats (Barlag and Hensel, 2015; Lenders *et al.*, 2015; Linhartová *et al.*, 2010; Meier *et al.*, 2007). Even though over 1000 RTX family members have been identified, only a few biological functions have been uncovered. The RTX toxins that have been characterized function as pore-forming leukotoxins, hemolysins, multifunctional enzymatic toxins, and hydrolytic enzymes (such as proteases and lipases), among others (Linhartová *et al.*, 2010; Oropeza-Wekerle *et al.*, 1989).

Additionally, it has been shown that these T1SS substrates, the RTX proteins, have the ability to bind calcium ions (Baumann et al., 1993) via both high and low affinity calcium binding sites (Linhartová et al., 2010). The calcium concentration within a bacterial cell is generally low, in the μM range, whereas extracellular calcium is in the mM range (Lecher *et al.*, 2012). As shown in Figure 7, it is thought that in the case of many T1SS substrates, substrate–calcium interactions prevent misfolding within the bacterial cell and promote secretion upon initial contact of the substrate with the extracellular space. Calcium ions interact with calcium-binding sites within the substrate, pulling away from the secretion complex and triggering/ aiding in substrate folding toward a functional conformation within the extracellular environment (Baumann *et al.*, 1993; Linhartová *et al.*, 2010). Hence, these ions are essential for the toxins' stability and functionality.

RTX proteins are generally pore-forming toxins and have a few characteristics in common: they are secreted via a T1SS, their activation requires both the amide-linked fatty acylation of internal lysine residues as well as interaction with calcium ions once exposed to the extracellular environment, and they contain a hydrophobic domain thought to be responsible for the formation of pores in target cell membranes (Linhartová *et al.*, 2010).

Pore-forming RTX cytotoxins are produced as inactive protoxins that then undergo activation via acylation within the bacterial cytosol, prior to export. The acylation reaction is catalyzed by acyltransferases that are co-expressed with the protoxin substrate (Linhartová *et al.*, 2010). In *E. coli*, for instance, the acyltransferase HlyC is able to use fatty acyl residues brought by an acyl-ACP protein to acylate HlyA at Lys 540 and Lys 648 (Linhartová *et al.*, 2010; Stanley *et al.*, 1994). *Bordetella pertussis* CyaA is another example of a well-studied acylated RTX toxin; this substrate becomes acylated on Lys860 and Lys983, and as expected, this post-translational modification (PTM) is necessary for CyaA activity. Thus, acylation confers upon CyaA its full capacity to bind to its $\alpha_M\beta 2$ integrin receptor (CD11b/CD18) triggering the toxin's integration into host cell membranes (Basar *et al.*, 2001; Masin *et al.*, 2005).

Although RTX toxin acylation is not essential for secretion, there is evidence of this PTM being critical to activating the toxicity of these proteins. For instance, it has been shown that acylation activates HlyA (Linhartová *et al.*, 2010; Stanley *et al.*, 1994) and is necessary for the hemolytic activity of the toxin, since it is the acyl chains that enable the toxin to penetrate the host cell membrane leading to pore formation (Stanley *et al.*, 1998). However, the exact role of this PTM in the mechanism of action of the toxins is not fully elucidated. Indeed, experiments showed that the non-acylated proHlyA and proCyaA are

able to form pores in lipid bilayers and in liposome membranes, although less effectively than their acylated counterparts. Hence although it seems like acylation is not absolutely required for toxin penetration, it may mediate the oligomerization required for the RTX toxins to bind target cell receptors, thereby leading to cytotoxicity (El-Azami-El-Idrissi *et al.*, 2003; Masin *et al.*, 2005).

Indeed, recent studies suggested that the highly potent RTX toxins, such as CyaA and HlyA, not only can recognize and bind the N-linked oligosaccharide of their β_2 integrin receptors, but also show detectable activity on cells other than their main target (e.g. cells other than erythrocytes in the case of HlyA), regardless of the absence of the specific protein receptors (such as $\alpha_M\beta 2$ integrin on leukocytes, which is used by CyaA) (Chenal *et al.*, 2009; El-Azami-El-Idrissi *et al.*, 2003).

Nonetheless, the exact mechanism whereby RTX toxins insert to form a pore within the host cell remains poorly understood, and whether pore formation requires oligomerization is controversial. The initial interaction between the RTX toxin and the host cell required for pore-formation is possible due to the ability of these toxins to interact with the plasma membrane itself. Hence, the toxin first adsorbs reversibly to the plasma membrane via electrostatic forces and then inserts irreversibly (i.e. it cannot be recovered from the membrane without detergent treatment (Linhartová *et al.*, 2010)) into the plasma membrane (Ostolaza *et al.*, 1997). Studies with HlyA showed that this toxin coordinates the adsorption to host cells through its calcium binding domains (Sanchez-Magraner *et al.*, 2007), other investigations suggest that a hydrophobic region within the toxin is responsible for the insertion event, and others that the toxins form cation-selective pores of a defined size and for only a few seconds (Hyland *et al.*, 2001; Schmidt *et al.*, 1996).

RTX proteins are the most abundant and diverse T1SS substrates and hence there are many predicted RTX proteins that remain to be characterized. Much is left to be uncovered regarding their different mechanisms of action.

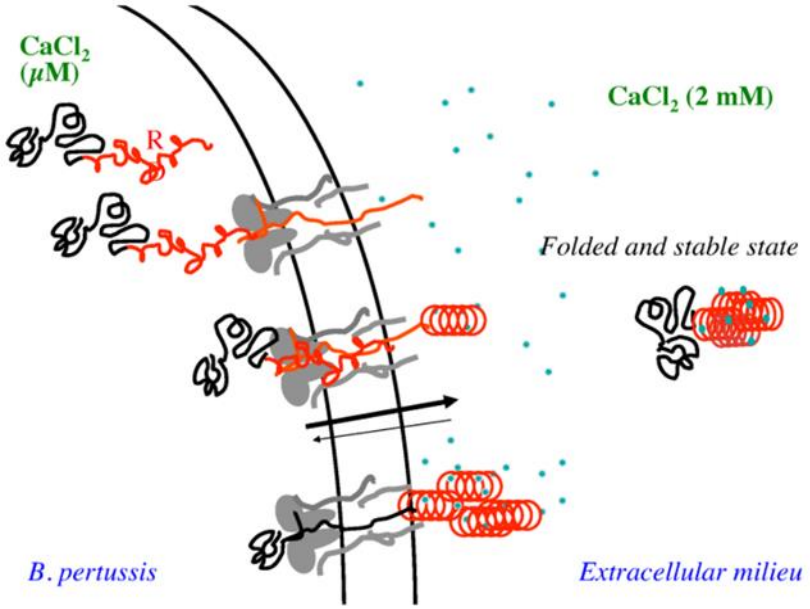


Figure 7. Model of the *B. pertussis* T1SS substrate, CyaA, an RTX toxin
The lower calcium concentration within the bacterial cell prevents premature substrate folding. Upon contact with higher calcium concentrations in the extracellular milieu, RTX toxins adopt a folded and stable conformation (Sotomayor-Pérez *et al.*, 2015).

1.4.2.1.2.2 Non-RTX proteins

The other group of T1SS substrates lack RTX repeats and hence are referred to as non-RTX proteins. These non-RTX proteins lack consensus sequences or unique characteristics, making them more challenging to identify and resulting in their poor characterization compared to RTX proteins.

Thus, only two T1SS substrates that are non-RTX proteins have been studied in detail, the hemophore HasA from *S. marcescens* and the bacteriocin Colicin V from *E. coli*.

HasA is secreted via the T1SS and is able to take heme from other heme-binding proteins within the host in order to transfer it to a receptor on the bacterial surface leading to its uptake (Létoffé *et al.*, 1994). This virulence factor appears to differ from other known T1SS substrates in that, prior to secretion, it interacts with the cytoplasmic chaperone, SecB. SecB aids in maintaining HasA in an unfolded, or loosely folded conformation necessary for secretion. Indeed, *in vitro* experiments have shown that in the absence of SecB, HasA undergoes quick and spontaneous folding, highlighting the crucial role of this chaperone in HasA secretion (Wolff *et al.*, 2003). Moreover, a delayed expression of the ABC transporter and MFP components of the *S. marcescens* T1SS leads to the accumulation of HasA in the cytoplasm, resulting in premature folding and secretion inhibition, thus suggesting that synthesis and secretion of this non-RTX T1SS substrate are coupled activities (Debarbieux and Wandersman, 2001). Furthermore, studies have suggested that in addition to the C-terminal secretion signal, HasA possesses extra T1SS-interacting, primary recognition, sites that become exposed via the HasA-SecB interaction, as shown in Figure 8 (Masi and Wandersman, 2010). Additionally, the first ten residues at the N-terminus of HasA appear to be important for efficient secretion. Indeed, experiments have shown that HasA lacking the C-terminal secretion signal retains its ability to interact with the ABC transporter, and thus trigger the full assembly of the T1SS (Cescau *et al.*, 2007; Sapriel *et al.*, 2002). Interestingly, and differing from most T1SS substrates, the C-terminus of HasA was originally thought to undergo unusual random cleavage at several sites during secretion (Letoffe *et al.*, 1994). More recent studies however, have shown that hemophores do undergo cleavage but it occurs in the extracellular medium, and likely through the activity of bacterial proteases. Regardless, both cleaved and uncleaved

hemophores seem to properly fold upon secretion, and retain heme-binding capabilities (Létoffé *et al.*, 2000).

The other well-studied non-RTX T1SS substrate, Colicin V, is a small antibiotic agent, a bacteriocin produced and secreted by some Gram-negative organisms to kill other bacteria with a similar cellular structure that compete for resources. This peptide antibiotic is able to kill competing bacteria by disrupting their membranes (Belkum *et al.*, 1997; Gérard *et al.*, 2005). Colicin V gains entry to bacterial cells by interacting with surface receptors dedicated to nutrient uptake, followed by active transport into the cell (Wu *et al.*, 2012). Bactericidal activity is only observed once Colicin V gains access to the bacterial inner membrane from the periplasmic space (Zhang, Fath, *et al.*, 1995), and it is thought that the formation of a disulfide bond within the oxidizing environment of the periplasm may be responsible for the polypeptide activation (Havarstein *et al.*, 1994). Interestingly, bacteria that produce Colicin V also produce a Colicin V immunity protein that protects them from its bactericidal activity (Wu *et al.*, 2012).

Although bacteriocins are secreted via the T1SS, they seem to differ from other T1SS substrates in that these polypeptides are virulence factors produced to target competing bacteria rather than a host. Additionally, bacteriocins contain an atypical T1SS signal located at the N-terminus that is cleaved during secretion, rather than the typical uncleaved C-terminus signal sequence found in most T1SS substrates (Wooldridge, 2009).

The T1SS responsible for Colicin V secretion is thought to bind the substrate via a substrate-binding site located within the ABC transporter component. This component possesses a Ca^{2+}-dependent cysteine proteolytic domain and ATP hydrolysis by the ATPase drives the cleavage of the leader peptide from the bacteriocin precursor and

simultaneous translocation (Havarstein *et al.*, 1995). Indeed it has been shown that both the NBD of the ABC transporter and the protease domain are crucial for Colicin V secretion (Wu and Tai, 2004; Zhong and Tai, 1998). The protease domain is thought to be part of the C39 peptidase superfamily containing the Cys-His-Asp sequence that acts as a catalytic triad for proteolysis, which results in the mature Colicin V ready to lyse competing bacterial cells (Wu *et al.*, 2012).

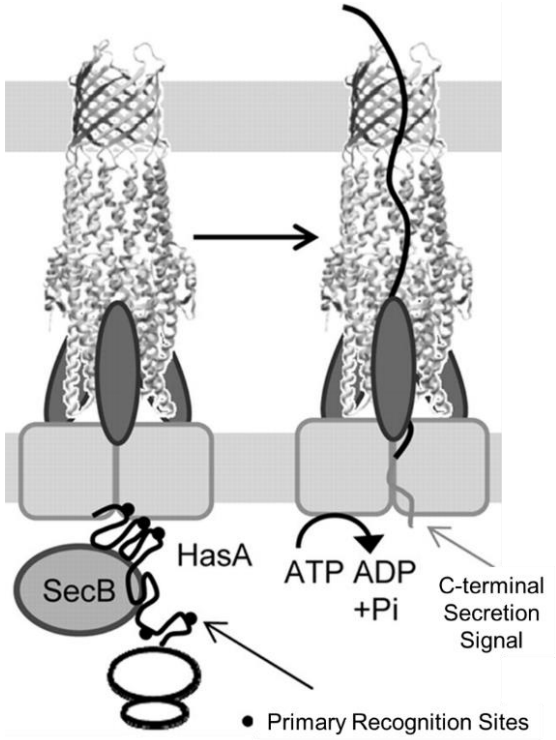


Figure 8. Model of T1SS secretion of *S. marcescens* HasA, a non-RTX protein. The C-terminal secretion signal and the primary recognition sites are shown, adapted from (Masi and Wandersman, 2010). Printed with permission.

1.4.2.1.3 The Type 1 Secretion signal

Secretion via the T1SS begins with the interaction of the substrate with the ABC and MFP components of the T1SS (Balakrishnan *et al.*, 2001; Thomas *et al.*, 1988; Zhang, Yin, *et al.*, 1995). Substrates of this system most commonly possess a C-terminal secretion signal, which differs from other secretory proteins and suggests that secretion occurs post-translationally. Additionally, this C-terminal sequence remains uncleaved throughout the secretion process.

The less common T1SS substrates bacteriocins differ from the canonical T1SS substrates in that they harbor their secretion signal at the N-terminus. This signal consists of a 15-30 residue leader peptide that is cleaved during transport at a conserved double glycine motif (Kanonenberg *et al.*, 2013). Indeed, the ABC transporter component of bacteriocin secretion systems usually possesses C39 peptidase activity that functions to cleave the leader N-terminal peptide characteristic of these substrates, resulting in substrate maturity (Dirix *et al.*, 2004; Lecher *et al.*, 2012).

Thus, generally there is no structure or sequence consensus and no real conservation between T1SS signals from different bacteria (Boyd *et al.*, 2014). However, a few characteristics are observed in a great number of substrates which is useful when screening for potential T1SS substrates. There is evidence showing that the T1SS signal is amphipathic in nature and located within the last 60 to 80 amino acids (Boyd *et al.*, 2014). The *E. coli* HlyA has been one of the few T1SS substrates studied in detail and is often used as a baseline to compare other putative T1SS substrates, since, as shown in Figure 9, HlyA possesses a clear amphiphilic α-helix within the last 60-70 residues at its C-terminus (Koronakis *et al.*, 1989).

…NKVSYVYGHDASTYGSQDN**LNPLINEISKIISAAGSF**DVKEERSAASLLQ
LSGNASDFSYGRNSITLTASA

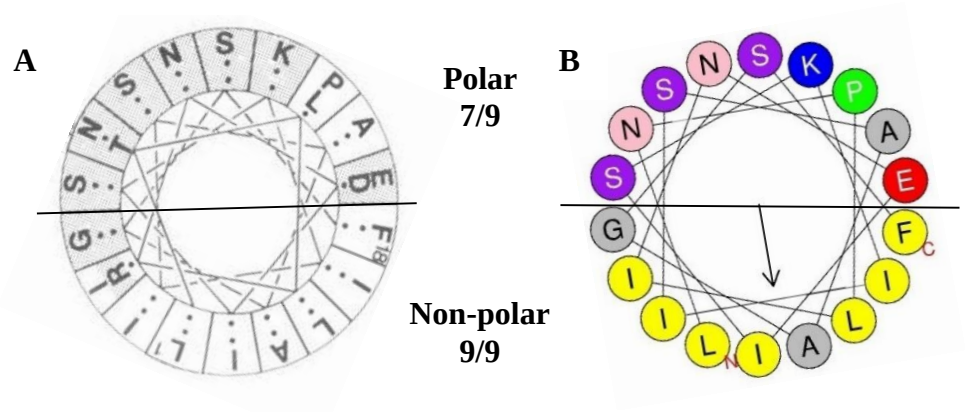


Figure 9. Amphiphilic α-helical wheel diagram of the consensus T1SS substrate, HlyA, from *E. coli.*

(A) Amphiphilic α-helix from *E. coli* HlyA, adapted from (Koronakis *et al.*, 1989), printed with permission, (B) Amphiphilic α-helix from *E. coli* HlyA as predicted by the software, HeliQuest Analysis. Note that diagrams were made using the underlined, bold portion from the HlyA sequence

1.4.2.1.4 Hybrid T1SS and cross-secretion

The lack of primary sequence conservation among T1SS substrates makes it challenging to identify and characterize T1SS substrates. However, this lack of conservation seems to allow for secretion of non-endogenous substrates by other Gram-negative bacteria harboring a T1SS. Hence, the type 1 secretion signal seems to allow a certain degree of cross-complementation in which a T1SS from a specific organism is able to secrete T1SS substrates from another Gram-negative bacterium. For instance, there is evidence that the hemolysin system from *E. coli* can secrete RTX toxins from other Gram-negative bacteria such as CyaA (Masure *et al.*, 1990) and LktA (Chang *et al.*, 1989), both substrates carrying their endogenous type I secretion signal. Additionally, if the secretion signal of the *E. coli* HlyA is exchanged for that of LktA, secretion of HlyA still occurs (Zhang *et al.*, 1993). The nodulation protein, NodO, can similarly be secreted via the

hemolysin and Prt systems from *E. coli* and *E. chrysanthemi*, respectively (Scheu *et al.*, 1992). The Prt T1SS from *Erwinia* has also been shown to secrete *E. coli*'s Colicin V (Fath *et al.*, 1991), and the substrate PrtC can be secreted by the Lip exporter from *S. marcescens* (Akatsuka *et al.*, 1997). Conversely, this cross-secretion between Gram-negative bacteria T1SSs and their substrates is weaker in other cases, such as the low PrtC secretion levels observed by the Hly transporter (Guzzo *et al.*, 1991). Nonetheless it is important to note that a low level of secretion still occurs, suggesting that substrate recognition does occur to a certain extent (Wooldridge, 2009).

1.4.2.2 The Type III Secretion System

The T3SS is a molecular syringe able to inject bacterial effector proteins directly into host cells (Beeckman and Vanrompay, 2010). This system is composed of approximately 25 proteins that together span the bacterial IM and OM, as well as the host cellular membrane. The basal body of the system consists of stacked rings that span both bacterial membranes, and a cytoplasmic IM-associated ATPase provides the energy required for substrate transport (Beeckman and Vanrompay, 2010). A needle-like structure projects from the bacterial surface toward the host cell and upon contact, a pore is built within the host cell membrane (Marlovits *et al.*, 2006; Marlovits *et al.*, 2004) through which T3SS effectors can be directly inserted into the host cytoplasm. The T3SS signal required for effector recognition and subsequent transport has been a matter of controversy. There is evidence suggesting that T3SS effectors possess an N-terminal sequence of 30 residues in length responsible for T3SS targeting (Samudrala *et al.*, 2009). However, other evidence supports the presence of a signal sequence in the mRNA rather than the polypeptide sequence (Sorg *et al.*, 2005). Once inside the host cell, these effectors disturb

host cell functions to favor pathogen invasion, colonization and survival within the host (Beeckman and Vanrompay, 2010; Costa *et al.*, 2015).

1.4.2.3 The Type IV Secretion System

The T4SS is unique among other secretion systems in that not only can it transport proteinaceous virulence factors, but it is also capable of transporting DNA into the environment, or directly into bacterial or eukaryotic target cells (Costa *et al.*, 2015) (Wallden *et al.*, 2010). This system is composed of 12 proteins that make up an OM complex connected to an IM complex by a central elongated structure, a stalk (Costa *et al.*, 2015). Additionally, a pilus extends into the extracellular space and three cytoplasmic ATPases provide the energy required for transport (Trokter *et al.*, 2014). Preparation of DNA for transport is carried out by accessory proteins that unwind, nick and deliver the DNA to the T4SS ATPase components that may lead to subsequent transport (Cascales and Christie, 2004; Wallden *et al.*, 2010). Thus, the T4SS mediates conjugation of DNA which results in the distribution of antibiotic resistance genes between bacteria (Costa *et al.*, 2015). Additionally, protein virulence factor secretion occurs in a similar way and directly into the cytoplasm of target cells, in a contact-dependent manner (Wallden *et al.*, 2010). Much about the T4SS remains to be elucidated, including the exact mechanism of substrate transport and interacting sites between the substrate and the system.

1.4.2.4 The Type VI Secretion System

The T6SS is one of the most recently discovered bacterial secretory machineries and it is structurally related to the injection apparatus of phage. This secretion system is composed of three parts, an IM complex, a nanotube structure built with hemolysin

coregulated protein (Hcp) monomers, and a spike-like feature composed of a trimer of valine–glycine repeat protein G (VgrG) at the tip of the complex (Pukatzki *et al.*, 2009). The IM complex appears to be responsible for the secretion of the VgrG, followed by Hcp monomers that assemble within the periplasm into hexameric rings that stack under the VgrG trimer to form the nanotube structure. Upon full elongation, the nanotube spans the periplasm and extends through the OM, to the extracellular milieu and into the host cell cytosol. It is thought that the VgrG structure at the tip of the complex is able to puncture the host cell membrane in a manner that appears to be a similar mechanism to that used by phage to insert DNA into bacterial cells. Once inserted into the host cell, the T6SS can deliver important virulence factors (Pukatzki *et al.*, 2009). Indeed, the T6SS has been involved in virulence such as lung infection, cell adherence and actin cross-linking (Pukatzki *et al.*, 2006). However only a few candidate T6SS substrates have been identified and the molecular mechanism of secretion remains to be fully elucidated (Pukatzki *et al.*, 2009).

1.5 Secretion systems in *T. pallidum*

Bacterial secretion systems are an important means through which bacteria interact with their environment and are an important field of study since they can be virulence determinants. Thus, understanding their mechanism of action can lead to the development of improved therapeutics and preventative vaccines. *T. pallidum* research has shown that this pathogen possesses a BamA homolog (Tp0326) suggesting the presence of a BAM complex that may be responsible for positioning rare OMPs into the OM (Desrosiers *et al.*, 2011). Interestingly, bioinformatic analyses of the *T. pallidum* proteome failed to identify homologs of all the LOL system components. *T. pallidum* is predicted to

possess a potential chaperone LolA (Tp0333), and a putative inner membrane complex LolCDE (Tp0580, Tp0581, Tp0582), but no LolB has been identified. Hence, the molecular mechanism whereby *T. pallidum* transports and inserts lipoproteins into the OM remains to be elucidated. Despite efforts, no other secretion systems have been identified and therefore the molecular mechanism whereby *T. pallidum* may secrete virulence factors that aid in its invasiveness and overall pathogenesis remains unknown.

1.5.1 Putative Type I Secretion System in *Treponema pallidum*

Bioinformatic analyses were recently carried out to investigate the Tp34 operon from *T. pallidum,* which consists of genes *tp0959* through *tp0972.* Thus, *tp0959* through *tp0961* are predicted to encode flagellar proteins, *tp0962* through *tp0969* gene products are predicted to form a transport complex, and the small coding region of *tp0970* is thought to be untranslated. Finally, the *tp0971* and *tp0972* gene products are proposed to be involved in metal homeostasis. Indeed, there is experimental evidence showing that Tp0971 (also called Tp34) binds specifically to the iron-carrying protein lactoferrin, suggesting a role in iron acquisition. Additionally, Tp0972 *in silico* predictions showed sequence similarity to the iron/lead transporter superfamily of Fe^{3+}/Pb^{2+} further supporting this potential role (Brautigam *et al.*, 2012).

The predicted transport system encoded by *tp0962* through *tp0969,* has been suggested to be involved in the export of small molecules such as lipids and peptides (Brautigam *et al.*, 2012). The Cameron lab proposes a more refined function, since this operon harbors all the components of a putative T1SS (Houston and Cameron, unpublished data): two permeases (*tp0962* and *tp0963*), an ATPase (*tp0964*), a membrane fusion protein (*tp0965*), and four TolC-like proteins (*tp0966* through *tp0969*) (Brautigam *et al.*,

2012). The putative *T. pallidum* T1SS differs from the canonical T1SS in that it is composed of eight proteins, instead of three. Thus a total of three proteins are predicted to form the ABC transporter found within the IM: two permeases (Tp0962 and Tp0963, which make up the TMD components), and a cytoplasmic ATPase (Tp0964, which contains the NBD). A single type of periplasmic protein is predicted to be the MFP (Tp0965); and four TolC-like proteins are predicted to form the OM channel (Tp0966, Tp0967, Tp0968, Tp0969).

The unusual number of components predicted to form this transport system is consistent with a "mix-and-match" mechanism that is believed to be adopted by *T. pallidum* (Brautigam *et al.*, 2012). Due to its minimalistic genome, it would be most efficient for this pathogen to possess multiple permeases and TolC-like proteins that could assemble in different combinations to form variations of a T1SS. This would represent a successful strategy to survive with its reduced genome, while keeping the versatility required for the recognition and secretion of multiple substrates using the same core structure (Brautigam *et al.*, 2012).

Indeed, some organisms have been reported to harbor T1SSs that possess additional features or differ from the traditional three component complex, such as the T1SS harbored by *Vibrio.* This system is responsible for the secretion of a subgroup of very large RTX proteins, the multifunctional autoprocessing RTX toxins (MARTX), and is composed of four proteins rather than three: two ABC transporter components RtxB and RtxE (each one with a NBD and a TMD), a MFP RtxD and a TolC-like protein. Surprisingly, both ABC transporter proteins seem to be necessary for protein secretion

(Boardman and Fullner Satchell, 2004), suggesting that although the basic core structure of most T1SSs is conserved, there are some systems that show compositional differences.

This putative non-canonical *T. pallidum* T1SS would be the first of its kind and its characterization would represent a breakthrough in the syphilis field since it would lead to a better understanding of the invasion and dissemination mechanisms employed by this pathogen. Moreover, the investigation and characterization of pathogen secretion systems and their substrates has been shown to be useful in the identification of vaccine targets for the prevention of other diseases. Indeed, evidence suggests that immunization against OM components and secretion systems substrates can provide protection against infection (Baldi *et al.*, 2012; Notti and Stebbins, 2016; Shrivastava and Miller, 2009). Thus, the characterization of the putative *T. pallidum* T1SS and potential substrates could contribute to the development of a successful vaccine as a preventative measure against syphilis infection.

1.5.1 Pallilysin, a potential *T. pallidum* T1SS substrate

The virulence factor, pallilysin, has been shown to be secreted from pallilysin-expressing *T. phagedenis* and *T. pallidum*, however, the complex in charge of its secretion remains to be elucidated. There is evidence suggesting that pallilysin is lipidated within *T. pallidum* and bioinformatic predictions of pallilysin previously indicated that this protein possesses a signal peptidase II cleavage site and thus suggested a putative lipidation site at Cys^{24} (Setubal *et al.*, 2006). Follow-up experiments with the culturable and pallilysin-expressing, *T. phagedenis*/pKMR/pallilysin model treponeme, tailored to confirm these *in silico* predictions, showed the palmitoylation of pallilysin in *T. phagedenis* (Houston *et al.*, 2011). These results further suggested the export of pallilysin and its membrane-

association within *T. phagedenis* and indirectly provided supporting evidence for membrane-association within *T. pallidum* (Houston *et al.*, 2011).

The putative *T. pallidum* secretory machineries identified thus far do not convey a complete picture on how virulence factors could get secreted during infection. Moreover, the presence of an incomplete LOL system within this pathogen, may suggest an alternative secretion mechanism involved in lipidated virulence factor secretion.

1.6 Research hypotheses and objectives

Uncovering the mechanisms whereby *T. pallidum* is able to rapidly disseminate and invade all of its host's tissues has proven to be challenging. However, the identification of virulence factors that may play an important role in pathogenesis has been a significant step toward defining key vaccine targets. Moreover, the bioinformatic characterization of a potential T1SS in *T. pallidum* represents an important breakthrough in elucidating the potential mechanisms for secretion of these virulence factors and provides an important direction to understanding the invasion mechanisms of this pathogen, hence bringing us a step closer to designing a successful preventative vaccine.

As mentioned above, the MFP component of other T1SSs, has been shown to provide substrate recognition (Balakrishnan *et al.*, 2001). Hence, the *T. pallidum* MFP, Tp0965, may interact with potential T1SS substrates. The putative T1SS may thus play a key role in exporting virulence factors, thereby contributing to the unique invasiveness and pathogenesis of this spirochete.

The objective of the studies presented in this thesis is to begin the characterization of the putative T1SS from *T. pallidum* by investigating the possible interaction between the previously identified *T. pallidum* virulence factor, pallilysin, and the MFP component

of the putative T1SS, Tp0965. These experiments are intended to shed light on the possibility of this pathogen using a T1SS to transport virulence factors into the host environment to aide in dissemination.

Moreover, since it has been shown that the closely related model treponeme *T. phagedenis* is able to heterologously express pallilysin and appropriately localize it on its surface, combined with the fact that some secreted T1SS substrates remain associated with the external face of the OM following secretion (Cameron *et al.*, 2008), it is hypothesized that the secretory machinery in charge of exporting pallilysin from *T. pallidum* and *T. phagedenis* might possess some degree of homology between the two organisms. Additionally, as described in Chapter 4, bioinformatic analyses performed on the *T. phagedenis* genome identified a similar operon harboring components of a putative T1SS. Thus, it is of further interest to use a *T. phagedenis* lysate in an unbiased, global pull-down experiment with recombinant *T. pallidum* Tp0965 to investigate potential protein-protein interactions between the *T. phagedenis* proteins and Tp0965. Proteomic methods and bioinformatic tools could then be used to identify putative novel T1SS substrates in *T. phagedenis* with sequence similarity to *T. pallidum* proteins that may be candidate virulence factors aiding in the pathogenesis of *T. pallidum*.

1.7 Experimental approach

The characterization of the potential *T. pallidum* T1SS was carried out using plate-based binding assays and pull-down assays to investigate the protein-protein interaction between the virulence factor pallilysin and Tp0965 (the MFP of the putative T1SS). Furthermore, to identify novel potential T1SS substrates, the model treponeme, *T. phagedenis*, was used in a non-biased, global pulldown approach to identify protein-

protein interactions with the recombinant *T. pallidum* protein Tp0965. Upon identification of a candidate T1SS substrate in *T. phagedenis*, bioinformatic approaches were used to scan the *T. pallidum* proteome in search of a protein with sequence similarity to the putative *T. phagedenis* T1SS substrate. Finally, confirmatory binding assays were carried out to show the interaction between the newly identified putative *T. pallidum* T1SS substrate / virulence factor and Tp0965.

Chapter 2: Materials and methods

2.1 Preliminary investigation of a potential *Treponema pallidum* T1SS via *in vitro* protein-protein interaction studies

2.1.1 Construct cloning

PCR amplification and cloning of tp0327, tp0750 and *pallilysin.* The negative control, *tp0327* (**I**^{23}-**S**^{172}) was previously cloned into pET28a and the virulence factors, *tp0750* (**G**^{23}-**D**^{223}) and *tp0751* (*pallilysin* HAXXH **C**^{24}-**P**^{237}) were previously cloned into pDEST-17 (Houston *et al.*, 2011; Houston *et al.*, 2014).

PCR amplification and cloning of tp0965. The DNA encoding the predicted MFP, Tp0965, amino acid residues **P**^{40}-**A**^{318}, was PCR amplified from *T. pallidum* subsp. *pallidum* (Nichols strain) genomic DNA using the forward primer 5'- CAGACTGCTAGC-CCTACCCTCAGTGTGAGTAAGG-3' and reverse primer 5'- CTGTCTGGCGGCCG-CCGCTGCACTTTGGTCTTTGACG-3'. The resulting amplicon was cloned into a pET28a-derived construct, pETHisTEVNTerm. This expression vector was designed and kindly provided to us by the Boulanger lab at the University of Victoria, Canada. The pETHisTEVNTerm expression vector is Isopropyl β-D-1-thiogalactopyranoside (IPTG)-inducible and adds a TEV-cleavable N-terminal hexahistidine-tag for easy purification of recombinant proteins.

2.1.2 Recombinant TEV protease expression and purification

Recombinant TEV protease was produced in-house by Charmaine Wetherell. *E. coli* strain BL21(DE3)-RIL (codon plus) transformed with the TEV protease-encoding plasmid, pRK792 was used. Successful soluble expression of recombinant TEV protease was achieved at an OD_{600} of 0.6 with a final IPTG concentration of 0.4 mM to induce

overexpression. TEV protease was then purified using Fast Protein Liquid Chromatography (FPLC) as well as immobilized metal ion affinity chromatography (IMAC) followed by size exclusion chromatography (SEC). Briefly, *E. coli* lysate was loaded onto 1 ml HisTrap FF affinity columns (GE Healthcare, Baie D'Urfe, QC) pre-packed with nickel sepharose; the columns were washed with binding buffer (20 mM HEPES pH 8.0, 500 mM NaCl, 30 mM imidazole, 10% glycerol) and bound recombinant protein was eluted with elution buffer (20 mM HEPES pH 8.0, 500 mM NaCl, 280 mM imidazole, 10% glycerol). Concentrated eluted protein was then transferred to the HiLoad 16/60 Superdex 75 (GE Healthcare) SEC column to further remove contaminants; the proteins were further washed and eluted using SEC buffer with 20 mM HEPES, 150 mM NaCl, 2.5 mM TCEP, 10% glycerol. Purified TEV protease was stored at -20 ^{o}C in SEC buffer.

2.1.3 Recombinant *T. pallidum* protein expression and purification

Tp0327, tp0750 and pallilysin. The recombinant negative control His-Tp0327 (**I**^{23}-**S**^{172}, H-Tp0327^{rec}) was expressed in the *E. coli* strain BL21 (*DE3) (Invitrogen). The positive control His-Tp0750 (**G**^{23}-**D**^{223}, H-Tp0750^{rec}), and the putative T1SS substrate His-tagged Tp0751 (HAXXH **C**^{24}-**P**^{237}, H-pallilysin^{mut}) were expressed in the *E. coli* strain BL21-AI (Invitrogen). All proteins were expressed under growth conditions described previously (Houston *et al.*, 2011; Houston *et al.*, 2014).

Tp0965. Recombinant His-tagged Tp0965 (**P**^{40}-**A**^{318}, H-Tp0965^{rec}) was expressed in the *E. coli* strain BL21 (*DE3) (Invitrogen) where successful soluble expression was achieved at an OD_{600} of 1.0 with a final IPTG concentration of 1 mM to induce

overexpression. The production of H-Tp0965^{rec} was successfully carried out by a former honours student in the Cameron lab, Bianca Jackson.

All recombinant *T. pallidum* proteins used were expressed solubly and purified using FPLC, IMAC and SEC, as carried out for the TEV protease purification mentioned in Section 2.1.2. As above, IMAC columns were washed with binding buffer (20 mM HEPES pH 7.5, 500 mM NaCl, 20 mM imidazole, 1% glycerol) and bound recombinant protein was eluted with elution buffer (20 mM HEPES pH 7.5, 500 mM NaCl, 500 mM imidazole, 1% glycerol). Concentrated eluted proteins were transferred to the SEC column to further remove contaminants (Figure S1-S5); the proteins were washed and eluted using SEC buffer with 20 mM HEPES pH 7.5, 150 mM NaCl, 1% glycerol. For H-pallilysin^{mut} purification, IMAC buffers were supplemented with 10 μM $ZnCl_2$ and 25 mM $CaCl_2$; and the SEC buffer with 25 mM $CaCl_2$.

When necessary, His-tags were removed using thrombin agarose (Thrombin CleanCleave kit, Sigma-Aldrich) for H-Tp0327^{rec}, and TEV protease for H-Tp0965^{rec}, added at a 1:25, TEV to H-Tp0965^{rec} ratio. TEV protease produced in-house as described above, was allowed to cleave for 16-18h at 16 ^{o}C in the presence of 0.5 mM Tris (2-carboxyethyl) phosphine (TCEP) and at a pH of 7.5. Upon cleavage, proteins were run through the SEC column once again and the purity of all recombinant proteins and removal of His-tag was assessed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis (Figure S3- S7).

2.1.4 Protein quantitation

Bicinchoninic acid assay (BCA). Bovine Serum Albumin (BSA) standards were prepared in 0.1% SDS/PBS to cover a protein concentration range of 10 to 250 μg/ml. Two-

fold serial dilutions were carried out in a 96-well plate for both the standards and the samples. The standards and samples had a final volume of 25 µl per well and were analyzed in triplicate. BCA reagents A and B (Thermo Scientific Pierce BCA Protein Assay Kit) were mixed at a 50:1 ratio and 200 µl of the prepared reagent were added to each well. Samples were incubated and absorbance read at 570 nm after 30 and 60 minutes using a BioTek plate reader (Fisher Scientific). Protein concentrations were then extrapolated according to the BSA standards.

NanoDrop. The NanoDrop® ND-1000 spectrophotometer was used as a second method to determine protein concentration by measuring the absorbance of each purified protein at 280 nm. Calculations were performed using the absorbance of a 0.1% (=1 g/l) solution of each protein, as obtained from the online software ExPASy ProtParam. Thus the absorbance at 280 nm, as measured by the NanoDrop, was divided by the absorbance value of a 0.1% obtained for each protein.

Spectrophotometer. The Beckman Coulter DU Series 700 Spectrophotometer was used as a third method for protein concentration determination by measuring absorbance at 280 nm. As with the NanoDrop, calculations were carried out using the absorbance 0.1% (=1 g/l) value of each protein. In this case, the instrument parameters required the absorbance measurements to be multiplied by 10 and divided by the absorbance of a 0.1% (=1 g/l) solution of each protein to calculate protein concentrations.

2.1.5 Plate-based binding assays with recombinant Tp0965 and pallilysin

Protein-protein interactions between $Tp0965^{rec}$ and the potential T1SS substrate pallilysin was investigated using plate–based binding assays built upon previously

described methodology (Cameron, 2003). Briefly, and as shown in Figure 10, Nunc-Immuno^{TM} 96 Microwell^{TM} solid plate MaxiSorp Surface (Thermo Fisher Scientific) wells were coated with 100 µl of 1 µM of Tp0965^{rec} in Tris-Buffered Saline (TBS) and incubated for 90 minutes at 37^{o}C. All wells were blocked with 1% BSA in TBS-0.1% Tween-20 (TBS-T) for 60 minutes at 37^{o}C, and then coated with 100 µl of 1 µM of either H-pallilysin^{mut} or H-Tp0327^{rec} (negative control) in TBS for 90 minutes at 37^{o}C. TBS-only was used as the blank sample. A 1:5,000 dilution of Ni-conjugated horseradish peroxidase (Ni-HRP, KPL, Gaithersburg MD) in 1%BSA/TBS-T was used for His-tag detection; plates were incubated 60 minutes at room temperature (RT, 25^{o}C) and wells were washed three times in between incubations using TBS-T. Finally, plates were developed with the Tetramethylbenzidine (TMB) peroxidase substrate system (Kirkegaard and Perry Laboratories) and read at 600 nm with a BioTek plate reader (Fisher Scientific). All statistical analyses were performed using the Student's two-tailed t-test in Microsoft Excel.

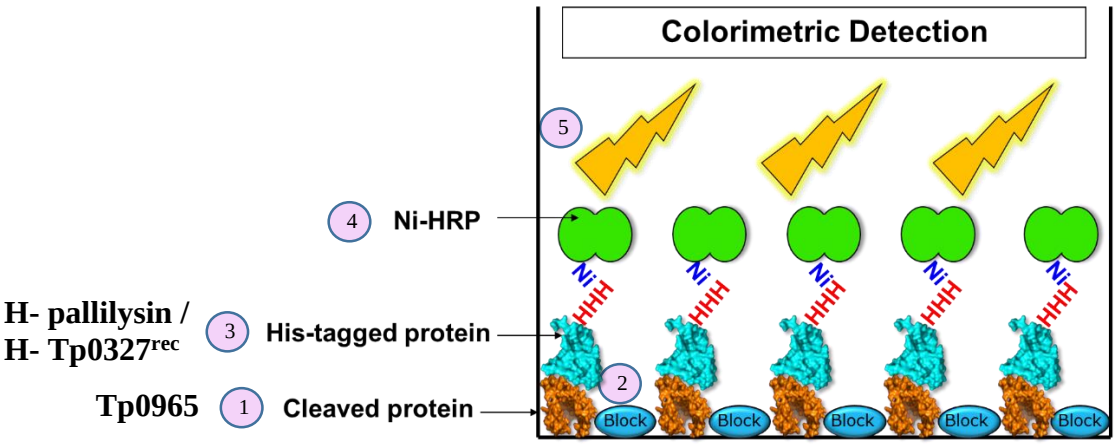


Figure 10. Diagram of plate-based binding assay methodology to investigate a potential pallilysin - Tp0965 interaction.

2.1.6 Pull-down assays with recombinant Tp0965 and pallilysin

The interaction between Tp0965 (the MFP of the putative T1SS) and pallilysin was investigated via *in vitro* pull down assays where 300 μg of purified "bait" protein, either recombinant *T. pallidum* H-pallilysin^{mut} or H-Tp0327^{rec}, were coupled to 100 μl Nickel-conjugated Nitrilotriacetic acid (Ni-NTA) beads (Qiagen) by incubating each set of beads with protein at 4 ^{o}C with gentle rotation for 1 hour. Unbound protein was removed with three gentle washing steps, 300 μg of purified, non-his-tagged "prey" protein, *T. pallidum* Tp0965^{rec}, was added to each bead-coupled bait protein sample (as well as a blank sample with Ni-NTA beads alone), and incubated for 2 hours at RT with gentle rotation. Another three washes were carried out prior to collecting elution fractions. All washes were performed with 20 mM Hepes pH 7.5, 150 mM NaCl, 20 mM imidazole, 1% glycerol, and elution was performed by increasing the imidazole concentration to 500 mM. Fractions were collected during every step of the experiment (including bait and prey flow through, washes and elution) and were analyzed visually via SDS-PAGE and Coomassie Brilliant Blue staining.

2.1.7 Cross-linking studies with recombinant Tp0965 and pallilysin

Recombinant *T. pallidum* proteins H-pallilysin^{mut} and Tp0965^{rec} were cross-linked using the cross-linker, cyanurbiotindipropionylsuccinimide (CBDPS) (Petrotchenko *et al.*, 2011). The *T. pallidum* recombinant proteins Tp0965^{rec} and H-pallilysin^{mut}, as well as pallilysin^{mut} and H-Tp0327^{rec} (negative control), were incubated at room temperature, cross-linked using varying concentrations of CBDPS (10, 25, 50, 75, 100 mM), and quenched using ammonium bicarbonate after 25 minute incubations. The cross-linking reactions were analyzed visually via SDS-PAGE and Coomassie Brilliant Blue staining to

assess the formation of possible complexes. Upon visualization of bands migrating higher on the SDS-PAGE gel and thus possibly containing potential complexes, they were excised, digested with trypsin, and subjected to mass spectrometric analysis.

2.1.7.1 In-gel trypsin digestion

Upon protein separation via SDS-PAGE, protein bands visualized above approximately 57 kDa in the titration gel containing H-pallilysin^{mut} and Tp0965^{rec}, and above approximately 42 kDa in the titration gel containing H-pallilysin^{mut} and H-Tp0750^{rec} were excised from the gels. Trypsin digestion using sequencing grade Trypsin (Promega) was carried out using an adapted version of the methodology described previously (Parker *et al.*, 2005). After the first lyophilisation step described in the mentioned protocol, 100 µl of 50 mM ammonium bicarbonate were added to each sample until gel pieces were hydrated, then the solution was removed. Then, 100 µl acetonitrile (ACN) was added, incubated for 5 minutes, followed by removal of the solution. This step was repeated twice. The gel pieces were cooled at 4 ^{o}C for 30 minutes prior to trypsin digestion carried out as outlined by Parker et al. On day 2, 50 µl ACN were added to each sample and incubated 10 minutes at RT before transferring the solution to a new, clean Eppendorf tube. Then, 30 µl of dH_2O were added to the gel pieces and incubated for 10 minutes, then 50 µl ACN were added on top and incubated another 10 minutes at RT. The solution was removed from the gel pieces and added to the Eppendorf tube which was then frozen at -80 ^{o}C and lyophilized. Finally, samples were reconstituted with 50% ACN, reduced with TCEP and acidified with 10% trifluoroacetic acid (TFA) prior to MS/MS analysis.

2.1.7.2 Identification of cross-links using mass spectrometry

The mass spectrometry (MS) experiments were carried out using a Thermo Scientific LTQ Velos Orbitrap MS mass spectrometer and the data generated was processed via the ICC-CLASS software developed by Petrotchenko et al. (Petrotchenko and Borchers, 2010).

2.2 Investigation of potential novel T1SS substrates via protein-protein interaction studies with a *T. phagedenis* lysate

2.2.1 Bacteria

Treponema phagedenis biotype Kazan previously transformed with the pallilysin-expressing construct *pallilysin*/pKMR4PEMCS (*pallilysin*/pKMR) (Cameron *et al.*, 2008) was grown in tryptone-yeast extract-gelatin-volatile fatty acids-serum (TYGVS) medium (Ohta *et al.*, 1986) supplemented with 10% heat-inactivated (56 ^{o}C for 30 minutes) rabbit serum (HI-NRS; Life Technologies, Frederick, MD) in the presence of 10 μg/ml rifampin and 40 μg/ml erythromycin (Sigma-Aldrich) at 37 °C in a custom Coy Laboratory Products anaerobic chamber (Mandel Scientific Company Inc., Guelph, ON).

T. phagedenis cultures were grown to a density of 3.5 x 10^8 cells/ml and a total estimated count of 3.5 x 10^{10} organisms were harvested. Bacteria were washed with TBS, lysed using gentle sonication and centrifuged. After centrifugation, the lysate was prepared in 20 mM Hepes pH 7.5, 0.1% Triton X-100, 20 mM imidazole and stored on ice to be used the same day for pull-down assays.

2.2.2 Recombinant protein expression and purification

The negative control, Tp0327^{rec}, the putative MFP of the potential *T. pallidum* T1SS, H-Tp0965^{rec}, and H-pallilysin^{mut} were produced as described in Section 2.1.1.

2.2.3 Pull-downs using recombinant *T. pallidum* Tp0965 and a *T. phagedenis* protein lysate

Pull-downs with recombinant *T. pallidum* proteins as "bait" and a *T. phagedenis* lysate as "prey" were carried out as summarized in Figure 11. First, 300 µg of purified "bait" protein, H-pallilysin^{mut} or negative control H-Tp0327^{rec}, were coupled to 100 µl Ni-NTA beads (Qiagen) by incubating together at 4 ^{o}C with gentle rotation for 1 hour. Unbound protein was removed with three gentle washing steps. The *T. phagedenis* lysate was added to the Ni-NTA bead-coupled bait proteins H-Tp0965^{rec}, negative control H-Tp0327^{rec}, as well as a blank sample of only Ni-NTA beads with no "bait" protein attached. All samples were incubated for 4 hours at RT with gentle rotation, washed gently and finally the "bait" protein was eluted along with putative interacting partners. All washes were performed with 20 mM Hepes pH 7.5, 150 mM NaCl, 20 mM imidazole, 1% glycerol and the elution was performed by increasing the imidazole concentration to 500 mM. Fractions were collected during every step of the experiment and were analyzed visually via SDS-PAGE followed by Coomassie Brilliant Blue staining for further analysis via mass spectrometry. Additionally, another set of gels were run and silver stained to achieve higher sensitivity using the BioRad Silver Stain Kit (catalog number 161-0443) to confirm the presence of a band in the elution fraction.

2.2.4 In-gel trypsin digestion

Upon protein separation via SDS-PAGE and staining with Coomassie Brilliant Blue, protein bands present on the H-Tp0965^{rec} – *T. phagedenis* gel, but not on the H-Tp0327^{rec} – *T. phagedenis* (negative control) or the Ni-NTA beads – only (blank) gels, were excised and digested with trypsin using the same protocol described in Section 2.1.7.1. The reduced, acidified samples were then spotted onto a matrix-assisted laser desorption/ionization (MALDI) plate and overlaid with α-Cyano-4-hydroxycinnamic acid (HCCA) matrix prior to tandem mass spectrometry (MS/MS) analysis.

2.2.5 Mass spectrometric analysis and protein identification

The mass spectrometry (MS) experiments were carried out using an Applied Biosystems 4800 MALDI-TOF/TOF mass spectrometer using data-dependent acquisition (DDA) to generate MS/MS spectra for the ten most intense peaks in the corresponding MS spectra. TS2 Mascot software (Matrix Science) was used to generate Mascot generic format (mgf) files. The generated mgf files were searched against the *T. phagedenis* proteome database kindly uploaded by Derek Smith at the UVic – Genome BC Proteomics Centre, and the Uniprot-SwissProt database using the following parameters: MS/MS Ion Search type, enzyme used was trypsin, variable modifications: Acetyl (K), Deamidated (NQ), Methyl (C-term), monoisotropic mass values, unrestricted protein mass, peptide mass tolerance of ±100 ppm, fragment mass tolerance of ±0.8 Da, maximum of 2 missed cleavages, and instrument type used was MALDI TOF-TOF.

2.2.6 Bioinformatic analyses

Protein-protein and Position-Specific Iterated Basic Local Alignment Search Tool (BLASTP, PSI-BLAST) (Altschul *et al.*, 1990) were used to search for a *T. pallidum* protein with sequence similarity to the *T. phagedenis* proteins identified via the pull-downs and mass spectrometry experiments described above (Sections 2.6 – 2.8). Upon identification, the *T. pallidum* proteins' primary sequences were aligned to the *T. phagedenis* originally found proteins, using Clustal Omega (Clustal W 2.1) Bioinformatics, to assess regions of sequence variation and similarity in more detail (Sievers *et al.*, 2011). Additionally, Phyre2 Protein Fold Recognition Server (Kelley *et al.*, 2015) was used to infer on the predicted structures of the proteins in question and as a complementary tool to confirm the functional predictions performed by BLAST and the Conserved Domains Database (CDD) (Marchler-Bauer and Bryant, 2004). Further, PRED-TMBB (Bagos *et al.*, 2004) and TMHMM Server, v. 2.0 (Sonnhammer *et al.*, 1998) online software were used to investigate the possibility of the proteins forming β-barrels or possess transmembrane α-helices, respectively. Finally, HeliQuest (Gautier *et al.*, 2008) was used to screen for putative amphiphilic α-helices at the C-terminus of potential T1SS substrates identified, and create helical wheel diagrams outlining physiochemical properties of individual residues.

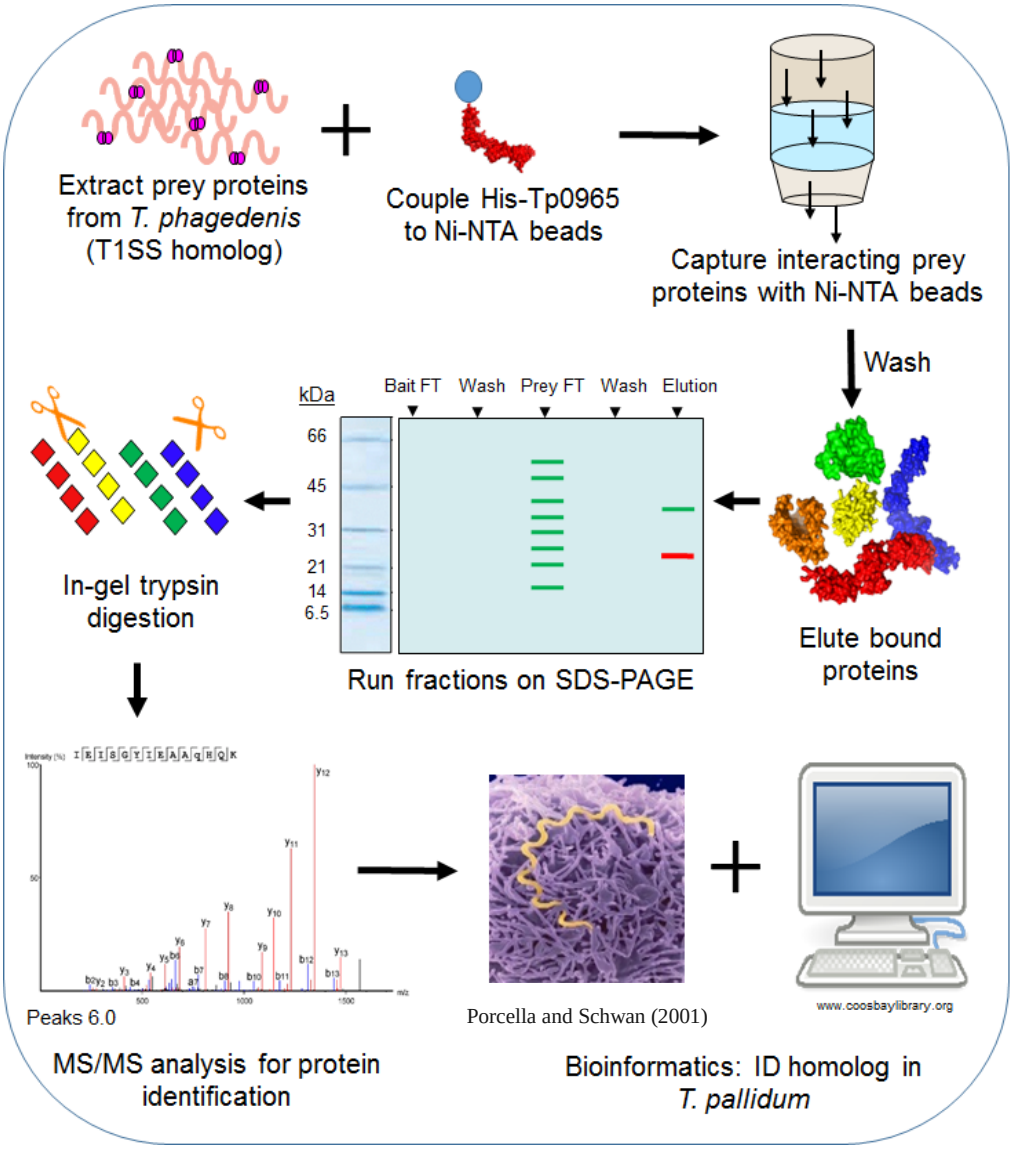


Figure 11. Summary of the unbiased, global pull-down approach to identify novel putative T1SS substrates in *T. phagedenis*

T. phagedenis containing a predicted T1SS was cultured *in vitro*, washed, harvested and lysed (prey protein). His-tagged *T. pallidum* recombinant H-Tp0965^{rec} was immobilized on nickel beads and used to pull out potential interacting *T. phagedenis* prey proteins. Upon gentle washes and elution, all the experiment fractions were run in an SDS-PAGE gel to obtain protein band separation. Bands from the elution lane were excised and digested using MS-grade trypsin. MS/MS analysis was carried out using a 4800 MALDI-TOF/TOF mass spectrometer and TS2 Mascot software (Matrix Science) was used to generate Mascot generic format (mgf) files. Extensive bioinformatic analyses were then carried out to search against the *T. pallidum* proteome for a protein with sequence similarity.

2.3 Confirmatory plate-based binding assays to explore an interaction between the novel putative T1SS substrate, Tp0854, and Tp0965

2.3.1 Construct cloning

Putative phosphatase domain of tp0854. The DNA encoding the predicted phosphatase domain of Tp0854, amino acid residues **D**^{1194}-**R**^{1533}, was PCR amplified from *T. pallidum* subsp. *pallidum* (Nichols strain) genomic DNA using the forward primer 5'-AGACTGCTAGCGATGATCACGTGGAGT-3' and reverse primer 5'-ATGTCTTG-CGGCCGCTCTCTTTCTTATTCCCAAGATAGTG-3'. The resulting amplicon was cloned into pETHisTEVNTerm, an expression vector described in Section 2.1.1.

2.3.2 Recombinant protein expression and purification

The *tp0854* (**D**^{1194}-**R**^{1533})/pETHisTEVNTerm construct was then transformed into the *E. coli* strain BL21 (*DE3) (Invitrogen) where successful soluble expression of His-tagged Tp0854 (**D**^{1194}-**R**^{1533}, H-Tp0854^{phos}) was achieved at an OD_{600} of 1.0 with a final IPTG concentration of 0.4 mM to induce overexpression.

Soluble recombinant protein was purified as above (Section 2.1.1) using FPLC, IMAC and SEC. As before, the columns were washed with binding buffer (20 mM HEPES pH 7.5, 500 mM NaCl, 20 mM imidazole, 1% glycerol) and protein was eluted with elution buffer (20 mM HEPES pH 7.5, 500 mM NaCl, 500 mM imidazole, 1% glycerol). Concentrated eluted proteins were transferred to the HiLoad 16/60 Superdex 75 (GE Healthcare) SEC column to further remove contaminants, were washed and eluted using SEC buffer with 20 mM HEPES pH 7.5, 150 mM NaCl, 1% glycerol.

The His-tag from H-Tp0854^{phos} was removed using TEV protease under the conditions mentioned above: at a 1:25, TEV to H-Tp0854^{phos} ratio, 16 ^{o}C overnight.

Proteins were run through the SEC column once again, and both recombinant protein purity as well as His-tag removal were assessed visually via SDS-PAGE analysis.

2.3.3 Plate-based binding assays with recombinant Tp0965 and Tp0854

Protein-protein interactions between Tp0965 and the novel potential substrate Tp0854 were investigated using plate-based binding assays as described in section 2.1.5. As shown in Figure 12, for these assays, Tp0854^{phos} was used to first coat the plate and Tp0327^{rec} was used as a negative control. H-Tp0965^{rec} was added as the second protein to both samples. Additionally, this assay was reversed and hence Tp0965^{rec} or the negative control Tp0327^{rec} were added to first coat the plate, followed by the addition of H-Tp0854^{phos} as the second protein.

The dose-dependent attachment of recombinant Tp0965 to the phosphatase domain of Tp0854 was also investigated by varying the molarity of either H-Tp0965^{rec} to Tp0854^{phos} or H-Tp0854^{phos} to Tp0965^{rec}, from 0.1 to 3 μM and 0.01 to 1 μM, respectively.

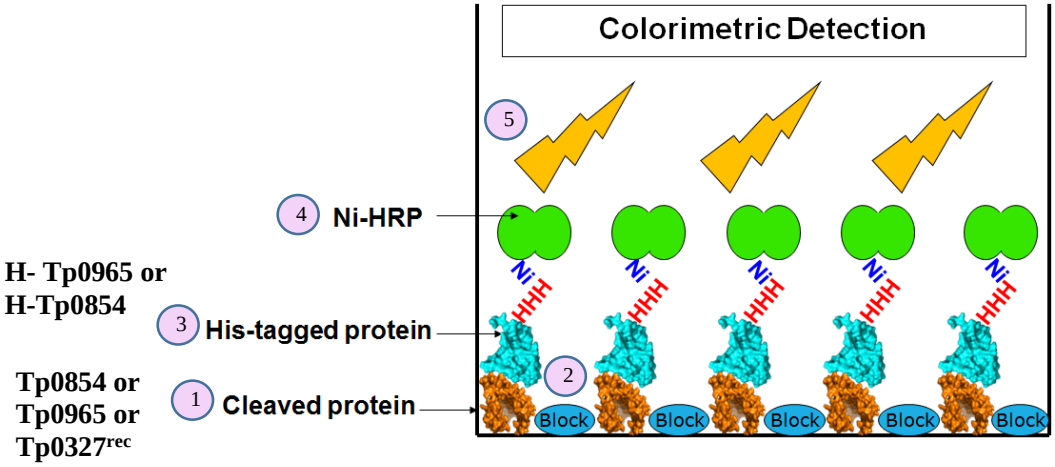


Figure 12. Diagram of plate-based binding assay methodology to confirm a Tp0854 - Tp0965 interaction.

Chapter 3: Direct investigation of potential Type I Secretion System substrates using recombinant *T. pallidum* proteins

3.1 Contributions to the data

The data presented in this chapter represents the first *in vitro* analysis of a previously characterized virulence factor, pallilysin, as a potential T1SS substrate in the syphilis spirochete, *Treponema pallidum.* Binding assays, previously developed by Dr. Caroline Cameron (Cameron, 2003) were adapted for experiments tailored to detect a potential interaction between pallilysin and Tp0965 (the MFP of the putative T1SS in *T. pallidum*). Moreover, traditional pull-down assays were carried out as a complementary technique to assess said interaction. Finally, Jason Serpa from the UVic – Genome BC Proteomics Centre kindly processed the trypsin-digested excised bands from the cross-linking experiments.

3.2 Introduction

The syphilis spirochete, *Treponema pallidum*, possesses highly invasive and disseminating properties that allow it to invade even the most privileged tissues of the human host, including ocular tissues, the CNS and the uterus (LaFond and Lukehart, 2006). *T. pallidum* is able to bind to varying host ECM components, endothelial cells *in vitro*, and disrupt the intercellular junctions and hence it is able to cross the host vascular endothelial barrier and the basement membrane (Blanco *et al.*, 1997; Fitzgerald *et al.*, 1984; Thomas *et al.*, 1988). Several characteristics contribute to *T. pallidum*'s invasiveness and host immune evasion; however, the exact molecular mechanism and virulence factors involved in dissemination have just recently begun to emerge. The *T. pallidum* virulence factor, pallilysin, was first identified in 2003 as a laminin-binding protein, via binding assays

carried out by Dr. Caroline Cameron (Cameron, 2003). This adhesin is thought to play an important role in *T. pallidum*'s pathogenesis by being one of the constituents that confer upon this pathogen the ability to bind to ECM components and hence aid in invasion and dissemination (Cameron *et al.*, 2008; Houston *et al.*, 2012). Moreover, previous studies in the Cameron lab suggested that wild-type pallilysin may possess proteolytic and self-activating properties (Houston *et al.*, 2012). Since there is evidence showing that pallilysin is expressed and exposed to the host cellular milieu under infection conditions (Cameron, 2003; Cameron *et al.*, 2008; Houston *et al.*, 2012), it is hypothesized that this virulence factor might be secreted via a putative *T. pallidum* T1SS.

The preliminary investigation of pallilysin as a putative T1SS substrate was initiated by studying the potential protein-protein interaction between the MFP of the putative T1SS (Tp0965) and a pallilysin mutant previously generated. This mutant pallilysin [Tp0751(**C**24-**P**237) HAXXH (E199A), pallilysin^{mut}] lacks proteolytic activity (Houston *et al.*, 2012) and could be used to study protein-protein interactions without compromising the *in vitro* integrity of this virulence factor and/or that of other recombinant proteins used during experiments. Hence, the initial protein-protein interaction studies between Tp0965 and pallilysin were carried out using two methods: *in vitro* binding assays and pull-down techniques. Moreover, in an attempt to identify specific protein-protein interaction sites, cross-linking experiments were performed. In all the experiments carried out, Tp0327, an unrelated *T. pallidum* chaperone, was used as a negative control.

3.3 Binding assays to investigate a potential interaction between Tp0965 and the virulence factor, pallilysin

Binding assays can be used successfully to study protein-protein interactions in a variety of scenarios and hence, it was determined to be an adequate initial step into the investigation of a putative interaction between the *T. pallidum* virulence factor pallilysin and Tp0965 from the potential T1SS. Binding assays were performed as described previously (Cameron, 2003) and as mentioned in section 2.1.5. Each triplicate sample was averaged and plotted as shown in Figure 13. It can be seen that the absorbance value for the negative control H-Tp0327^{rec} was found to be considerably lower than the absorbance value for pallilysin. The student t-test statistical analysis, showed a p-value lower than 0.005 between the negative control and the Tp0965^{rec} - H-pallilysin^{mut} binding levels, suggesting significantly more binding of H-pallilysin^{mut} to Tp0965^{rec} in comparison with the binding of H-Tp0327^{rec} to Tp0965^{rec}. The greater interaction between H-pallilysin^{mut} and Tp0965^{rec} suggests that the MFP of the putative T1SS may indeed interact with this key *T. pallidum* virulence factor and could indicate secretion via this system.

Even though statistical analysis showed that the absorbance of the Tp0965^{rec} - H-pallilysin^{mut} sample was significantly higher than that of the negative control, H-Tp0327^{rec} to Tp0965^{rec}, the overall absorbance values for both samples were low, possibly due to the potentially transient nature of the interaction, and therefore it was decided that further investigation using alternative techniques were required to verify the detected potential interaction.

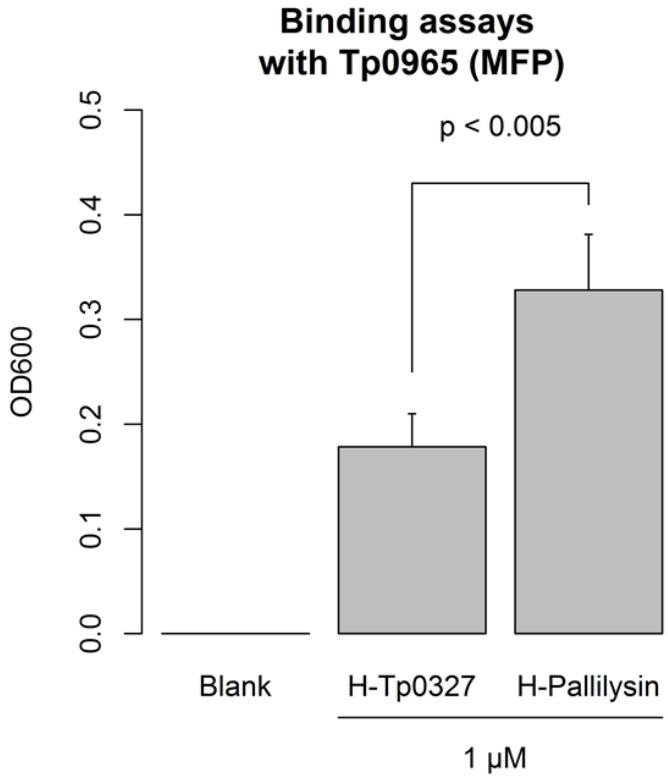


Figure 13. Plate-based binding studies between recombinant *T. pallidum* pallilysin and Tp0965, using Tp0327 as a negative control.

3.4 Pull-down assays to further investigate a potential interaction between Tp0965 and the virulence factor, pallilysin

Pull-down assays are commonly used in the biological sciences to investigate protein-protein interactions and the presence of a His-tag in our recombinant proteins made this technique a relatively simple approach to further explore our hypothesis. Therefore, the second technique used in an attempt to show an interaction between pallilysin and Tp0965, was *in vitro* pull-downs with recombinantly-produced *T. pallidum* H-pallilysin^{mut} and Tp0965^{rec}, as mentioned in section 2.1.6. Upon collection of washes and elution fractions, all samples were run on SDS-PAGE gels and stained, leading to the results shown in Figure 14. Neither the blank, Tp0965^{rec} incubated with beads alone, nor the negative control, Tp0965^{rec} with H-Tp0327^{rec}, showed a visible Tp0965 band on the elution fraction lane upon gel staining. The Tp0965^{rec} with H-pallilysin^{mut} sample on the other hand, did show a Tp0965^{rec} band on the elution fraction. There is a significant loss of Tp0965^{rec} during the washes in all cases, including the gel showing an interaction between pallilysin and Tp0965^{rec}. It is thought that the reason why Tp0965^{rec} does not remain completely attached to the His-pallilysin sample is because the hypothesized interaction between pallilysin and Tp0965^{rec} is expected to be transient. If pallilysin were indeed secreted via this system, the interaction between the MFP, Tp0965 and this virulence factor would be short lived. Even though the interaction between a T1SS and its substrate would be expected to be of transient nature, *in vitro* conditions and the presence of excess protein make it possible to visually assess such weak interactions.

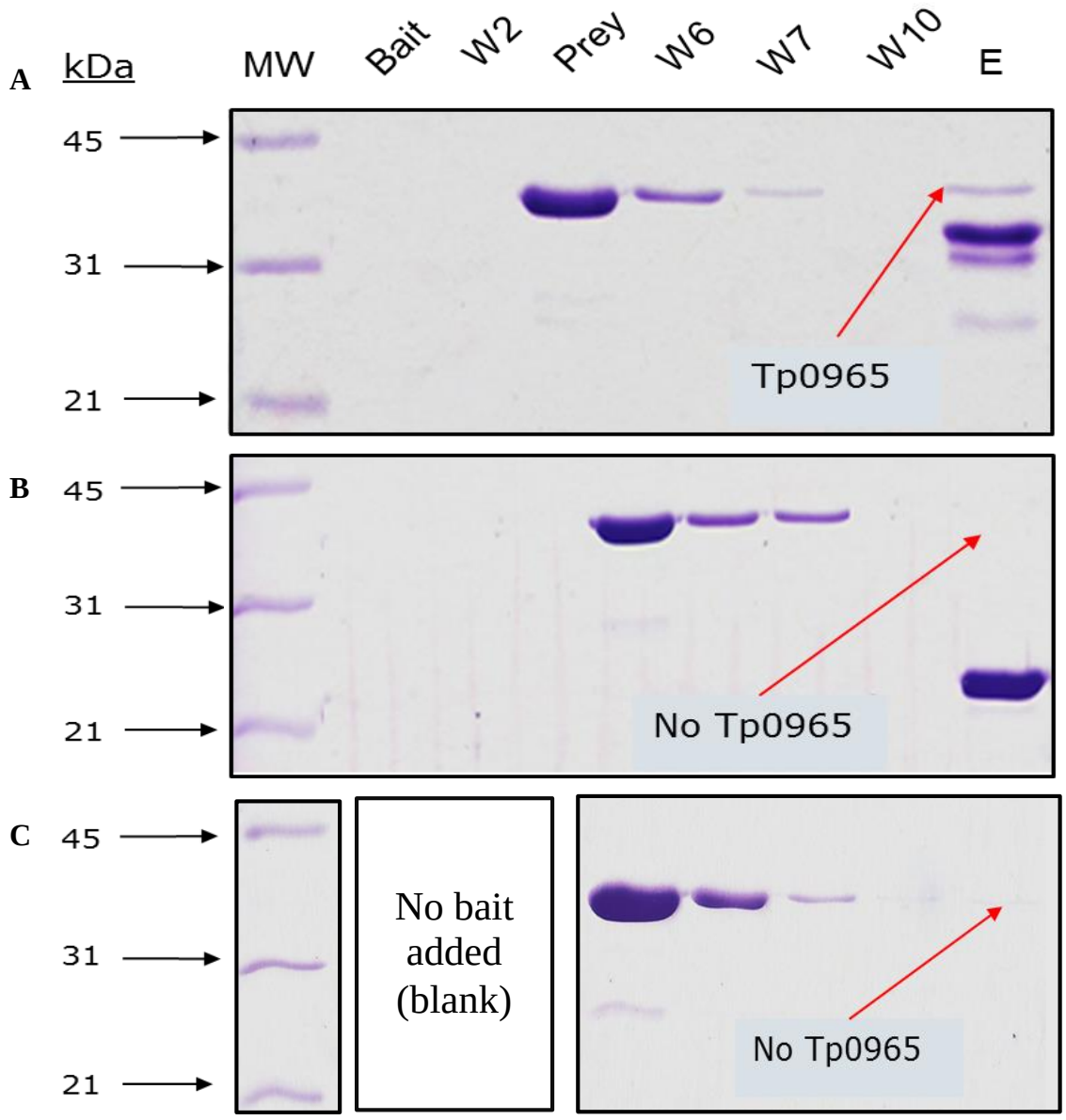


Figure 14. Coomassie Brilliant Blue stain of fractions obtained from pull-down assays between the virulence factor H-pallilysin^{rec}, and the MFP Tp0965^{rec}

(A) Fractions collected during pull-down assay between Tp0965^{rec} and H-pallilysin^{mut}, showing a visible Tp0965^{rec} band in the elution fraction lane. (B) Fractions collected for the negative control sample Tp0965 and H- Tp0327^{rec} showing no visible Tp0965^{rec} band in the elution fraction lane. (C) Fractions collected for the blank sample (no bait protein) show no visible Tp0965^{rec} band in the elution fraction lane. All samples were treated equally, with the same amount of lysate and same amount of bait protein (except in the blank), and incubated under the same conditions. MW: Molecular Weight protein ladder, Bait: bait flow through, W2: Wash #2, last wash after bait flow through, Prey: prey flow through, W6: Wash #6, W7: Wash #7, W10: Wash #10, last wash after prey flow through, E: Elution fraction.

3.5 Cross-linking experiments between Tp0965 and pallilysin to identify specific interacting sites

Due to the transient nature of the interactions observed between secretory machineries and their substrates, novel techniques have been developed in attempts to capture such contacts between proteins. One such technique is cross-linking, which is successfully used to detect interactions that would otherwise remain unknown. Furthermore, cross-linking, followed by mass spectrometry allows for the identification of the specific residues from two interacting proteins whose interaction becomes covalently linked due to reaction with the cross-linker. For instance, the novel homobifunctional primary amine reactive (primarily lysines and N-termini) cross-linker with a 14 Å spacer arm, CBDPS, has been developed to be isotopically coded, biotinylated to allow for an enrichment step prior to mass spectrometric analysis, collision-induced dissociation (CID)-cleavable and possesses isotopic MS/MS signatures for dead-end and inter-peptide cross-links (Petrotchenko *et al.*, 2011). The versatility and spacer arm flexibility, as well as the mass spectrometry-friendly characteristics of this cross-linker made it suitable for *in vitro* cross-linking studies between Tp0965 and pallilysin. H-pallilysin^{mut} with Tp0965^{rec}, the negative control H-Tp0327^{rec}, with Tp0965^{rec}, and the positive control H-Tp0750^{rec} with H-pallilysin^{mut} were cross-linked and run on an SDS-PAGE gel. The results in Figure 15A shows a similar banding pattern between the titration of H-pallilysin^{mut} with Tp0965^{rec} and that of Tp0965^{rec} alone with 100 mM CBDPS. However, it can be seen that as the concentration of CBDPS increases, the pallilysin band seems to decrease in intensity, suggesting that it might not be migrating to its monomeric location but rather might be interacting with Tp0965^{rec} and hence appear at an additive molecular weight higher on the gel.

Furthermore, the sample containing H-pallilysin^{mut} alone and the highest concentration of CBDPS (100 mM) shows a fainter band than the sample lacking any CBDPS, indicating that the conditions at which cross-linking was carried out might have resulted in a certain degree of pallilysin precipitation (loss), hence preventing its full migration into and through the SDS-PAGE gel. Moreover, since Tp0965 is predicted to be the MFP component of the putative *T. pallidum* T1SS, it is expected that this protein might form trimeric or hexameric complexes, as is the case for the MFP, HlyD, from the *E. coli* hemolysin system (shown in Figure 5). Hence, the banding pattern of the Tp0965^{rec} sample cross-linked with 100 mM CBDPS on its own is to be expected, with monomeric, dimeric, trimeric, tetrameric, pentameric and hexameric forms at approximately 35, 72, 93, above 125, above 165, and below 240 kDa, respectively. It was thought that perhaps the band observed in the titration sample lanes at approximately 60 kDa could be a complex between pallilysin^{mut} and Tp0965^{rec} and due to the multimeric nature of Tp0965 it was unclear if one or more of the Tp0965^{rec} complex bands were masking pallilysin^{mut}, thereby making it unclear whether there was indeed an interaction between pallilysin^{mut} and Tp0965^{rec}. Thus, the bands highlighted in red in Figure 15A were excised, digested with trypsin and subjected to mass spectrometric analysis as described in section 2.1.4. Interestingly it was found that pallilysin^{mut} and Tp0965^{rec} were present together in the four sample bands, and Table S1 shows the high sequence coverage obtained for both proteins in each band. Unfortunately, exact cross-links could not be identified. Although CBDPS has been shown to be a powerful reagent, it is mainly reactive with the primary amines on lysines and N-termini, as mentioned above. Even though the primary sequence of Tp0965 contains 21 lysines, pallilysin contains only two, substantially reducing the probability of CBDPS

reacting with this virulence factor, hence if only minimal reactivity between the two proteins in question occurred, it might explain why specific cross-links were not identified upon mass spectrometric analysis.

Similarly, in the case of the positive control shown in Figure 15B, a band at approximately 50 kDa can be seen in the titration samples of H-Tp0750^{rec} (which has 12 lysines) with pallilysin. However, the lane showing H-Tp0750^{rec} with 100 mM CBDPS alone also shows a band at the same migration distance, suggesting that the appearing band in the titration samples could simply be a dimer of H-Tp0750^{rec}. Moreover, the lane showing pallilysin with 100 mM CBDPS alone also shows a very faint putative dimer at approximately 57 kDa, suggesting a potential pallilysin dimer that interestingly does not seem to appear in the titration samples.

The interaction between Tp0750 and pallilysin was previously investigated via dynamic light scattering (DLS) techniques by the Cameron lab, and results showed a positive interaction between these two virulence factors (Houston *et al.*, 2014). Hence, it was decided that Tp0750 would be an adequate positive control for cross-linking experiments with pallilysin. The fact that the interaction between H-Tp0750^{rec} and H-pallilysin^{mut} was not detected via cross-linking experiments using CBDPS once again suggests that this reagent is inadequate for experiments involving pallilysin under the experimental conditions used.

The negative control in Figure 15C, H-Tp0327^{rec} (containing 11 lysines) with Tp0965^{rec} showed the interesting dimerization and trimerization of H-Tp0327^{rec} alone with 100 mM CBDPS, which can also be seen throughout the titration samples. The Tp0965^{rec} with 100 mM CBDPS sample again shows the multimeric pattern described above, and the

titration samples show no clear interaction between Tp0965^{rec} and H-Tp0327^{rec}, since all the bands that appear can be accounted for by bands that show up in the H-Tp0327^{rec} with 100 mM CBDPS alone as well as the Tp0965^{rec} with 100 mM CBDPS alone.

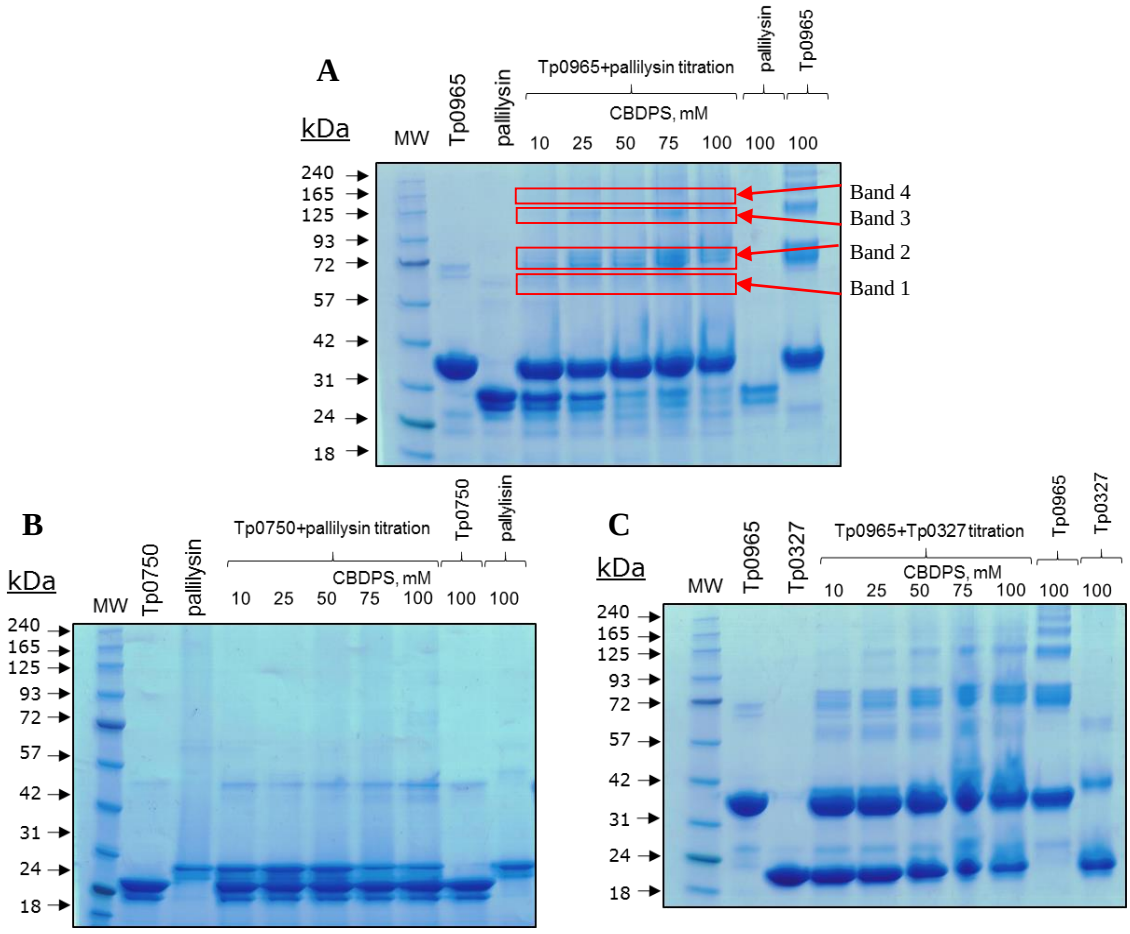


Figure 15. Cross-linking experiments between the virulence factor H-pallilysin^{mut} and the MFP H-Tp0965^{rec}

(A) Cross-linking titration of Tp0965 and pallilysin; from left to right: Protein ladder, H-Tp0965^{rec} 0 mM CBDPS, pallilysin 0 mM CBDPS, titration of H-Tp0965^{rec} with pallilysin: 10 mM to 100 mM CBDPs, pallilysin 100 mM CBDPS, H-Tp0965^{rec} 100 mM CBDPS. (B) Cross-linking titration of H-Tp0750^{rec} and pallilysin (positive control); from left to right: Protein ladder, H-Tp0750^{rec} 0 mM CBDPS, pallilysin 0 mM CBDPS, titration of H-Tp0750^{rec} with pallilysin: 10 mM to 100 mM CBDPs, H-Tp0750^{rec} 100 mM CBDPS, pallilysin 100 mM CBDPS. The red boxes show the band excisions subject to mass spectrometric analysis. (C) Cross-linking titration of H-Tp0965^{rec} and H-Tp0327^{rec} (negative control); from left to right: Protein ladder, H-Tp0965^{rec} 0 mM CBDPS, H-Tp0327^{rec} 0 mM CBDPS, titration of H-Tp0965^{rec} with H-Tp0327^{rec}: 10 mM to 100 mM CBDPS, H-Tp0965^{rec} 100 mM CBDPS, H-Tp0327^{rec} 100 mM CBDPS.

3.6 Conclusions

Three different techniques were used for the preliminary investigation of the putative T1SS in *Treponema pallidum* that might be responsible for the secretion of virulence factors that confer upon this pathogen its highly invasive properties. The MFP component (Tp0965) of the putative T1SS and potential T1SS substrate, virulence factor pallilysin, were used for binding studies, pull-down assays and cross-linking experiments. The binding studies and pull-down assays carried out suggested a weak interaction between the putative MFP $Tp0965^{rec}$ and $pallilysin^{mut}$. The cross-linking experiments further showed potential co-migration of $Tp0965^{rec}$ and $pallilysin^{mut}$ along the SDS-PAGE gel, and mass spectrometric analysis suggested the presence of both $Tp0965^{rec}$ and $pallilysin^{mut}$ in all the bands analyzed, which appeared to have the molecular weight of multimeric complexes. Unfortunately, no specific cross-links were identified and it was concluded that the cross-linker used might not have been appropriate for the discovery of interactions involving the lysine-devoid virulence factor pallilysin. Regardless, the three experimental approaches successfully showed a potential transient interaction between the MFP Tp0965 and the *T. pallidum* virulence factor pallilysin. Further studies are required in order to obtain more details regarding the functionality of this system as an actual T1SS and more investigations should lead to the discovery of novel putative T1SS substrates.

Chapter 4: Bioinformatic analyses show a predicted T1SS homolog in *Treponema phagedenis*

4.1 Contributions to the data

In this chapter, bioinformatic techniques were used to screen the proteome of the closely related spirochete, *Treponema phagedenis*, to explore the possibility of it possessing a T1SS homologous to the putative T1SS in *T. pallidum*. My contribution to the data is the novel investigation of an operon in *T. phagedenis*, which might encode a homolog corresponding to each and every component of the putative *T. pallidum* T1SS.

4.2 Introduction

The spirochete *T. phagedenis* is a poorly characterized organism that has been recognized as being part of the normal human flora, a genitalia commensal and thus non-pathogenic to humans. However, a recent study showed that the causative agent of Bovine papillomatous digital dermatitis (PDD), a *Treponema phagedenis*-like spirochete, which affects cattle and represents a great economic burden to society, shows no significant differences at the genomic and phenotypic levels between several isolates from cows and the human commensal *T. phagedenis* (Wilson-Welder *et al.*, 2013). Like the human commensal *T. phagedenis*, the causative agent of Bovine PDD is a strict anaerobe that requires serum and volatile fatty acids to grow. Hence, technically speaking, *Treponema phagedenis* is a pathogen that does not affect humans and as such is able to live in a commensal relationship with the human host.

Unfortunately, although the genome of *Treponema phagedenis* has been completely sequenced via whole genome shotgun approaches by the Human Microbiome Project, it has not yet been fully annotated. Nonetheless, the 16S rRNA sequence has been

previously used to assess phylogenetic relationships between a diverse number of spirochetes (Paster *et al.*, 1991; Woese, 1987). Interestingly, these studies showed a 91.1% similarity between the 16S rRNA from *T. pallidum* and *T. phagedenis* (Paster *et al.*, 1991), indicating a high degree of sequence conservation and great similarity between these genomes, but confirmatory whole genome analyses remain to be carried out.

Due to the similarities between *T. pallidum* and *T. phagedenis*, and the ability of the latter to express and appropriately lipidate and position pallilysin on its surface when transformed with the *pallilysin*/pKMR vector (Cameron *et al.*, 2008), it was of interest to use bioinformatic analyses to scan the *Treponema phagedenis* proteome in search for the presence of a T1SS homologous to the putative T1SS in *T. pallidum* that might be responsible for the secretion and appropriate localisation of not only pallilysin, but also other virulence factors from within these treponemes.

4.3 The predicted T1SS homolog in *T. phagedenis*

As mentioned earlier in this thesis, the T1SS can be found in a wide variety of Gram-negative microorganisms. Bacteria that produce bacteriocins, such as Colicin V, have been shown to use T1SSs to secrete these virulence factors to the extracellular space, targeting competing microorganisms (Gérard *et al.*, 2005; Wooldridge, 2009). Moreover, recent studies showed that even non-pathogenic microorganisms, such as the Antarctic Gram-negative bacterium *Marinomonas primoryensis*, can carry T1SSs to secrete specialized proteins that aid in survival, as is the case of the MpAFP antifreeze protein adhesin secreted by *M. primoryensis* and used to bind to the underside of ice covering lake surfaces, allowing for the bacteria to remain in a favorable position with access to oxygen

and nutrients from nearby photosynthetic organisms (Vance *et al.*, 2014). However, T1SSs are most commonly found in pathogenic bacteria.

Interestingly, bioinformatic analyses using BLAST and Phyre2 identified in *T. phagedenis* protein homologs to the putative T1SS protein components in *T. pallidum* and an illustration comparing both operons is shown in Figure 16. Unlike most T1SSs known, which are made up of three proteins, there were five proteins identified that could make up the potential T1SS in *T. phagedenis*: two permeases, an ABC transporter, a membrane fusion protein and a TolC-like protein. Table 2 shows the proteins from the predicted T1SS in *T. pallidum* and their corresponding potential homologs in *T. phagedenis.*

In these bioinformatic studies, the two *T. phagedenis* permeases, denoted as efflux ABC transporter permeases, and with accession numbers WP_002695293 and WP_002695296, showed high homology to the *T. pallidum* Tp0962 and Tp0963, with amino acid sequence identity of 42% and 43%, respectively, as well as query cover of 99% in both cases. As seen in Table 2, the predicted structures of Tp0962 and Tp0963 show great resemblance to their *T. phagedenis* homologs, and transmembrane helix predictions (TMHMM Server, v. 2.0, (Sonnhammer *et al.*, 1998)) further showed that the *T. phagedenis* proteins are predicted to contain transmembrane helices and hence it is thought they might form a pore within the inner membrane, similar to the predicted *T. pallidum* T1SS permeases.

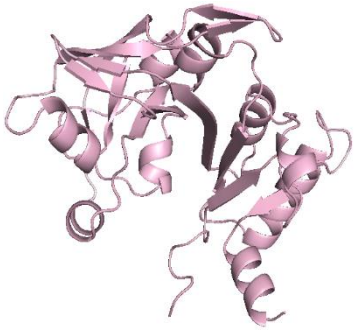
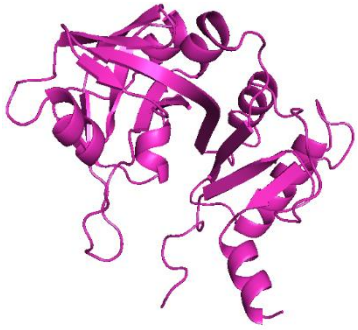
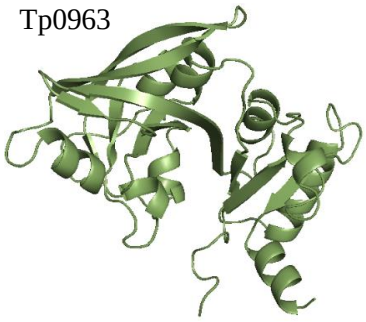
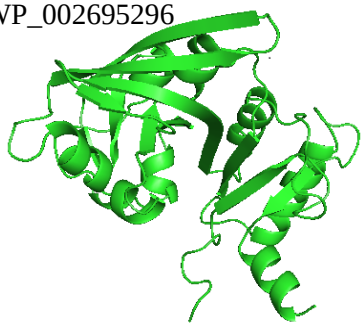
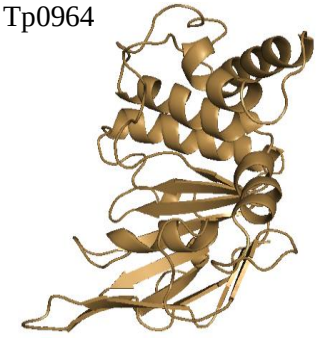
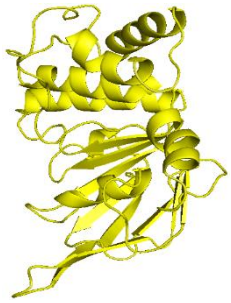
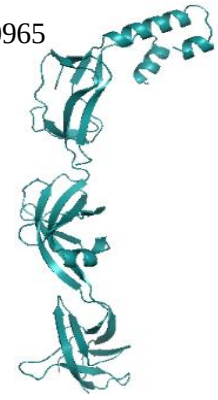
The ATPase protein component of the putative T1SS in *T. pallidum*, Tp0964, showed 71% identity to the *T. phagedenis* ABC transporter ATP binding protein, accession number WP_002695297, with 97% query cover, suggesting great homology. Table 2

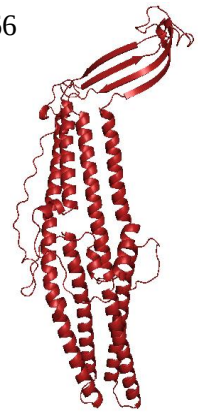
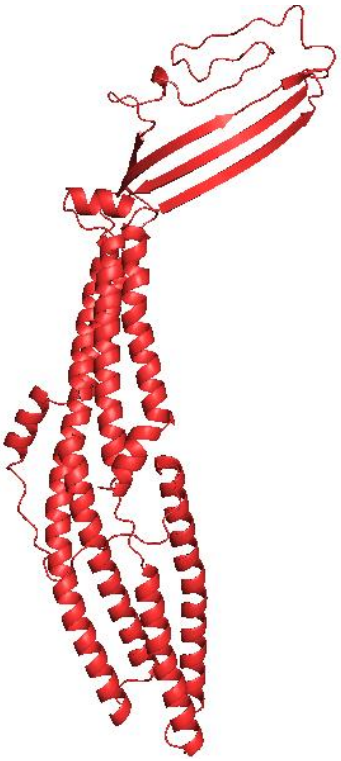
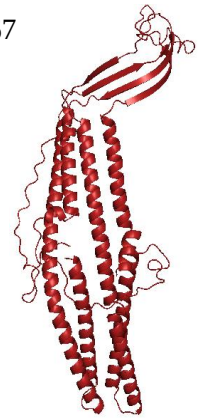
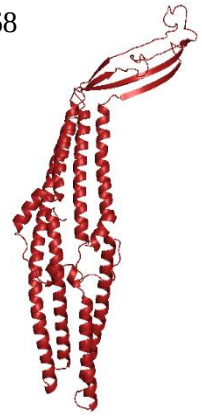
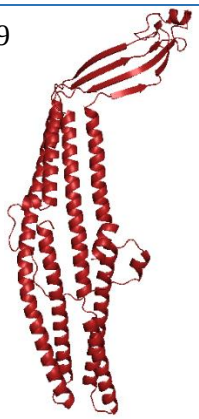
shows the clear structural identity in the models predicting both ATPases from the putative T1SSs of these spirochetes. This ATP-binding protein is predicted to be cytoplasmic, and would be essential to supply the energy required for substrate transport across both the inner and outer membranes.

Importantly, the membrane fusion protein component of the putative *T. pallidum* T1SS, Tp0965 showed 59% identity to the *T. phagedenis* protein with accession number WP_002695300, and 92% query cover suggesting a high level of homology, also clearly represented in the Phyre2 predicted structures showing clear similarities.

Finally, unlike the *T. pallidum* potential T1SS, which contains four predicted TolC-like proteins, the *T. phagedenis* putative T1SS contained only one predicted TolC-like protein with accession number WP_002695302. The reason why there might be such a difference in the number of TolC-like proteins between these two spirochetes could be due to the fact that *Treponema pallidum* has a characteristically small genome, encoding for just over 1000 proteins, whereas the *Treponema phagedenis* genome is substantially larger, about three times the size of *T. pallidum*'s. Thus, *T. phagedenis* possesses a significantly higher number of proteins that could be employed to assemble a greater number of secretory machineries, whereas the limited number of proteins encoded by *T. pallidum*'s genome would favor a "mix-and-match" mechanism where the same core structure of a secretion system is used and components of the system can be interchanged (e.g. the four TolC-like proteins in the case of the putative T1SS). In fact, proteins in *T. pallidum* have previously shown multi-functional properties such as the proteolytic and adhesin abilities of Tp0750 and pallilysin (Houston *et al.*, 2011; Houston *et al.*, 2014).

Table 2. Structure predictions of the putative *T. pallidum* T1SS components and potential *T. phagedenis* homologs

Phyre2		BLAST	
T. pallidum	***T. phagedenis***	**Identity**	**Query Cover**
Tp0962 	WP_002695293 	42%	99%
Tp0963 	WP_002695296 	43%	99%
Tp0964 	WP_002695297 	71%	97%
Tp0965 	WP_002695300 	59%	92%

Tp0966 	WP_002695302 	27%	81%
Tp0967 			
Tp0968 			
Tp0969 			

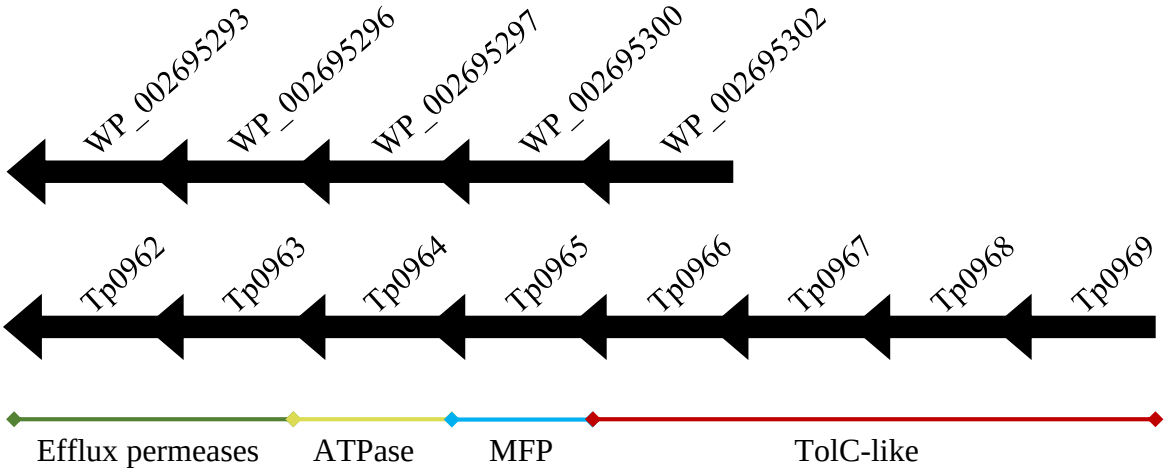


Figure 16. Schematic representation of the *T. pallidum* and *T. phagedenis* operons encoding putative T1SSs

Top: The *T. phagedenis* operon encoding a putative T1SS, showing two permeases, an ATPase, a MFP and a single TolC-like protein. Bottom: The *T. pallidum* operon encoding a putative T1SS, showing two permeases, an ATPase, a MFP and four TolC-like proteins.

4.4 Conclusions

Despite the fact that the *Treponema phagedenis* genome has not been fully annotated we were able to find in its genome an operon that encodes for all the components of a potential T1SS homologous to the putative *T. pallidum* T1SS. All the potential components of the *T. phagedenis* putative T1SS show sequence similarity to the corresponding potential T1SS components in *T. pallidum*. This suggested that an experiment with recombinant proteins and a *T. phagedenis* lysate may lead to novel discoveries due to a possible cross-recognition between the T1SS components from *T. pallidum* and *T. phagedenis* putative T1SS substrates as described in the next chapter.

Chapter 5: Global and unbiased pull-down approach to detect putative novel Type I Secretion System substrates using the model spirochete *Treponema phagedenis*

5.1 Contributions to the data

This chapter describes the results obtained from a global and unbiased pull-down approach used to screen for novel *T. pallidum* putative T1SS substrates using recombinant *T. pallidum* proteins and the model spirochete *T. phagedenis.* My contribution to the data is the successful identification of a potential interaction between Tp0965 and a novel, uncharacterized *T. phagedenis* protein that may constitute a T1SS substrate in *Treponema phagedenis.*

5.2 Introduction

Research in the syphilis field is severely hindered by the inability to culture the syphilis spirochete, *T. pallidum.* Direct work with this pathogen requires the passage of bacteria using rabbits, and upon harvest, the bacteria remain viable and adequate for experimentation for only a few hours. The model treponeme *T. phagedenis*, on the other hand, although also delicate in its own way, is relatively easy to work with since it is culturable and not pathogenic to humans.

As described in the previous chapter, bioinformatic analysis predicted that *T. phagedenis* possesses homolog proteins to each of the components of the putative *T. pallidum* T1SS. The non-invasive nature of *T. phagedenis,* however, would suggest that the core function of the putative *T. pallidum* T1SS may have been augmented to facilitate a role in pathogenesis. Regardless, due to the difficulties of working directly with *T. pallidum*, and the possibility of *T. phagedenis* possessing a homologous T1SS, a global and unbiased pull-down approach was implemented. Hence, the degree of homology

between these two systems would allow to use the culturable spirochete *T. phagedenis* in an experiment tailored to the identification of novel potential T1SS substrates. The *T. phagedenis* novel putative T1SS substrates could then be used to screen the *T. pallidum* proteome to find a homolog that might be a virulence factor conferring upon this spirochete its characteristic pathogenesis.

As shown in Table 2, the MFP homolog in *T. phagedenis* is predicted to have a 59% identity to the *T. pallidum* MFP Tp0965, over a 92% query cover. This high degree of sequence similarity suggested that perhaps these MFPs are similar enough for Tp0965 to cross-interact with substrates from the *T. phagedenis* potential T1SS machinery. A global and unbiased analysis via pull-down experiments was performed using the *T. pallidum* MFP (Tp0965) from the putative T1SS as bait, to investigate protein-protein interactions with a *T. phagedenis* lysate prey. These studies were carried out with the goal of identifying more potential T1SS substrates in *T. phagedenis* via mass spectrometric analysis. Bioinformatic analyses were then used to find proteins with sequence and/or structure similarity in *T. pallidum* that might be a T1SS effector in the syphilis spirochete.

5.3 Potential interaction of the putative *T. pallidum* T1SS MFP component, Tp0965, with a novel *T. phagedenis* protein

As shown in Figure 17, the SDS-PAGE gels resulting from the global and unbiased pull-down assays carried out with recombinant *T. pallidum* proteins and a *T. phagedenis* lysate, showed a protein band in the elution fraction in the gel run with the H-Tp0965^{rec} fractions (Figure 17A), whereas the negative control H-Tp0327^{rec} and the blank (Figure 17B and C, respectively) did not show such a band.

Under the conditions tested, most of the *T. phagedenis* proteins did not seem to interact with H-Tp0965^{rec}. Thus, as shown in Figure 17A, most of the proteins from the *T. phagedenis* lysate flowed through, past the beads (prey flow through) and were washed off leading to little to no protein through consecutive washes 4, 5 and 6. However, the elution fraction showed a new protein band, band X, at approximately 45 kDa, that was absent in the fraction containing the last wash, wash 6 and was therefore concluded to be a potential H-Tp0965^{rec} -interacting protein from *T. phagedenis*. There are three additional faint bands in this elution fraction, however, the H-Tp0965^{rec}-alone lane also shows three bands at the same approximate molecular weights, indicating that these are not proteins from the *T. phagedenis* lysate, but rather H-Tp0965^{rec} contaminants.

The negative control pull-down between H-Tp0327^{rec} and a *T. phagedenis* lysate showed the same trend where all the lysate seemed to be washed off throughout the prey flow through and the washing steps. Furthermore, as shown in Figure 17B, the elution fraction did not show any bands other than that corresponding to H-Tp0327^{rec}, thus it appears that H-Tp0327^{rec}, as expected, did not interact with any *T. phagedenis* proteins under the conditions tested.

Similarly, the blank sample containing just beads and a *T. phagedenis* lysate, did not show any protein bands in the elution fraction (Figure 17C), indicating that band X observed in the elution fraction in Figure 17A is not due to non-specific binding of the lysate to the beads, but rather a specific H-Tp0965^{rec}-interacting protein from *T. phagedenis*.

The results presented indicate a possible interaction between the *T. pallidum* H-Tp0965^{rec} and a novel *T. phagedenis* protein, suggesting cross-interaction between the

T. pallidum putative T1SS MFP Tp0965 and putative T1SS substrates from *T. phagedenis*. Coomassie Brilliant Blue stain was used to visually analyze all the gels run and to be able to further analyze gel excisions via mass spectrometry. Due to the low intensity of band X observed in Figure 17A, another set of SDS-PAGE gels was run and stained using the more sensitive Silver stain kit (BioRad) to show bands of higher intensity (Figure 18). Unfortunately, the samples were stored at -20 ^{o}C for approximately 10 months before the silver stains were run, potentially resulting in significant protein degradation. Thus, the band intensity observed in the silver stained gels was found to be very similar to the intensity seen in the Coomassie stained gels. Nonetheless, the presence of band X in the pull-down assay fractions obtained from the H-Tp0965^{rec} samples was shown in both sets of gels, stained via both silver staining methods and Coomassie staining methods, thereby increasing the confidence in our findings. Mass spectrometric analysis of excisions from the Coomassie stained gels was then carried out to identify the protein present in band X.

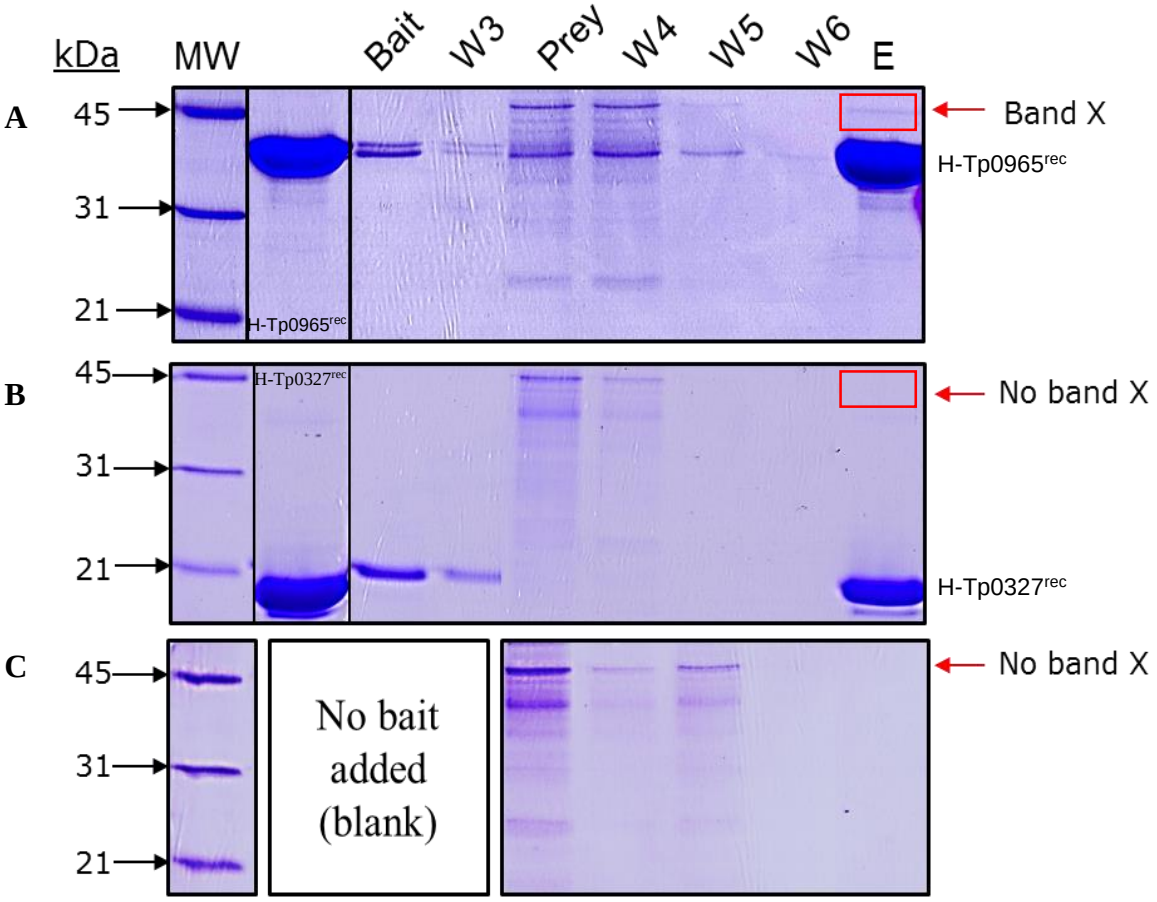


Figure 17. Coomassie Brilliant Blue stained SDS-PAGE gels for analysis of fractions from pull-down assays between recombinant *T. pallidum* proteins and a *T. phagedenis* lysate.

(A) Pull-down assay between H-Tp0965^{rec} and a *T. phagedenis* lysate; shows a visible putative H-Tp0965^{rec}-interacting *T. phagedenis* protein band, band X, in elution fraction. (B) Negative control sample H-Tp0327^{rec} and a *T. phagedenis* lysate; shows no visible band X in elution fraction. (C) Blank sample, only Ni-NTA beads with a *T. phagedenis* lysate and no bait; shows no visible band X in elution fraction. MW: Molecular Weight protein ladder, Bait: bait flow through, W3: Wash #3, Prey: prey flow through, W4: Wash #4, W5: Wash #5, W6: Wash #6, last wash after prey flow through, E: Elution fraction. Red box: band excisions for mass spectrometric analysis.

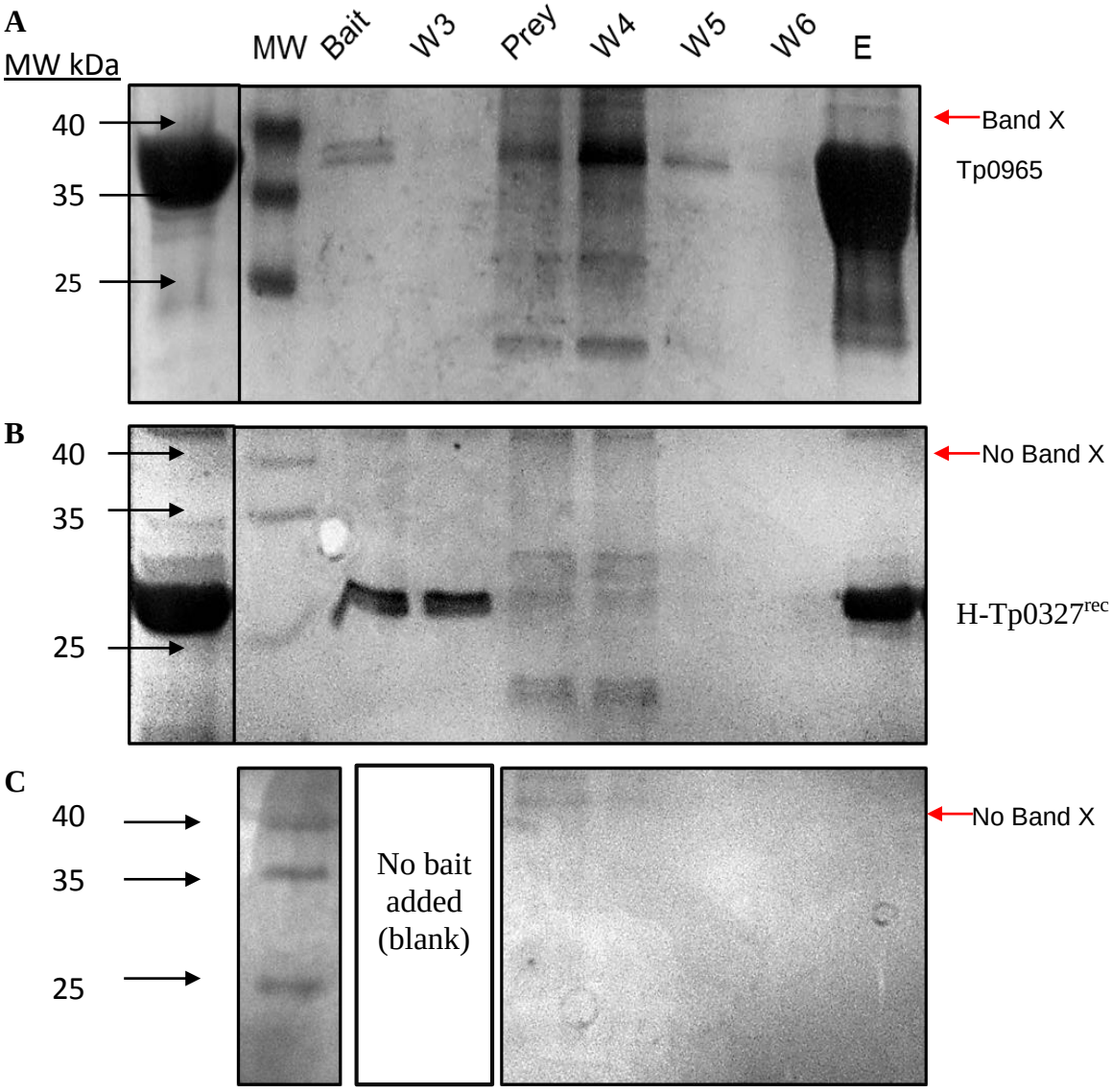


Figure 18. Silver stained SDS-PAGE gels for analysis of fractions from pull-down assays between recombinant *T. pallidum* proteins and a *T. phagedenis* lysate

(A) Pull-down assay between H-Tp0965^{rec} and a *T. phagedenis* lysate; shows a visible putative H-Tp0965^{rec}-interacting *T. phagedenis* protein band, band X, in elution fraction. (B) Negative control sample H-Tp0327^{rec} and a *T. phagedenis* lysate; shows no visible band X in elution fraction. (C) Blank sample, only Ni-NTA beads with a *T. phagedenis* lysate and no bait; shows no visible band X in elution fraction. MW: Molecular Weight protein ladder, Bait: bait flow through, W3: Wash #3, Prey: prey flow through, W4: Wash #4, W5: Wash #5, W6: Wash #6, last wash after prey flow through, E: Elution fraction. MW: Thermo Fisher Scientific Spectra Multicolor Broad Range Protein Ladder (catalog number 26634).

5.4 Mass spectrometric analysis of band X identified a novel *T. phagedenis* protein that may interact with the MFP from the putative *T. pallidum* T1SS.

Mass spectrometric analysis was used to find the identity of the novel *T. phagedenis* protein pulled-out with recombinant *T. pallidum* H-Tp0965^{rec}. Band X, visualized at approximately 45kDa was excised from the Coomassie-stained SDS-PAGE gel shown in Figure 17A, digested with trypsin as described in section 2.2.4 and then subjected to mass spectrometric analysis as mentioned in section 2.2.5. Two peptides, unique to *T. phagedenis* were successfully obtained via MALDI-TOF/TOF MS and upon searches against the MASCOT database, a match was found. As shown in Table 3, band X was identified as a *T. phagedenis* protein labelled as "conserved exported protein", with accession number CEM60944 with a protein sequence coverage of 9%. The high MASCOT score of 235, and the predicted molecular weight of 43 kDa, which is highly similar to the observed 45 kDa in the SDS-PAGE gels shown in Figure 17A, and Figure 18A give high confidence in the protein identification. For simplicity, this *T. phagedenis* protein will be referred to as TphBIg (for *T. phagedenis* protein with a BIg domain, as explained in the following sections).

Moreover, the portion of the negative control gel H-Tp0327^{rec}, shown in a red box in Figure 17B, and corresponding to the absence of band X was also subjected to trypsin digestion and mass spectrometric analysis. However, no proteins were identified in this sample, supporting the unique presence of band X in the H-Tp0965^{rec} - *T. phagedenis* lysate pull-down elution fraction, as shown in Figure 17A.

Figure 19A shows the primary sequences of the two peptides used for protein identification, highlighted within the TphBIg protein sequence in red and blue, and Figure

19B shows the corresponding MS/MS spectra. The peptide sequence coverage was 100% in both cases, further increasing confidence in the *T. phagedenis* protein identification.

Table 3. MASCOT results obtained for a novel potential Tp0965-interacting partner from *T phagedenis*.

Bait protein	Observed MW (SDS-PAGE)	MASCOT Hit	Hit MW	MASCOT Score	Peptide Hits	Protein Sequence Coverage
H-Tp0965^{rec}	45kDa	Conserved Exported Protein Accession: CEM60944	43kDa	235	RYKNPFQILEEGAFDISYRG KEAVYTVFPSDSVVDAEGRI	9%
H-Tp0327^{rec}	N/A	N/A	N/A	N/A	N/A	N/A
N/A (beads only)	N/A	N/A	N/A	N/A	N/A	N/A

A

MIKKTLLGTACAVFFFSAWAQAPQTIGTDYQRVATHYSDGKKEFVNSDVFFKLNASDKETGLDFVEFSLDGTKFM**RY**
KNPFQILEEGAFDISYRGLDNSRNLEVPKTLSVVVDNTPPKAEIETTEPVYRKGLTTYCSANTKWYVSASDNLTGAGV
AGTYIGTNLQALELRGKGKEAEDAYFSFESEGPAKLYYTALDNVGNLTPIALTSVIVDMTPPVIYLENSDRLIN**KEAVY**
TVFPSDSVVDAEGRIIISTSEAIAFGAKDELSGLDAIYIKINDAEYTKYVEPIKFNTEDVYTIEVKAIDNVGNVSEPVTYTF
YVDKINPASSVEMIDRSGNKLDTITPDGTIPAEGSAEVTPVESDIETAPAEEAVEVTPEEEETDVVPAEDEVPVLE

B

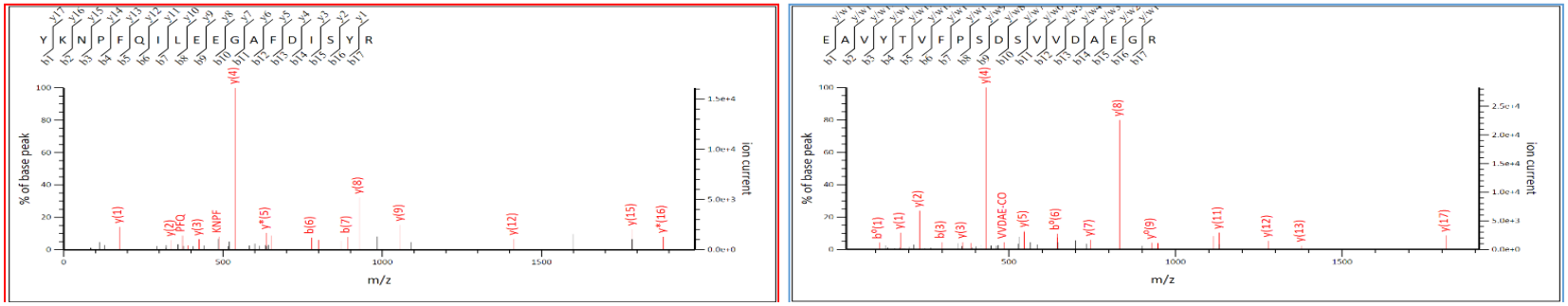


Figure 19. Mass spectrometric analysis results for the identification of a *T. phagedenis* protein that showed a potential interaction with H-Tp0965^{rec}

(A) Sequence of the newly identified *T. phagedenis* labelled as "conserved exported protein" (labelled TphBIg in this thesis); (B) mass spectra showing the complete breakage of the two unique peptides (blue and red) used for its identification upon finding a potential interaction with H-Tp0965^{rec}

5.5 Bioinformatic analysis shows that TphBIg may possess characteristics of a T1SS substrate

Upon the identification of the novel *T. phagedenis* protein, TphBIg, it was of interest to carry out bioinformatic analysis to characterize this protein *in silico* and determine if it could potentially be a T1SS substrate within this model treponeme.

A series of analyses showed that TphBIg is predicted to have sequence similarity to the bacterial immunoglobulin-like (BIg) domain. A Conserved Domain (CD) search with the National Center for Biotechnology Information (NCBI) showed that TphBIg contains, close to its C-terminus, a sequence of 35 residues in length corresponding to a putative BIg domain, group 3, shown in Figure 20A. BIg domains can be found in a variety of putative bacterial surface proteins with a wide variety of functions, including proteins that play roles in pathogenesis (Matsunaga *et al.*, 2003). For instance, virulence factors such as intimin and invasin from *E. coli* and *Yersinia pseudotuberculosis*, respectively, have been shown to play a role in host cell invasion, and Lig proteins from *Leptospira interrogans*, seem to function as adhesins that aid in bacterial host cell-binding (Matsunaga *et al.*, 2003; Wang *et al.*, 2013). Furthermore, BIg-domain repeats in some virulence factors are thought to be linkers that project a C-terminal domain away from the bacterial cell to act as the host-cell receptor required for interaction (Luo *et al.*, 2000). Interestingly, the BIg domain was also recently documented as a novel, strong Ca^{2+}-binding module, suggesting a role in calcium-dependent processes. Structural studies of the *Streptococcus pneumoniae* potential surface protein SP0498, showed an NMR structure of its BIg domain forming a β-barrel-like fold able to bind calcium (Wang *et al.*, 2013).

As shown in Figure 20B, a search for similar domain architectures to the novel TphBIg protein via the Conserved Domain Architecture Retrieval Tool, NCBI, suggested that the architecture of TphBIg shows similarities to a wide array of BIg-domain-containing proteins, including a number of predicted biofilm-promoting factors and adhesins. Some of these are putative T1SS substrates, such as the Biofilm-Associated Protein (BAP), found in *Acinetobacter baumannii* (De Gregorio *et al.*, 2015; Goh *et al.*, 2013). Others have been shown to be T1SS substrates, such as Bap/RTX cell surface protein (BpfA), a Ca^{2+}-binding protein mediator of biofilm formation found in the soil opportunistic pathogen *Shewanella oneidensis* (Theunissen *et al.*, 2010).

Furthermore, a Phyre2 structure prediction of TphBIg, depicted in Figure 20C, showed highest architectural similarity to SiiE from *Salmonella enterica*, with 68% coverage, 14% identity and 99.8% confidence. SiiE, is a giant Ca^{2+}-binding and BIg-domain-containing adhesin known to be secreted via the T1SS. The crystal structure of a fragment from SiiE was recently published and provided further evidence of the ability of BIg-domains to interact with calcium. SiiE has been shown to be secreted via a T1SS and promotes initial contact of *Salmonella* to host cells (Griessl *et al.*, 2013). The predicted structural similarity between TphBIg and SiiE suggests a potentially shared function and perhaps even shared means of secretion.

Finally, the TphBIg primary sequence was further analyzed with the goal to identify clear T1SS substrate characteristics. Although TphBIg did not show any of the typically nonapeptide glycine-rich RTX repeats (consensus sequence GGXGXDXUX), some of the known T1SS substrates, such as HasA from *S. marcescens* (Arnoux *et al.*,

1999), also lack clear similarities to the consensus T1SS substrate characteristics, such as RTX repeats, or a perfect amphiphilic α-helix.

However, TphBIg did show an amphipathic α-helix close to the C-terminus, as depicted in Figure 20D. The amphipathic α-helix in TphBIg showed 7/9 non-polar residues on one side, and 6/9 polar residues on the other side, which is close to the consensus amphipathic α-helix in the known T1SS substrate, HlyA, from *E. coli*. This amphipathic α-helix contains 9/9 and 7/9 non-polar and polar amino acids, respectively. This observation further suggests that TphBIg might be a T1SS substrate in *T. phagedenis*.

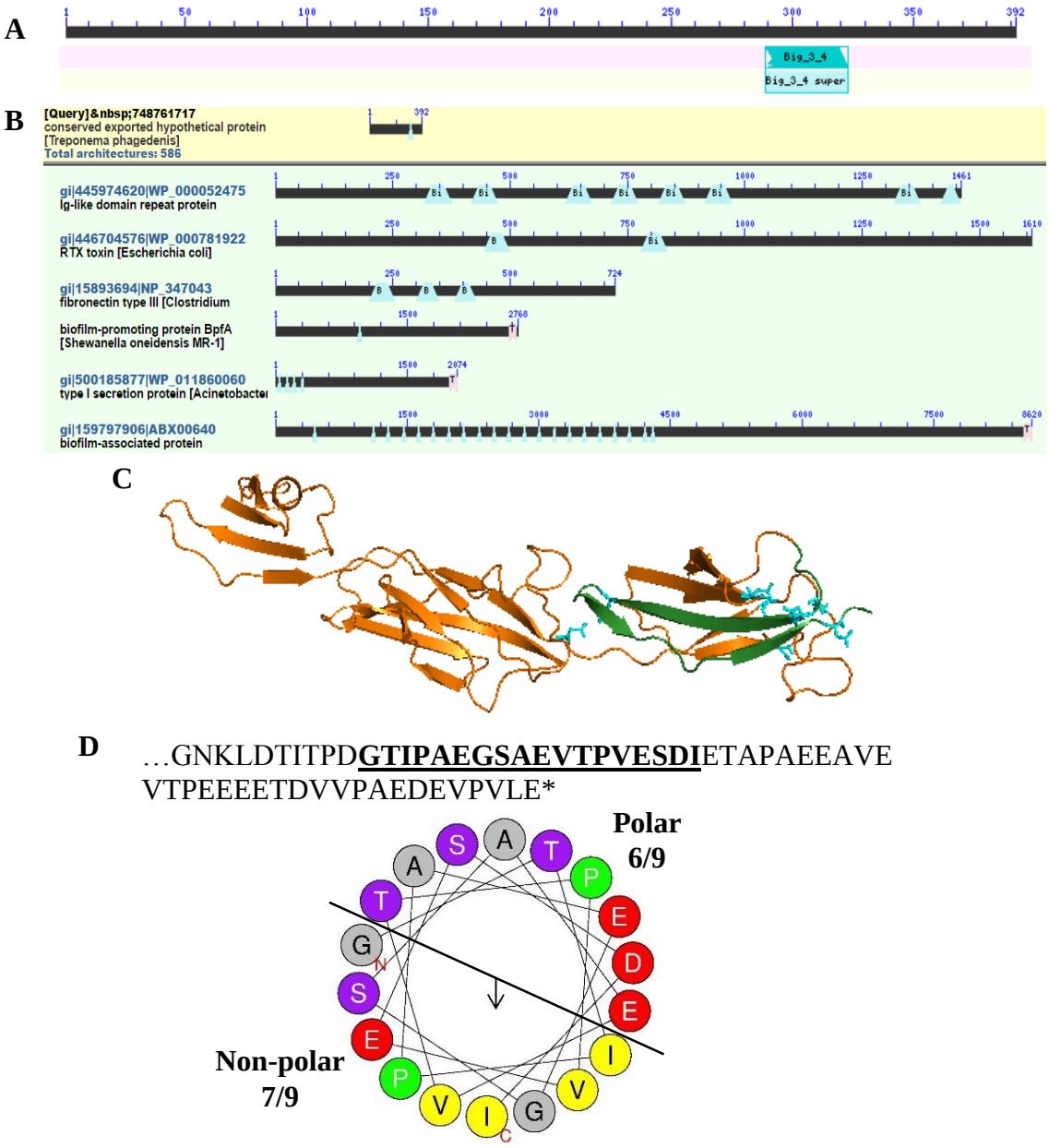


Figure 20. Bioinformatic analyses of a novel putative T1SS substrate from *T. phagedenis*

(A) Single BIg-domain found in TphBIg novel *T. phagedenis* protein (Marchler-Bauer and Bryant, 2004). (B) Results obtained from a search in the Conserved Domain Architectural Retrieval Tool, showing similar domain architectures between TphBIg and T1SS substrates including RTX toxins, and BIg-domain containing proteins, among others; the score obtained was 1 since there was a single domain match (i.e. the BIg-domain). (C) Phyre2-predicted structure of TphBIg, BIg-domain highlighted in green, potential Ca^{2+}-interacting aspartic acid residues within and around the BIg-domain, shown in cyan. (D) α-helical wheel prediction of the TphBIg C-terminus showing a pseudo-amphiphilic α-helix (HeliQuest analysis); linear primary sequence in bold and underlined; note that this sequence is not present in the Phyre2 prediction (Kelley *et al.*, 2015).

5.6 Conclusions

The bioinformatic preliminary characterization of a putative T1SS in both *T. phagedenis* and *T. pallidum*, as well as the predicted similarities between the MFP components from both systems, was essential in providing a strong foundation to carry on with *in vitro* pull-down assays between a *T. phagedenis* lysate and the *T. pallidum* MFP Tp0965. Thus, global and unbiased pull-down assays between the putative MFP of a potential T1SS in *T. pallidum*, Tp0965, and a *T. phagedenis* lysate followed by mass spectrometric analysis resulted in the successful and confident identification of a novel putative T1SS substrate in *T. phagedenis*, TphBIg.

Detailed *in silico* analysis of the Tp0965-interacting TphBIg, using bioinformatic tools such as BLAST, CD searches, Phyre2, HeliQuest, as well as manual bioinformatic analyses showed that TphBIg has a BIg-domain, and possesses clear structural similarities to T1SS substrates. Hence, overall *in silico* analysis of the novel TphBIg suggests that there is a high possibility of this protein being a T1SS substrate, giving additional confidence to the previous findings that indicated a potential interaction with Tp0965.

Although *in silico* studies suggest that TphBIg might be a T1SS substrate, *in vitro* analysis using recombinantly produced TphBIg would be necessary to confirm the putative interaction with Tp0965 (or the *T. phagedenis* homolog WP_002695300.). However, since *T. phagedenis* is not a pathogen that affects humans, but rather a model used to study pathogens, such as the non-culturable *T. pallidum*, follow-up studies to characterize the putative *T. phagedenis* T1SS and its substrates is of minor interest at this moment. Nonetheless, the cross-interaction between Tp0965 and TphBIg does support the presence of a putative T1SS both in *T. phagedenis* and in *T. pallidum*. Furthermore, these findings

represent an important step towards the indirect identification of novel *T. pallidum* putative T1SS substrates, as well as a key step toward the development of a methodology for the direct identification of T1SS substrates in *T. pallidum.*

Chapter 6: Indirect identification and *in silico* characterization of Tp0854, a novel putative T1SS substrate from *T. pallidum*

6.1 Contributions to the data

This is the first study showing a potential T1SS substrate in *T. pallidum* that was indirectly identified via a global, unbiased pull-down assay between the recombinant *T. pallidum* protein Tp0965 (**P**^{40}-**A**^{318}) and a *T. phagedenis* lysate followed by bioinformatic analyses to identify potential T1SS substrates. My contribution to the data is the finding of a novel *T. pallidum* putative T1SS substrate via an indirect approach using a lysate from the culturable, model treponeme *T. phagedenis*, as well as the development of a new methodology applicable to *T. phagedenis* and possibly directly to *T. pallidum* for *de novo* substrate identification of predicted secretory complexes and/or protein-protein interactions in general.

6.2 Introduction

Bioinformatic analyses in the field of molecular biology have become powerful tools to characterize novel proteins. These tools are especially important in fields where experimental procedures are challenging and expensive. For instance, it is thought that *T. pallidum* virulence factors might be key in the highly invasive nature of this pathogen. However, the inability to culture and genetically modify *T. pallidum* makes it impossible to study the specific functionalities of each protein *in vivo.* Hence, the availability of a model treponeme, *T. phagedenis* as well as powerful bioinformatic tools and recombinant protein expression systems are key to enable scientists to indirectly study the pathogenesis of *T. pallidum*.

The preliminary characterization of a putative T1SS in *T. pallidum* and potential substrates began with *in vitro* assays using recombinant *T. pallidum* proteins, but the identification of more putative T1SS substrates continued with an indirect approach, using *T. phagedenis* as a bridging tool to identify *T. pallidum*'s T1SS substrates that may confer upon this spirochete, properties that contribute to its characteristic pathogenesis.

Upon identifying a Tp0965-interacting *T. phagedenis* protein, denoted TphBIg, it was of interest to scan the *T. pallidum* proteome in search for a homolog within the syphilis spirochete that might possess sequence similarity and hence be a potential T1SS substrate.

6.3 Bioinformatic identification of a *T. pallidum* protein with sequence similarity to the putative Tp0965-interacting, TphBIg

Bioinformatic analysis showed that TphBIg possesses sequence similarity to Tp0854 from *T. pallidum*. It is important to mention that during our investigations, the *tp0854* nucleotide sequence was found to differ from that in the literature. As shown in Figure S6, DNA sequencing analyses carried out revealed that Tp0854 contains an aspartic acid residue at amino acid position 492, instead of a glycine residue as is suggested in the Tp0854 sequence available in the literature. Hence, bioinformatic analyses were carried out using the corrected sequence (G→D).

The amino acid sequence of the newly identified TphBIg from *T. phagedenis* was submitted to BLAST and searched against the *T. pallidum* proteome in hopes to identify a protein with sequence similarity that might be a putative T1SS substrate. The highest BLAST score obtained was 30.8 with an E-value of 0.19, 9% query cover, 34% identity and corresponded to the *T. pallidum* protein Tp0854. This *T. pallidum* protein is 1533

residues long, whereas TphBIg is only 392, and at first glance they appear to possess different functionalities.

The amino acid sequences of TphBIg and Tp0854 were then aligned using the Clustal W 2.1 software (Larkin *et al.*, 2007), resulting in a calculated 20.95% identity matrix between the two protein sequences. The lower percent identity value obtained from the Clustal W 2.1 software can be explained by the differences in the alignment methods used by BLAST (local alignment method) compared to Clustal W 2.1 (global alignment method). Thus, percent identity obtained by BLAST was calculated from the 9% protein coverage stretch of 38 residues that were found to be similar between the two proteins. Clustal W 2.1 however, calculated the percent identity based on the full primary sequence, taking into account gaps and mismatches along the whole sequences, and resulting in a different value from that obtained via BLAST analysis.

Tp0854, like TphBIg was predicted to have a BIg-domain, highlighted in green in Figure 21, however, Clustal W 2.1 failed to align the BIg domain sequences of these two proteins. Interestingly, as shown in Figure 21, the strongest similarity was instead found to be between part of the BIg-domain from TphBIg and a Tp0854 uncharacterized stretch of amino acids with no predicted domain.

Since the majority of T1SS substrates possess a C-terminal secretion signal, the amino acid sequences of the C-termini of TphBIg and Tp0854 were analyzed. Clustal W 2.1 was used to align the last 90 residues of these proteins, as shown in Figure 22, leading to a percent identity matrix of 23.08%. Although this sequence similarity is weak, there has been evidence showing that the T1SS signal has no sequence consensus and is not conserved between different bacteria (Boyd *et al.*, 2014). Furthermore, proteins such as

HasA show extra T1SS-interacting sites dispersed throughout the protein, which remains unfolded within the cytoplasm exposing these primary recognition sites, ready to contact the T1SS (Masi and Wandersman, 2010). Moreover, secondary structure characteristics have been reported to be involved in secretion (Boyd *et al.*, 2014; Filloux, 2010). Thus, it is unclear at this point how both TphBIg and Tp0854 could interact with Tp0965, but it can be hypothesized that it may involve subtle shared primary and mostly secondary structures, perhaps including the BIg domain secondary structure, which is shared between the two proteins, and described in Chapter 6.5. Moreover, the hybrid cross-secretion observed in many T1SS, and as described in Section 1.4.2.1.4, provides evidence of this possibility. Nonetheless, in depth studies using techniques such as crosslinking, would need to be carried out to determine the specific interacting sites and means of secretion for both TphBIg and Tp0854.

```
TphBIg      ----------------MIKKTLLGTACAVFFFSAWAQAPQTIGTDYQRVATHYS-DGKKE
Tp0854      GNTWSAPVPVTQPDEYHNQRPFLDRLSDDRFAVTWERSERTS-TRYEMCYAELDRYGRK-
                           :: :*   .   *  :* :: :*  * *:   :. .  *:*

TphBIg      FVNSDVFFKL-NASDKETGLDFVEFSLDGTKFMRYKNPFQI--------LEEGAF-----
Tp0854      -IG--TTLRLAEPSDRLITPNF--VHIDGTTFCVWAGESAGLNTIFLAQKKEGAWSTTAV
             :   . ::* : **:    :*  . :***.*  :               :***:

TphBIg      -------------------DISYRGLD----------NSRNLEVPKTLSVVVDNTPPKAE
Tp0854      RSSEDALLFPHAVRVDNHLEVFWQEGEGARARVMRLRPDQSVQPPT---LIAENFSPNAV
                           :: ::  :           .:.:: *.   ::.:*  *:*

TphBIg      IET----TEPVYRKGLTTYCSANTKWYVSASDNLTGA-----------GVAGTYIGTNLQ
Tp0854      RKGTRARVRIVFPRDSSGIAGYNYAWQCGVQPAAPPDYVAHFPDKPQIELEATQDGTWFL
             :     .. *: :  :  .. *  *  ...                  : .*  ** :

TphBIg      ALE-LRGKGKEAEDAYFSFESEGPAKLYYTALDNVGNLTPIALTSVIVDMTPPVIYLENS
Tp0854      AVTVWDFAGNKSAPAYLSYTRDTT---------------PAARPQ----LQTP--LLENT
            *:      *:::  **:*:  :                 * *  .    :  *   ***:

TphBIg      DRLINKEAVYTVFPSDSVVDAEGRIIIS-TSEAIAFGAKDELSGLDAIYI----------
Tp0854      HALKS--NTFTLSWNQPSTDAQGNEERDHTSFLWSLQQVAPLSALTSLRVDTDVRTFEEF
            . * .   .:*:  .:  .**:*.   . **   ::     **.* :: :

TphBIg      ------------KINDAEYTKYVEPIKFNTEDVYTIEVKAIDNVGNVSEPVTYTFYVDKI
Tp0854      QQRCVRAFPIPVDVHGTRSRQSSVSFTNKENGIYRFSVYALDRSGNVSEPAVVFFALRHF
                        .:. :.  :    :. : : :* :.* *:*. ******..  * : ::

TphBIg      NPASSVEMIDRSGNKLDTIT---PDGTIPAEGSAEVTPVESDIETAPAEEAVE-----VT
Tp0854      VPYTAIRYVDVKKDPAGSLQMSIVGNGFRAQGTVSQVYIDRDR-KAPYDLVLHAQEFAVG
             * :::. :* . :   ::        : *:*:.. . :: *  .** : .:.     *

TphBIg      PEEEE---------------------TDVVPAEDEVPVLE--------------------
Tp0854      SDNLISDIHIDNLKKGSYHVGVWHPARGVHFAESRVTVSEMGTVKFGAYDYEHQVRWSIP
             ::                          *  **..* * *
```

Figure 21. Clustal W 2.1 amino acid sequence alignment of *T. phagedenis* TphBIg and *T. pallidum* Tp0854 (G241-F732).

Predicted BIg-domains are highlighted in green; note that the red highlighted residue has been changed from the literature glycine to an aspartic acid as per new sequencing data. Most significant alignment shown in blue box and agrees with BLAST finding (Larkin et al., 2007).

```
Tp0854      SCEGNLDEAVLALVAVEQVF-----RMYKHPRATNLDKIRVDKKVDMFLA------R-YF
TphBIg      DNVGNVSEPVTYTFYVDKINPASSVEMID-RSGNKLDTITPDGTIPAEGSAEVTPVESDI
            .  **:.* *   . *:::      .* .   ..:**.*  * .:    :      .  :

Tp0854      VQYPEYCARKEVNSEYEEYLYYTFIKEDDQYDDLTILGIRKR
TphBIg      ETAPAE-EAVEVTPEEEET---DVVPAE---DEVPVLE----
               *      **. * **     .:  :   *:: :*
```

Figure 22. Clustal W 2.1 amino acid sequence alignment of the C-termini of *T. pallidum* Tp0854 and *T. phagedenis* TphBIg

6.4 A Conserved Domain search shows that Tp0854 may be a multi-domain, multi-functional protein

Upon identifying Tp0854 as the protein in *T. pallidum* with highest similarity to TphBIg, its amino acid sequence was analyzed using several bioinformatic platforms as well as manual analysis to characterize this novel *T. pallidum* protein *in silico* and infer on the possibility of it being a virulence factor secreted via the putative T1SS.

The corrected full-length Tp0854 amino acid sequence was submitted to a CD search that predicted a multi-domain and possibly a multi-functional protein. As shown in Figure 23, Tp0854 from *T. pallidum* is predicted to have a sialidase domain close to its N-terminus, followed by a BIg-domain, similar to TphBIg. Further toward the C-terminus it is predicted to have a dimerization interface with a Histidine kinase, Adenylyl cyclase, Methyl-accepting protein, and Phosphatase (HAMP) domain and finally a predicted phosphatase domain at the C-terminus (Marchler-Bauer and Bryant, 2004).

Sialidases are hydrolases able to cleave glycosidic linkages of terminal sialic acid residues in oligosaccharides, glycoproteins and glycolipids, among other substrates and have been widely reported in the literature as virulence factors used by a broad spectrum of pathogens as well as non-pathogens (Gaskell *et al.*, 1995; Kim, Oh, *et al.*, 2011). Sialic acids are a family of monosaccharides present on all cell surfaces of vertebrates and on some vertebrate-interacting bacteria. Due to their abundance and diversity in structure and location, they possess multiple roles in immunity and pathogenesis. Sialic acids are involved in many molecular and cellular recognition events and act as receptors or masking structures that cover recognition sites. For instance, host sialic acid provides an effective means of communication between cells, including those from the immune system. Sialic acid-binding proteins on the surface of immune cells recognize sialic acid attached to host

cells, leading to immune evasion and host cell survival (Varki and Gagneux, 2012). Furthermore, Factor H from the complement system is able to bind sialic acid found on the surface of host cells and mediate cleavage of C3b, a crucial step toward inhibiting complement activation on host cells (Klein *et al.*, 2005; Varki and Gagneux, 2012). The importance of sialic acids is not limited to host-host cellular interactions, but can also be used as receptors by pathogens such as bacteria, fungi, protozoa and toxins. Indeed, some pathogens possess sialidases that are able to cleave sialic acid from host cell surfaces in order to use sialic acid as a nutrient, to expose glycan-binding sites found under sialic acids, to disrupt host cellular communication, or to coat their surfaces with sialic acid as an immune evasion strategy (Varki and Gagneux, 2012). Further, the Tp0854 putative sialidase domain is predicted to be followed by a BIg-domain which, as predicted in TphBIg, could also be involved in calcium-binding and host-cell interactions.

The predicted HAMP domain in the second half of Tp0854 is situated between the BIg-domain and the phosphatase domain, and might be involved in cell signaling. These domains have been documented to connect extracellular sensory structures with intracellular signaling domains in proteins such as histidine kinases, chemotaxis proteins, adenylyl cyclases and phosphatases (Hulko *et al.*, 2006). The CD search predicts a dimerization interface within the HAMP sequence (Marchler-Bauer and Bryant, 2004), suggesting that if this prediction is true, Tp0854 may dimerize to carry out its function, which may be to transmit a conformational change in the domain downstream, the phosphatase, and thus aid in phosphorylation regulation.

Finally, Tp0854 is predicted to have a serine/threonine PP2C protein phosphatase domain closest to the C-terminus. Phosphatases are important regulators of many signaling

and metabolic processes (Cohen, 1989). Specifically, the PP2C serine phosphatases, such as RsbP and RsbU, which are required for response to energy and environmental stresses, respectively, regulate transcription in bacteria by interacting with σ transcription factors (Vijay *et al.*, 2000). PP2C-type phosphatases seem to have an absolute requirement for Mg^{2+} ions and are also involved in signaling pathways in eukaryotic cells (Schweighofer *et al.*, 2004). For instance, in mammals, the PP2C phosphatases have been implicated in stress response signaling (Fujimoto *et al.*, 2005) and have been shown to act directly on activated Mitogen-Activated Protein Kinases (MAPKs), which have a key role in cell signaling and are essential in eukaryotic physiology (Martín *et al.*, 2005; Saxena and Mustelin, 2000). Thus, due to the ubiquitous nature of phosphatases, and the potential for Tp0854 to be a secreted *T. pallidum* virulence factor with phosphatase activity, it is possible that it may act to regulate host processes (described in more detail in Section 6.8).

Additionally, the full, corrected Tp0854 amino acid sequence was subject to the TMHMM online tool, which resulted in the prediction of three putative transmembrane helices at positions 753-775, 816-838 and 1085-1107, with a calculated probability close to 100%. Thus, if Tp0854 is indeed a T1SS substrate secreted from *T. pallidum*, and like most T1SS substrates mentioned before, interacts with host cells, the three predicted transmembrane helices would suggest membrane association. Hence, one of the Tp0854 domains would remain in the extracellular space and the other could be inserted into the host cell. Thus, it is thought that upon secretion, the potential sialidase domain could remain in contact with the extracellular space and be involved in cleaving sialic acid from host cells, while the phosphatase domain becomes inserted into the host cell where it can regulate cellular processes (more details in section 6.8).

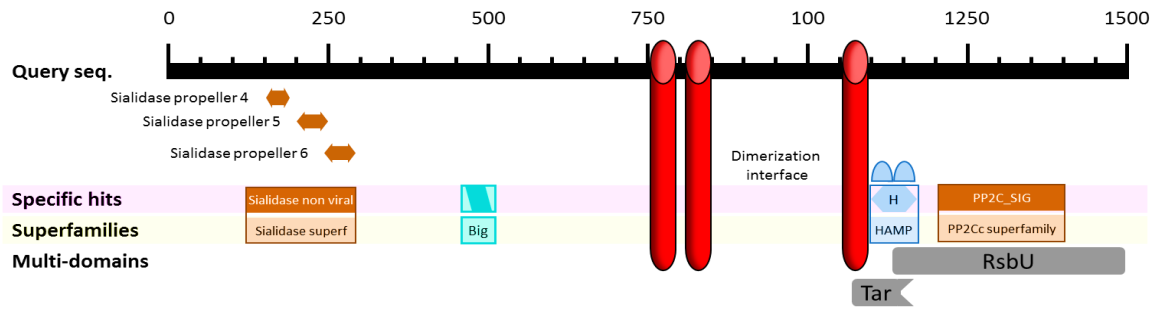


Figure 23. Tp0854 Conserved Domain search results.

Tp0854 is predicted to possess a sialidase domain closest to the N-terminus, followed by a BIg-domain, a HAMP domain and a PP2C-type phosphatase closest to the C-terminus (Marchler-Bauer and Bryant, 2004). The three predicted transmembrane helices are shown in red (Krogh *et al.*, 2001).

6.5 Phyre2 structure predictions agree with the Tp0854 domain structure found by the Conserved Domain search

Phyre2 predictions provided supporting evidence showing that Tp0854 may have a sialidase domain, a BIg-domain, and a regulatory domain with potential PP2C-like phosphatase activity. As shown in Figure 24A, the predicted sialidase structure shows characteristics common of bacterial sialidases, which have been shown to have six β-propeller structures, an immunoglobulin module and a galactose-binding domain (Gaskell *et al.*, 1995). The Tp0854 sialidase structure prediction shows six β-propeller structures containing Asp-box-motifs, shown in Figure 24A in pink, which is another characteristic of bacterial sialidases. These motifs have the consensus sequence Ser/Thr-X-Asp-[X]-Gly-X-Thr- Trp/Phe (Gaskell *et al.*, 1995), and the Tp0854 Asp-boxes possess the sequences Ser-Glu-Asp-Gly-Glu-Arg-Trp, Ser-Phe-Asp-Gln-Gly-Asn-Thr-Trp, and Ser-Asp-Asp-Arg-Phe-Ala-Val-Thr-Trp and are found at positions 156-190, 200-250, and 258-293, respectively. The immunoglobulin domain characteristic of sialidases is also predicted to be part of Tp0854, as shown in Figure 24B in cyan. This Tp0854 BIg- domain, like in the case of TphBIg, was predicted to show structural similarity to SiiE from *Salmonella*

enterica and thus might be involved in Ca^{2+}-binding. In bacterial sialidases, an immunoglobulin-like fold serves as a linker to an additional domain closer to the C-terminus, which has previously been documented as a galactose-binding domain (Gaskell *et al.*, 1995). In the case of Tp0854, however, it is predicted to be a phosphatase, as depicted in Figure 24C. To our knowledge, this is the first putative sialidase with a potential phosphatase domain documented.

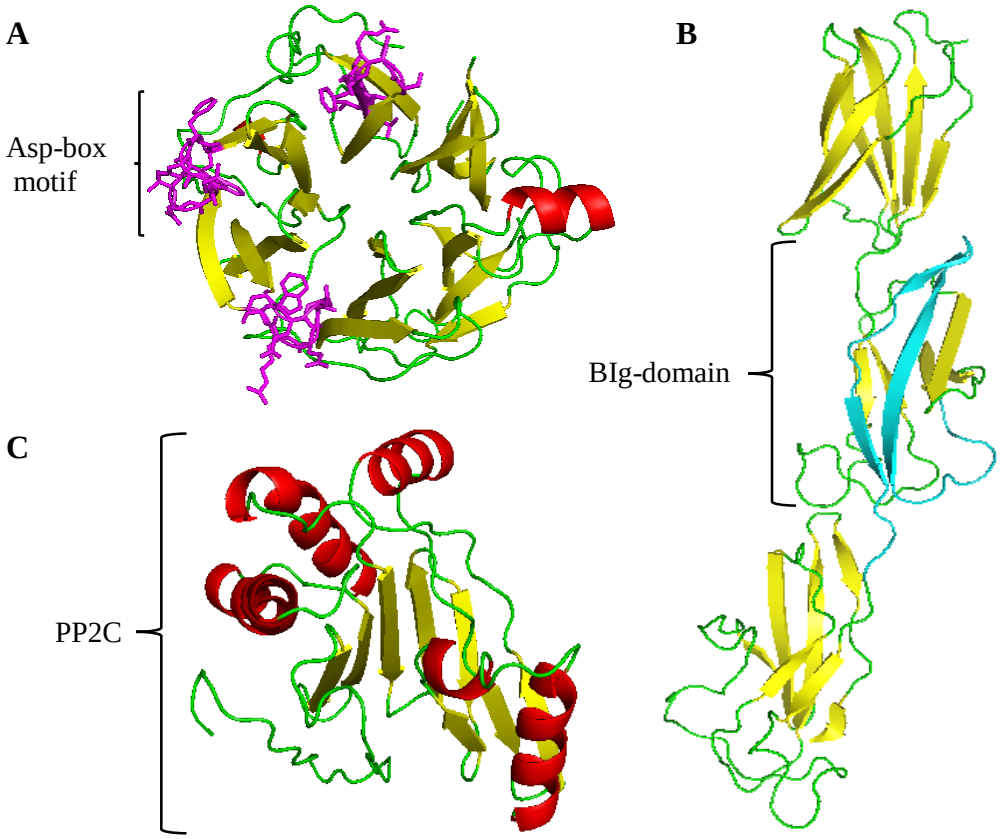


Figure 24. Tp0854 structure prediction according to Phyre2.

(A) Predicted structure of the sialidase domain found closest to the N-terminus, (B) predicted structure of the middle BIg-domain, (C) predicted structure of the PP2C phosphatase found closest to the C-terminus. Modeled using Phyre2 (Marchler-Bauer and Bryant, 2004).

6.6 Manual bioinformatic analysis of Tp0854 show characteristics of a T1SS substrate

The complete and corrected Tp0854 amino acid sequence was further analyzed manually in search for potential T1SS substrate characteristics resulting in the identification of a putative RTX motif, as well as a putative amphipathic α-helix.

Tp0854 has a potential RTX-motif with sequence Gly-Gly-Phe-Gly-Val-Glu-Thr-Leu-Thr, highlighted in yellow in Figure 25A. The consensus RTX motif contains an aspartic acid residue instead of a glutamic acid residue, which plays a key role in the RTX toxin known Ca^{2+}-binding capabilities. However, there is evidence of interactions between glutamic acid and calcium ions as well (Haber-Pohlmeier *et al.*, 2007; Vavrusova and Skibsted, 2013), suggesting that the Ca^{2+}-binding function of this motif should be conserved and that the presence of a glutamic acid instead of an aspartic acid in Tp0854 might retain a Ca^{2+}-binding function. Moreover, the presence of a BIg-domain could aid in conferring upon Tp0854 Ca^{2+}-binding properties. Nonetheless, *in vitro* experiments would need to be carried out to confirm any calcium-binding properties predicted *in silico.*

Furthermore, Tp0854 contains clusters of positively and negatively charged residues, such as aspartic acid, glutamic acid, arginine and lysine, right at the C-terminus, as shown in red in Figure 25A. These clusters at the C-terminus are also seen in some T1SS substrates, (Filloux, 2010).

Finally, as highlighted in Figure 25A, the last 60 residues at the C-terminus of Tp0854 were submitted to HeliQuest to create α-helical wheel diagrams and determine if it may form an amphiphilic α-helix. Unfortunately, this analysis did not provide compelling evidence for the presence of a C-terminal amphiphilic α-helix within this protein. As shown in Figure 25B, the helix possesses a non-polar half that is interrupted by the presence of a

polar, but uncharged glutamine. However, the other half of the helix with mostly polar residues is interrupted by a few non-polar residues including a proline, alanine, valine and methionine. Thus, the predicted α-helix at the C-terminus of Tp0854 does not seem to be truly amphipathic, indicating that if Tp0854 is indeed a T1SS substrate, other recognition signals, which are not clear at this time, may be required for secretion.

A

…AQTSAAGVFPSFMIDMK**GG**F**G**V**E**T**L**TLRTGDVLFLYTDGIEEAKRLFRN
KRFELVLCQEQGLAHDAPHETHTVGQAGEELGAERVSSIIESVFLRKGFSLQ
KWHNPVEGEKFEFDFSSCEGNLDEAVLALVAVEQVFRMYKHPRATNLDKI
RVDKKVDMFLARYFVQYPEYCARKEVNSEYEEYLYYTFIKEDDQYDDLTIL
GIRKR

B

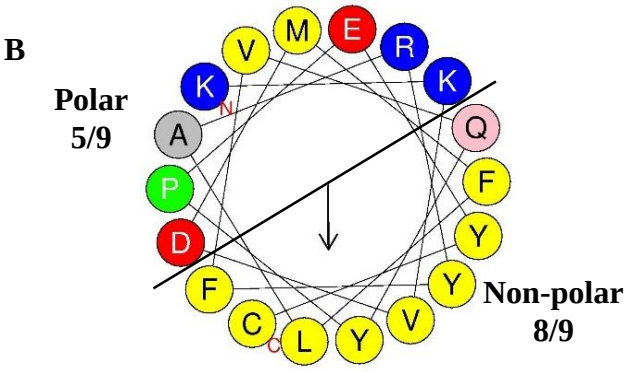


Figure 25. Predicted α-helix and RTX-motif in the *T. pallidum* protein Tp0854, a putative T1SS substrate.

(A) Portion of the Tp0854 primary sequence showing the last 60 residues at the C-terminus highlighted in green; a putative RTX motif highlighted in yellow; the charged residues toward the C-terminus in red; and the underlined sequence was used to generate the predicted α-helix. (B) HeliQuest-predicted α-helix where 8/9 residues are non-polar and 5/9 are polar, showing no amphiphilic nature.

6.7 Confirmatory binding assays show a potential interaction between recombinant *T. pallidum* Tp0854 and Tp0965

Since the *in silico* characterization of Tp0854 predicted potential T1SS substrate properties, recombinant Tp0854 was produced to carry out confirmatory *in vitro* analyses. The multi-domain nature and the large predicted size of Tp0854 suggested that soluble expression in *E. coli* would be difficult. Therefore, it was determined that the C-terminus phosphatase domain, which was predicted to contain the RTX repeat, would be the best domain to express solubly and carry out plate-based binding assays with recombinant Tp0965, as described in section 2.3.3.

Due to the fact that the MFP of other T1SSs is thought to form trimeric or hexameric complexes, it was of interest to test the interaction of H-Tp0965^{rec} with Tp0854^{phos} at a 3:1 ratio. The results in Figure 26A, showed a significantly higher absorbance value for the interaction of H-Tp0965^{rec} with Tp0854^{phos} compared to that of H-Tp0327^{rec} with Tp0965^{rec}, the negative control. Indeed, a student t-test performed for the values obtained from both samples resulted in a p-value lower than 0.005, suggesting a significant difference between the negative control, H-Tp0965 with Tp0327^{rec}, and the sample of interest, H-Tp0965 with Tp0854^{phos}. Furthermore, the interaction between H-Tp0965 and Tp0854^{phos} showed characteristics of a dose-response relationship, as shown in Figure 26B, where a constant concentration of 1.0 μM Tp0854^{phos} was tested with varying concentrations of H-Tp0965^{rec}, resulting in an evident decrease in absorbance with decreasing amounts of H-Tp0965^{rec}.

Equimolar concentrations were used to further assess an interaction between H-Tp0965^{rec} and Tp0854^{phos} (Figure 26C). Similar to the trend observed at a 3:1 ratio, the 1:1 molar ratio of H-Tp0965^{rec} to Tp0854^{phos} showed a much higher absorbance value

compared to the negative control, H-Tp0965^{rec} with Tp0327^{rec}, suggesting again a significant difference between the binding of H-Tp0965^{rec} to the negative control Tp0327^{rec} and the binding of H-Tp0965^{rec} to Tp0854^{phos}. This significant difference was confirmed by a student t-test which resulted in a p-value lower than 0.005. Hence, another experiment to test the dose-response of this interaction was carried out by using a constant concentration of 1.5 μM Tp0854^{phos} and varying H-Tp0965^{rec} concentrations to see a more detailed response. As shown in Figure 26D, an almost perfect dose-response was obtained with decreasing amounts of H-Tp0965^{rec}, confirming the results from Figure 26B.

Finally, the binding assay was reversed and H-Tp0854^{phos} rather than H-Tp0965^{rec} was used as the second protein, to investigate whether the trend of interaction observed would remain. As shown in Figure 26E, the absorbance resulting from the binding of H-Tp0854^{phos} to Tp0965^{rec} was twice as high as the absorbance observed for the interaction between H-Tp0854^{phos} and Tp0327^{rec}. Statistical analysis via a student t-test calculated a p-value below 0.05 indicating a significant difference between the binding of H-Tp0854^{phos} to Tp0965^{rec} compared to the binding of H- Tp0854^{phos} to Tp0327^{rec} (negative control). Moreover, the corresponding dose-response experiment, shown in Figure 26F, with a constant concentration of 1 μM Tp0965^{rec} and varying concentrations of H-Tp0854^{phos} showed a decrease in signal corresponding to a decreased amount of H-Tp0854^{phos}.

Overall, *in vitro* confirmatory binding studies using recombinant *T. pallidum* proteins suggested a specific interaction between the MFP of the putative T1SS (Tp0965), and the phosphatase domain of Tp0854. In comparison, there was minimal interaction between Tp0965 and Tp0327 (negative control). Thus, the results presented in this section support the *in silico* predictions mentioned above, suggesting that *T. pallidum* Tp0854 may

contain characteristics of a T1SS substrate. Thus Tp0854 appears to interact with the MFP of the potential T1SS, which could indicate secretion of this putative virulence factor.

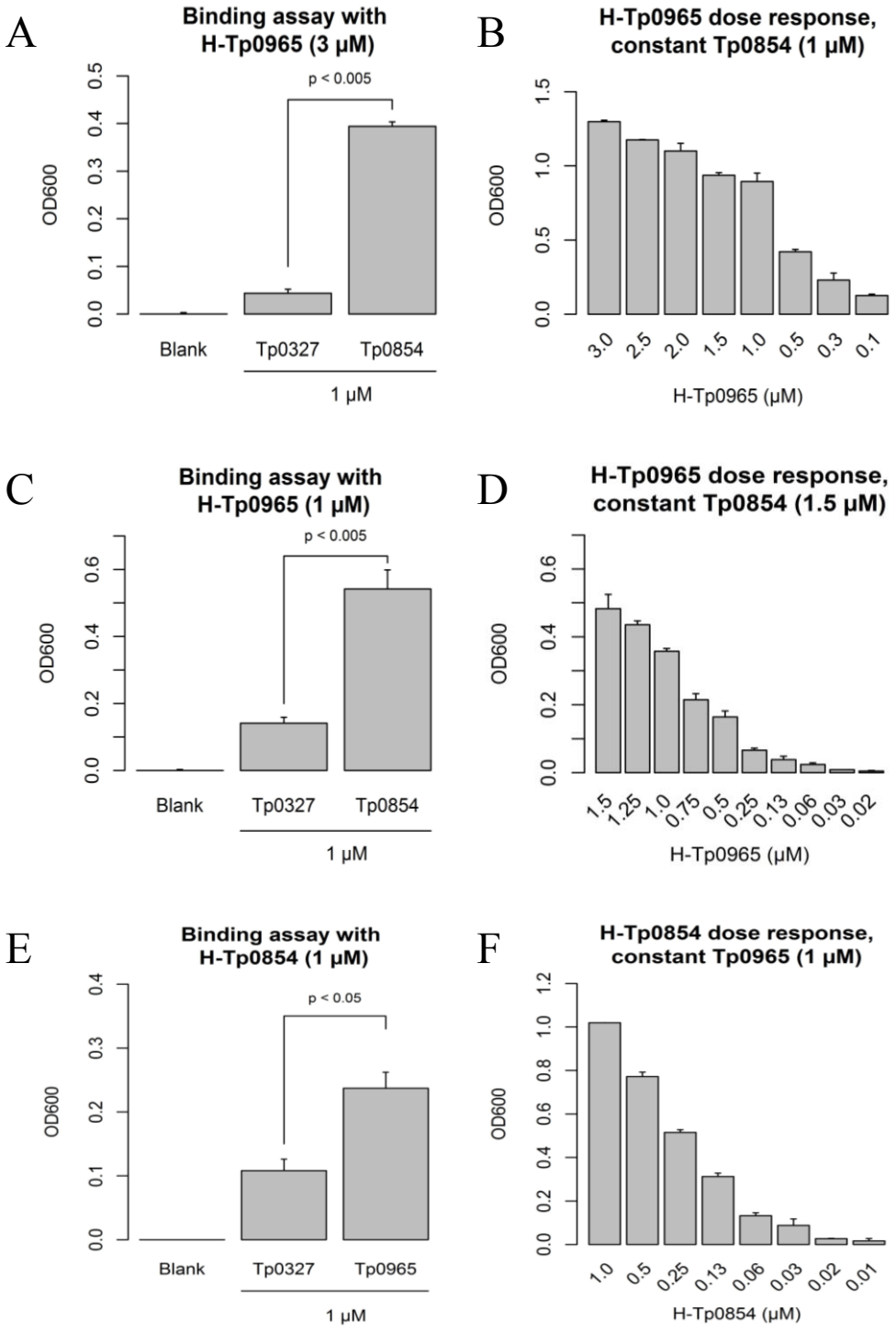


Figure 26. Confirmatory plate-based binding assays between recombinant *T. pallidum* Tp0854 and Tp0965, using Tp0327 as a negative control.

6.8 The potential role of Tp0854 in the context of *T. pallidum* pathogenesis

Tp0854 is predicted to have a sialidase and a regulatory phosphatase domain, both domain types of which have been shown to be used by other pathogens to favour survival. For instance, it was recently shown that the human parasite *T. brucei gambiense* possesses sialidase activity and is able to activate endothelial cells via cell signaling through the NF-κB pathway (Ammar *et al.*, 2013). *Clostridium chauvoei*, the causative agent of blackleg in cattle and sheep, has also been shown to contain the sialidase gene, *nanA.* NanA from this bacterium is thought to be secreted and associate in dimers to become active (Vilei *et al.*, 2011). Sialidases and toxins produced by *C. chauvoei* are believed to play a key role in the spread of the pathogen as well as lesions that result from invasion (Useh *et al.*, 2003). The epidemic strains of *Vibrio cholerae* have been shown to secrete a sialidase, key to achieve infection by facilitating entry into epithelial cells by targeting the host protective mucosal coat (Jermyn and Boyd, 2002; Ohyashiki *et al.*, 1987). Furthermore, this sialidase has been shown to cleave sialic acid from the receptor responsible for interacting with cholera toxin on epithelial cells, resulting in an increased number of available receptors (Merritt *et al.*, 1994). Finally, bacteria that are transmitted sexually have also shown to encode sialidases. *Gardnerella vaginalis,* found in women with bacterial vaginosis, has been shown to produce sialidases that aid in dissemination and that have been linked to premature birth and other adverse pregnancy outcomes (Cauci *et al.*, 2003; Lopes dos Santos Santiago *et al.*, 2011). The highly invasive and disseminating properties characteristic of *T. pallidum* have been well documented across the literature (LaFond and Lukehart, 2006; Magnuson *et al.*, 1948; Mahoney and Bryant, 1933) and thus, it is thought that the predicted sialidase domain of Tp0854 may function in sialic acid cleavage to either

expose receptors, break cellular communication, provide sialic acid monomers for nutrition, or as a mimicry and immune evasion strategy by coating its surface with host components. Furthermore, investigations of the activation of complement by *T. pallidum* showed that this spirochete is able to bind to sialic acid (Fitzgerald, 1987). The study compared the non-pathogen *T. vincentii* with *T. pallidum* and showed that *T. pallidum* bound to four times more sialic acid than *T. vincentii*. Upon addition of an exogenous sialidase, sialic acid was enzymatically removed from the surface of *T. pallidum*, resulting in a significant increase in complement activation. Thus it was concluded that *T. pallidum* may use sialic acid to coat its surface and delay complement-mediated damage, further contributing to the poor antigenicity and limited reactivity that the characteristic paucity of OMPs provide to *T. pallidum*. Importantly, syphilis is linked to premature births and adverse pregnancy outcomes (De Santis *et al.*, 2012; World Health Organization, 2015), which might also be related to the presence of sialidase, like in the case of *Gardnerella vaginalis*.

The regulatory domain found in Tp0854, a putative serine/threonine phosphatase, may be involved in host cellular regulation. Some pathogenic bacteria have been shown to produce and secrete regulatory proteins, such as phosphatases that are able to interact with the host and regulate immune responses. For instance, *Porphyromonas gingivalis*, a major pathogen of periodontal disease, secretes a phosphatase, SerB, that interacts with the transcriptional regulator NF-κB resulting in suppression of the production of the pro-inflammatory chemokine, IL-8, by epithelial cells. Thus the innate immune response is dysregulated and recruitment of neutrophils is impaired, favoring bacterial survival (Takeuchi *et al.*, 2013). Moreover, this phosphatase can also facilitate invasion of

P. gingivalis by influencing host actin remodelling (Moffatt *et al.*, 2012; Takeuchi *et al.*, 2013). *Mycobacterium tuberculosis* produces and secretes a number of phosphatases: PtpA, PtpB and SapM. PtpA inhibits V-ATPase trafficking to the mycobacterial phagosome and blocks phagosome-lysosome fusion (Wong *et al.*, 2011). PtpB might play a role in downregulating the host immune response (Zhou *et al.*, 2010), and SapM inhibits the production of phosphatidylinositol3-phosphate, thus inhibiting host signaling pathways as well as arresting phagosome maturation favoring the survival of *M. tuberculosis* (Vergne *et al.*, 2005; Wong *et al.*, 2013). It is easier to imagine how the regulatory proteins from the pathogens discussed above would have an impact in the host regulatory pathways, since these are intracellular pathogens, however, in the case of *T. pallidum*, there is no evidence of any cellular invasion.

Bordetella pertussis however, possesses a regulatory protein that although not a phosphatase, is used to regulate host processes and may present a more likely scenario applicable to the extracellular *T. pallidum*. *B. pertussis* is the causative agent of whooping cough (Goodwin and Weiss, 1990). One of the key virulence factors secreted by *Bordetella* during early respiratory tract colonization, is an adenylate cyclase toxin-hemolysin (CyaA, AC-Hly, or ACT). CyaA is a T1SS substrate that is able to form pores and insert its adenylate cyclase domain into host cells, resulting in cell lysis, as well as signal transduction by the unregulated conversion of ATP to cAMP. This toxin is synthesized as a single polypeptide of 1706 residues, that contains RTX repeats located within the middle of the protein sequence, and an adenylyl cyclase domain toward the C-terminus (Glaser *et al.*, 1988). CyaA toxin enables *Bordetella* to escape the host immune system by promoting macrophage apoptosis (Hanski, 1989). Specific studies interested in the interaction of

CyaA with the host membrane have shown that, as depicted in Figure S 9, CyaA may exist in solution as two isomeric conformations that upon insertion into the host cell membrane become either a precursor for subsequent translocation of the adenylyl cyclase domain or a pore precursor that is prone to CyaA oligomerization within the host cell membrane resulting in larger pore formation (Basler *et al.*, 2007). CyaA has been shown to target mostly cells that possess $\alpha_M\beta 2$ integrin receptors, however its activity can be seen in other cell types lacking such a receptor (Guermonprez *et al.*, 2001). Unfortunately, there is no structural data available for CyaA and thus its membrane-insertion mechanism along with membrane pore formation and its ability to transfer its adenylyl cyclase domain across the host cytoplasmic membrane have not been fully elucidated. There is however, evidence suggesting that the translocation of the adenylyl cyclase domain and pore-formation requires the involvement of two predicted transmembrane α-helices containing a pair of glutamate residues each, within the pore-forming domain of CyaA (Basler *et al.*, 2007; Osičková *et al.*, 1999). It is thought that the negative charges of the glutamate residues provide a means to bring into the host cell the overall slightly positively charged adenylyl cyclase domain (Basler *et al.*, 2007). It is unclear at this moment if Tp0854 from *T. pallidum* could use a similar mechanism to insert into the host cell and translocate its phosphatase domain into the cell. The predicted Tp0854 transmembrane helices however, are not predicted to contain any negatively charged residues, but rather positively charged residues. The helix at position 753-775 may include two arginine residues and the helix at position 816-838 may contain an arginine and lysine residue, thus, similar electrostatic interactions allowing for the translocation of the overall negatively phosphatase domain into the host cell may occur.

The secretion of Tp0854 by *T. pallidum* could result in a double advantage to this pathogen. The putative sialidase domain of Tp0854 could cleave off sialic acid from host surfaces, to be used by *T. pallidum* as an immune evasion strategy, or to expose receptors required for interactions between the host and *T. pallidum*. Additionally, the predicted Tp0854 phosphatase domain could be translocated inside the host cell leading to a disruption of host cellular processes, all whilst forming a pore that could eventually lead to host cell lysis and tissue damage.

6.9 Conclusions

In silico analysis of a novel *T. phagedenis* protein, TphBIg, allowed us to indirectly identify a novel *T. pallidum* protein that might be a virulence factor secreted via a putative T1SS. Thus, bioinformatic investigations suggested weak sequence and structural similarities between TphBIg and *T. pallidum* Tp0854. Further *in silico* analysis indicated that *T. pallidum* Tp0854 may possess a sialidase domain, a BIg domain, a regulatory HAMP domain and a phosphatase domain. Within the phosphatase domain and closer to the C-terminus, Tp0854 appears to also possess a putative RTX motif, which is characteristic of a certain type of T1SS substrates, the RTX proteins. The phosphatase domain of Tp0854, containing the RTX motif and the full C-terminus, was successfully cloned and expressed recombinantly in *E. coli*.

In vitro experimental approaches tailored to detect an interaction between recombinant Tp0965, the MFP of the potential T1SS, and the recombinant *T. pallidum* phosphatase domain of Tp0854, were able to provide strong supporting evidence that agreed with the *in silico* analyses carried out. Plate-based binding assays showed that the binding between Tp0854 and Tp0965 is significantly higher compared to the binding of

Tp0327 and Tp0965. It was not possible to carry out pull-down assays as a confirmatory method since the recombinant Tp0854 and Tp0965 proteins produced possess a similar molecular weight, making it difficult to visually assess the presence or absence of their respective protein bands.

Overall, the bioinformatic analysis, supported by the confirmatory binding assays carried out, suggested that Tp0965 and Tp0854 might interact and thus Tp0854 may be a T1SS substrate. Thus, these investigations supported the hypothesis that Tp0854 may interact with the putative Tp0965, and thus might get secreted via the putative T1SS predicted via bioinformatic analysis in *T. pallidum*.

The significance of secreting a multifunctional protein such as Tp0854, is that having a minimalistic genome and hence a minimalistic proteome, *T. pallidum* would be able to efficiently perform more than one function with a single protein. If the predicted functions of Tp0854 are true, this protein may allow for both immune-evasion by coating the surface of *T. pallidum* with sialic acid, and host damage both via host cell signaling disruption, as well as host cell lysis via pore formation.

Chapter 7: Significance and future directions

7.1 Significance

Syphilis is the most significant disease in fetuses worldwide, it has shown a strong link to HIV infection and the number of affected patients is increasing despite the availability of effective antibiotic treatment (World Health Organization, 2015). It is strongly believed that a successful preventative measure is required to finally eradicate this old disease (Cameron and Lukehart, 2014; LaFond and Lukehart, 2006). However, studying the causative agent of syphilis, *T. pallidum*, is extremely challenging. *T. pallidum* is one of the most invasive pathogens, certainly the most invasive spirochete; it is an obligate human pathogen and possesses a minimalistic genome with a delicate outer membrane, which makes it impossible to culture it *in vitro* and hence, impossible to manipulate its DNA. Investigations carried out directly with *T. pallidum* require the intratesticular inoculation of rabbits, which is challenging both ethically and economically. Furthermore, upon harvesting the live bacteria, these are only viable and adequate for experimentation for a few hours (LaFond and Lukehart, 2006). Thus, the syphilis field is hindered by the inability to carry out modern techniques to study this simplistic pathogen and therefore the molecular mechanisms whereby *T. pallidum* is able to undergo rapid and widespread dissemination remain to be fully elucidated.

The significance of the studies presented in this thesis is that despite the inability to successfully culture *T. pallidum in vitro*, the availability of the model treponeme, *T. phagedenis*, makes it possible to carry out preliminary investigations of protein-protein interactions. This non-pathogenic, culturable spirochete shares genomic features with *T. pallidum* and the availability of a transformed strain that is able to appropriately localize

one of the *T. pallidum* virulence factors, pallilysin, suggests shared features in their export machineries. Thus, protein-protein interactions can be studied and optimized in an indirect way by using a *T. phagedenis* lysate, avoiding unnecessary costs and ethical obstacles.

Moreover, modern computational technologies allow scientists to carry out experiments *in silico* to avoid unnecessary expenses and waste of materials, providing an important tool to predict the potential outcome of an *in vitro / in vivo* experiment and plan accordingly. Hence, in this thesis we showed that the treponeme model *T. phagedenis* as well as bioinformatic analyses, can be successfully used to study protein-protein interactions. Thus, an experimental approach using first a *T. phagedenis* lysate and *T. pallidum* recombinant proteins to study protein-protein interactions was developed. Upon identification of an interacting protein, we showed that bioinformatic analyses can be used to scan the *T. pallidum* proteome and find proteins with sequence similarity.

Hence, we propose that with this technique it is possible to work around the inability to culture *T. pallidum*, by using the model treponeme *T. phagedenis* as a first step in studying protein-protein interactions. This technique can be successfully applied to general protein-protein interaction studies, and specifically in this case, the interaction between the putative MFP from the predicted T1SS and a *T. phagedenis* lysate was used to identify a potential novel T1SS substrate both in *T. phagedenis* and in *T. pallidum*.

Additionally, the observed interaction between a component of the putative T1SS and a novel *T. pallidum* protein, Tp0854, not only indicated the presence of a potentially functional T1SS, but also the identification of a putative virulence factor that may play a key role in *T. pallidum* immune evasion, invasion and overall pathogenesis.

Overall, this constitutes the first investigation of a putative T1SS and potential substrates in the syphilis spirochete, *T. pallidum*. The interaction between Tp0965 with both pallilysin and the recombinantly expressed phosphatase domain of Tp0854 may indicate export via this secretory machinery.

Hence, the presence of a T1SS in *T. pallidum* would suggest that this system may be crucial in exporting key *T. pallidum* virulence factors that confer upon this pathogen its high invasiveness and dissemination properties. Even though much remains to be studied, the insight gained from these investigations have led to a whole new path in the syphilis investigations that will eventually lead to a better understanding of the mechanisms facilitating *T. pallidum* host invasion and may reveal new potential vaccine targets that may successfully prevent bacterial dissemination and chronic infection.

7.2 Future directions

The novel bioinformatic characterization and identification of Tp0854 as a potential T1SS substrate represents the first step toward identifying more virulence factors that may be of key importance for *T. pallidum* invasion and dissemination. Thus, it would be of interest to use crosslinking techniques and find out specific binding sites between Tp0854 and Tp0965 which would provide supporting evidence for the interaction previously observed via plate-based binding assays.

Furthermore, in order to provide strong evidence for the secretion of Tp0854 and pallilysin via a potential T1SS system, it would be necessary to heterologously express Tp0854 or pallilysin in an *E. coli* strain harboring a T1SS and test for secretion. Moreover, the reconstruction of the *T. pallidum* putative T1SS in liposomes would allow us to confirm such a secretion of Tp0854 as well as pallilysin.

Regarding the virulent action of Tp0854 in the context of *T. pallidum* pathogenesis, it would be necessary to carry out a full functional and structural characterization of Tp0854, including applying sialidase and phosphatase assays to assess enzymatic activity. If indeed Tp0854 has functional sialidase and phosphatase domains, it would then be of interest to determine the substrate specificity of the sialidase domain, and to uncover which host proteins may be dephosphorylation targets of the phosphatase domain. Furthermore, the interaction of Tp0854 with host cells would need to be studied, as well as the detailed mechanism of potential insertion and pore formation.

Finally, the technique developed in this thesis could be optimized to carry out unbiased pull-down assays tailored to detect more protein-protein interactions within *T. pallidum.* Direct *T. pallidum* lysate with recombinant *T. pallidum* proteins could allow for the identification of a wider range of virulence factors that might be secreted via this system.

Appendix

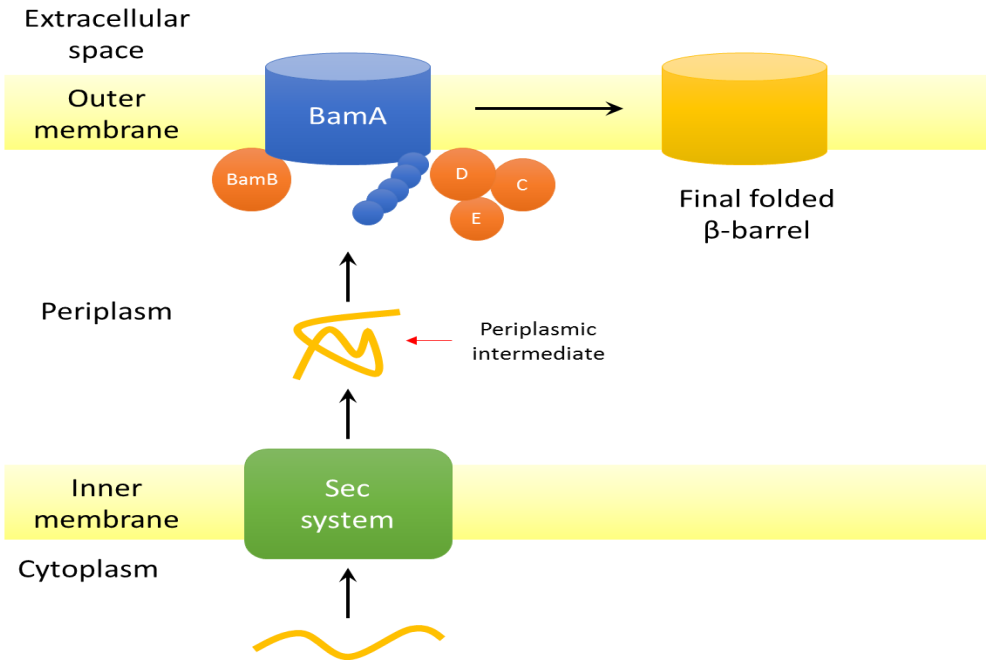


Figure S 1. β-barrel assembly machinery (BAM complex) from *E. coli*; mechanism to process outer membrane β-barrel proteins

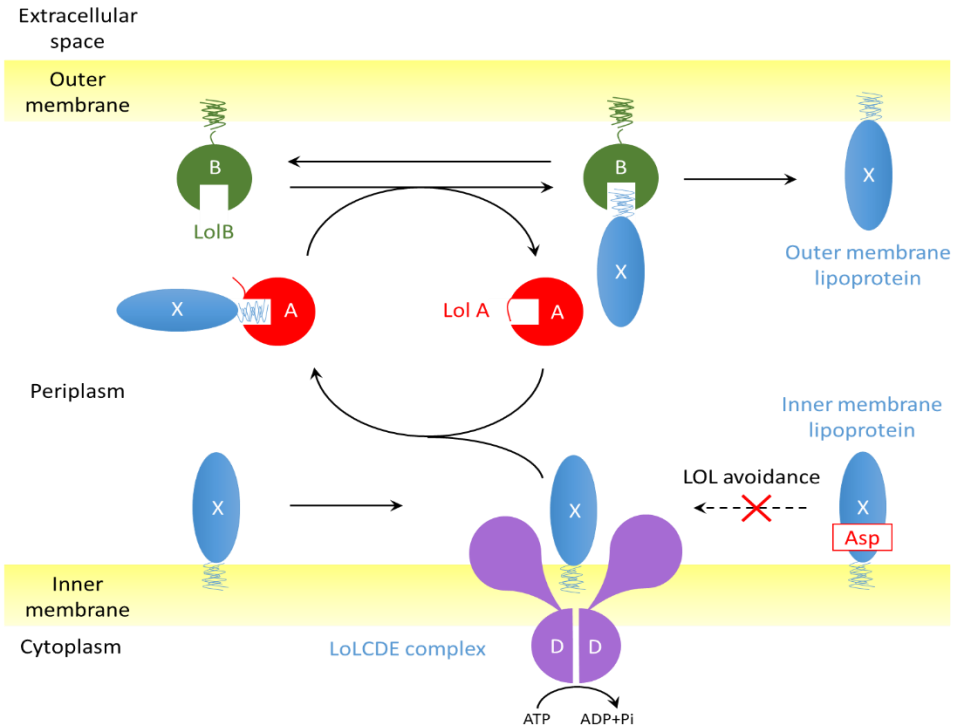


Figure S 2. Localization of Lipoprotein (LOL) system from *E. coli*; mechanism to process bacterial lipoproteins

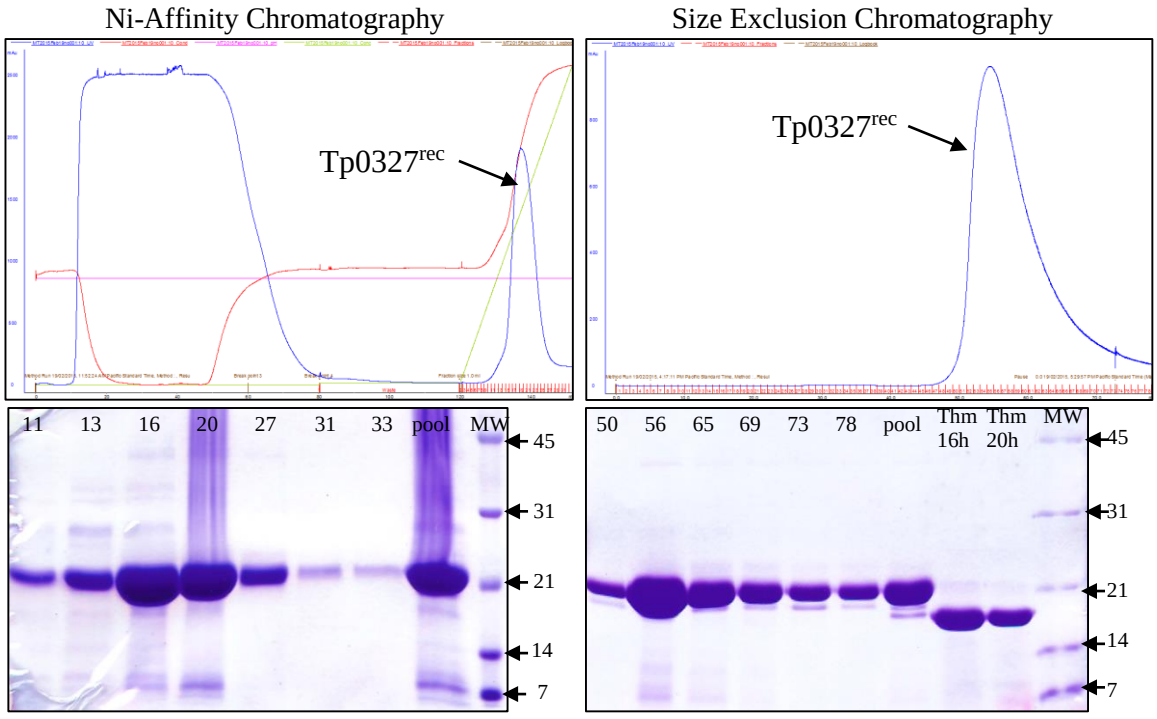


Figure S 3. FPLC Ni-affinity and SEC chromatograms, with SDS-PAGE gels with fractions of recombinant *T. pallidum* Tp0327^{rec} purification steps.

Thm: Thrombin cleavage for His-tag removal

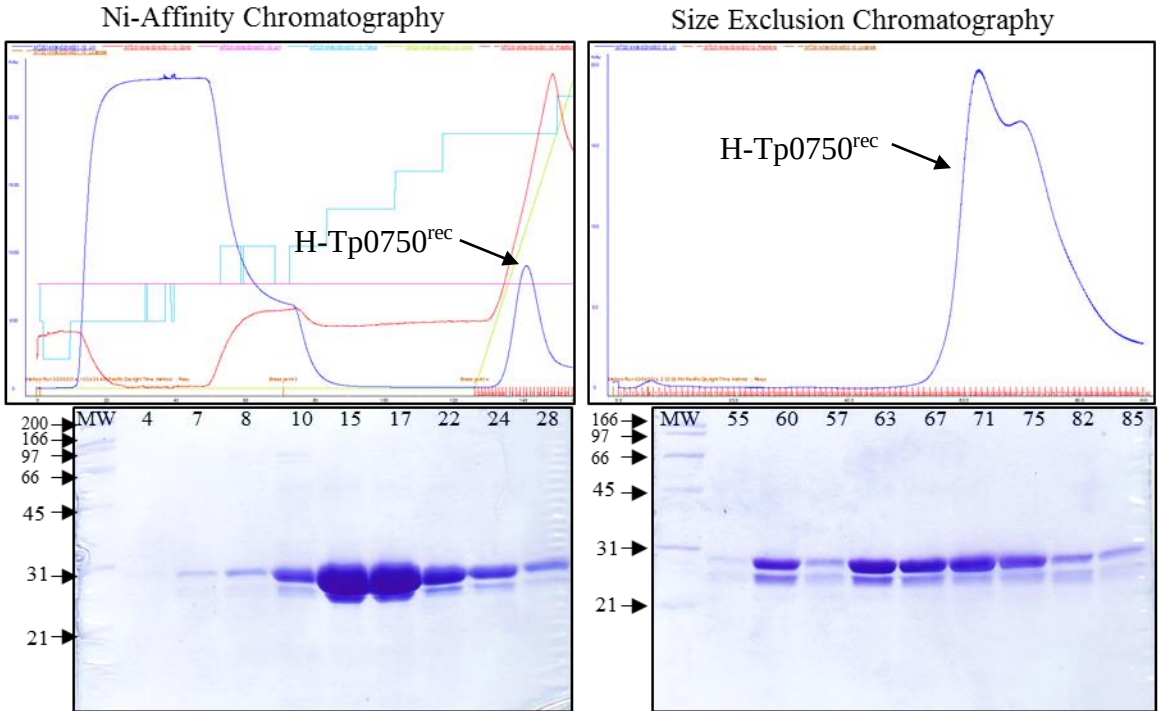


Figure S 4. FPLC Ni-affinity and SEC chromatograms with SDS-PAGE gels with fractions of recombinant *T. pallidum* H- Tp0750^{rec} purification steps.

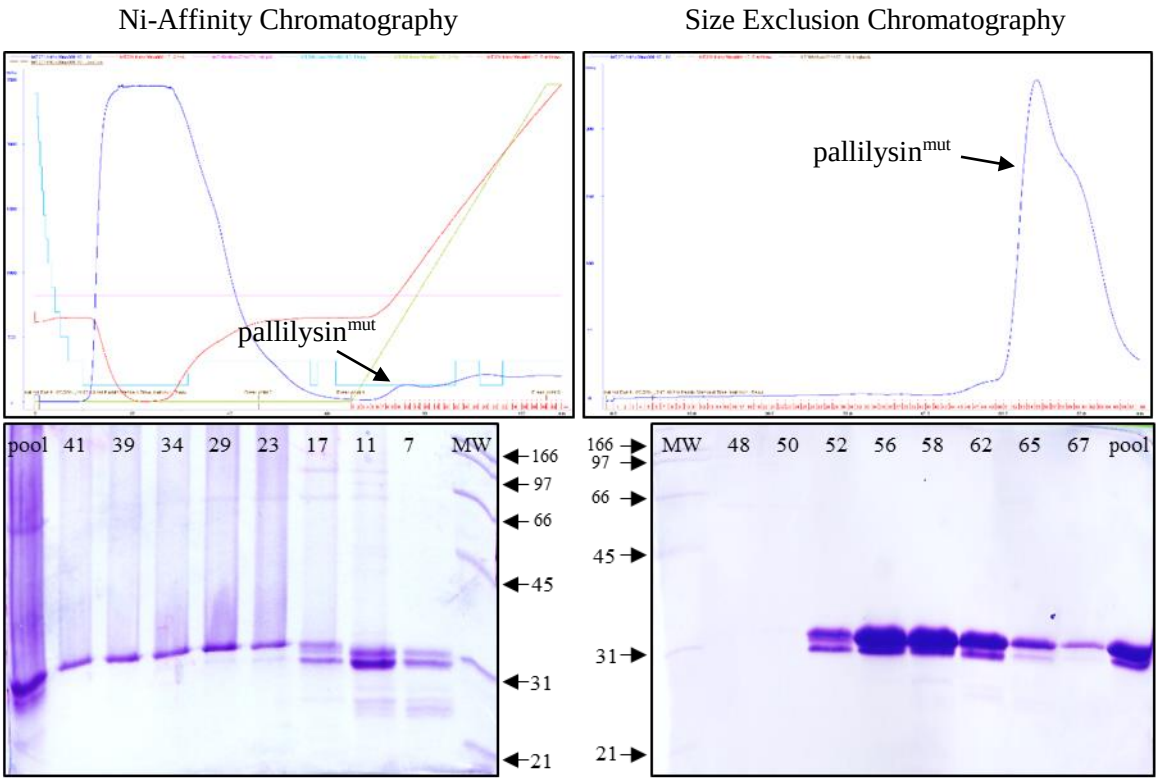


Figure S 5. FPLC Ni-affinity and SEC chromatograms with SDS-PAGE gels with fractions of recombinant *T. pallidum* H-pallilysin^{mut} purification steps.

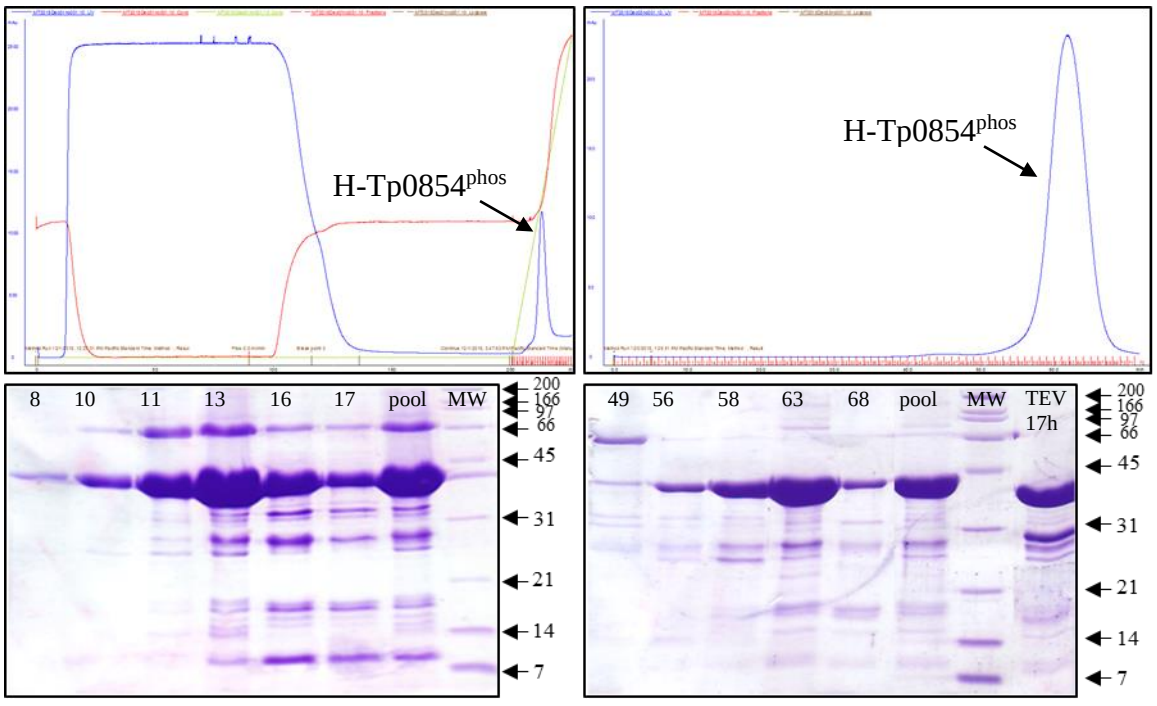


Figure S 6. FPLC Ni-affinity and SEC chromatograms with SDS-PAGE gels with fractions of recombinant *T. pallidum* Tp0854^{rec} purification steps

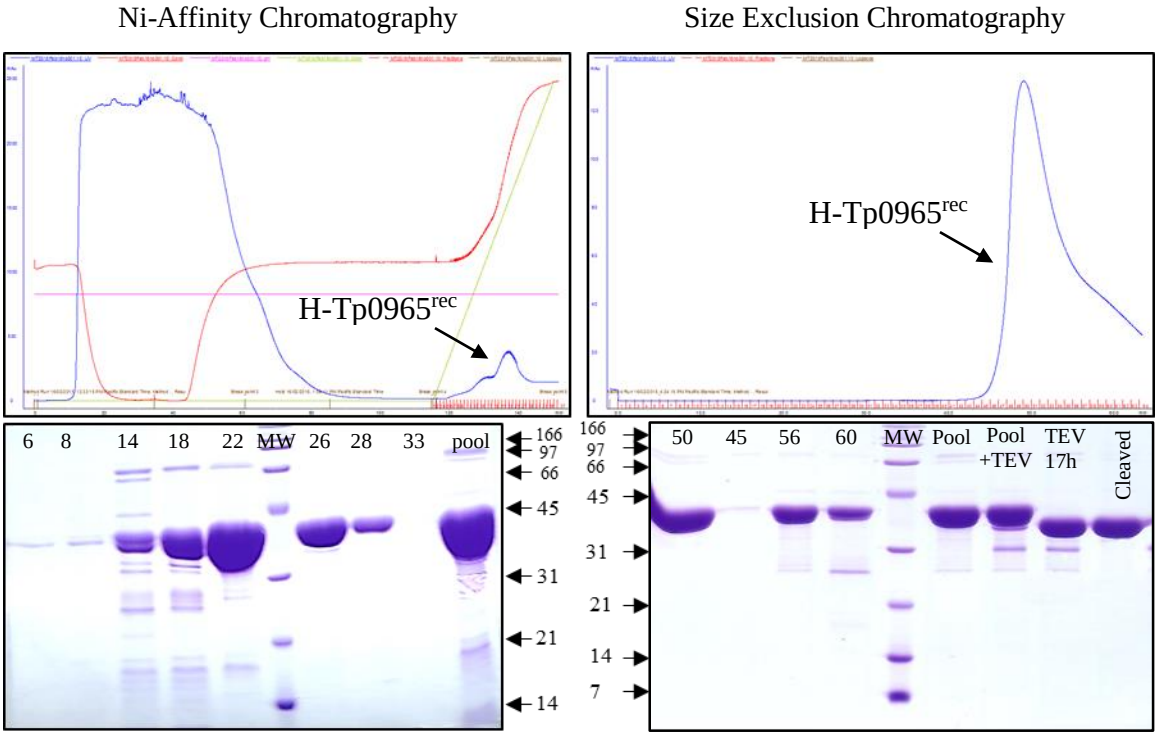


Figure S 7. FPLC Ni-affinity and SEC chromatograms with SDS-PAGE gels with fractions of recombinant *T. pallidum* H-Tp0965^{rec} purification steps.

Table S 1. Protein coverage (in red) obtained from MS analysis of cross-linked recombinant pallilysin and Tp0965

Band	Recombinant Pallilysin MS coverage	Recombinant Tp0965 MS coverage
1	MSYYHHHHHHLESTLYKKAGCCFQHGHVPPRRIPPHD TFGALPTAALPSNARDTAAHPSDTADNTSGSSTTTDPRS HGNAPPAPVGGAAQTHTQPPVQTAMRIALWNRATHGE QGALQHLLAGLWIQTEISPNSGDIHPLLFFDREHAEITFS RASVQEIFLVDSAHTHRKTVSFLTRNTAISSIRRRLEVTF ESHAVIHVRAVEDVARLKIGSTSMWDGQYTRYHAGPA SAPSP	GSMASPTLSVSKEVVLNRIEISGYIEAAQHQKLESPGEGIVRT VRVQEGDTVKKGQLLFSLENSHQQLDLAEHEFAIEQEEINGV SKKMEIMKLKRNMLQKRLRERYVTAQFDGVVAAFKLSPGQ YAKPQDYFGTLIDRSYFKANVEIPEVDASRLKVGQRVEISFPA EPSVKAVGSVTSYPSIARVTSVGRTVVDASIRIDELPEILPGYS FSGAIVAGEQEEILVLKQDGLRYEKGAPFVDRVLPSGKIKSVP VTVEPYVPGFVKIISGLGAGDRVKDQSAAAAA
2	MSYYHHHHHHLESTLYKKAGCCFQHGHVPPRRIPPHD TFGALPTAALPSNARDTAAHPSDTADNTSGSSTTTDPRS HGNAPPAPVGGAAQTHTQPPVQTAMRIALWNRATHGE QGALQHLLAGLWIQTEISPNSGDIHPLLFFDREHAEITFS RASVQEIFLVDSAHTHRKTVSFLTRNTAISSIRRRLEVTF ESHAVIHVRAVEDVARLKIGSTSMWDGQYTRYHAGPA SAPSP	GSMASPTLSVSKEVVLNRIEISGYIEAAQHQKLESPGEGIVRT VRVQEGDTVKKGQLLFSLENSHQQLDLAEHEFAIEQEEINGV SKKMEIMKLKRNMLQKRLRERYVTAQFDGVVAAFKLSPGQ YAKPQDYFGTLIDRSYFKANVEIPEVDASRLKVGQRVEISFPA EPSVKAVGSVTSYPSIARVTSVGRTVVDASIRIDELPEILPGYS FSGAIVAGEQEEILVLKQDGLRYEKGAPFVDRVLPSGKIKSVP VTVEPYVPGFVKIISGLGAGDRVKDQSAAAAA
3	MSYYHHHHHHLESTLYKKAGCCFQHGHVPPRRIPPHD TFGALPTAALPSNARDTAAHPSDTADNTSGSSTTTDPRS HGNAPPAPVGGAAQTHTQPPVQTAMRIALWNRATHGE QGALQHLLAGLWIQTEISPNSGDIHPLLFFDREHAEITFS RASVQEIFLVDSAHTHRKTVSFLTRNTAISSIRRRLEVTF ESHAVIHVRAVEDVARLKIGSTSMWDGQYTRYHAGPA SAPSP	GSMASPTLSVSKEVVLNRIEISGYIEAAQHQKLESPGEGIVRT VRVQEGDTVKKGQLLFSLENSHQQLDLAEHEFAIEQEEINGV SKKMEIMKLKRNMLQKRLRERYVTAQFDGVVAAFKLSPGQ YAKPQDYFGTLIDRSYFKANVEIPEVDASRLKVGQRVEISFPA EPSVKAVGSVTSYPSIARVTSVGRTVVDASIRIDELPEILPGYS FSGAIVAGEQEEILVLKQDGLRYEKGAPFVDRVLPSGKIKSVP VTVEPYVPGFVKIISGLGAGDRVKDQSAAAAA
4	MSYYHHHHHHLESTLYKKAGCCFQHGHVPPRRIPPHD TFGALPTAALPSNARDTAAHPSDTADNTSGSSTTTDPRS HGNAPPAPVGGAAQTHTQPPVQTAMRIALWNRATHGE QGALQHLLAGLWIQTEISPNSGDIHPLLFFDREHAEITFS RASVQEIFLVDSAHTHRKTVSFLTRNTAISSIRRRLEVTF ESHAVIHVRAVEDVARLKIGSTSMWDGQYTRYHAGPA SAPSP	GSMASPTLSVSKEVVLNRIEISGYIEAAQHQKLESPGEGIVRT VRVQEGDTVKKGQLLFSLENSHQQLDLAEHEFAIEQEEINGV SKKMEIMKLKRNMLQKRLRERYVTAQFDGVVAAFKLSPGQ YAKPQDYFGTLIDRSYFKANVEIPEVDASRLKVGQRVEISFPA EPSVKAVGSVTSYPSIARVTSVGRTVVDASIRIDELPEILPGYS FSGAIVAGEQEEILVLKQDGLRYEKGAPFVDRVLPSGKIKSVP VTVEPYVPGFVKIISGLGAGDRVKDQSAAAAA

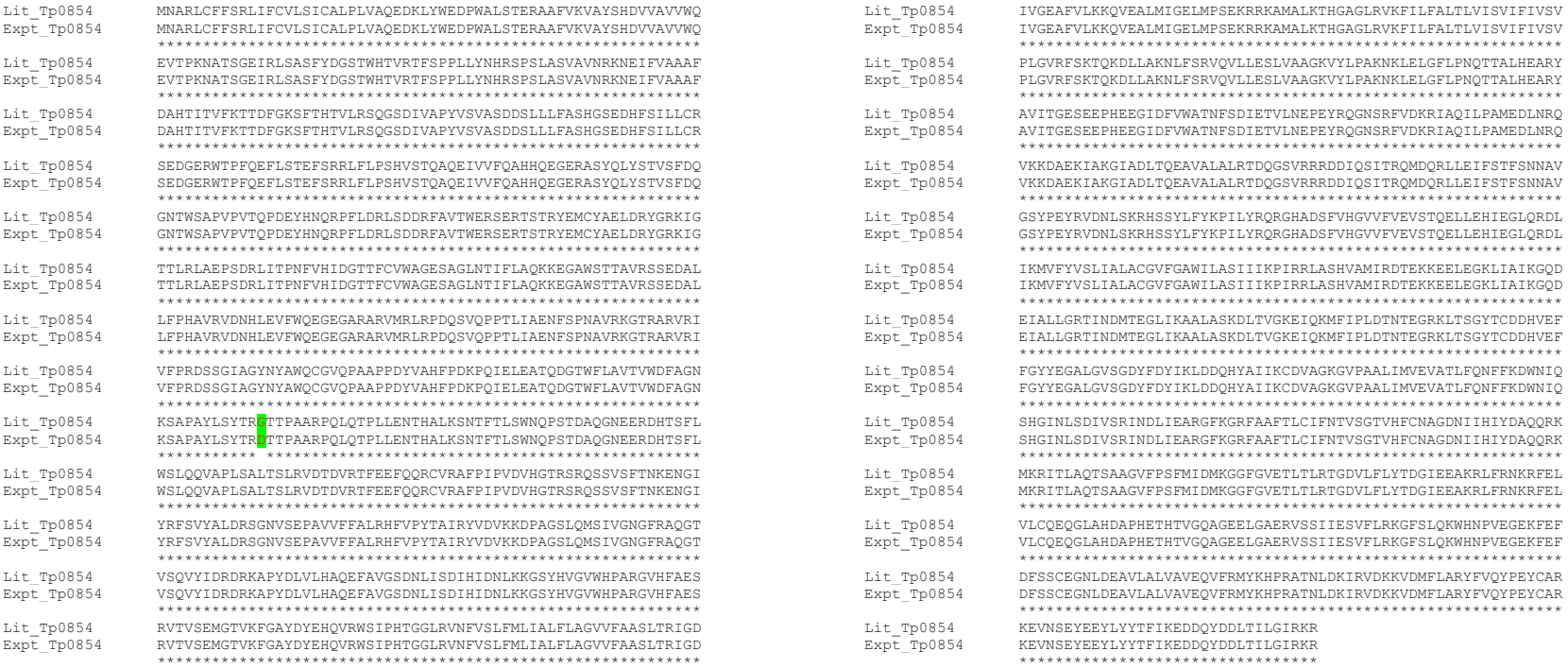


Figure S 8. Alignment of the Tp0854 amino acid literature sequence with the experimental sequence.
The glycine residue at position 492 has been corrected to an aspartic acid residue; correction is shown highlighted in green.

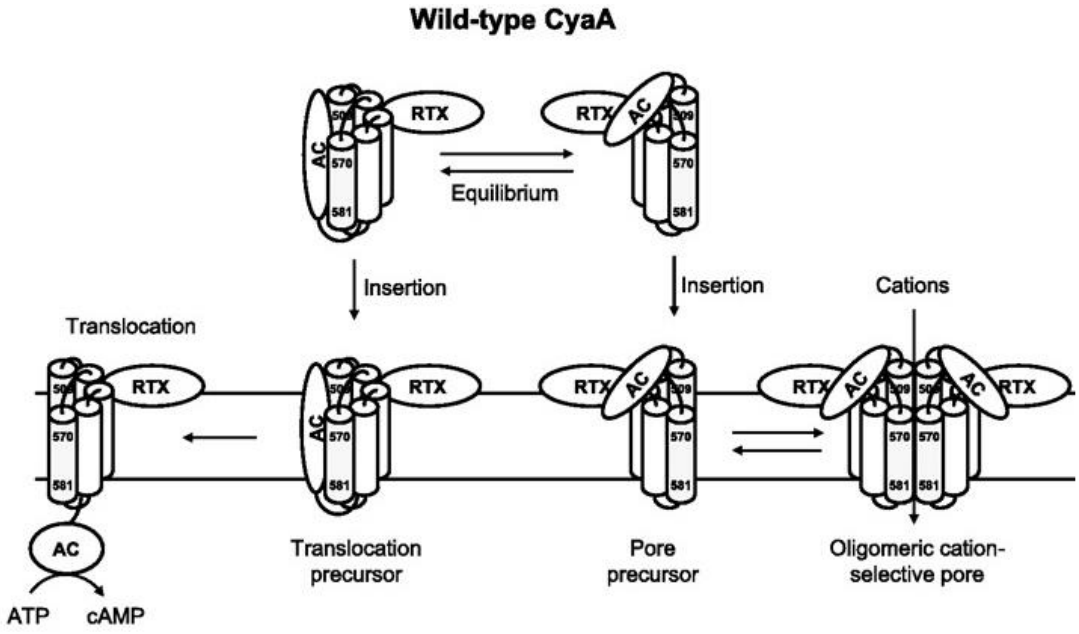


Figure S 9. Model for the mechanism of action of CyaA from *Bordetella pertussis.*

CyaA toxin is thought to exist in equilibrium between two conformations, one that favors an adenylyl cyclase translocation precursor and another one that favors oligomerization within the host cellular membrane and thus pore formation. Figure adapted from (Basler *et al.*, 2007).

Bibliography

Akatsuka, H., Binet, R., Kawai, E., Wandersman, C., & Omori, K. (1997). Lipase secretion by bacterial hybrid ATP-binding cassette exporters: molecular recognition of the LipBCD, PrtDEF, and HasDEF exporters. *J Bacteriol, 179*(15), 4754-4760.

Albrecht, R., & Zeth, K. (2011). Structural basis of outer membrane protein biogenesis in bacteria. *Journal of Biological Chemistry, 286*(31), 27792-27803.

Alder, J., Friess, L., Tengowski, M., & Schell, R. (1990). Phagocytosis of opsonized *Treponema pallidum* subsp. *pallidum* proceeds slowly. *Infection and immunity, 58*(5), 1167-1173.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol, 215*(3), 403-410.

Ammar, Z., Plazolles, N., Baltz, T., & Coustou, V. (2013). Identification of Trans-Sialidases as a Common Mediator of Endothelial Cell Activation by African Trypanosomes. *PLoS Pathog, 9*(10), e1003710.

Andersen, C., Koronakis, E., Bokma, E., Eswaran, J., Humphreys, D., Hughes, C., *et al.* (2002). Transition to the open state of the TolC periplasmic tunnel entrance. *Proc Natl Acad Sci U S A, 99*(17), 11103-11108.

Angelos, J. A., Hess, J. F., & George, L. W. (2003). An RTX operon in hemolytic *Moraxella bovis* is absent from nonhemolytic strains. *Vet Microbiol, 92*(4), 363-377.

Arnoux, P., Haser, R., Izadi, N., Lecroisey, A., Delepierre, M., Wandersman, C., *et al.* (1999). The crystal structure of HasA, a hemophore secreted by *Serratia marcescens*. [10.1038/9281]. *Nat Struct Mol Biol, 6*(6), 516-520.

Bagos, P. G., Liakopoulos, T. D., Spyropoulos, I. C., & Hamodrakas, S. J. (2004). PRED-TMBB: a web server for predicting the topology of β-barrel outer membrane proteins. *Nucleic Acids Research, 32*(Web Server issue), W400-W404.

Bailey, M. J. A., Koronakis, V., Schmoll, T., & Hughes, C. (1992). *Escherichia coli* HlyT protein, a transcriptional activator of hemolysin synthesis and secretion, is encoded by the RFAH (SFRB) locus required for expression of sex factor and lipopolysaccharide genes. *Molecular Microbiology, 6*(8), 1003-1012.

Bakás Laura, M. S., Vazquez Romina and Herlax Vanesa (2012). *E. coli* Alpha Hemolysin and Properties. *Biochemistry, Prof. Deniz Ekinci (Ed.),* .

Bakkes, P. J., Jenewein, S., Smits, S. H., Holland, I. B., & Schmitt, L. (2010). The rate of folding dictates substrate secretion by the *Escherichia coli* hemolysin type 1 secretion system. *J Biol Chem, 285*(52), 40573-40580.

Balakrishnan, L., Hughes, C., & Koronakis, V. (2001). Substrate-triggered recruitment of the TolC channel-tunnel during type I export of hemolysin by *Escherichia coli*. *Journal of Molecular Biology, 313*(3), 501-510.

Baldi, D. L., Higginson, E. E., Hocking, D. M., Praszkier, J., Cavaliere, R., James, C. E., *et al.* (2012). The Type II Secretion System and Its Ubiquitous Lipoprotein Substrate, SslE, Are Required for Biofilm Formation and Virulence of Enteropathogenic *Escherichia coli*. *Infection and Immunity, 80*(6), 2042-2052.

Balsalobre, C., Silvan, J. M., Berglund, S., Mizunoe, Y., Uhlin, B. E., & Wai, S. N. (2006). Release of the type I secreted alpha-haemolysin via outer membrane vesicles from *Escherichia coli*. *Molecular Microbiology, 59*(1), 99-112.

Barlag, B., & Hensel, M. (2015). The Giant Adhesin SiiE of *Salmonella enterica*. *Molecules, 20*(1), 1134.

Basar, T., Havlíček, V. r., Bezoušková, S., Hackett, M., & Šebo, P. (2001). Acylation of Lysine 983 Is Sufficient for Toxin Activity of *Bordetella pertussis* Adenylate Cyclase: Substitutions of alanine 140 modulate acylation site selectivity of the toxin acyltransferase CyaC. *Journal of Biological Chemistry, 276*(1), 348-354.

Basler, M., Knapp, O., Masin, J., Fiser, R., Maier, E., Benz, R., *et al.* (2007). Segments Crucial for Membrane Translocation and Pore-forming Activity of *Bordetella* Adenylate Cyclase Toxin. *Journal of Biological Chemistry, 282*(17), 12419-12429.

Baughn, R. E. (1987). Role of fibronectin in the pathogenesis of syphilis. *Rev Infect Dis, 9 Suppl 4*, S372-385.

Baughn, R. E., & Musher, D. M. (2005). Secondary Syphilitic Lesions. *Clinical Microbiology Reviews, 18*(1), 205-216.

Baumann, U., Wu, S., Flaherty, K. M., & McKay, D. B. (1993). Three-dimensional structure of the alkaline protease of *Pseudomonas aeruginosa*: a two-domain protein with a calcium binding parallel beta roll motif. *The EMBO Journal, 12*(9), 3357-3364.

Beeckman, D. S. A., & Vanrompay, D. C. G. (2010). Bacterial Secretion Systems with an Emphasis on the Chlamydial Type III Secretion System. *Current Issues in Molecular Biology, 12*, 17-41.

Belkum, M. J., Worobo, R. W., & Stiles, M. E. (1997). Double-glycine-type leader peptides direct secretion of bacteriocins by ABC transporters: Colicin V secretion in *Lactococcus lactis*. *Molecular microbiology, 23*(6), 1293-1301.

Bernadac, A., Gavioli, M., Lazzaroni, J.-C., Raina, S., & Lloubès, R. (1998). *Escherichia coli* tol-pal mutants form outer membrane vesicles. *Journal of bacteriology, 180*(18), 4872-4878.

Bishop, N. H., & Miller, J. N. (1976). Humoral Immunity in Experimental Syphilis II. The Relationship of Neutralizing Factors in Immune Serum to Acquired Resistance. *The Journal of Immunology, 117*(1), 197-207.

Blanco, D. R., Miller, J. N., & Lovett, M. A. (1997). Surface antigens of the syphilis spirochete and their potential as virulence determinants. *Emerging Infectious Diseases, 3*(1), 11-20.

Boardman, B. K., & Fullner Satchell, K. J. (2004). *Vibrio cholerae* Strains with Mutations in an Atypical Type I Secretion System Accumulate RTX Toxin Intracellularly. *Journal of Bacteriology, 186*(23), 8137-8143.

Boyd, C. D., Smith, T. J., El-Kirat-Chatel, S., Newell, P. D., Dufrêne, Y. F., & O'Toole, G. A. (2014). Structural Features of the *Pseudomonas fluorescens* Biofilm Adhesin LapA Required for LapG-Dependent Cleavage, Biofilm Formation, and Cell Surface Localization. *Journal of Bacteriology, 196*(15), 2775-2788.

Brautigam, C. A., Deka, R. K., Ouyang, Z., Machius, M., Knutsen, G., Tomchick, D. R., *et al.* (2012). Biophysical and Bioinformatic Analyses Implicate the *Treponema pallidum* Tp34 Lipoprotein (Tp0971) in Transition Metal Homeostasis. *Journal of Bacteriology, 194*(24), 6771-6781.

Brown, D. L., & Frank, J. E. (2003). Diagnosis and management of syphilis. [Article]. *American Family Physician, 68*(2), 283-290.

Buchacz, K., Patel, P., Taylor, M., Kerndt, P. R., Byers, R. H., Holmberg, S. D., *et al.* (2004). Syphilis increases HIV viral load and decreases CD4 cell counts in HIV-infected patients with new syphilis infections. *AIDS, 18*(15), 2075-2079.

Cameron, C. E. (2003). Identification of a *Treponema pallidum* laminin-binding protein. *Infection and Immunity, 71*(5), 2525-2533.

Cameron, C. E., Brouwer, N. L., Tisch, L. M., & Kuroiwa, J. M. Y. (2005). Defining the interaction of the *Treponema pallidum* adhesin Tp0751 with laminin. *Infection and Immunity, 73*(11), 7485-7494.

Cameron, C. E., Kuroiwa, J. M., Yamada, M., Francescutti, T., Chi, B., & Kuramitsu, H. K. (2008). Heterologous expression of the *Treponema pallidum* laminin-binding adhesin Tp0751 in the culturable spirochete *Treponema phagedenis*. *J Bacteriol, 190*(7), 2565-2571.

Cameron, C. E., & Lukehart, S. A. (2014). Current status of syphilis vaccine development: Need, challenges, prospects. *Vaccine, 32*(14), 1602-1609.

Carlson, J. A., Dabiri, G., Cribier, B., & Sell, S. (2011). The immunopathobiology of syphilis: the manifestations and course of syphilis are determined by the level of delayed-type hypersensitivity. *The American Journal of dermatopathology, 33*(5), 433-460.

Cascales, E., & Christie, P. J. (2004). Definition of a bacterial type IV secretion pathway for a DNA substrate. *Science, 304*(5674), 1170-1173.

Cauci, S., Thorsen, P., Schendel, D. E., Bremmelgaard, A., Quadrifoglio, F., & Guaschino, S. (2003). Determination of Immunoglobulin A against *Gardnerella vaginalis* Hemolysin, Sialidase, and Prolidase Activities in Vaginal Fluid: Implications for Adverse Pregnancy Outcomes. *Journal of Clinical Microbiology, 41*(1), 435-438.

Cescau, S., Debarbieux, L., & Wandersman, C. (2007). Probing the *in vivo* dynamics of type I protein secretion complex association through sensitivity to detergents. *Journal of bacteriology, 189*(5), 1496-1504.

Chang, Y. F., Young, R., Moulds, T. L., & Struck, D. K. (1989). Secretion of the *Pasteurella* leukotoxin by *Escherichia coli*. *FEMS Microbiol Lett, 51*(1), 169-173.

Chen, S., & Zückert, W. R. (2011). Probing the *Borrelia burgdorferi* Surface Lipoprotein Secretion Pathway Using a Conditionally Folding Protein Domain. *Journal of Bacteriology, 193*(23), 6724-6732.

Chenal, A., Guijarro, J. I., Raynal, B., Delepierre, M., & Ladant, D. (2009). RTX calcium binding motifs are intrinsically disordered in the absence of calcium: implication for protein secretion. *J Biol Chem, 284*(3), 1781-1789.

Clavel, T., Germon, P., Vianney, A., Portalier, R., & Lazzaroni, J. C. (1998). TolB protein of *Escherichia coli* K-12 interacts with the outer membrane peptidoglycan-associated proteins Pal, Lpp and OmpA. *Molecular microbiology, 29*(1), 359-367.

Cohen, P. (1989). The structure and regulation of protein phosphatases. *Annu Rev Biochem, 58*, 453-508.

Costa, T. R. D., Felisberto-Rodrigues, C., Meir, A., Prevost, M. S., Redzej, A., Trokter, M., *et al.* (2015). Secretion systems in Gram-negative bacteria: structural and mechanistic insights. [Review]. *Nat Rev Micro, 13*(6), 343-359.

Cox, D. L., Chang, P., McDowall, A. W., & Radolf, J. D. (1992). The outer membrane, not a coat of host proteins, limits antigenicity of virulent *Treponema pallidum*. *Infection and Immunity, 60*(3), 1076-1083.

Davidson, A. L., Dassa, E., Orelle, C., & Chen, J. (2008). Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiology and Molecular Biology Reviews, 72*(2), 317-364.

Dawson, R. J. P., & Locher, K. P. (2006). Structure of a bacterial multidrug ABC transporter. [10.1038/nature05155]. *Nature, 443*(7108), 180-185.

De Gregorio, E., Del Franco, M., Martinucci, M., Roscetto, E., Zarrilli, R., & Di Nocera, P. P. (2015). Biofilm-associated proteins: news from *Acinetobacter*. *BMC Genomics, 16*, 933.

De Santis, M., De Luca, C., Mappa, I., Spagnuolo, T., Licameli, A., Straface, G., *et al.* (2012). Syphilis infection during pregnancy: fetal risks and clinical management. *Infectious diseases in obstetrics and gynecology, 2012*.

Debarbieux, L., & Wandersman, C. (2001). Folded HasA inhibits its own secretion through its ABC exporter. *The EMBO journal, 20*(17), 4657-4663.

Delepelaire, P. (2004). Type I secretion in gram-negative bacteria. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research, 1694*(1–3), 149-161.

Delepelaire, P., & Wandersman, C. (1998). The SecB chaperone is involved in the secretion of the *Serratia marcescens* HasA protein through an ABC transporter. *The EMBO Journal, 17*(4), 936-944.

Desrosiers, D. C., Anand, A., Luthra, A., Dunham-Ems, S. M., LeDoyt, M., Cummings, M. A., *et al.* (2011). TP0326, a *Treponema pallidum* beta-barrel assembly machinery A (BamA) orthologue and rare outer membrane protein. *Mol Microbiol, 80*(6), 1496-1515.

Dirix, G., Monsieurs, P., Dombrecht, B., Daniels, R., Marchal, K., Vanderleyden, J., *et al.* (2004). Peptide signal molecules and bacteriocins in Gram-negative bacteria: a genome-wide *in silico* screening for peptides containing a double-glycine leader sequence and their cognate transporters. *Peptides, 25*(9), 1425-1440.

Ehrmann, M., Ehrle, R., Hofmann, E., Boos, W., & Schlösser, A. (1998). The ABC maltose transporter. *Molecular microbiology, 29*(3), 685-694.

El-Azami-El-Idrissi, M., Bauche, C., Loucka, J., Osicka, R., Sebo, P., Ladant, D., *et al.* (2003). Interaction of *Bordetella pertussis* adenylate cyclase with CD11b/CD18: Role of toxin acylation and identification of the main integrin interaction domain. *J Biol Chem, 278*(40), 38514-38521.

Endo, T., Kawano, S., & Yamano, K. (2011). BamE structure: the assembly of β-barrel proteins in the outer membranes of bacteria and mitochondria. *EMBO reports, 12*(2), 94-95.

Eswaran, J., Hughes, C., & Koronakis, V. (2003). Locking TolC entrance helices to prevent protein translocation by the bacterial type I export apparatus. *J Mol Biol, 327*(2), 309-315.

Fath, M. J., Skvirsky, R. C., & Kolter, R. (1991). Functional complementation between bacterial MDR-like export systems: Colicin V, alpha-hemolysin, and *Erwinia* protease. *J Bacteriol, 173*(23), 7549-7556.

Filloux, A. (2004). The underlying mechanisms of type II protein secretion. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research, 1694*(1–3), 163-179.

Filloux, A. (2010). Secretion Signal and Protein Targeting in Bacteria: a Biological Puzzle. *Journal of Bacteriology, 192*(15), 3847-3849.

Fitzgerald, T. J. (1987). Activation of the classical and alternative pathways of complement by *Treponema pallidum* subsp. *pallidum* and *Treponema vincentii*. *Infect Immun, 55*(9), 2066-2073.

Fitzgerald, T. J., Cleveland, P., Johnson, R. C., Miller, J. N., & Sykes, J. A. (1977). Scanning electron microscopy of *Treponema pallidum* (Nichols strain) attached to cultured mammalian cells. *J Bacteriol, 130*(3), 1333-1344.

Fitzgerald, T. J., Johnson, R. C., Miller, J. N., & Sykes, J. A. (1977). Characterization of the attachment of *Treponema pallidum* (Nichols strain) to cultured mammalian cells and the potential relationship of attachment to pathogenicity. *Infect Immun, 18*(2), 467-478.

Fitzgerald, T. J., Repesh, L. A., Blanco, D. R., & Miller, J. N. (1984). Attachment of *Treponema pallidum* to fibronectin, laminin, collagen IV, and collagen I, and blockage of attachment by immune rabbit IgG. *Br J Vener Dis, 60*(6), 357-363.

Fraser, C. M., Norris, S. J., Weinstock, C. M., White, O., Sutton, G. G., Dodson, R., *et al.* (1998). Complete genome sequence of *Treponema pallidum*, the syphilis spirochete. *Science, 281*(5375), 375-388.

French, P. (2007). Syphilis. *BMJ : British Medical Journal, 334*(7585), 143-147.

Frith, J. (2012). Syphilis - Its Early History and Treatment Until Penicillin, and the Debate on its Origins. [Article]. *Journal of Military & Veterans' Health, 20*(4), 49-58.

Fujimoto, H., Onishi, N., Kato, N., Takekawa, M., Xu, X. Z., Kosugi, A., *et al.* (2005). Regulation of the antioncogenic Chk2 kinase by the oncogenic Wip1 phosphatase. *Cell Death Differ, 13*(7), 1170-1180.

Gaskell, A., Crennell, S., & Taylor, G. (1995). The three domains of a bacterial sialidase: a β-propeller, an immunoglobulin module and a galactose-binding jelly-roll. *Structure, 3*(11), 1197-1205.

Gauthier, A., & Finlay, B. B. (2001). Bacterial pathogenesis: the answer to virulence is in the pore. *Curr Biol, 11*(7), R264-267.

Gautier, R., Douguet, D., Antonny, B., & Drin, G. (2008). HeliQuest: a web server to screen sequences with specific α-helical properties. *Bioinformatics, 24*(18), 2101-2102.

Gérard, F., Pradel, N., & Wu, L.-F. (2005). Bactericidal Activity of Colicin V Is Mediated by an Inner Membrane Protein, SdaC, of *Escherichia coli*. *Journal of Bacteriology, 187*(6), 1945-1950.

Gerlach, R. G., Jäckel, D., Stecher, B., Wagner, C., Lupas, A., Hardt, W.-D., *et al.* (2007). *Salmonella* Pathogenicity Island 4 encodes a giant non-fimbrial adhesin and the cognate type 1 secretion system. *Cellular Microbiology, 9*(7), 1834-1850.

Glaser, P., Sakamoto, H., Bellalou, J., Ullmann, A., & Danchin, A. (1988). Secretion of cyclolysin, the calmodulin-sensitive adenylate cyclase-haemolysin bifunctional protein of *Bordetella pertussis*. *Embo j, 7*(12), 3997-4004.

Goh, H. M. S., Beatson, S. A., Totsika, M., Moriel, D. G., Phan, M.-D., Szubert, J., *et al.* (2013). Molecular Analysis of the *Acinetobacter baumannii* Biofilm-Associated Protein. *Applied and Environmental Microbiology, 79*(21), 6535-6543.

Goodwin, M. S., & Weiss, A. A. (1990). Adenylate cyclase toxin is critical for colonization and pertussis toxin is critical for lethal infection by *Bordetella pertussis* in infant mice. *Infection and Immunity, 58*(10), 3445-3447.

Griessl, M. H., Schmid, B., Kassler, K., Braunsmann, C., Ritter, R., Barlag, B., *et al.* (2013). Structural insight into the giant Ca(2)(+)-binding adhesin SiiE: implications for the adhesion of *Salmonella enterica* to polarized epithelial cells. *Structure, 21*(5), 741-752.

Guermonprez, P., Khelef, N., Blouin, E., Rieu, P., Ricciardi-Castagnoli, P., Guiso, N., *et al.* (2001). The Adenylate Cyclase Toxin of *Bordetella pertussis* Binds to Target Cells via the αMβ2 Integrin (Cd11b/Cd18). *The Journal of Experimental Medicine, 193*(9), 1035-1044.

Guzzo, J., Duong, F., Wandersman, C., Murgier, M., & Lazdunski, A. (1991). The secretion genes of *Pseudomonas aeruginosa* alkaline protease are functionally related to those of *Erwinia chrysanthemi* proteases and *Escherichia coli* alpha-haemolysin. *Mol Microbiol, 5*(2), 447-453.

Haber-Pohlmeier, S., Abarca-Heidemann, K., Körschen, H. G., Dhiman, H. K., Heberle, J., Schwalbe, H., *et al.* (2007). Binding of Ca(2+) to Glutamic Acid-Rich Polypeptides from the Rod Outer Segment. *Biophysical Journal, 92*(9), 3207-3214.

Hagan, C. L., Kim, S., & Kahne, D. (2010). Reconstitution of outer membrane protein assembly from purified components. *Science, 328*(5980), 890-892.

Hanski, E. (1989). Invasive adenylate cyclase toxin of *Bordetella pertussis*. *Trends in Biochemical Sciences, 14*(11), 459-463.

Hardy Jr, P., & Nell, E. E. (1957). Study of the antigenic structure of *Treponema pallidum* by specific agglutination. *American journal of hygiene, 66*(2), 160-172.

Havarstein, L. S., Diep, D. B., & Nes, I. F. (1995). A family of bacteriocin ABC transporters carry out proteolytic processing of their substrates concomitant with export. *Molecular Microbiology, 16*(2), 229-240.

Havarstein, L. S., Holo, H., & Nes, I. F. (1994). The leader peptide of Colicin V shares consensus sequences with leader peptides that are common among peptide bacteriocins produced by gram-positive bacteria. *Microbiology, 140 (Pt 9)*, 2383-2389.

Hayashi, S., & Wu, H. C. (1990). Lipoproteins in bacteria. *Journal of bioenergetics and biomembranes, 22*(3), 451-471.

Hege, T., & Baumann, U. (2001). Protease C of *Erwinia chrysanthemi*: the crystal structure and role of amino acids Y228 and E189. *J Mol Biol, 314*(2), 187-193.

Henderson, I. R., & Nataro, J. P. (2001). Virulence Functions of Autotransporter Proteins. *Infection and Immunity, 69*(3), 1231-1243.

Hinsa, S. M., Espinosa-Urgel, M., Ramos, J. L., & O'Toole, G. A. (2003). Transition from reversible to irreversible attachment during biofilm formation by *Pseudomonas fluorescens* WCS365 requires an ABC transporter and a large secreted protein. *Mol Microbiol, 49*(4), 905-918.

Ho, E. L., & Lukehart, S. A. (2011). Syphilis: using modern approaches to understand an old disease. *J Clin Invest, 121*(12), 4584-4592.

Holland, I. B., Schmitt, L., & Young, J. (2005). Type 1 protein secretion in bacteria, the ABC-transporter dependent pathway (Review). *Molecular Membrane Biology, 22*(1-2), 29-39.

Hook, E. W. (1989). Syphilis and HIV Infection. *Journal of Infectious Diseases, 160*(3), 530-534.

Houston, S., Hof, R., Francescutti, T., Hawkes, A., Boulanger, M. J., & Cameron, C. E. (2011). Bifunctional Role of the *Treponema pallidum* Extracellular Matrix Binding Adhesin Tp0751. *Infection and Immunity, 79*(3), 1386-1398.

Houston, S., Hof, R., Honeyman, L., Hassler, J., & Cameron, C. E. (2012). Activation and Proteolytic Activity of the *Treponema pallidum* Metalloprotease, Pallilysin. *Plos Pathogens, 8*(7).

Houston, S., Russell, S., Hof, R., Roberts, A. K., Cullen, P., Irvine, K., *et al.* (2014). The multifunctional role of the pallilysin-associated *Treponema pallidum* protein, Tp0750, in promoting fibrinolysis and extracellular matrix component degradation. *Mol Microbiol, 91*(3), 618-634.

Hulko, M., Berndt, F., Gruber, M., Linder, J. U., Truffault, V., Schultz, A., *et al.* (2006). The HAMP Domain Structure Implies Helix Rotation in Transmembrane Signaling. *Cell, 126*(5), 929-940.

Hyland, C., Vuillard, L., Hughes, C., & Koronakis, V. (2001). Membrane Interaction of *Escherichia coli* Hemolysin: Flotation and Insertion-Dependent Labeling by Phospholipid Vesicles. *Journal of bacteriology, 183*(18), 5364-5370.

Izard, J., Renken, C., Hsieh, C. E., Desrosiers, D. C., Dunham-Ems, S., La Vake, C., *et al.* (2009). Cryo-electron tomography elucidates the molecular architecture of *Treponema pallidum*, the syphilis spirochete. *J Bacteriol, 191*(24), 7566-7580.

Jermyn, W. S., & Boyd, E. F. (2002). Characterization of a novel *Vibrio* pathogenicity island (VPI-2) encoding neuraminidase (nanH) among toxigenic *Vibrio cholerae* isolates. *Microbiology, 148*(11), 3681-3693.

Jorgensen, S. E., Hammer, R. F., & Wu, G. K. (1980). Effects of a single hit from the alpha hemolysin produced by *Escherichia coli* on the morphology of sheep erythrocytes. *Infect Immun, 27*(3), 988-994.

Jumpertz, T., Schmitt, L., Jenewein, S., Zaitseva, J., & Holland, I. B. (2005). H662 is the linchpin of ATP hydrolysis in the nucleotide-binding domain of the ABC transporter HlyB. *The EMBO Journal, 24*(11), 1901-1910.

Kanonenberg, K., Schwarz, C. K., & Schmitt, L. (2013). Type I secretion systems - a story of appendices. *Res Microbiol, 164*(6), 596-604.

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. [Protocol]. *Nat. Protocols, 10*(6), 845-858.

Kerr, I. D. (2002). Structure and association of ATP-binding cassette transporter nucleotide-binding domains. *Biochimica et Biophysica Acta (BBA) - Biomembranes, 1561*(1), 47-64.

Kim, J.-S., Song, S., Lee, M., Lee, S., Lee, K., & Ha, N.-C. (2016). Crystal Structure of a Soluble Fragment of the Membrane Fusion Protein HlyD in a Type I Secretion System of Gram-Negative Bacteria. *Structure*.

Kim, K. H., Aulakh, S., & Paetzel, M. (2011). Crystal Structure of beta-Barrel Assembly Machinery BamCD Protein Complex. [Article]. *Journal of Biological Chemistry, 286*(45), 39116-39121.

Kim, K. H., Aulakh, S., & Paetzel, M. (2012). The bacterial outer membrane β-barrel assembly machinery. *Protein Science : A Publication of the Protein Society, 21*(6), 751-768.

Kim, S., Malinverni, J. C., Sliz, P., Silhavy, T. J., Harrison, S. C., & Kahne, D. (2007). Structure and function of an essential component of the outer membrane protein assembly machine. *Science, 317*(5840), 961-964.

Kim, S., Oh, D. B., Kang, H. A., & Kwon, O. (2011). Features and applications of bacterial sialidases. *Appl Microbiol Biotechnol, 91*(1), 1-15.

Klein, R. J., Zeiss, C., Chew, E. Y., Tsai, J.-Y., Sackler, R. S., Haynes, C., *et al.* (2005). Complement factor H polymorphism in age-related macular degeneration. *Science, 308*(5720), 385-389.

Koronakis, V., Koronakis, E., & Hughes, C. (1989). Isolation and analysis of the C-terminal signal directing export of *Escherichia coli* hemolysin protein across both bacterial membranes. *The EMBO Journal, 8*(2), 595-605.

Koronakis, V., Li, J., Koronakis, E., & Stauffer, K. (1997). Structure of TolC, the outer membrane component of the bacterial type I efflux system, derived from two-dimensional crystals. *Mol Microbiol, 23*(3), 617-626.

Koronakis, V., Sharff, A., Koronakis, E., Luisi, B., & Hughes, C. (2000). Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. *Nature, 405*(6789), 914-919.

Korotkov, K. V., Sandkvist, M., & Hol, W. G. J. (2012). The type II secretion system: biogenesis, molecular architecture and mechanism. *Nature reviews. Microbiology, 10*(5), 336-351.

Krogh, A., Larsson, B., von Heijne, G., & Sonnhammer, E. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol, 305*(3), 567 - 580.

LaFond, R. E., & Lukehart, S. A. (2006). Biological basis for syphilis. *Clinical Microbiology Reviews, 19*(1), 29-+.

Larkin, M. A., Blackshields, G., Brown, N., Chenna, R., McGettigan, P. A., McWilliam, H., *et al.* (2007). Clustal W and Clustal X version 2.0. *Bioinformatics, 23*(21), 2947-2948.

Latasa, C., Roux, A., Toledo-Arana, A., Ghigo, J.-M., Gamazo, C., Penadés, J. R., *et al.* (2005). BapA, a large secreted protein required for biofilm formation and host colonization of *Salmonella enterica serovar Enteritidis*. *Molecular Microbiology, 58*(5), 1322-1339.

Lecher, J., Schwarz, Christian K. W., Stoldt, M., Smits, Sander H. J., Willbold, D., & Schmitt, L. (2012). An RTX Transporter Tethers Its Unfolded Substrate during Secretion via a Unique N-Terminal Domain. *Structure, 20*(10), 1778-1787.

Lenders, M. H. H., Weidtkamp-Peters, S., Kleinschrodt, D., Jaeger, K.-E., Smits, S. H. J., & Schmitt, L. (2015). Directionality of substrate translocation of the hemolysin A Type I secretion system. [Article]. *Scientific Reports, 5*, 12470.

Letoffe, S., Delepelaire, P., & Wandersman, C. (1996). Protein secretion in Gram-negative bacteria: Assembly of the three components of ABC protein-mediated exporters is ordered and promoted by substrate binding. *Embo Journal, 15*(21), 5804-5811.

Letoffe, S., Ghigo, J. M., & Wandersman, C. (1994). Secretion of the *Serratia marcescens* HasA protein by an ABC transporter. *Journal of Bacteriology, 176*(17), 5372-5377.

Létoffé, S., Ghigo, J. M., & Wandersman, C. (1994). Iron acquisition from heme and hemoglobin by a *Serratia marcescens* extracellular protein. *Proceedings of the National Academy of Sciences, 91*(21), 9876-9880.

Létoffé, S., Omori, K., & Wandersman, C. (2000). Functional characterization of the HasAPF hemophore and its truncated and chimeric variants: determination of a region involved in binding to the hemophore receptor. *Journal of bacteriology, 182*(16), 4401-4405.

Leyton, D. L., Sevastsyanovich, Y. R., Browning, D. F., Rossiter, A. E., Wells, T. J., Fitzpatrick, R. E., *et al.* (2011). Size and Conformation Limits to Secretion of Disulfide-bonded Loops in Autotransporter Proteins. *The Journal of Biological Chemistry, 286*(49), 42283-42291.

Limberger, R. J. (2004). The Periplasmic Flagellum of Spirochetes. *Journal of Molecular Microbiology and Biotechnology, 7*(1-2), 30-40.

Linhartová, I., Bumba, L., Mašín, J., Basler, M., Osička, R., Kamanová, J., *et al.* (2010). RTX proteins: a highly diverse family secreted by a common mechanism. *Fems Microbiology Reviews, 34*(6), 1076-1112.

Lopes dos Santos Santiago, G., Deschaght, P., El Aila, N., Kiama, T. N., Verstraelen, H., Jefferson, K. K., *et al.* (2011). *Gardnerella vaginalis* comprises three distinct genotypes of which only two produce sialidase. *American Journal of Obstetrics and Gynecology, 204*(5), 450.e451-450.e457.

Lukehart, S. A., Hook, I. I. I. E. W., Baker-Zander, S. A., Collier, A. C., Critchlow, C. W., & Handsfield, H. H. (1988). Invasion of the Central Nervous System by *Treponema pallidum*: Implications for Diagnosis and Treatment. *Annals of Internal Medicine, 109*(11), 855-862.

Luo, Y., Frey, E. A., Pfuetzner, R. A., Creagh, A. L., Knoechel, D. G., Haynes, C. A., *et al.* (2000). Crystal structure of enteropathogenic *Escherichia coli* intimin-receptor complex. [10.1038/35016618]. *Nature, 405*(6790), 1073-1077.

Mackman, N., & Holland, I. B. (1984). Functional characterization of a cloned haemolysin determinant from *E. coli* of human origin, encoding information for the secretion of a 107K polypeptide. *Mol Gen Genet, 196*(1), 129-134.

Mackman, N., Nicaud, J. M., Gray, L., & Holland, I. B. (1985). Identification of polypeptides required for the export of haemolysin 2001 from *E. coli*. *Mol Gen Genet, 201*(3), 529-536.

Magnuson, H. J., Hill, C., Eagle, H., & Fleischman, R. (1948). The Minimal Infectious Inoculum of *Spirochaeta pallida* (Nichols Strain), and a consideration of its Rate of Multiplication in Vivo. *American Journal of Syphilis, 32*(1), 1-18.

Mahoney, J., & Bryant, K. (1933). Contact infection of rabbits in experimental syphilis. *Am. J. Syphilis, 17*, 188-193.

Marchler-Bauer, A., & Bryant, S. H. (2004). CD-Search: protein domain annotations on the fly. *Nucleic Acids Res, 32*(Web Server issue), W327-331.

Marlovits, T. C., Kubori, T., Lara-Tejero, M., Thomas, D., Unger, V. M., & Galán, J. E. (2006). Assembly of the inner rod determines needle length in the type III secretion injectisome. *Nature, 441*(7093), 637-640.

Marlovits, T. C., Kubori, T., Sukhan, A., Thomas, D. R., Galán, J. E., & Unger, V. M. (2004). Structural insights into the assembly of the type III secretion needle complex. *Science, 306*(5698), 1040-1042.

Martín, H., Flández, M., Nombela, C., & Molina, M. (2005). Protein phosphatases in MAPK signalling: we keep learning from yeast. *Molecular Microbiology, 58*(1), 6-16.

Masi, M., & Wandersman, C. (2010). Multiple Signals Direct the Assembly and Function of a Type 1 Secretion System. *Journal of Bacteriology, 192*(15), 3861-3869.

Masin, J., Basler, M., Knapp, O., El-Azami-El-Idrissi, M., Maier, E., Konopasek, I., *et al.* (2005). Acylation of lysine 860 allows tight binding and cytotoxicity of *Bordetella* adenylate cyclase on CD11b-expressing cells. *Biochemistry, 44*(38), 12759-12766.

Masure, H. R., Au, D. C., Gross, M. K., Donovan, M. G., & Storm, D. R. (1990). Secretion of the *Bordetella pertussis* adenylate cyclase from *Escherichia coli* containing the hemolysin operon. *Biochemistry, 29*(1), 140-145.

Matsunaga, J., Barocchi, M. A., Croda, J., Young, T. A., Sanchez, Y., Siqueira, I., *et al.* (2003). Pathogenic *Leptospira* species express surface-exposed proteins belonging to the bacterial immunoglobulin superfamily. *Molecular microbiology, 49*(4), 929-946.

Matsuyama, S. i., Yokota, N., & Tokuda, H. (1997). A novel outer membrane lipoprotein, LolB (HemM), involved in the LolA (p20)-dependent localization of lipoproteins to the outer membrane of *Escherichia coli*. *The EMBO journal, 16*(23), 6947-6955.

Meier, R., Drepper, T., Svensson, V., Jaeger, K.-E., & Baumann, U. (2007). A Calcium-gated Lid and a Large β-Roll Sandwich Are Revealed by the Crystal Structure of Extracellular Lipase from *Serratia marcescens*. *Journal of Biological Chemistry, 282*(43), 31477-31483.

Merritt, E. A., Sarfaty, S., Akker, F. V. D., L'Hoir, C., Martial, J. A., & Hol, W. G. (1994). Crystal structure of cholera toxin B-pentamer bound to receptor GM1 pentasaccharide. *Protein Science, 3*(2), 166-175.

Moayeri, M., & Welch, R. A. (1994). Effects of temperature, time, and toxin concentration on lesion formation by the *Escherichia coli* hemolysin. *Infect Immun, 62*(10), 4124-4134.

Moffatt, C. E., Inaba, H., Hirano, T., & Lamont, R. J. (2012). *Porphyromonas gingivalis* SerB-mediated dephosphorylation of host cell cofilin modulates invasion efficiency. *Cellular microbiology, 14*(4), 577-588.

Morgan, E., Bowen, A. J., Carnell, S. C., Wallis, T. S., & Stevens, M. P. (2007). SiiE Is Secreted by the *Salmonella enterica Serovar Typhimurium* Pathogenicity Island 4-Encoded Secretion System and Contributes to Intestinal Colonization in Cattle. *Infection and Immunity, 75*(3), 1524-1533.

Moskophidis, M., & Muller, F. (1984). Molecular characterization of glycoprotein antigens on surface of *Treponema pallidum*: comparison with nonpathogenic *Treponema phagedenis* biotype *Reiter*. *Infect Immun, 46*(3), 867-869.

Narita, S.-i., & Tokuda, H. (2007). Amino acids at positions 3 and 4 determine the membrane specificity of *Pseudomonas aeruginosa* lipoproteins. *Journal of Biological Chemistry, 282*(18), 13372-13378.

Natale, P., Brüser, T., & Driessen, A. J. M. (2008). Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane—Distinct translocases and mechanisms. *Biochimica et Biophysica Acta (BBA) - Biomembranes, 1778*(9), 1735-1756.

Nelson, R., & Mayer, M. (1949). J. exp. *Med, 89*, 369.

Nicaud, J. M., Mackman, N., Gray, L., & Holland, I. B. (1985). Regulation of haemolysin synthesis in *E. coli* determined by HLY genes of human origin. *Mol Gen Genet, 199*(1), 111-116.

Nikaido, H. (1998). Multiple antibiotic resistance and efflux. *Current opinion in microbiology, 1*(5), 516-523.

Norris, S. J., Cox, D. L., & Weinstock, G. M. (2001). Biology of *Treponema pallidum*: correlation of functional activities with genome sequence data. *Journal of molecular microbiology and biotechnology, 3*(1), 37-62.

Notti, R. Q., & Stebbins, C. E. (2016). The Structure and Function of Type III Secretion Systems. *Microbiology spectrum, 4*(1), 10.1128/microbiolspec.VMBF-0004-2015.

Nusbaum, M. R., Wallace, R. R., Slatt, L. M., & Kondrad, E. C. (2004). Sexually transmitted infections and increased risk of co-infection with human immunodeficiency virus. *J Am Osteopath Assoc, 104*(12), 527-535.

Ohta, K., Makinen, K., & Loesche, W. J. (1986). Purification and characterization of an enzyme produced by *Treponema denticola* capable of hydrolyzing synthetic trypsin substrates. *Infection and immunity, 53*(1), 213-220.

Ohyashiki, T., Taka, M., & Mohri, T. (1987). Effect of neuraminidase treatment on the lipid fluidity of the intestinal brush-border membranes. *Biochimica et Biophysica Acta (BBA) - Biomembranes, 905*(1), 57-64.

Okuda, S., & Tokuda, H. (2011). Lipoprotein sorting in bacteria. *Annu Rev Microbiol, 65*, 239-259.

Oropeza-Wekerle, R. L., Muller, E., Kern, P., Meyermann, R., & Goebel, W. (1989). Synthesis, inactivation, and localization of extracellular and intracellular *Escherichia coli* hemolysins. *J Bacteriol, 171*(5), 2783-2788.

Osičková, A., Osička, R., Maier, E., Benz, R., & Šebo, P. (1999). An Amphipathic α-Helix Including Glutamates 509 and 516 Is Crucial for Membrane Translocation of Adenylate Cyclase Toxin and Modulates Formation and Cation Selectivity of Its Membrane Channels. *Journal of Biological Chemistry, 274*(53), 37644-37650.

Ostolaza, H., Bakas, L., & Goni, F. M. (1997). Balance of electrostatic and hydrophobic interactions in the lysis of model membranes by *E. coli* alpha-haemolysin. *J Membr Biol, 158*(2), 137-145.

Parker, C. E., Warren, M. R., Loiselle, D. R., Dicheva, N. N., Scarlett, C. O., & Borchers, C. H. (2005). Identification of components of protein complexes. *Methods Mol Biol, 301*, 117-151.

Paster, B. J., Dewhirst, F. E., Weisburg, W. G., Tordoff, L. A., Fraser, G. J., Hespell, R. B., *et al.* (1991). Phylogenetic analysis of the spirochetes. *Journal of Bacteriology, 173*(19), 6101-6109.

Peterman, T. A., Newman, D. R., Maddox, L., Schmitt, K., & Shiver, S. (2014). High risk for HIV following syphilis diagnosis among men in Florida, 2000-2011. *Public Health Rep, 129*(2), 164-169.

Petrotchenko, E. V., & Borchers, C. H. (2010). ICC-CLASS: isotopically-coded cleavable crosslinking analysis software suite. *BMC Bioinformatics, 11*, 64.

Petrotchenko, E. V., Serpa, J. J., & Borchers, C. H. (2011). An isotopically coded CID-cleavable biotinylated cross-linker for structural proteomics. *Mol Cell Proteomics, 10*(2), M110.001420.

Pimenta, A. L., Racher, K., Jamieson, L., Blight, M. A., & Holland, I. B. (2005). Mutations in HlyD, part of the type 1 translocator for hemolysin secretion, affect the folding of the secreted toxin. *Journal of Bacteriology, 187*(21), 7471-7480.

Porcella, S. F., & Schwan, T. G. (2001). *Borrelia burgdorferi* and *Treponema pallidum*: a comparison of functional genomics, environmental adaptations, and pathogenic mechanisms. *The Journal of Clinical Investigation, 107*(6), 651-656.

Pukatzki, S., Ma, A. T., Sturtevant, D., Krastins, B., Sarracino, D., Nelson, W. C., *et al.* (2006). Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the Dictyostelium host model system. *Proceedings of the National Academy of Sciences, 103*(5), 1528-1533.

Pukatzki, S., McAuley, S. B., & Miyata, S. T. (2009). The type VI secretion system: translocation of effectors and effector-domains. *Current Opinion in Microbiology, 12*(1), 11-17.

Radolf, J. D., Blanco, D. R., Miller, J. N., & Lovett, M. A. (1986). Antigenic interralationship between endoflagella of *Treponema phagedenis* biotype *reiter* and *Treponema pallidum* (Nichols) - Molecular characterization of endoflagellar proteins. *Infection and Immunity, 54*(3), 626-634.

Radolf, J. D., & Lukehart, S. A. (2006). *Pathogenic Treponema: molecular and cellular biology*: Horizon Scientific Press.

Radolf, J. D., & Norgard, M. V. (1988). Pathogen specificity of *Treponema pallidum* subsp. *pallidum* integral membrane proteins identified by phase partitioning with Triton X-114. *Infection and Immunity, 56*(7), 1825-1828.

Ribet, D., & Cossart, P. (2015). How bacterial pathogens colonize their hosts and invade deeper tissues. *Microbes and Infection, 17*(3), 173-183.

Ruiz, N., Kahne, D., & Silhavy, T. J. (2009). Transport of lipopolysaccharide across the cell envelope: the long road of discovery. [10.1038/nrmicro2184]. *Nat Rev Micro, 7*(9), 677-683.

Samudrala, R., Heffron, F., & McDermott, J. E. (2009). Accurate Prediction of Secreted Substrates and Identification of a Conserved Putative Secretion Signal for Type III Secretion Systems. *PLoS Pathog, 5*(4), e1000375.

Sanchez-Magraner, L., Viguera, A. R., Garcia-Pacios, M., Garcillan, M. P., Arrondo, J. L., de la Cruz, F., *et al.* (2007). The calcium-binding C-terminal domain of *Escherichia coli* alpha-hemolysin is a major determinant in the surface-active properties of the protein. *J Biol Chem, 282*(16), 11827-11835.

Sandoval, C. M., Baker, S. L., Jansen, K., Metzner, S. I., & Sousa, M. C. (2011). Crystal structure of BamD: an essential component of the β-barrel assembly machinery of Gram-negative bacteria. *Journal of molecular biology, 409*(3), 348-357.

Sankaran, K., & Wu, H. C. (1994). Lipid modification of bacterial prolipoprotein. Transfer of diacylglyceryl moiety from phosphatidylglycerol. *Journal of Biological Chemistry, 269*(31), 19701-19706.

Sapriel, G., Wandersman, C., & Delepelaire, P. (2002). The N Terminus of the HasA Protein and the SecB Chaperone Cooperate in the Efficient Targeting and Secretion of HasA via the ATP-binding Cassette Transporter. *Journal of Biological Chemistry, 277*(8), 6726-6732.

Satchell, K. J. (2011). Structure and function of MARTX toxins and other large repetitive RTX proteins. *Annu Rev Microbiol, 65*, 71-90.

Saxena, M., & Mustelin, T. (2000). Extracellular signals and scores of phosphatases: All roads lead to MAP kinase. *Seminars in Immunology, 12*(4), 387-396.

Scheu, A. K., Economou, A., Hong, G. F., Ghelani, S., Johnston, A. W., & Downie, J. A. (1992). Secretion of the *Rhizobium leguminosarum* nodulation protein NodO by haemolysin-type systems. *Mol Microbiol, 6*(2), 231-238.

Schmidt, H., Maier, E., Karch, H., & Benz, R. (1996). Pore-Forming Properties of the Plasmid-Encoded Hemolysin of Enterohemorrhagic *Escherichia coli* 0157: H7. *European Journal of Biochemistry, 241*(2), 594-601.

Schmitt, L., Benabdelhak, H., Blight, M. A., Holland, I. B., & Stubbs, M. T. (2003). Crystal Structure of the Nucleotide-binding Domain of the ABC-transporter Haemolysin B: Identification of a Variable Region Within ABC Helical Domains. *Journal of Molecular Biology, 330*(2), 333-342.

Schulze, R. J., Chen, S., Kumru, O. S., & Zuckert, W. R. (2010). Translocation of *Borrelia burgdorferi* surface lipoprotein OspA through the outer membrane requires an unfolded conformation and can initiate at the C-terminus. *Mol Microbiol, 76*(5), 1266-1278.

Schulze, R. J., & Zuckert, W. R. (2006). *Borrelia burgdorferi* lipoproteins are secreted to the outer surface by default. *Mol Microbiol, 59*(5), 1473-1484.

Schweighofer, A., Hirt, H., & Meskiene, I. (2004). Plant PP2C phosphatases: emerging functions in stress signaling. *Trends in Plant Science, 9*(5), 236-243.

Setubal, J. C., Reis, M., Matsunaga, J., & Haake, D. A. (2006). Lipoprotein computational prediction in spirochaetal genomes. *Microbiology, 152*(1), 113-121.

Sharff, A., Fanutti, C., Shi, J., Calladine, C., & Luisi, B. (2001). The role of the TolC family in protein transport and multidrug efflux. From stereochemical certainty to mechanistic hypothesis. *Eur J Biochem, 268*(19), 5011-5026.

Shrivastava, R., & Miller, J. F. (2009). Virulence Factor Secretion by *Bordetella* Species. *Current opinion in microbiology, 12*(1), 88-93.

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., *et al.* (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology, 7*(1).

Silhavy, T. J., Kahne, D., & Walker, S. (2010). The bacterial cell envelope. *Cold Spring Harbor perspectives in biology, 2*(5), a000414.

Sklar, J. G., Wu, T., Gronenberg, L. S., Malinverni, J. C., Kahne, D., & Silhavy, T. J. (2007). Lipoprotein SmpA is a component of the YaeT complex that assembles outer membrane proteins in *Escherichia coli*. *Proceedings of the National Academy of Sciences, 104*(15), 6400-6405.

Sonnhammer, E. L., von Heijne, G., & Krogh, A. (1998). A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc Int Conf Intell Syst Mol Biol, 6*, 175-182.

Sorg, J. A., Miller, N. C., & Schneewind, O. (2005). Substrate recognition of type III secretion machines –testing the RNA signal hypothesis. *Cellular Microbiology, 7*(9), 1217-1225.

Sotomayor-Pérez, A.-C., Ladant, D., & Chenal, A. (2015). Disorder-to-Order Transition in the CyaA Toxin RTX Domain: Implications for Toxin Secretion. *Toxins, 7*(1), 1.

Stanley, P., Koronakis, V., & Hughes, C. (1998). Acylation of *Escherichia coli* hemolysin: A unique protein lipidation mechanism underlying toxin function. *Microbiology and Molecular Biology Reviews, 62*(2), 309-+.

Stanley, P., Packman, L. C., Koronakis, V., & Hughes, C. (1994). Fatty acylation of two internal lysine residues required for the toxic activity of *Escherichia coli* hemolysin. *Science, 266*(5193), 1992-1996.

Takeuchi, H., Hirano, T., Whitmore, S. E., Morisaki, I., Amano, A., & Lamont, R. J. (2013). The Serine Phosphatase SerB of *Porphyromonas gingivalis* Suppresses IL-8 Production by Dephosphorylation of NF-?B RelA/p65. *PLoS Pathog, 9*(4), e1003326.

Tanaka, S.-y., Narita, S.-i., & Tokuda, H. (2007). Characterization of the *Pseudomonas aeruginosa* Lol system as a lipoprotein sorting mechanism. *Journal of Biological Chemistry, 282*(18), 13379-13384.

Terada, M., Kuroda, T., Matsuyama, S.-i., & Tokuda, H. (2001). Lipoprotein sorting signals evaluated as the LolA-dependent release of lipoproteins from the cytoplasmic membrane of *Escherichia coli*. *Journal of Biological Chemistry, 276*(50), 47690-47694.

Thanabalu, T., Koronakis, E., Hughes, C., & Koronakis, V. (1998). Substrate-induced assembly of a contiguous channel for protein export from *E-coli*: reversible bridging of an inner-membrane translocase to an outer membrane exit pore. *Embo Journal, 17*(22), 6487-6496.

Theunissen, S., De Smet, L., Dansercoer, A., Motte, B., Coenye, T., Van Beeumen, J. J., *et al.* (2010). The 285 kDa Bap/RTX hybrid cell surface protein (SO4317) of *Shewanella oneidensis* MR-1 is a key mediator of biofilm formation. *Research in Microbiology, 161*(2), 144-152.

Thomas, D. D., Navab, M., Haake, D. A., Fogelman, A. M., Miller, J. N., & Lovett, M. A. (1988). *Treponema pallidum* Invades Intercellular Junctions of Endothelial Cell Monolayers. *Proceedings of the National Academy of Sciences of the United States of America, 85*(10), 3608-3612.

Tokuda, H. (2009). Biogenesis of outer membranes in Gram-negative bacteria. *Bioscience, biotechnology, and biochemistry, 73*(3), 465-473.

Tokuda, H., & Matsuyama, S.-i. (2004). Sorting of lipoproteins to the outer membrane in *E. coli*. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research, 1693*(1), 5-13.

Trokter, M., Felisberto-Rodrigues, C., Christie, P. J., & Waksman, G. (2014). Recent advances in the structural and molecular biology of type IV secretion systems. *Current opinion in structural biology, 27*, 16-23.

Useh, N. M., Nok, A. J., & Esievo, K. A. (2003). Pathogenesis and pathology of blackleg in ruminants: the role of toxins and neuraminidase. A short review. *Vet Q, 25*(4), 155-159.

van Ulsen, P., Rahman, S. u., Jong, W. S. P., Daleke-Schermerhorn, M. H., & Luirink, J. (2014). Type V secretion: From biogenesis to biotechnology. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research, 1843*(8), 1592-1611.

Vance, Tyler D. R., Olijve, Luuk L. C., Campbell, Robert L., Voets, Ilja K., Davies, Peter L., & Guo, S. (2014). Ca(2+)-stabilized adhesin helps an Antarctic bacterium reach out and bind ice. *Bioscience Reports, 34*(4), e00121.

Varki, A., & Gagneux, P. (2012). Multifarious roles of sialic acids in immunity. *Annals of the New York Academy of Sciences, 1253*(1), 16-36.

Vavrusova, M., & Skibsted, L. H. (2013). Calcium binding to dipeptides of aspartate and glutamate in comparison with orthophosphoserine. *J Agric Food Chem, 61*(22), 5380-5384.

Vergne, I., Chua, J., Lee, H.-H., Lucas, M., Belisle, J., & Deretic, V. (2005). Mechanism of phagolysosome biogenesis block by viable *Mycobacterium tuberculosis*. *Proceedings of the National Academy of Sciences of the United States of America, 102*(11), 4033-4038.

Vijay, K., Brody, M. S., Fredlund, E., & Price, C. W. (2000). A PP2C phosphatase containing a PAS domain is required to convey signals of energy stress to the σB transcription factor of *Bacillus subtilis*. *Molecular Microbiology, 35*(1), 180-188.

Vilei, E. M., Johansson, A., Schlatter, Y., Redhead, K., & Frey, J. (2011). Genetic and functional characterization of the NanA sialidase from *Clostridium chauvoei*. *Veterinary Research, 42*(1), 2-2.

Walker, E., Zampighi, G., Blanco, D., Miller, J., & Lovett, M. (1989). Demonstration of rare protein in the outer membrane of *Treponema pallidum* subsp. *pallidum* by freeze-fracture analysis. *Journal of bacteriology, 171*(9), 5005-5011.

Wallden, K., Rivera-Calzada, A., & Waksman, G. (2010). Type IV secretion systems: versatility and diversity in function. *Cellular Microbiology, 12*(9), 1203-1212.

Wang, T., Zhang, J., Zhang, X., Xu, C., & Tu, X. (2013). Solution structure of the BIg domain from *Streptococcus pneumoniae* reveals a novel Ca2+-binding module. [Article]. *Scientific Reports, 3*, 1079.

Wasserheit, J. N. (1992). Epidemiological synergy. Interrelationships between human immunodeficiency virus infection and other sexually transmitted diseases. *Sex Transm Dis, 19*(2), 61-77.

Wilson-Welder, J. H., Elliott, M. K., Zuerner, R. L., Bayles, D. O., Alt, D. P., & Stanton, T. B. (2013). Biochemical and molecular characterization of *Treponema phagedenis*-like spirochetes isolated from a bovine digital dermatitis lesion. *BMC microbiology, 13*(1), 1.

Woese, C. R. (1987). Bacterial evolution. *Microbiological reviews, 51*(2), 221.

Wolff, N., Sapriel, G., Bodenreider, C., Chaffotte, A., & Delepelaire, P. (2003). Antifolding activity of the SecB chaperone is essential for secretion of HasA, a quickly folding ABC pathway substrate. *Journal of Biological Chemistry, 278*(40), 38247-38253.

Wong, D., Bach, H., Sun, J., Hmama, Z., & Av-Gay, Y. (2011). *Mycobacterium tuberculosis* protein tyrosine phosphatase (PtpA) excludes host vacuolar-H+–

ATPase to inhibit phagosome acidification. *Proceedings of the National Academy of Sciences, 108*(48), 19371-19376.

Wong, D., Chao, J. D., & Av-Gay, Y. (2013). *Mycobacterium tuberculosis*-secreted phosphatases: from pathogenesis to targets for TB drug development. *Trends in Microbiology, 21*(2), 100-109.

Wooldridge, K. (2009). Bacterial secreted proteins: secretory mechanisms and role in pathogenesis. Norfolk, UK: Caister Academic.

World Health Organization. (2015). Sexually Transmitted Infections, 2015. *WHO Press*.

Wu, K.-H., Hsieh, Y.-H., & Tai, P. C. (2012). Mutational Analysis of Cvab, an ABC Transporter Involved in the Secretion of Active Colicin V. *PLoS ONE, 7*(4), e35382.

Wu, K.-H., & Tai, P. C. (2004). Cys32 and His105 are the critical residues for the calcium-dependent cysteine proteolytic activity of CvaB, an ATP-binding cassette transporter. *Journal of Biological Chemistry, 279*(2), 901-909.

Wu, T., Malinverni, J., Ruiz, N., Kim, S., Silhavy, T. J., & Kahne, D. (2005). Identification of a multicomponent complex required for outer membrane biogenesis in *Escherichia coli*. *Cell, 121*(2), 235-245.

Yakushi, T., Masuda, K., Narita, S., Matsuyama, S., & Tokuda, H. (2000). A new ABC transporter mediating the detachment of lipid-modified proteins from membranes. *Nature Cell Biology, 2*(4), 212-218.

Yasuda, M., Iguchi-Yokoyama, A., Matsuyama, S.-i., Tokuda, H., & Narita, S.-i. (2009). Membrane topology and functional importance of the periplasmic region of ABC transporter LolCDE. *Bioscience, biotechnology, and biochemistry, 73*(10), 2310-2316.

Zaitseva, J., Jenewein, S., Wiedenmann, A., Benabdelhak, H., Holland, I. B., & Schmitt, L. (2005). Functional Characterization and ATP-Induced Dimerization of the Isolated ABC-Domain of the Haemolysin B Transporter. *Biochemistry, 44*(28), 9680-9690.

Zaitseva, J., Oswald, C., Jumpertz, T., Jenewein, S., Wiedenmann, A., Holland, I. B., *et al.* (2006). A structural analysis of asymmetry required for catalytic activity of an ABC-ATPase domain dimer. *The EMBO Journal, 25*(14), 3432-3443.

Zhang, F., Greig, D. I., & Ling, V. (1993). Functional replacement of the hemolysin A transport signal by a different primary sequence. *Proceedings of the National Academy of Sciences of the United States of America, 90*(9), 4211-4215.

Zhang, F., Yin, Y., Arrowsmith, C. H., & Ling, V. (1995). Secretion and circular dichroism analysis of the C-terminal signal peptides of HlyA and LktA. *Biochemistry, 34*(13), 4193-4201.

Zhang, L. H., Fath, M. J., Mahanty, H. K., Tai, P. C., & Kolter, R. (1995). Genetic Analysis of the Colicin V Secretion Pathway. *Genetics, 141*(1), 25-32.

Zhong, X., & Tai, P. C. (1998). When an ATPase is not an ATPase: at low temperatures the C-terminal domain of the ABC transporter CvaB is a GTPase. *Journal of bacteriology, 180*(6), 1347-1353.

Zhou, B., He, Y., Zhang, X., Xu, J., Luo, Y., Wang, Y., *et al.* (2010). Targeting *Mycobacterium* protein tyrosine phosphatase B for antituberculosis agents. *Proceedings of the National Academy of Sciences, 107*(10), 4573-4578.

Zolnerciks, J. K., Andress, E. J., Nicolaou, M., & Linton, K. J. (2011). Structure of ABC transporters. *Essays In Biochemistry, 50*, 43-61.

Zuckert, W. R. (2014). Secretion of bacterial lipoproteins: through the cytoplasmic membrane, the periplasm and beyond. *Biochim Biophys Acta, 1843*(8), 1509-1516.