Characterizing a potential β-barrel assembly machinery (BAM) complex in *Treponema pallidum*

by

Michael Cummings
Bachelor of Science, University of Victoria, 2008

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

Characterizing a potential β-barrel assembly machinery (BAM) complex in *Treponema pallidum*

by

Michael Cummings
Bachelor of Science, University of Victoria, 2008

Supervisory Committee

Dr. Caroline Cameron, Department of Biochemistry and Microbiology
Supervisor

Dr. Francis Nano, Department of Biochemistry and Microbiology
Departmental Member

Dr. Fraser Hof, Department of Chemistry
Outside Member

Abstract

Supervisory Committee
Dr. Caroline Cameron, Department of Biochemistry and Microbiology
Supervisor
Dr. Francis Nano, Department of Biochemistry and Microbiology
Departmental Member
Dr. Fraser Hof, Department of Chemistry
Outside Member

Previous experimentation using differential immunological screening identified Tp0326, a protein predicted to be located in the outer membrane (OM) of the bacterium *Treponema pallidum*. This protein is homologous to BamA members of the β-barrel assembly machinery (BAM) family of proteins, which are conserved throughout pathogenic Gram-negative bacteria. In *Escherichia coli* the BAM proteins are found as a complex composed of five proteins: BamA, which is an integral membrane protein, and four accessory lipoproteins, BamB - BamE, which localize to the inner leaflet of the outer membrane. In *E. coli* BamA has been shown to mediate the insertion and assembly of proteins in the OM via interaction with the BAM complex and periplasmic chaperones (SurA, Skp, and DegP). We hypothesize that a similar OMP translocation complex exists within *T. pallidum* and that this complex is responsible for ushering *T. pallidum* OMPs to the bacterial surface. Characterization of the putative *T. pallidum* OMP transport machinery was performed by bioinformatic analyses and protein-protein interaction studies. Protein-protein interaction studies included screening a *T. pallidum* Lambda genomic expression library with recombinant *T. pallidum* protein Tp0326 and Far-Western blotting techniques. Using bioinformatic analyses we have identified putative *T. pallidum* homologues of the *E. coli* lipoproteins BamD (Tp0622) and BamB (Tp0133) as

well as putative homologues of the *E. coli* chaperone proteins Skp (Tp0327) and DegP (Tp0773). The *T. pallidum* Lambda genomic expression library screen identified the putative *E. coli* BamD homologue (Tp0622), which was originally discovered through bioinformatic analyses. The expression library screen also identified two putative *T. pallidum* OMPs (Tp0750 and Tp0751) as potential interaction partners of Tp0326. Combined bioinformatic analyses and protein-protein interaction studies provide evidence a BAM complex may exist within *T. pallidum*, and similar to *E. coli*, this complex may be involved in ushering *T. pallidum* OMPs to the bacterial surface.

Table of Contents

Supervisory Committee ii
Abstract iii
Table of Contents v
List of Tables vii
List of Figures viii
List of Abbreviations ix
Acknowledgments xii
Dedication xiii
Chapter 1: Introduction 1
1.1 Syphilis 1
1.2 *Treponema pallidum* subspecies *pallidum* 5
1.3 *Treponema pallidum* pathogenesis involving host interactions 7
1.4 Bacterial cell envelopes 9
1.4.1 *Treponema pallidum* cell envelope ultrastructure 11
1.5 Bacterial outer membrane proteins 13
1.5.1 *Treponema pallidum* outer membrane proteins 14
1.5.2 Identification of *Treponema pallidum* outer membrane proteins 15
1.6 Biogenesis of outer membrane proteins 17
1.6.1 Transportation of OMPs across the inner membrane 19
1.6.2 Transportation of OMPs across the periplasm 21
1.7 The β-barrel assembly machinery (BAM) complex 22
1.7.1 BamA structure and function in *Escherichia coli* 23
1.7.2 The BAM accessory lipoproteins 25
1.7.3 The BAM associated periplasmic chaperones 26
1.7.4 Mechanism of OMP insertion 28
1.7.5 Identification of *Treponema pallidum* protein Tp0326 32
1.8 Research hypotheses and objectives 33
1.9 Experimental approach 34
Chapter 2: Materials and methods 35
2.1 Materials 35
2.2 General methods 39
2.2.1 DNA manipulation 39
2.2.2 Small scale protein expression 43
2.2.3 Soluble protein expression and purification 44
2.2.4 Western and Far-Western blotting 47
2.3 Bioinformatics 49
2.4 LambdaZAP II assay 52
2.4.1 LambdaZAP library 52
2.4.2 Performing LambdaZAP plaque lifts 54
2.4.3 Screening the plaques 55
2.4.4 Sequencing the plaques 57
Chapter 3: Characterizing the putative *Treponema pallidum* BAM complex using bioinformatics 59

3.1 Contributions to the data 59
3.2 Introduction 59
3.3 *Treponema pallidum* protein Tp0326 (BamA) 60
3.4 *Treponema pallidum* BAM complex lipoproteins 64
3.4.1 Predicted structure of *Treponema pallidum* hypothetical protein Tp0622 66
3.4.2 Predicted structure of *Treponema pallidum* hypothetical protein Tp0133 67
3.5 *Treponema pallidum* BAM complex chaperones 67
3.5.1 Investigating *Treponema pallidum* protein Tp0327 69
3.5.2 Investigating *Treponema pallidum* protein Tp0773 70
3.5.3 Investigating *Treponema pallidum* protein Tp0862 71
3.6 Conclusions 72
Chapter 4: Characterizing the putative *Treponema pallidum* BAM complex using protein-protein interaction studies 76
4.1 Contributions to the data 76
4.2 Introduction 76
4.3 Cloning and protein expression for Tp0326-POTRA1-5 79
4.4 Lambda ZAP assay using Tp0326-POTRA1-5 81
4.4.1 Introduction 81
4.4.2 Screening the Lambda ZAP library 83
4.4.3 Lambda ZAP DNA sequencing results 86
4.5 Recombinant soluble protein expression and purification 90
4.5.1 Expression of *Treponema pallidum* protein Tp0327 90
4.5.2 Expression of *Treponema pallidum* protein Tp0773 91
4.5.3 Expression of *Treponema pallidum* protein Tp0622 92
4.6 Protein-protein interactions through Far-Western blot analysis 94
4.6.1 Introduction 94
4.6.2 Far-Western blot analysis 95
4.6.3 Analyzing protein-protein interactions 98
4.7 Conclusions 100
Chapter 5: Overview of the putative *Treponema pallidum* BAM complex 103
5.1 Connecting bioinformatic and protein-protein interaction data 104
5.1.1 Putative *Treponema pallidum* BAM associated lipoproteins 104
5.1.2 Putative *Treponema pallidum* BAM associated chaperones 107
5.1.3 Putative *Treponema pallidum* BAM interacting membrane proteins 110
5.2 Constructing the putative *Treponema pallidum* BAM complex 112
5.3 Conclusions 114
5.4 Future directions 116
Bibliography 119

List of Tables

Table 1. BLAST results for *Treponema pallidum* protein Tp0326 61
Table 2. Delineation of *Treponema pallidum* Tp0326 POTRA domains..... 63
Table 3. *Treponema pallidum* BAM lipoprotein BLASTp results 65
Table 4. *Treponema pallidum* BAM accessory chaperone BLASTp results..... 69
Table 5. Lambda ZAP DNA sequencing and NCBI BLAST results..... 87
Table 6. Information on proteins used in Far-Western blots 97

List of Figures

Figure 1. Schematic representation of the stages of untreated syphilis 3
Figure 2. Infectious syphilis case reports and rates in British Columbia..... 4
Figure 3. Electron microscopy image of *Treponema pallidum* 6
Figure 4. Schematic representation of a typical Gram-negative cell envelope..... 11
Figure 5. Outer membrane fractures of *Escherichia coli* and *Treponema pallidum* 15
Figure 6. Sec factors and translocation processes..... 20
Figure 7. Schematic representation of OMP biogenesis in *Escherichia coli*..... 31
Figure 8. Three dimensional model of Tp0326-POTRA-2..... 62
Figure 9. Schematic represenation of the predicted structure of Tp0326 64
Figure 10. Clustal-W multiple sequence alignment of Skp and Tp0327..... 70
Figure 11. Gel filtration chromatograph for Tp0326-POTRA1-5 80
Figure 12. SDS-PAGE analysis of Tp0326-POTRA1-5..... 80
Figure 13. Map of the Lambda ZAP insertion vector..... 82
Figure 14. Lambda ZAP pBluescript SK(-) phagemid map 82
Figure 15. AP developed Lambda ZAP nitrocellulose membrane 85
Figure 16. SDS-PAGE analysis of recombinant Tp0327 91
Figure 17. SDS-PAGE analysis of recombinant Tp0622 94
Figure 18. Far-Western blot analysis of Tp0326-POTRA1-5 97
Figure 19. Level of protein binding from Far-Western blot assay 99
Figure 20. Schematic representation of the Treponema pallidum BAM complex 114

List of Abbreviations

~	approximately
1º	primary
2º	secondary
º C	degrees Celsius
A	amps
a.a	amino acid
Ab	antibody
Amp	ampicillin
AP	alkaline phosphatase
BAM	β-barrel assembly machinery
BCIP	5-Bromo-4-chloro-3-indolyl phosphate
BLAST	Basic Local Alignment Search Tool
bp	base pairs
BSA	bovine serum albumin
CET	cryo-electron tomography
Cys, C	cysteine
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
DTT	dithiothreitol
ECM	extracellular matrix
EDTA	ethylenediaminetetraacetic acid
EGF	epidermal growth factor
EM	electron microscopy
EtBr	ethidium bromide
EXPASY	Expert Protein Analysis System
FPLC	fast protein liquid chromatography
FTA-ABS	fluorescent treponemal antibody absorption
GI	geninfo identifier number
GIB	Genome Information Broker
GTOP	Genomes to Protein structures and functions
6X His tag	hexahistidine tag
HIV	Human Immunodeficiency Virus
HMM	hidden markov model
IPTG	isopropyl-beta-D-thiogalactopyranoside
kan	kanamycin
kDa	kilodalton
KEGG	Kyoto Encyclopedia of Genes and Genomes
L	liter
LB	Luria-Bertani medium
LPS	lipopolysaccharide
M	Molar
mA	milliamps

MBGD	Microbial Genome Database
mg	milligram
Mg^{2+}	magnesium
min	minutes
ml	milliliter
mM	millimolar
MSCRAMM	microbial surface components recognizing adhesive matrix molecules
MW	molecular weight
MWCO	molecular weight cut-off
NBT	nitro blue tetrazolium chloride
NCBI	National Center for Biotechnology Information
NMR	nuclear magnetic resonance
OD_{600}	optical density at 600nm
OM	outer membrane
OMP	outer membrane protein
O/N	over night
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDB	protein data base
Pfu	Pfu DNA polymerase enzyme
pfu	plaque forming unit
pg	picogram
pmol	picomole
POTRA	polypeptide transport associated
PPI	protein-protein interaction
PPIase	peptidy-prolyl *cis/trans* isomerase
PVDF	polyvinylidene fluoride
rpm	revolutions per minute
RT	room temperature
SAXS	small angle X-ray scattering
sdH_2O	sterile distilled water
SDS	sodium dodecyl sulphate
sec	second
sp.	species
TBS	tris buffered saline
TBS-T	tris buffered saline with tween
TEMED	tetramethylethylenediamine
tet	tetracycline
TIGR	The Institute for Genomic Research
TPHA	*Treponema pallidum* hemaglutination
TPR	tetratricopeptide repeat
µg	microgram
µl	microliter

V	volts
VDRL	Venereal Disease Research Laboratory
v/v	volume to volume ratio
w/v	weight to volume ratio
WHO	World Health Organization
WT	wild type

Acknowledgments

I would like to start off by giving a huge thanks to my supervisor, Dr. Caroline Cameron, for her support and confidence in me throughout the past three years. Dr. Cameron has been a great mentor for me and I could not have asked for a better supervisor. I would also like to thank my committee members, Dr. Francis Nano and Dr. Fraser Hof for their advice and discussion surrounding my thesis. My co-workers have been invaluable towards my work in the lab, specifically Rebecca Hof, Teresa Brooks, and Charmaine Wetherell. Last but certainly not least, I would like to thank my fellow Cameron lab graduate students, Azad Eshghi, and Brenden Smith, and Cameron lab post-doctoral fellow, Dr. Simon Houston for the countless discussions about our projects.

Dedication

Man's journey through science is both manic and enigmatic. I have found the trick to becoming a successful scientist is not strictly through success in the lab, but also through success as a person. Achieving balance can yield extraordinary results both in the lab and in life. For that reason I owe a great deal of gratitude to my friends and family. I could have never maintained such balance without with the continued love and support from those around me. There is nothing that compares to coming home, after a mentally exhausting day in the lab, to my loving fiancée Amanda and my pup Baxter. To all of my closest friends and family, I dedicate this thesis.

Chapter 1: Introduction

1.1 Syphilis

Syphilis is a sexually and congenitally transmitted disease caused by infection with the bacterium *Treponema pallidum* subsp. *pallidum.* The route of transmission of the bacterium is usually either through direct sexual contact, from mother to fetus *in utero*, or in more rare circumstances from either direct physical contact or via blood transfusion (Pickering, 2006). Syphilis is a disease that can establish a lifelong chronic infection in the absence of appropriate antibiotic treatment (Stebeck *et al.*, 1997).

Syphilis is a multistage disease, with localized, disseminated, and chronic phases of infection (Cullen & Cameron, 2006). The disease can fluctuate between symptomatic stages and prolonged asymptomatic stages due to the ability of *T. pallidum* to evade the host immune response and remain latent for extended periods of time (Cameron, 2006). Syphilis can produce a systemic infection which, if left untreated, can cause serious damage to any organ system, including the heart, aorta, eyes, brain, and bones (Cameron, 2003).

Different manifestations of the disease occur depending on the stage of the infection. A schematic diagram of the stages of untreated syphilis can be seen in Figure 1. *Treponema pallidum* is typically acquired through direct sexual contact with lesions of an infected individual (Pickering, 2006), however, it can also be transmitted from mother to fetus *in utero*, or via blood transfusions. Transmission of the bacterium can lead to the development of primay stage syphilis. During primary stage syphilis a single primary lesion, or chancre, typically forms at the site of contact and heals within 4-6 weeks; the

individual otherwise remains asymptomatic (British Columbia Center for Disease Control website, Syphilis overview). Secondary syphilis normally occurs 2 months after the primary infection and results in a body rash. In most areas of the body the rash can develop into flat, broad, whitish lesions, which are all infectious (World Health Organization website, Disease watch: Syphilis), however, most individuals only acquire the rash on their trunk, the soles of their feet, and the palms of their hands. Other symptoms common at this stage include fever, sore throat, malaise, weight loss, headache, and enlarged lymph nodes (British Columbia Center for Disease Control website, Syphilis overview). Latent syphilis is defined as having serologic proof of infection without signs or symptoms of disease (World Health Organization website, Disease watch: Syphilis). If left untreated, latent syphilis can potentially develop into tertiary syphilis, leading to neuro- and cardiovascular syphilis, which can be fatal (Pickering, 2006). Syphilis currently lacks a vaccine and penicillin, in the form of penicillin G, is currently the suggested treatment for all manifestations of the disease (Workowski & Berman, 2006).

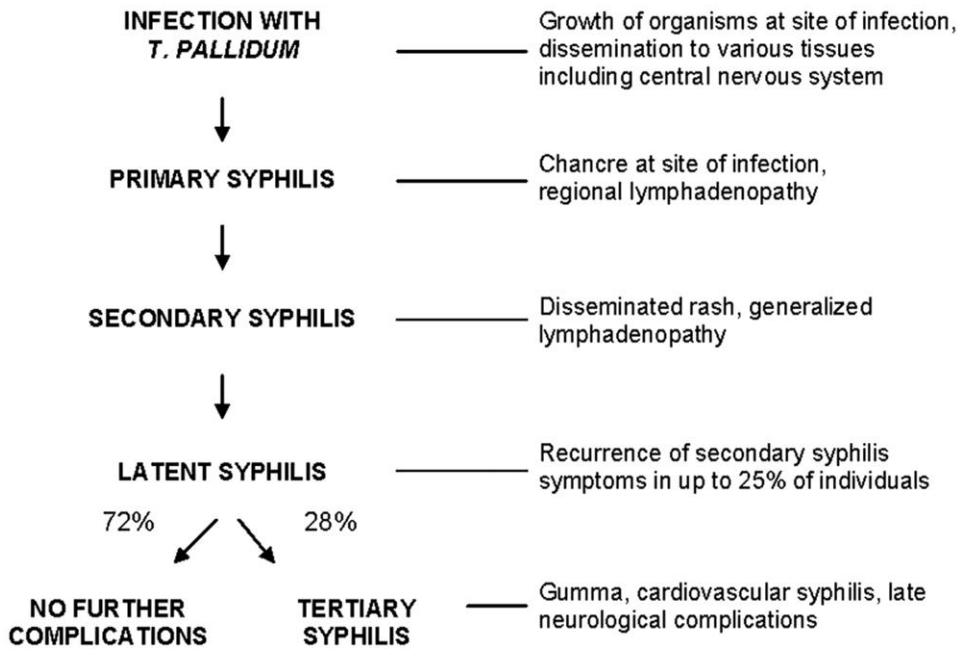
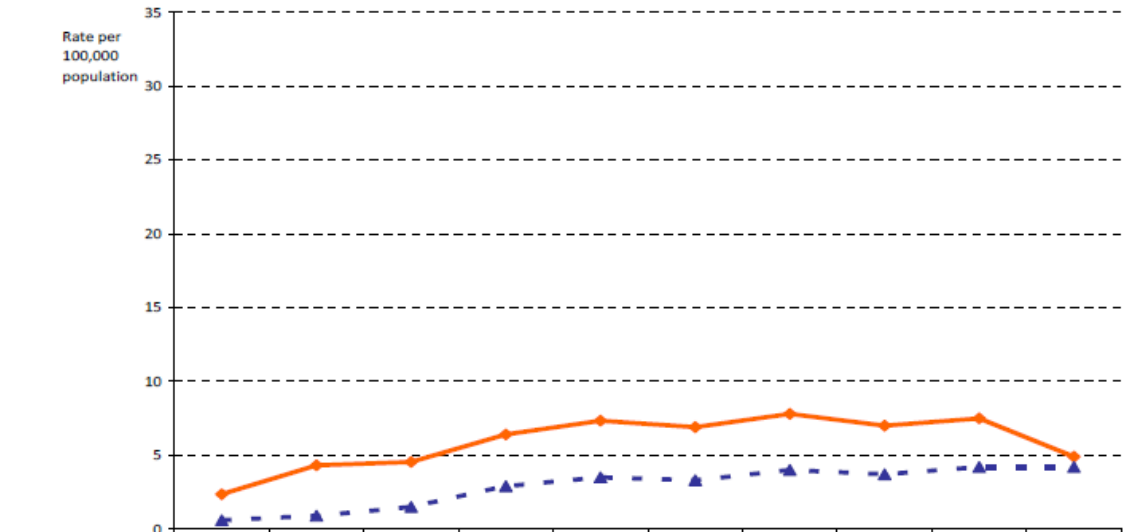


Figure 1. Schematic representation of the stages of untreated syphilis
Natural history of untreated syphilis, according to Gjestland (Gjestland, 1955).

Despite the availability of antibiotic treatment syphilis still remains a public health concern worldwide. The World Health Organization estimated 12 million new cases worldwide in 2001, with more than 90% occurring in developing countries (World Health Organization website, Disease watch: Syphilis). The number of cases in the United States and Eastern Europe has steadily been increasing, with a 12.4% increase in cases between the years 2001 and 2002 in the United States (World Health Organization, 2003). The number of cases in Canada has also been steadily increasing, with the number of cases nearly doubling from the year 2002 to 2003 (Public Health Agency of Canada website, Syphilis 2008). Within Canada, British Columbia lead the country with the highest infectious syphilis rate in 2008 at 7.4 per 100 thousand people (British Columbia Center for Disease Control, 2009). The rate of infectious syphilis in British Columbia has since dropped to 4.9 cases per 100 thousand people in 2009, which is close

to the Canadian rate (Figure 2) (British Columbia Center for Disease Control, 2009). Congenital syphilis is of particular concern in developing nations, where the lack of prenatal testing and antibiotic treatment of infected pregnant women results in congenital infection of the fetus (Lafond & Lukehart, 2006). Congenital syphilis causes spontaneous abortion, stillbirth, death of the neonate, or disease in the infant (Lafond & Lukehart, 2006). Further concern for public health is the fact that syphilis infection leads to an increased risk of transmission, and acquisition, of the human immunodeficiency virus (Nusbaum *et al.*, 2004). There is currently no preventative vaccine for syphilis, highlighting the need for a greater understanding of the mechanisms of *T. pallidum* pathogenesis (Cullen & Cameron, 2006).



	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009*
BC Reports	95	176	186	262	305	288	331	300	328	216
BC Rate	2.4	4.3	4.5	6.4	7.3	6.9	7.8	7.0	7.5	4.9
Canadian Rate	0.6	0.9	1.5	2.9	3.5	3.3	4.0	3.7	4.2	4.2

*2009 Canadian rate is projected and is subject to change (Public Health Agency of Canada, 2010).

Figure 2. Infectious syphilis case reports and rates in British Columbia
Infectious syphilis case reports and rates in British Columbia, Canada from 2000 – 2009 (British Columbia Center for Disease Control, 2009).

1.2 *Treponema pallidum* subspecies *pallidum*

Treponema pallidum are spirochete bacteria belonging to the family Spirochaetaceae, with subspecies that cause treponemal diseases such as syphilis, bejel, pinta, and yaws (Antal *et al*., 2002). *Treponema pallidum* subspecies *pallidum* is the causative agent of syphilis, the only subspecies which causes a venereal disease. These bacteria can not be visualized by a traditional Gram stain because they are too thin. They can, however, be viewed with a special stain called Dieterle stain which is used to visualize *Mycobacterium tuberculosis* (Brady, 1998), and they can also be visualized by dark-field microscopy, which is the common method for visualization. *Treponema pallidum* is also detected by the non-specific VDRL and Rapid plasma reagin (RPR) tests, as well as treponemal antibody tests (FTA-ABS, *T. pallidum* immobilization reaction (TPI) and Syphilis TPHA test).

Treponema pallidum subspecies *pallidum* (hereafter referred to as *Treponema pallidum*) are helical or spiral in shape, and are approximately 6-15 microns long and 0.2 microns wide (Lafond & Lukehart, 2006). These bacteria are very thin; in comparison typical *Escherichia coli* cells are 0.5 microns wide (Kubitschek, 1990). *Treponema pallidum* are highly motile, the helical structure allows them to move in a corkscrew like motion helping them penetrate though viscous mucous and tissue membranes within the host (Lafond & Lukehart, 2006). An electron microscopy slide depicting a single *T. pallidum* organism can be seen in Figure 3.

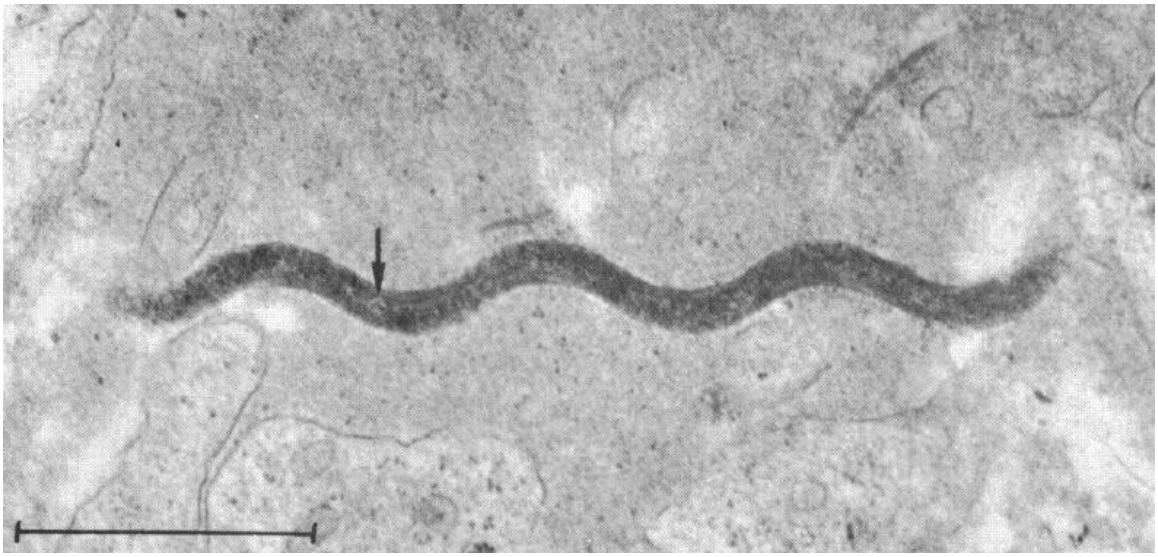


Figure 3. Electron microscopy image of *Treponema pallidum*
Treponema pallidum located in amorphous extracellular ground substance, but proximal to areas of collagen. The magnification marker represents 1µm (Drusin, 1969).

The complete genome of *T. pallidum* has been sequenced and was confirmed by the Genome Sequencing Project to be 1.14 Mb and to encode 1,041 putative proteins (Fraser *et al.*, 1998). There are a total of 1041 predicted ORFs which comprise 93% of the total genomic DNA (Fraser *et al*., 1998). The genome of *T. pallidum* is extremely small in comparison to conventional gram-negative (*E. coli* K-12, 4.6 Mb) and gram-positive (*Bacillus subtilis*, 4.2 Mb) bacteria and has been shown to lack metabolic capabilities (Lafond & Lukehart, 2006). *Treponema pallidum* lacks tricarboxylic acid cycle enzymes and lacks an electron transport chain (Fraser *et al*., 1998). Amino acid and fatty acid synthesis pathways are also lacking, but *T. pallidum* does carry enzymes for the interconversion of amino acids and fatty acids (Fraser *et al*., 1998). With limited metabolic machinery it is believed that *T. pallidum* derives most essential macromolecules from the host mammalian tissues (Lafond & Lukehart, 2006).

Research surrounding the molecular mechanisms of *T. pallidum* virulence has been troubled by the fact that the organism can not be continuously cultured *in vitro*. *Treponema pallidum* does not survive outside the mammalian host; infectious capability is lost within a few hours or days of harvest (Lafond & Lukehart, 2006). This could, in part, be due to the fact that *T. pallidum* is sensitive to oxygen (Lafond & Lukehart, 2006). Researchers have been unable to propagate *T. pallidum* in tissue culture more than 100-fold, an equivalent of about seven generations (Fieldsteel *et al*., 1981). Difficulties surrounding *in vitro* culturing are also hampered by the extremely slow generation time of *T. pallidum in vitro* which is approximately 30 to 50 hours (Fieldsteel *et al*., 1981). Because of the difficulties surrounding *in vitro* culturing, in order to obtain sufficient organisms for experimental manipulation, *T. pallidum* must be propagated in rabbits (Turner & Hollander, 1957). The generation time of *T. pallidum* in rabbits is also extremely slow with doubling time of 30 to 33 h *in vivo* (Cumberland & Turner, 1949). The inability of *T. pallidum* to survive and multiply outside the mammalian host is believed to be the largest obstacle surrounding syphilis research (Lafond & Lukehart, 2006).

1.3 *Treponema pallidum* pathogenesis involving host interactions

One of the initial steps in establishing an infection by a pathogen, such as *T. pallidum*, is attaching to host components, such as basement membranes of epithelial and endothelial cell layers. All stages of syphilis involve interaction with the host vascular system, specifically, perivascular areas in infected tissues (Lukehart *et al*., 1980). *Treponema pallidum* has been shown to interact specifically with vascular endothelial cell layers (Lee *et al.*, 2003), and can move through various cell junctions which aids in

dissemination (Fitzgerald, 1983). The basement membrane of vascular endothelial cell layers is one site of attachment for *T. pallidum* (Fitzgerald *et al*., 1984), and crucial for its dissemination and sustained infection. Attachment to host cells is critical for bacterial pathogenesis. Bacterial outer membrane proteins are potentially involved in the attachment process, and therefore, remain a key area of research in bacterial pathogenesis. Understanding the attachment process of the organism is required not only for a better understanding of its pathogenesis, but also in order to create successful vaccines.

Treponema pallidum initially gains entry to the host through intact mucosal barriers or microscopic epidermal abrasions (Pike, 1976), via sexual contact with another infected individual. Due to the pathogen's limited toxigenic properties, it is believed to rely on a strong host inflammatory response to cause massive tissue destruction (Lafond & Lukehart, 2006). Recent research has shown, however, that *T. pallidum* may possess a protein on its outer surface involved in degrading host ECM components (Houston *et al*., unpublished findings). Tissue destruction, as well as the pathogens ability to penetrate intact membranes and cell monolayers (Riviere *et al*., 1989), allows it to enter and disseminate through the blood stream and various tissues, resulting in a widespread bacterial infection. The highly motile nature of *T. pallidum* also aids in dissemination.

The outer membrane of *T. pallidum* has very little antigenic reactivity (Radolf *et al*., 1989), allowing it to go relatively undetected by the host's acquired immune response. The poorly antigenic nature of the pathogens outer membrane increases invasiveness and has earned *T. pallidum* the designation of "stealth" pathogen (Salazar *et al*., 2002).

Adherence to, and colonization of, host epithelial surfaces are strategies employed by numerous pathogens in order to initiate infection. One manner in which this can occur is through the interaction of bacterial surface adhesins with host extracellular matrix (ECM) molecules. The outer membrane proteins of pathogens which recognize host ECM components have been termed MSCRAMMs (microbial surface components recognizing adhesive matrix molecules)(Patti *et al.*, 1994). Host ECM components which interact with MSCRAMMs include fibronectin, fibrinogen, collagens, laminins, vitronectin, and heparan sulfate (Patti *et al*., 1994). Adherence to, and colonization of, host epithelial surfaces are indeed the primary events in the pathogenesis of *T. pallidum* (Beachey, 1981). Host cell attachment is also believed to be a critical event in the ability of *T. pallidum* to invade and disseminate through the bloodstream. Studies have shown the specific interaction of *T. pallidum* with host ECM components fibronectin and laminin (Cameron, 2003; Fitzgerald *et al*., 1984). A specific laminin-binding adhesin has been identified in *T. pallidum* (Tp0751) through bioinformatic analyses performed on the *T. pallidum* genome and subsequent ECM attachment assays (Cameron, 2003). The *T. pallidum* adhesin Tp0751 is expressed during infection and exhibits a strong affinity for laminin (Cameron, 2003). It is believed that the ability of *T. pallidum* to attach to the host component laminin possesses important implications for bacterial dissemination (Cameron *et al.*, 2005).

1.4 Bacterial cell envelopes

Bacteria are commonly grouped based on the chemical composition of their cell envelope. There are two main types of bacterial cell envelopes, Gram-positive and Gram-negative, which are differentiated by their Gram staining characteristics. As

previously stated, *T. pallidum* can not be classified in this traditional way because its cell envelope is too thin to be properly Gram stained. Research has shown that the cell envelope ultrastructure of *T. pallidum* differs significantly from that of both traditional Gram-positive and Gram-negative bacteria.

Peptidoglycans (mucopeptides, glycopeptides, and mureins) are the structural elements of almost all bacterial cell envelopes. They constitute almost 95% of the cell envelope in some Gram-positive bacteria and as little as 5-10% of the cell envelope in Gram-negative bacteria (Wilson, 2002). The cell envelope of Gram-positive bacteria consists of a thick layer of peptidoglycan with small amounts of teichoic acid dispersed which is separated from the cell membrane by a small periplasmic space (Wilson, 2002).

Unlike the Gram-positive cell envelope, the Gram-negative cell envelope contains a thin peptidoglycan layer which resides in a space between the cytoplasmic membrane and an additional outer membrane (Wilson, 2002). The outer membrane of the Gram-negative cell envelope is composed of phospholipids and lipopolysaccharides, which faces into the external environment (Wilson, 2002). A schematic representation of a typical Gram-negative cell envelope can be seen in Figure 4. As the lipopolysaccharides are highly-charged, the Gram-negative cell envelope has an overall negative charge. The chemical structure of the outer membrane lipopolysaccharides is often unique to specific bacterial strains (i.e. sub-species) and is responsible for many of the antigenic properties of these strains (Wilson, 2002).

The outer membrane functions as a permeability barrier protecting the bacterium from harmful compounds, such as antibiotics and bile salts (Tommassen, 2010). Most nutrients pass this barrier via a family of integral outer-membrane proteins (OMPs),

collectively called porins (Tommassen, 2010). Other OMPs have more specialized transport functions, such as the secretion of proteins and the extrusion of drugs, or function as enzymes or structural components of the outer membrane (Koebnik *et al.*, 2000). Besides integral OMPs, the membrane also contains lipoproteins, which are attached to the membrane via an N-terminal lipid moiety (Tommassen, 2010).

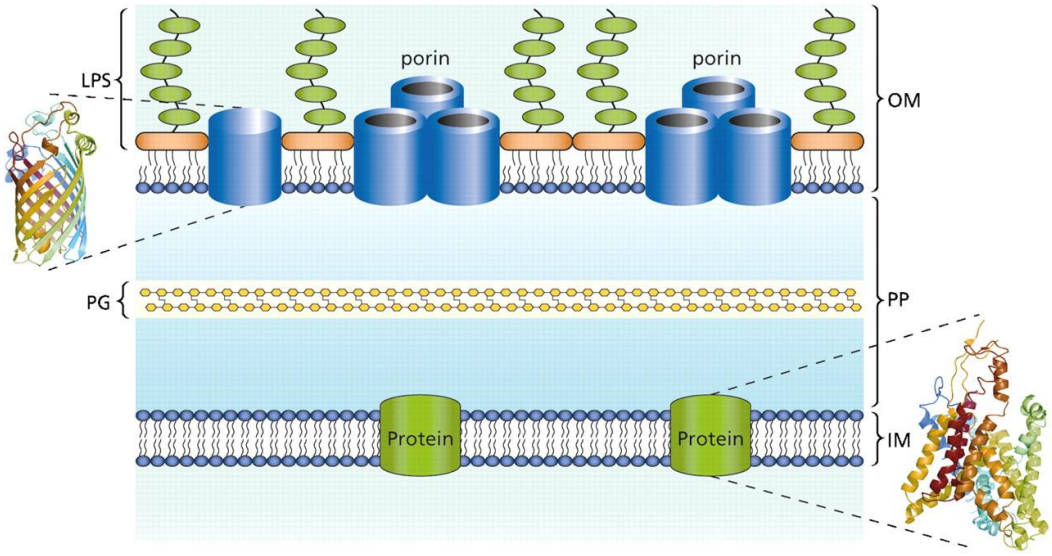


Figure 4. Schematic representation of a typical Gram-negative cell envelope
Structure of the Gram-negative cell envelope. Outer membrane (OM) containing LPS in the outer leaflet of the bilayer and porins as the major protein components; periplasm (PP) containing the peptidoglycan layer (PG); inner membrane (IM) (Tommassen, 2010). Examples of a typical β-barrel structure of an OMP (PDB file 3FID) (Rutten *et al.*, 2009), and of a typical α-helical inner-membrane protein (PDB file 2ZQP) (Tsukazaki *et al.*, 2008), are shown on the left and the right, respectively.

1.4.1 *Treponema pallidum* cell envelope ultrastructure

The cell envelope of *T. pallidum* differs greatly from both Gram-positive and Gram-negative bacteria, but it most closely resembles that of Gram-negative and is thus referred to as 'Gram-negative like'.

Treponema pallidum morphologically consists of an outer membrane that surrounds the periplasmic endoflagella, the cytoplasmic membrane, and the protoplasmic cylinder (Holt, 1978). The most significant difference between the outer membrane of *T. pallidum* and that of a traditional Gram-negative bacterium is that it does not contain lipopolysaccharide (Radolf & Norgard, 1988).

In contrast to traditional Gram-negative bacteria, it was originally believed that the thin peptidoglycan layer in *T. pallidum* directly overlied the cytoplasmic membrane (Holt, 1978). This created a large periplasmic space which housed the internal flagella responsible for the characteristic corkscrew motility (Jepsen *et al.*, 1968). It has recently been shown by cryo-electron tomography (CET) that the peptidoglycan layer in *T. pallidum* divides the periplasmic space into two distinct regions; above the cytoplasmic membrane and below the outer membrane (Izard *et al*., 2009). The region between the peptidoglycan and the outer membrane contains the flagellar particles (Izard *et al*., 2009), soluble polypeptides (chaperones and antioxidant enzymes)(Mulay *et al*., 2007; Shevchenko *et al*., 1997), and the protein moieties of a presumably small number of lipoproteins anchored to the inner leaflet of the OM (Hazlett *et al*., 2005).

A major difference in the cell envelope ultrastructure between *T. pallidum* and traditional Gram-negative bacteria is that in Gram-negative bacteria the peptidoglycan is tightly linked to the outer membrane. Covalent linkage occurs via Braun's lipoprotein and noncovalent linkage occurs via interactions with lipoproteins, such as peptidoglycan-associated lipoprotein, and numerous membrane-spanning proteins, most notably, porins and OmpA (De Mot & Vanderleyden, 1994; Nikaido, 1996; Parsons *et al*., 2006). In

contrast there has been no demonstrated biochemical linkage between the peptidoglycan and the outer membrane in *T. pallidum*.

Another notable difference between *T. pallidum* and traditional Gram-negative bacteria is that research using freeze-fracture and freeze-etch electron microscopy has shown that the outer membrane of *T. pallidum* contains a very small number of transmembrane proteins in comparison to traditional Gram-negative bacteria, roughly $1/100^{th}$ the amount than that of *E. coli* (Radolf *et al*., 1989).

Because of the absence of lipopolysaccharide, and the lack of stabilization between the peptidoglycan and the outer membrane, the outer membrane of *T. pallidum* is much more fragile than that of traditional Gram-negative bacteria. The outer membrane of *T. pallidum* can easily be disrupted by low concentrations of detergents and by physical manipulations such as centrifugation, resuspension, and washing (Penn *et al.*, 1985; Radolf & Norgard, 1988). Because of the fragility of its outer membrane and the fact that it can not be cultured *in vitro*, *T. pallidum* is genetically intractable (Lafond & Lukehart, 2006). *Treponema pallidum* can not be genetically manipulated by traditional experimental methods which use recombinant DNA. This is a major hinderance surrounding research of this pathogen. Heterologous expression in related organisms such as *Treponema denticola* (Chi *et al.*, 1999) and more recently *Treponema phagedenis* (Cameron *et al*., 2008) may be the most practical way to study *T. pallidum* genes and advance our understanding of this elusive pathogen.

1.5 Bacterial outer membrane proteins

Outer membrane proteins (OMPs) are a class of proteins resident at the outer membrane of Gram-negative and Gram-negative like bacteria cells; they are either

attached to the outer or the inner leaflet of the outer membrane. Outer membrane proteins often reside on the surface of bacteria and thus are often involved in initial bacterium/host interactions and are often crucial to the pathogenesis of the organism. For example the pathogenesis of *Helicobacter pylori* has been shown to be greater in strains which possess the OMP VacA compared to strains that do not (Keenan *et al*., 2000). As well, the OMP OmpA in *E. coli* has been shown to be critical for its pathogenesis (Weiser & Gotschlich, 1991).

Outer membrane proteins are often of medical importance; because they are exposed at the bacterial surface they often represent vaccine candidates. Recombinant OMPs from *Pseudomonas aeruginosa* have proven to work as successful vaccines against sepsis in humans (von Specht *et al*., 1996). As well, nasal immunization with *Burkholderia multivorans* OMPs has proven to give protection against subsequent *B. multivorans* lung infections (Bertot *et al*., 2007).

1.5.1 *Treponema pallidum* outer membrane proteins

Bacterial OMPs reside on the bacterial surface and are often the targets of host adaptive immunity. Early researchers noted that antibodies in serum from infected animals did not readily bind to intact treponemes (Deacon *et al*., 1957). This suggests that there are few antigenic targets on the surface of the organisms. Freeze fracture EM studies by Radolf *et al*. and Walker *et al*. confirmed the paucity of integral OMPs in *T. pallidum*, a characteristic that may help the organism escape immune detection and that has inspired researchers to call *T. pallidum* "the stealth pathogen" (Salazar *et al*., 2002). The rare *T. pallidum* OMPs are likely very important in interactions with the host and the host immune system and as such could likely constitute an effective syphilis vaccine. For

these reasons, their identities have been the subject of intense research over the last two decades (Lafond & Lukehart, 2006).

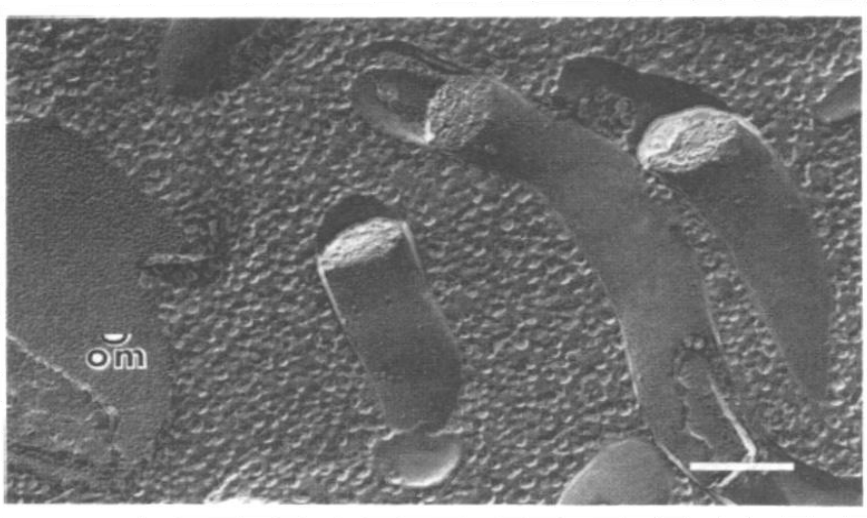


Figure 5. Outer membrane fractures of *Escherichia coli* and *Treponema pallidum*
The concave outer membrane (om) of *E. coli* shows a uniformly dense distribution of intramembranous particles in sharp contrast to the scarce particles in the fracture faces of the *T. pallidum* outer membrane, Bar = 0.2 µm (Radolf *et al*., 1989).

1.5.2 Identification of *Treponema pallidum* outer membrane proteins

Many techniques have been used in attempt to explore and identify the rare *T. pallidum* outer membrane proteins. Past studies have used phase partitioning with various detergents (Cunningham *et al.*, 1988; Penn *et al*., 1985), separation of membranes with acid (Stamm & Bassford, 1985), or density gradient ultracentrifugation of organisms lysed in a hypotonic solution (Alderete & Baseman, 1980). These methods revealed a number of proteins that were initially believed to be surface-exposed proteins, however, further studies indicated that these proteins were not surface exposed but were more likely to be anchored in the inner membrane with portions extending into the periplasm (Hsu *et al*., 1989). It was then later discovered by Cox *et al.* that physical manipulations

such as centrifugation and washing, or treatment with detergents, damaged the *T. pallidum* outer membrane (Cox *et al*., 1992).

Because of the unusual ultrastructure of *T. pallidum*, the fragility of the outer membrane and the inability to cultivate the organism *in vitro*, conventional techniques used to identify OMPs are irrelevant when it comes to *T. pallidum* OMP identification. Researchers then began to use molecular methods utilizing *E. coli* to identify *T. pallidum* OMPs (Blanco *et al*., 1991). Blanco *et al*. created a *T. pallidum* genomic expression library of *T. pallidum*-alkaline phosphatase (AP) fusion proteins. Triton X-114 detergent phase partitioning in *E. coli* of individual *T. pallidum*-AP fusions revealed several clones whose AP activity partitioned preferentially into the hydrophobic detergent phase (Blanco *et al*., 1991). These clones were identified to possess cleavable N-terminal signal sequences and were predicted to be OMPs (Blanco *et al*., 1991). The concern with these types of molecular methods is that they provide only indirect evidence for surface exposure. Because of the inherent differences between *E. coli* and *T. pallidum*, one has to exercise a certain level of caution when translating *E. coli* expression data to biological meaning in *T. pallidum*.

Since the genome of *T. pallidum* was sequenced in 1998, researchers have been implementing bioinformatic analysis to identify proteins that may be exposed on the surface of the organism. Several proteins that are predicted to have a cleavable signal sequence, transmembrane domains, and other characteristics of proteins that span the outer membrane have been identified (Cameron, 2003; Centurion-Lara *et al*., 1999). Three of these identified proteins, Tp0155, Tp0483, and Tp0751, have been shown to bind to ECM components and are candidate host-binding molecules (Cameron, 2003;

Cameron *et al*., 2004). More recently another putative *T. pallidum* OMP which binds host ECM components has been identified, Tp0136, with its role as binding to fibronectin (Brinkman *et al.*, 2008). It appears that *T. pallidum* possesses multiple ECM binding adhesins, which supports the belief that host cell attachment is important for the organism's pathogenesis. Another *T. pallidum* protein, Tp0897 (TprK), is believed to be located in the outer membrane, as it has been shown to be preferentially expressed during infection, and is a target of opsonic antibodies (Centurion-Lara *et al*., 1999). As well, TprK has recently been shown to undergo antigenic variation, a common strategy employed by bacterial pathogens to escape the host adaptive immune response (Giacani *et al.*, 2010).

These findings do suggest that *T. pallidum* possesses OMPs that are important for its pathogenesis. However, because of the indirect research methods needed to study *T. pallidum* OMPs, no OMP has been definitely identified in *T. pallidum* to date.

1.6 Biogenesis of outer membrane proteins

In Gram-negative bacteria every component of the outer membrane is synthesized in the cytoplasm or at the inner leaflet of the inner membrane (Tommassen, 2010). Exactly how these components are transported and assembled into the outer membrane remains a hot area of research. Understanding the process of how OMPs are synthesized and transported into the outer membrane would be invaluable in furthering understanding of the pathogenesis of the organism and could potentially lead to vaccine development. Model organisms such as *E. coli* and *Neisseria meningitidis* have been the focus of study in the aim of understanding OMP biogenesis in Gram-negative bacteria.

Bacteria have a number of different pathways available for the export of proteins to their final destination depending on the chemical nature of the protein. As *T. pallidum* lacks LPS, its outer membrane is much simpler than that of typical Gram-negative bacteria. Only the transporting of phospholipids, β-barrel proteins, and lipoproteins occurs in *T. pallidum*. The exact composition of the outer membrane in *T. pallidum* still remains unknown and it is unclear whether it contains lipoproteins.

Lipoproteins are involved in various biological activities in the cell envelope. Lipoproteins have been shown to be involved in outer membrane sorting of β-barrel proteins (Ruiz *et al*., 2006), and lipoproteins (Matsuyama *et al*., 1997), and are often essential to the organism (Tokuda *et al*., 2007). Lipoproteins, in Gram-negative bacteria, are localized on the periplasmic side of the inner or outer membrane, or on the outer leaflet of the outer membrane (Tokuda, 2009). They are anchored to the inner or outer membrane through acyl chains attached to an N-terminal cysteine residue (Sankaran & Wu, 1994). Outer membrane β-barrel proteins (OMPs) are often associated with basic physiological functions, virulence, and drug resistance, and therefore play a fundamental part in the maintenance of cellular viability (Bos *et al*., 2007). Bacterial OMPs span the outer membrane by forming a β-barrel structure with amphipathic β-strands, which possess alternating hydrophobic residues (Tokuda, 2009).

Bacteria utilize the general secretory pathway (GSP) to transport OMPs across the inner membrane and several different pathways for transportation across the periplasm and to the outer membrane. Outer membrane β-barrel proteins are transported through the periplasm to the outer membrane by the BAM chaperone-usher pathway, type II, or type V secretion systems (Wilson, 2002), whereas lipoproteins are transported

to the outer membrane by the Lol system (Matsuyama *et al.*, 1997). Certain pathogenic Gram-negative bacteria have also developed various protein secretion systems which transport proteins directly from the cytosol into the extracellular milieu or host cells; these include the type I, type III, type IV, and type VI secretion systems.

1.6.1 Transportation of OMPs across the inner membrane

Bacterial β-barrel OMPs are synthesized in the cytoplasm as precursors with an N-terminal signal sequence, or leader peptide, which marks them for transport across the inner membrane via the Sec dependent pathway (Papanikou *et al.*, 2007). The signal sequence is cleaved by a signal peptidase enzyme during translocation across the membrane. The protein-conducting channel of the Sec system is composed of the integral membrane proteins SecY, SecE, and SecG (Driessen & Nouwen, 2008). The factors that comprise the Sec complex and the translocation process can be seen in Figure 7 (Mori & Ito, 2001). *Treponema pallidum* possesses homologues to each component of the Sec pathway with the exception of the non-essential SecB protein (Fraser *et al.*, 1998).

The Sec pathway is also utilized in the assembly of integral inner-membrane proteins (Tommassen, 2010). When large hydrophobic protein segments are inserted into the Sec translocon, the channel opens laterally allowing for the insertion of these proteins into the inner membrane (Driessen & Nouwen, 2008). The presence of large hydrophobic segments in OMPs would prevent them from reaching the outer membrane, while the amphipathic β-strands that constitute the transmembrane segments of OMPs are compatible with transport via the Sec pathway to the periplasm (Tommassen, 2010).

Bacterial lipoproteins are also synthesized in the cytosol, however, in contrast to β-barrel OMPs, their precursors possess a consensus lipobox sequence around the signal peptide cleavage site (Hayashi & Wu, 1990). Lipoproteins are translocated across the inner membrane by the Sec translocon, however, processing of the protein to its mature form is catalyzed by three well-conserved enzymes, Lgt (phosphatidylglycerol:prolipoprotein diacylglyceryl transferase), LspA (prolipoprotein signal peptidase), and Lnt (phospholipid:apolipoprotein transacylase)(Tokuda, 2009). The three lipoprotein processing enzymes are conserved in Gram-negative bacteria.

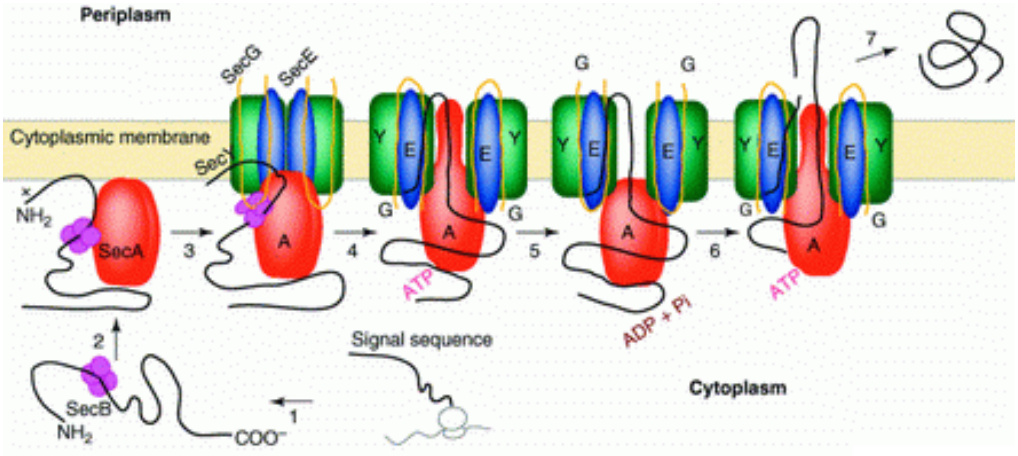


Figure 6. Sec factors and translocation processes
The preprotein is represented by a black line. Steps 1–3, targeting. A signal sequence is recognized by the Sec machinery. SecB, the Sec-system-specific chaperone, channels the preprotein to the Sec translocation pathway and, additionally, actively targets the bound precursor to the translocase by its ability to bind SecA. The preprotein-bearing SecA then binds to the membrane, at a high-affinity SecA-binding site. SecY, SecE and SecG form a hetero-trimeric complex, SecYEG, which constitutes a channel for polypeptide movement. Steps 4 and 5, initiation. The initiation step requires ATP but not its hydrolysis. Step 6, continuation. Continued translocation requires cycles of ATP hydrolysis and/or proton-motive force across the membrane. Translocation is thought to occur in a step-wise fashion with a step of 20–30 amino acid residues. Step 7, completion. As yet, little is known about the completion process, which occurs on the periplasmic side, leading to the release and/or folding of the substrate protein into the periplasmic space (Mori & Ito, 2001).

1.6.2 Transportation of OMPs across the periplasm

The vast majority of bacterial outer membrane β-barrel proteins are transported to the outer membrane by means of the BAM chaperone-usher pathway (Knowles *et al.*, 2009). The model Gram-negative organisms *E. coli* and *N. meningitidis* have been the focus of study for this process, and most research in this area has been done in these two organisms.

Upon cleavage of the OMP signal peptide by the signal peptidase in the Sec translocon, the nascent OMP associates with periplasmic chaperones (Knowles *et al*., 2009). These chaperones transport the nascent OMPs across the periplasmic space to the outer membrane (Sklar *et al*., 2007) where they then interact with a core complex at the outer membrane known as the β-barrel assembly machinery (BAM) complex. The BAM complex facilitates proper OMP folding and translocation into the outer membrane (Wu *et al*., 2005).

In contrast to β-barrel OMPs, lipoproteins are transported to the outer membrane by the Lol system. The Lol system is composed of an ABC transporter LolCDE complex, periplasmic chaperone LolA, and outer membrane receptor LolB (Tokuda, 2009). The process initiates with LolCDE catalyzing the release of the lipoprotein from the inner membrane to LolA, forming a water-soluble LolA-lipoprotein complex. The LolA protein then ushers the lipoprotein across the periplasm where it interacts with LolB at the inner leaflet of the outer membrane, releasing the lipoprotein to LolB where it is then transported into the inner leaflet of the outer membrane (Tokuda, 2009). The Lol proteins are conserved in most Gram-negative bacteria; however, *T. pallidum* lacks a homologue to LolB (Fraser *et al*., 1998). It is therefore unclear whether the *T. pallidum*

lipoprotein sorting mechanism occurs in the same manner as that in *E. coli* (Tokuda, 2009).

1.7 The β-barrel assembly machinery (BAM) complex

The outer membranes of Gram-negative bacteria possess OMPs that are associated with basic physiological functions, virulence, and drug resistance and therefore play a crucial part in the pathogenesis of the organisms and in maintaining cell viability (Bos *et al*., 2007). Understanding how OMPs are targeted and folded into the outer membrane remains invaluable to understanding the pathogenesis of Gram-negative bacteria and could yield medical benefits such as vaccine production. During the past decade much research, in *E. coli* and *N. meningitidis,* has been devoted to understanding this phenomenon. The core bacterial complex responsible for trafficking and folding proteins into the outer membrane of Gram-negative bacteria is now known as the β-barrel assembly machinery (BAM) complex (Knowles *et al*., 2009). In *E. coli* the core BAM complex is comprised of five proteins: YaeT (BamA), an integral membrane protein, and four accessory lipoproteins, YfgL (BamB), NlpB (BamC), YfiO (BamD), and SmpA (BamE), which localize to the inner leaflet of the outer membrane (Onufryk *et al*., 2005; Sklar *et al*., 2007; Wu *et al*., 2005). As well three periplasmic chaperones, including SurA, Skp, and DegP, are believed to be involved in transporting nascent OMPs across the periplasmic space to the BAM complex in the outer membrane of *E. coli* (Rizzitello *et al*., 2001; Sklar *et al*., 2007).

1.7.1 BamA structure and function in *Escherichia coli*

The core integral membrane protein that makes up the BAM complex is BamA. This protein was originally identified as Omp85 by Voulhoux *et al*., where they showed that *omp85* was an essential gene in *N. meningitidis* and depletion of Omp85 resulted in the accumulation of unfolded OMP aggregates in the periplasm (Voulhoux *et al*., 2003). Studies conducted by Doerrler and Raetz showed that a mutant BamA strain of *E. coli* lacked OMPs in the outer membrane in comparison with wild type BamA strains (Doerrler & Raetz, 2005). These findings show that BamA plays a central role in the proper folding and assembly of OMPs into the outer membrane.

The BamA protein is found in all Gram-negative and Gram-negative like bacteria, which correlates appropriately with its function. The BamA protein consists of two major components: a set of five independently folded polypeptide transport-associated (POTRA) domains, which reside in the periplasm and a transmembrane β-barrel domain, which resides in the outer membrane (Sanchez-Pulido *et al.*, 2003).

The structures of the *E. coli* BamA POTRA domains have been solved by NMR, SAXS, and X-ray crystallography (Gatzeva-Topalova *et al*., 2008; Kim *et al*., 2007; Knowles *et al.*, 2008). The individual POTRA domains have a low sequence identity (<17%), however, they adopt a common fold that is comprised of a three stranded β-sheet overlaid by a pair of antiparallel α-helices (Knowles *et al*., 2009). The five POTRA domains exist in an extended conformation within the periplasm and possess a substantial amount of conformational freedom, which could yield functional implications during the OMP folding pathway (Gatzeva-Topalova *et al*., 2008). A small interface with little inter-domain interaction is observed between POTRA2 and 3, suggesting this could be a

hinge point which provides periplasmic flexibility for BamA. This flexibility could allow for two disctinct conformations of the POTRA domains depending on whether OMP substrates are present or absent (Gatzeva-Topalova *et al*., 2008). The structure of the BamA β-barrel has not yet been solved.

It is unclear exactly how BamA functions in OMP assembly in the outer membrane. Evidence shows that the POTRA domains could have a role in binding unfolded OMPs (Robert *et al*., 2006; Sanchez-Pulido *et al.*, 2003) and that they recognize a specific recognition motif encoded in the C-terminal β-strand of OMPs (Robert *et al*., 2006). The targeting motif, however, appears to differ between different bacteria, suggesting that OMP sorting is species specific (Robert *et al*., 2006). It was also shown that β-strand peptides not possessing the C-terminal targeting motif could also interact with the BamA POTRA domains, suggesting that the role of the POTRA domains is strictly to bind β-barrel OMPs and guide them to the rest of the core BAM complex (Knowles *et al*., 2008).

It has been shown that the POTRA domains interact with folding substrates using β-strand pairing (β-augmentation) (Kim *et al.*, 2007; Knowles *et al*., 2008), which is a non-covalent protein-protein interaction mechanism that involves the donation of a β-strand in the ligand to a β-sheet in the receptor (Remaut & Waksman, 2006). POTRA 3 appears to be essential for this purpose as it possesses unique characteristics in comparison with the other POTRA domains. There is a surface groove located between the β-sheet and the long α-helix that is deeper and more hydrophobic than in other POTRA domains (Gatzeva-Topalova *et al*., 2008). As well, the groove is approximately

30Å in length, which is comparable to the average height of an outer membrane β-barrel protein (Gatzeva-Topalova *et al*., 2008).

Deletion studies from *E. coli* BamA have found that the three most C-terminal POTRA domains, POTRA 3-5, are essential, whereas removal of POTRA 1 and 2 compromise growth (Kim *et al.*, 2007). Similar research done by Kim *et al*. has also shown that the POTRA domains act as a scaffold for the binding of the accessory complexing lipoproteins. Deletion of POTRA 5 leads to disengagement of all accessory lipoproteins, whereas deletion of POTRA 1 maintains the binding of all accessory lipoproteins (Kim *et al*., 2007). A deletion study done in *N. meningitidis*, showed that correct folding of large OMPs correlated with the number of POTRA domains present, which supports the theory that OMPs slide along the POTRA domains to the β-barrel region of the protein (Bos *et al.*, 2007).

1.7.2 The BAM accessory lipoproteins

In *E. coli* there are four accessory lipoproteins which form a tight complex with BamA; these are Bam B-E (Sklar *et al*., 2007; Wu *et al*., 2005). The BAM lipoproteins interact specifically with the POTRA domains of BamA and not the β-barrel region of the protein, with the exception of BamE which does not possess the ability to interact directly with BamA (Sklar *et al*., 2007). All of the BAM lipoproteins have roles in OMP biogenesis, as the depletion of each leads to varying degrees of OMP assembly defects.

The only lipoprotein which has been shown to be essential for cell viability is BamD, and it is also ubiquitous in Gram-negative bacteria (Malinverni *et al*., 2006). BamD and its homologues are predicted to contain up to six tetratricopeptide repeat motifs (tpr) that form tandem helix-loop-helix structures and they are believed to be

involved with protein-protein interactions (Blatch & Lassle, 1999; D'Andrea & Regan, 2003). As well, BamD has been shown to interact with BamA directly in an interaction which requires the fifth POTRA domain (Malinverni *et al*., 2006; Sklar *et al*., 2007). The BamB lipoprotein is also highly conserved among many Gram-negative bacteria, however absent from some genomes such as *N. meningitidis* and *N. gonorrhoeae* (Knowles *et al*., 2009). These proteins are predicted to possess a β-propeller fold with seven or eight blades (Vuong *et al*., 2008) and are proposed to interact with the BamA POTRA domains and/or nascent OMPs through β-augmentation (Gatsos *et al*., 2008). The BamB lipoprotein can bind BamA independently of the other lipoproteins in an interaction that requires POTRA 2-5 (Kim *et al*., 2007; Vuong *et al*., 2008). The BamC lipoprotein is not ubiquitous throughout Gram-negative bacteria and lacks significant similarity to any known protein structures; its role in the process of OMP biogenesis is unknown (Knowles *et al*., 2009), however deletion mutants possess minor defects in outer membrane permeability (Onufryk *et al*., 2005). The BamC lipoprotein does possess the individual ability to bind BamA; however, it requires the C-terminus of BamD (Malinverni *et al*., 2006). The BamE lipoprotein, like BamC, is not found in all Gram-negative bacteria, however, deletion mutants appear to possess OMP folding defects and increased sensitivity to rifampicin and SDS (Sklar *et al*., 2007). The BamE lipoprotein has not been shown to possess the ability to bind to BamA; however, it has been shown to stabilize the binding of BamD to BamA (Sklar *et al*., 2007).

1.7.3 The BAM associated periplasmic chaperones

It was originally predicted by de Cock *et al*., that proteins destined for the outer membrane were shuttled across the periplasm and delivered to an outer membrane

assembly site by periplasmic chaperones (de Cock *et al*., 1996). Since this prediction, numerous periplasmic chaperones have been identified (Duguay & Silhavy, 2004). The majority of these chaperones can be classified into three distinct groups: those that catalyze the formation of disulfide bonds (Nakamoto & Bardwell, 2004), peptidyl-prolyl *cis/trans* isomerases (PPIases), and those with general chaperone activity such as Skp, DegP, and SurA (Duguay & Silhavy, 2004). Since the identification of these periplasmic chaperones, it has been identified that Skp, DegP, and SurA are the major factors involved with ushering OMPs to the BAM complex in the outer membrane (Ruiz *et al*., 2006). All three of the chaperones implicated to be involved in the process of OMP biogenesis have their genes regulated by the σ^E envelope stress response (Raivio & Silhavy, 2001; Rhodius *et al*., 2006), which is activated in response to unfolded OMPs (Mecsas *et al*., 1993; Walsh *et al*., 2003).

The periplasmic chaperone DegP has both protease and chaperone activity and is regulated in a temperature-dependent manner (Lipinska *et al*., 1990; Spiess *et al*., 1999). The SurA protein is a member of the PPIase family, but also has general chaperone activity; upon depletion of SurA the outer membrane has permeability defects (Behrens *et al*., 2001). Cells that lack SurA have also been shown to contain reduced levels of OMPs (Rouviere & Gross, 1996). The Skp chaperone is a member of the general chaperone family of periplasmic chaperones. As well, Skp has been shown to bind denatured OMPs but not denatured periplasmic or cytosolic proteins and depletion of Skp leads to decreased levels of OMPs in the outer membrane (Chen & Henning, 1996). The *skp* gene is also located directly downstream of the *yaeT* (*bamA*) gene in *E. coli* (Voulhoux & Tommassen, 2004), and both are regulated by the σ^E envelope stress response (Rhodius *et*

al., 2006). In fact this genetic organization is seen in the majority of bacteria which possess a BamA and Skp homologue.

The precise role that each individual periplasmic chaperone plays in the process of OMP delivery to the BAM complex is unknown. Double knock-out experiments have revealed functional redundancy among the three chaperones and it has been suggested that Skp and DegP function in one pathway, whereas SurA acts in a separate pathway (Sklar *et al*., 2007). Both SurA and Skp have been shown to interact directly with OMPs as they leave the Sec translocon (Harms *et al*., 2001; Ureta *et al.*, 2007). Only SurA has been shown to interact with the BAM complex, the interaction occurs through the POTRA domains of BamA, however it is not known whether it is a direct interaction or through a substrate protein (Sklar *et al*., 2007). It has been suggested that, in *E. coli*, SurA acts in a primary pathway responsible for the assembly of most OMPs, whereas Skp and DegP act in a secondary pathway (Sklar *et al.*, 2007).

1.7.4 Mechanism of OMP insertion

The precise series of events in OMP biogenesis that make up the pathway from inner membrane translocation to outer membrane deposition are unknown. It is clear from previous research that OMPs destined for the BAM complex are first targeted to the Sec translocon and then interact in some way with one or more of the periplasmic chaperones SurA, Skp, and DegP, before being ushered to the BAM complex and inserted into the outer membrane (Knowles *et al.*, 2009). A number of possible mechanisms exist for this process including the pore-folding model, complex pore folding model, barrel-folding model, chaperone-folding model, and accessory folding model. All potential models are shown in Figure 7.

In the pore-forming model OMPs are inserted into the outer membrane through the β-barrel pore formed in BamA. In the complex pore-forming model a pore is formed by multiple BamA proteins oligomerizing to form a central pore through which the OMP can be transported. The barrel-folding model suggests that BamA provides a surface where the interacting OMP can properly fold in the vicinity of the outer membrane (Knowles *et al.*, 2009). The final two models were derived from the finding that the periplasmic chaperone DegP can form multimeric cage-like structures (Krojer *et al.*, 2008). It is unclear, however, whether DegP forms these multimeric structures *in vitro*. In the chaperone-folding model, DegP binds the OMP directly at the Sec translocon and in the accessory-folding model DegP binds the OMP after it has come into contact with the BAM complex (Knowles *et al.*, 2009). Krojer *et al*. theorize that the DegP cage like structures could form a macropore spanning the entire periplasm and that OMPs could diffuse directly from the inner to outer membrane (Krojer *et al.*, 2008).

There is currently no evidence that directly supports a given model of OMP insertion. Studies have shown that DegP is not essential to the process of OMP insertion. Deletion of DegP alone does not yield decreased levels of properly folded OMPs in the outer membrane and does not lead to increased numbers of unproperly folded OMPs in the periplasm (Sklar *et al*., 2007). These findings support all models that do not include DegP as an essential factor in the process of OMP insertion. There has been no direct evidence reported that BamA multimerizes to form a large pore which weakens the theory of the complex pore-forming model. The finding that POTRA 3 of BamA potentially possesses a β-barrel protein binding pocket (Gatzeva-Topalova *et al*., 2008) supports the theory that the POTRA domains of BamA are necessary for the proper

folding of nascent OMPs. It seems likely that BamA is involved with the proper folding of nascent OMPs, however further research is needed in order to better understand how the properly folded OMPs are deposited into the outer membrane.

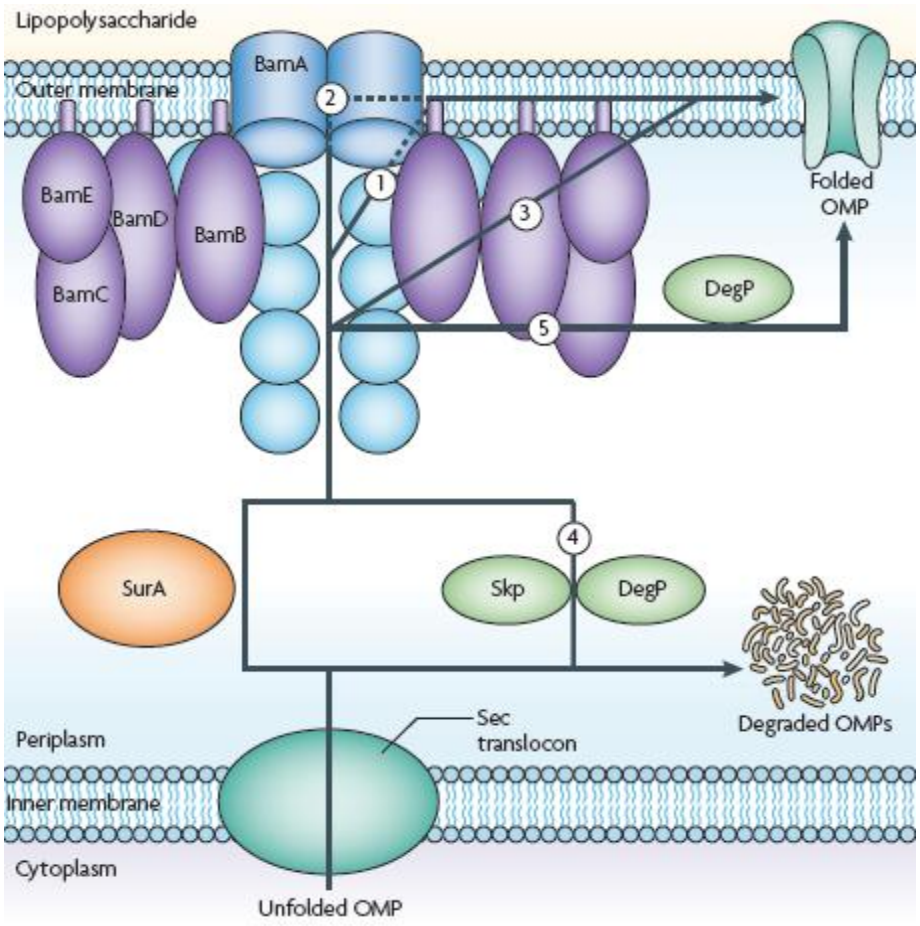


Figure 7. Schematic representation of OMP biogenesis in *Escherichia coli*
1) Pore-folding model, the β-barrel of BamA offers its pore for insertion of the nascent OMP into the membrane, and the POTRA domains and/or lipoproteins act to thread the OMP into the pore. 2) Complex pore-folding model, the central core is formed by a multimeric BAM complex that acts as the point of insertion into the membrane. Release of the OMP could then occur by dissociation of the multimeric BAM complexes. 3) Barrel-folding model suggests that the β-barrel of BamA provides a template for barrel folding in the vicinity of the BAM complex. 4) Chaperone-folding model, the periplasmic chaperones act to fold the protein and protect it from degradation during passage through the periplasm. The BAM complex thus functions only to insert the protein into the membrane. 5) Accessory folding model, the BAM complex functions to fold the nascent OMP but does not have a function in membrane insertion. The folded OMP is then released to DegP in a quality-control mechanism to remove incorrectly folded OMPs. The protein is then inserted into the membrane either by DegP or by some as-yet-unknown mechanism that could involve the BAM complex (Knowles *et al*., 2009).

1.7.5 Identification of *Treponema pallidum* protein Tp0326

The primary clearance mechanism responsible for removal of *T. pallidum* from syphilitic chancres is believed to be antibody mediated opsonisation followed by phagocytosis and macrophage killing. This has been supported by research done by Lukehart *et al*. and Baker-Zander *et al*. (Baker-Zander & Lukehart, 1992; Lukehart & Miller, 1978). As well, opsonic antibody has been shown to appear directly before bacterial clearance in the *T. pallidum* rabbit model (Baker-Zander *et al*., 1993). The target *T. pallidum* antigens of opsonic antibody are presumed to reside on the bacterial surface, allowing opsonisation of intact bacteria followed by phagocytosis (Cameron *et al*., 2000).

In the year 2000, Cameron *et al*. identified a 92kD putative *T. pallidum* OMP, Tp92, now designated Tp0326, using a differential immunological screen. The screen involved a *T. pallidum* λ genomic expression library that was incubated with opsonic rabbit serum and separately with non-opsonic rabbit serum in order to identify putative surface antigens (Cameron *et al*., 2000).

Previous research had shown that antiserum from rabbits which were immunized with heat killed *T. pallidum* failed to opsonize *T. pallidum* (Lukehart, unpublished data), however, antiserum from rabbits which were immunized with live *T. pallidum* would opsonize the bacteria leading to phagocytosis by macrophages (Baker-Zander & Lukehart, 1992; Lukehart & Miller, 1978). Cameron *et al.* used these two different sources of serum in a comparative immunological screen against a *T. pallidum* genomic expression library (Cameron *et al*., 2000). They identified one antigen which was specifically reactive to the opsonic rabbit serum; *T. pallidum* protein Tp0326. Using

bioinformatic analysis, Cameron *et al*. found that the protein was predicted to possess an N-terminal signal sequence and a β-barrel with exposed loops, which are both characteristic of OMPs (Cameron *et al*., 2000).

At the time of discovery not much was known about Tp0326, other than the fact that it was predicted to be located in the outer membrane. Recent research surrounding the identification and characterization of *E. coli* BamA, coupled with the availability of bioinformatic similarity search tools have led to the belief that Tp0326 is homologous to BamA.

Using bioinformatic analysis, it has been shown that *T. pallidum* protein Tp0326 is predicted to contain a C-terminal transmembrane β-barrel region and an N-terminal periplasmic region. The periplasmic region is predicted to contain 5 independent POTRA domains which each possess the typical β-sheet-α-helix-α-helix-β-sheet-β-sheet motif. As well, directly downstream of the *tp0326* gene lays *tp0327*, whose gene product is homologous to *E. coli* periplasmic chaperone Skp (Cameron, unpublished findings). As previously discussed, this genetic organization is conserved in bacteria that possess a BamA protein (Voulhoux & Tommassen, 2004).

1.8 Research hypotheses and objectives

Escherichia coli protein BamA has been shown to interact with various lipoproteins and periplasmic chaperones forming a complex involved in ushering OMPs through the periplasmic space and mediating insertion in the outer membrane.

The hypothesis of my research is that an OMP translocation complex similar to that found in *E. coli* exists within *T. pallidum* and that this complex is responsible for ushering *T. pallidum* OMPs to the bacterial surface.

The objectives of my research are to gain an understanding of the putative *T. pallidum* OMP transport machinery and identify OMPs that have, until this point, eluded discovery in *T. pallidum*.

1.9 Experimental approach

In order to identify and characterize the *T. pallidum* BAM complex we have utilize a two-pronged approach: bioinformatic analysis and protein-protein interaction studies. Initially, bioinformatic analysis was used in order to identify potential *T. pallidum* BAM complex proteins. As well, a *T. pallidum* lambda genomic expression library was screened against recombinant *T. pallidum* protein Tp0326 in order to identify putative *T. pallidum* BAM complex proteins that interact with Tp0326, as well as putative *T. pallidum* OMPs. The recombinant Tp0326 protein used in the library screen was comprised of only the Tp0326 POTRA domains, as it has been shown in *E. coli* that the BAM accessory lipoproteins and periplasmic chaperones interact with BamA though its five periplasmic POTRA domains. The findings from the two approaches were compared, and putative *T. pallidum* BAM complex proteins were verified for their interaction with the Tp0326 POTRA domains by means of Far-Western blot analysis.

Chapter 2: Materials and methods

2.1 Materials

Molecular cloning

Primers (IDT, Coralville, Iowa); PCR grade water (INVITROGEN, Carlsbad, California, CA.10977-015); dNTPs (FERMENTAS, Burlington, Ontario, CA.R0181); Pfu enzyme (FERMENTAS, CA.EP0501); Pfu buffer (FERMENTAS, CA.EP0501); Agarose (EMD, Gibbstown, New Jersey, CAS.9012-36-6); Low-melting point agarose (INVITROGEN, CA.15510-019); Ethidium bromide (AMRESCO, Solon, Ohio, CA.0492-5G); TAE buffer: Trizma® base (SIGMA, St. Louis, Missouri CA.T6066-1KG), Glacial acetic acid (ANACHEMIA, Sparks, Nevada, CA.00598-468), EDTA (ACP, Montreal, Quebec, CA.E-4320); TBE buffer: Trizma® base (SIGMA, CA.T6066-1KG), Boric acid (ACP, CA.B-2940), EDTA free acid (CALEDON, Georgetown, Ontario, CA.3466-5); GeneJET® gel-extraction kit (FERMENTAS, CA.K0691); CloneJET® PCR cloning kit (FERMENTAS, CA.K1232); One-shot® chemically competent *Escherichia coli* TOP10 cells (INVITROGEN, CA.404003); Bacto® yeast extract (BD BIOSCIENCES, Mississauga, Ontario, CA.288620); Bacto® tryptone (BD BIOSCIENCES, CA.211701); Bacto® agar (BD BIOSCIENCES, CA.214530); Sodium-chloride (EMD, CA.SX0420-1); Ampicillin (BIO BASIC, Amherst, New York, CA.99501-750); GeneJET® plasmid mini-prep kit (FERMENTAS, CA.K0503); Restriction enzymes: XhoI, NdeI, NcoI (NEB, Ipswich, Massachusettes, CA.R0146, CA.R0111, CA.R0913); Restriction enzyme buffers (NEB); T4 DNA ligase (NEB, CA.M0202); T4 DNA ligase buffer (NEB, CA.M0202); Kanamycin sulphate (BIO BASIC, CA.DB0286); Glycerol (ACP, CA.G-

3700); pET28a and pET32a plasmid DNA (NOVAGEN, Gibbstown, New Jersey, CA.69864-3, CA.69015-3)

Small-scale protein expression

One-shot® Chemically competent *Escherichia coli* Star™ BL21-DE3 cells (INVITROGEN, CA.601003); IPTG (INVITROGEN, CA.15529-019); Bug-Buster® (NOVAGEN, CA.70584-3); SDS-PAGE sample buffer: Tris-HCl pH 6.8 (EMD, CA.9310), DTT (BIO BASIC, CA.DB0058), Electrophoresis grade SDS (THERMO-FISHER, Waltham, Massachusetts, CA.BP166-100), Bromophenol blue (SIGMA, CA.161-0404), Glycerol (ACP, CA.G-3700); SDS-PAGE broad range unstained MW marker (BIO-RAD, Hercules, California, CA.161-0158); SDS-PAGE gels: Acrylamide mix 30% (BIO-RAD, CA.161-0317), Tris-HCl (EMD, CA.9310), SDS (FISHER, CA.9310), Ammonium persulfate (SIGMA, CA.215589), TEMED (BIO BASIC, CA.TB0508); SDS-PAGE running buffer: Trizma® base (SIGMA, CA.T6066-1KG), Glycine (CALEDON, CA.3880-5), SDS (FISHER, CA.9310); Coomassie stain: Coomassie Blue R-250 dye (OMNIPUR, Caldwell, Ohio, CA.3340), Methanol (VWR, Mississauga, Ontario, CA.CABDH6200-4), Glacial acetic acid (ANACHEMIA, CA.00598-468); Coomassie de-stain: Methanol (VWR, CA.CABDH6200-4), Glacial acetic acid (ANACHEMIA, CA.00598-468); Gel-drying solution: Methanol (VWR, CA.CABDH6200-4), Glycerol (ACP, CA.G-3700), Acetic acid (ANACHEMIA, CA.00598-468); Gel-drying film (PROMEGA, Madison, Wisconsin, CA.V7131)

Large-scale protein expression

Native purification buffer: Sodium phosphate (EMD, CA.8180-1), Sodium chloride (EMD, CA.SX0402-1), Imidazole (SIGMA, CA.I202); Protease inhibitor cocktail Set III, EDTA-Free (CALBIOCHEM, La Jolla, California, CA.539134); Syringe-driven PVDF 0.45µM filters (MILLIPORE, Billerica, Massachusettes, CA.SLHV033RS); Native elution/wash buffer: Sodium phosphate (EMD, CA.8180-1), Sodium chloride (EMD, CA.SX0402-1), Imidazole (SIGMA, CA.I202); SnakeSkin pleated dialysis tubing, 10000 MWCO (FISHER, CA.68100); Amicon Ultra-15 10, 000 MWCO centrifugal filter device (MILLIPORE, CA.UFC901024); BCA protein assay kit (PIERCE, Waltham, Massachusetts, CA.23235); Bovine serum albumin (SIGMA, CA.A7906); Syringes 10mL (TERUMO, Elkton, Maryland, CA.SS-10L); Needles 18G $1^{1/2}$ (BD BIOSCIENCES, CA.305196); PBS buffer: Sodium chloride (EMD), Potassium chloride (ACP, CA.P2940), Sodium phosphate (EMD), Potassium phosphate (ACP, CA.P4550); HisTrap™ FF columns (GE HEALTHCARE, Uppsala, Sweden, CA.17-5319-01); D-(+)-Trehalose dihydrate (SIGMA, CA.T0167)

Western and Far-Western blotting

Immobilon-P nitrocellulose membrane (MILLIPORE, CA.IPVH00010); Spectra™ broad-range multicolour prestained MW marker (FERMENTAS, CA.SM1842); Cathode solution: 6-Aminohexanoic acid (SIGMA, CA.A2504); Anode solution 1 and 2: Trizma® base (SIGMA), Methanol (VWR); Whatman filter paper (WHATMAN, Uppsala, Sweden, CA.1001917); Saranwrap (SC-JOHNSON, Brantford, Ontario); Heat-sealable pouch 6.5” X 8” (AMPAC FLEXIBLES, Cincinnati, Ohio, CA.01-812-16); TBS buffer: Sodium chloride (EMD), Potassium chloride (ACP), Trizma® base (SIGMA); Tween®-

20 (CALEDON, CA.9390-1); SEA BLOCK® blocking buffer (FISHER, CA.37527); AP reagent: Tris-HCl (EMD, CA.9310), Trizma® base (SIGMA), Sodium chloride (EMD), Magnesium chloride (CALEDON, CA.4720-1); NBT (PROMEGA, CA.S380C); BCIP (PROMEGA, CA.S381C); Rabbit-anti-chicken IgG (H&L) IRDye800® conjugated antibody (ROCKLAND IMMUNOCHEMICALS, Gilbertsville, Pennsylvania, CA.603-432-002); goat-anti-rabbit IgG-AP conjugated (SIGMA, CA.A3687)

LambdaZAP II assay

Lambda ZAP II Predigested *EcoR* I/CIAP-Treated Vector Kit (STRATAGENE, Santa Clara, California, CA.236211); ExAssist® helper phage (STRATAGENE, CA.200253); *E. coli* XL1 Blue MRF' cells (STRATAGENE, CA.200309); *E. coli* SOLR- cells (STRATAGENE, CA.200253); NitroPure, supported, pure nitrocellulose membranes, 45micron, 137mm (GE OSMONICS, Uppsala, Sweden, CA.WP4HY13700); NZ-amine-A casein hydrolysate (SIGMA, CA.C0626); Tetracycline hydrochloride (DUCHEFA BIOCHEMIE, Haarlem, Netherlands, CA.T0150); LB broth supplements: Maltose (ALFA AESAR, Ward Hill, Massachusettes, CA.24703), Magnesium sulphate (CALEDON, CA.MX0070-1); SM buffer: Sodium chloride (EMD), Magnesium sulphate (CALEDON), Gelatin (EMD, CA.1.04070.0500), Tris-HCl (EMD); Heat-sealable pouch 10" X 12" (AMPAC FLEXIBLES, CA.01-812-25FF); Chicken-anti-Tp92 antibodies (AVES); Chloroform (CALEDON, CA.CX-1055-2)

Materials contributed by others

Treponema pallidum subspecies *pallidum* DNA (Nichols strain) was provided by Dr. Caroline Cameron. The *T. pallidum* lambdaZAP genomic expression library was provided by Dr. Caroline Cameron. *E. coli* SA85 1.1 lysate was provided by Dr. Caroline Cameron. *Treponema pallidum* recombinant proteins Tp0751 and Tp0453 were provided by Rebecca Hof. *Treponema pallidum* recombinant protein Tp0750 was provided by Teresa Brooks. *Treponema pallidum* Tp0326-POTRA1-5 pET28a plasmid DNA was provided by Dr. Justin Radolf and Dr. Dan Desrosiers.

2.2 General methods

2.2.1 DNA manipulation

PCR reactions

Each PCR reaction contained 25pmol of each forward and reverse primer, 0.4ng of *T. pallidum* genomic DNA, 20mM of dNTPs, 10µL of 10X Pfu buffer with Mg^{2+}, and 79µL of PCR grade water (total reaction volume, 100µL). Hot start PCR was utilized by adding 5µL of Pfu enzyme to each PCR reaction during the initial hold step of the PCR reaction cycle. PCR cycling conditions included a single cycle of 3 mins at 95 ºC, 3 mins at 75 ºC, 2 mins at 50-65 ºC (depending on specific annealing temperature), 2 mins at 72 ºC, followed by 35 cycles of 1 min at 95 ºC, 1 min at 50-65 ºC, 2 mins at 72 ºC, followed lastly by a single cycle of 10 mins at 72 ºC. After the PCR cycle was complete the samples were stored at 4 ºC. The PCR analyses were performed using a DNA Thermocycler from Perkin Elmer (Waltham, Massachusetts). PCR products were initially visualized on a 1% (w/v) agarose gel stained with ethidium bromide (EtBr),

which ran at 100V for 1 hour in 1X TBE buffer (89 mM Tris base, 89 mM boric acid, 2 mM EDTA). Positive PCR reactions were then purified on a 1% (w/v) low-melting point agarose gel stained with EtBr, which ran at 50V, for 2.5 hours in 1X TAE buffer (40mM Tris base, 40 mM acetic acid, 1 mM EDTA). PCR products were then excised from the agarose gels and purified using GeneJET® gel extraction kit according to the manufacturer's directions.

Cloning PCR products into pJET

Purified PCR products were blunt-end ligated into the pJET plasmid using the GeneJET® cloning kit, following the manufacturer's directions. The ligation mixtures were incubated at RT for 1 hour and then directly transformed into chemically competent *E. coli* TOP10 cells.

Transformation of plasmid DNA

Thawed chemically competent cells (25µL) were incubated with 2µL of ligation mixture on ice for 15 minutes (mins), followed by a heat shock at 42 ºC for 90 seconds (secs) and further incubation on ice for an additional 2 mins. LB broth (10g Bacto-tryptone, 5g Bacto-yeast extract, 10g NaCl per 1L, pH 7.0) (500µL) warmed to 37 ºC was added to each transformation mixture and incubated with shaking at 37 ºC for 1 hour. Following incubation, the transformation mixtures were individually plated (75µL) on pre-warmed LB agar medium (10g Bacto-tryptone, 5g Bacto-yeast extract, 10g NaCl, 15g Bacto-agar per 1L) with appropriate antibiotic (100 µg/mL amp or 50 µg/mL kan) and incubated overnight (O/N) at 37 ºC. After O/N incubation the plates were stored at 4 ºC.

Isolation of plasmid DNA

Successfully transformed plasmids were identified by growing bacterial cultures in 5mL LB with appropriate antibiotic O/N at 37 ºC with shaking, followed by preparation of plasmid DNA using the GeneJET® miniprep kit following the manufacturer's directions, and finally by restriction enzyme digest and visualization by agarose gel electrophoresis.

Restriction enzyme digest

Isolated plasmid DNA was analyzed by restriction enzyme digest followed by visualization by agarose gel electrophoresis. Plasmid DNA (3µL) was digested with 0.5µL of each appropriate enzyme (NdeI/XhoI for pET28a and NcoI/XhoI for pET32a) in 10X NEBuffer with sdH_2O (10µL total reaction volume). The reaction was incubated at 37 ºC for 1 hour. Following incubation, correct digestion was confirmed by separation of products on a 1% (w/v) agarose gel, ran at 100V, for 1 hour in 1X TBE buffer and stained with EtBr.

DNA sequencing

Plasmid DNA (pJET) which was determined to possess the correct DNA insert through restriction enzyme digest and agarose gel electrophoresis was then sequenced by the DNA sequencing facility at the University of Washington DNA sequencing facility (Seattle, WA) using pJET forward and reverse primers (15pmol of each primer and 1200ng of DNA). DNA sequence data was analyzed by comparison to the appropriate *T. pallidum* DNA using the Clustal-W alignment tool (Larkin *et al*., 2007). Once positive clones were verified, they were then restriction digested out of the pJET plasmid and cloned into the pET28a and pET32a plasmids.

Cloning into pET28a and pET32a

Positive clones were restriction enzyme digested out of the pJET plasmid and separated out on a 1% (w/v) low-melting point agarose gel, ran at 50V, for 2.5 hours in 1X TAE buffer, and stained with EtBr. The digested insert DNA was excised from the agarose gel and purified using the GeneJET® gel extraction kit following the manufacturer's directions. Purified DNA was then ligated to pre-digested pET28a or pET32a vector DNA. Directly following ligation, the plasmids were transformed into chemically competent *E. coli* TOP10 cells for storage and glycerol stock production. As well, following ligation, the plasmids were transformed into chemically competent *E. coli* Star^{TM} BL21 DE3 cells for protein expression.

Ligation reaction

Ligation of the purified and digested plasmid DNA to pre-digested vector DNA was carried out using a 7:1 molar excess of insert to vector, 0.5µL of T4 DNA ligase, and 10X T4 DNA ligase buffer diluted in sdH_2O, to a final volume of 10µL. The ligation mixture was incubated at 16 ºC O/N.

Glycerol stock production

Positive clones were transformed into chemically competent *E. coli* TOP10 cells. Bacterial cultures were grown O/N at 37 ºC with shaking in 5mL LB with appropriate antibiotic. Glycerol stocks (20%) were created by adding 800µL of cells to 200µL of glycerol, flash freezing in liquid nitrogen, and storing at -80 ºC.

2.2.2 Small scale protein expression

Protein expression

All constructs were first tested by means of small scale protein expression before being scaled up for general production, in order to obtain optimal growth conditions. Initially, pET28a constructs were transformed into chemically competent *E. coli* Star™ BL21 DE3 cells, plated on LB + kan agar plates, and incubated at 37 ºC O/N. Bacterial cultures were then grown O/N at 37 ºC with shaking in 5mL LB + kan. After the initial cultures have grown, 1mL was inoculated into a 250mL Erlenmeyer flask containing 50mL LB + kan. Several individual flaks were inoculated with culture and were grown under different conditions. Cultures were grown with shaking at 37 ºC, 23 ºC, and 16 ºC. The cultures were monitored for OD_{600} using an Ultrospec 10 from Amersham Biosciences (Uppsala, Sweden) and protein expression was induced with 0.4mM IPTG at OD 0.8, 1.2, and 1.6. Two 0.5mL samples from each culture were taken at the time of protein induction and designated 'uninduced'. The samples were centrifuged using a 5415D centrifuge from Eppendorf (Mississauga, Ontario) at 13,000 rpm for 10 mins and then the supernatant was removed and the pellets were stored at -20 ºC for further analysis. After the induced 37 ºC culture was incubated for 3 hours, two more 0.5mL samples were taken and processed in the same manner as the 'uninduced' samples; these are the 'induced' samples. Two more 'induced' samples were taken for the 37 ºC culture at 5 hours induction. After the induced 23 ºC and 16 ºC cultures had incubated O/N, two more 0.5mL 'induced' samples were taken for each culture. The samples were then analyzed by SDS-PAGE.

SDS-PAGE protein analysis

The small scale protein expression samples were all analyzed by SDS-PAGE using the Bio-Rad system (Hercules, California). One sample from each different condition was resuspended in 150µL of 1X SDS-PAGE sample buffer (0.25M Tris-HCl pH 6.8, 2.5% glycerol, 1% SDS, and 50mM DTT); this is the 'total lysate'. The other sample, from each different condition, was resuspended in 75µL of BugBuster® and incubated on a Labquake from Thermo-Fisher (Waltham, Massachusetts) for 20 mins at RT. Following incubation the samples were centrifuged at 13,000 rpm for 10 mins at 4 ºC. The supernatant was then removed and added to 75µL of 2X SDS-PAGE sample buffer (0.5M Tris-HCl pH 6.8, 5% glycerol, 2% SDS, and 100 mM DTT); this is the 'soluble' protein. The pellet was resuspended in 150µL of 1X SDS-PAGE sample buffer; this is the 'insoluble' protein. All samples were boiled at 100 ºC in a heating block from VWR (Mississauga, Ontario) for 5 mins before being run alongside an unstained broad-range MW marker on a 15% SDS-PAGE gel for 1 hour at 200V, in 1X SDS-PAGE running buffer (25 mM Tris-base, 192 mM glycine, 0.1% SDS). The SDS-PAGE gels were then stained with Coomassie Blue R-250 for 1 hour and then destained O/N.

2.2.3 Soluble protein expression and purification

Soluble protein expression

Once small scale expression was successfully completed and conditions had been optimized for soluble protein expression, the growth scale was shifted to large scale. Large scale protein expression was done in 6L of LB medium compared to 50mL of LB medium in small scale protein expression. Bacterial cultures were grown at 37 ºC with

shaking for 6 hours in 5mL + kan. After the initial growth, 1mL was inoculated into 4 separate 250mL Erlenmeyer flasks, each containing 50mL LB + kan, and grown O/N at 37 ºC with shaking. After the O/N incubation, 6 separate 4L Erlenmeyer flasks, containing 1L of LB + kan, were each inoculated with 25mL of culture and grown at the optimal temperature and under the optimal conditions which were determined from small scale expression. Once the flasks had reached the appropriate OD_{600} they were induced with 0.4mM IPTG and grown under the optimal conditions.

Soluble protein purification

After large scale protein expression was complete, the cultures were centrifuged at 3,000g using an Avanti JE centrifuge from Beckman Coulter (Brea, California) for 15 mins at 4 ºC. After centrifugation the supernatant was discarded and the bacterial pellet was resuspended in a total of 20mL native binding buffer (50mM $NaHPO_4$, 500mM NaCl, 20mM imidazole, pH 8.0) with protease inhibitor cocktain set III (30µL per L of culture) for all 6L of bacterial culture. The cells are then lysed by sonication using a Sonicator 3000 from Misonix (Farmingdale, New York) for 30 secs dispersed over three bursts of sonicating, waiting 10 secs between bursts. Following sonication, the insoluble protein is pelleted by centrifugation at 15,000g for 30 mins at 4 ºC. The supernatant, which contains the soluble protein, is carefully aspirated into a clean tube. The supernatant was then filtered through a 0.45µm syringe filter before it was loaded onto a HisTrap® FF column and purified by column chromatography using an AKTAprime plus FPLC from GE Healthcare (Uppsala, Sweden). The protein was eluted off the HisTrap column using native elution buffer (50mM $NaHPO_4$, 500mM NaCl, 500mM imidazole, pH 8.0), which contains a high concentration of imidazole. The fractions from the FPLC

purification which contain the purified protein were identified using AKTAprime PrimeView software from GE Healthcare (Uppsala, Sweden). The fractions containing the soluble protein were then pooled and dialyzed overnight in 1X PBS (137mM NaCl, 2.7mM KCl, 4.3mM Na_2HPO_4, 1.47mM KH_2PO_4 per 1L, pH 7.4) using SnakeSkin® pleated dialysis tubing (10,000 MWCO). After dialysis the purified protein was concentrated down then subjected to gel filtration chromatography in order to remove any contaminating proteins.

Gel-filtration chromatography

Soluble purified protein was concentrated to 1.5mL using an Amicon Ultra-15 10, 000 MWCO centrifugal filter device from Millipore (Billerica, Massachusetts) by centrifugation at 4000g using a 5810R centrifuge from Eppendorf (Mississauga, Ontario). The concentrated protein was then centrifuged at 10,000rpm using a 5415D centrifuge from Eppendorf for 20 mins at 4 ºC to remove any precipitated protein. The supernatant, containing the soluble protein, was applied to a HiLoad 16/60 Superdex 75 prep-grade gel-filtration column from GE Healthcare (Uppsala, Sweden) which had been equilibrated with 1X PBS. The protein applied was separated out on the gel-filtration column, eluting with 1X PBS. The fractions from the gel-filtration column which contain the desired purified protein were identified using AKTAprime PrimeView software from GE Healthcare and were further analyzed using SDS-PAGE. The fractions containing the desired protein were pooled and the cryo-protectant trehalose was added to a final concentration of 7.5% (De Carlo *et al*., 1999). Protein concentration was determined using the Pierce BCA protein concentration kit following the manufacturer's

directions. The protein was then aliquoted, flash frozen in liquid nitrogen, and stored at -20 ºC for further use.

2.2.4 Western and Far-Western blotting

Transfer to nitrocellulose membrane

Proteins of interest, including a pre-stained broad range MW marker, are separated by SDS-PAGE and then transferred to a nitrocellulose membrane at 200 mA for 1 hour, using a semi-dry blotter. The membranes were then placed in blocking buffer (20% seablock®/TBS, 0.1% Tween-20, pH 7.4) O/N at 4 ºC for further analysis. The membranes can stay in blocking buffer for up to 7 days at 4ºC.

Western blotting

Following sufficient blocking, the membranes were washed for 2 x 10 min with TBS-T, pH 7.4. The membranes were then incubated with the primary antibody of interest, which was diluted with 20% seablock/TBS-T, for 1 hour at RT, shaking periodically. The membranes were incubated in 6.5" x 8" heat sealable bags from Ampac flexibles (Cincinnati, Ohio) and the volume of diluted antibody added was 2mL. For larger nitrocellulose membranes, 10" x 12" heat sealable bags were used with 5mL volume added. Following incubation, the membranes were washed 4 x 10min with TBS-T. The membranes were incubated with the appropriate secondary antibody, diluted in 20% seablock/TBS-T, for 1 hour at RT, shaking periodically. The membranes were washed 4 x 10min with TBS-T. The membranes were developed using the appropriate method.

AP development

All western blots performed in this thesis utilize alkaline phosphatase (AP) development, with the exception of Far-Western blots, which used Odyssey development. In order to utilize AP development, the appropriate AP-conjugated secondary antibody was incubated with the nitrocellulose membranes as described above. The AP reagent was prepared by adding 5mL of AP substrate (0.1M Tris, 100mM NaCl, 5mM $MgCl_2$, pH 9.5) to 33µL of NBT and 16.5µL of BCIP. Following blotting, the membranes were placed, protein side up, on a clean large petri dish. AP reagent was then added to the membranes and the membranes were allowed to develop. Complete development was determined by comparing intensities to the positive and negative control lanes, this was normally 5 minutes. When development was complete, the membranes were placed in water to stop development and remove excess reagent. The membranes were then air-dried in the dark O/N and then stored in SaranWrap. After the membranes were dry they were analyzed visually by eye.

Odyssey development

Odyssey development utilizes fluorescence technology and is a highly sensitive, quantitative development method. This method used the secondary antibody rabbit-anti-chicken IgG (H&L) IRDye800® conjugated from Rockland Immunochemicals (Gilbertsville**,** Pennsylvania) at a 1:10 000 dilution. After incubation of the membranes with the specific secondary antibody and subsequent washing, the membranes were developed using an Odyssey® infrared imaging system from Li-Cor Biosciences (Lincoln, Nebraska). The excitation wavelength was 778nm and the emission

wavelength was 806nm. The blots were viewed and analyzed using the Odyssey® infrared imaging system software from Li-Cor Biosciences (Lincoln, Nebraska).

Far-Western blotting

The procedure for Far-Western blotting is the same as the above outlined Western blotting protocol except for the inclusion of an additional incubation of the nitrocellulose membrane with a protein of interest (bait protein) before incubation with the antibodies. Following blocking, and the initial washing of the membrane, the membrane was incubated with a bait protein of interest for 2 hours at RT, shaking periodically. The bait protein was diluted with 1X PBS at a concentration of 50-250µg/ml, depending on the application. Following bait protein incubation, the membranes were washed 4 x 10min with TBS-T and then incubated with the appropriate primary antibody for the bait protein. The procedure then follows that of a traditional Western blot.

2.3 Bioinformatics

Protein/nucleotide sequence information

Protein and nucleotide sequence information was obtained using the Microbial Genome Database (MBGD) (Uchiyama, 2003) and verified using the National Center for Biotechnology Information (NCBI) GenBank database (Benson *et al.*, 2003). The MBGD database lists the name of the gene of interest and corresponding protein product, description of function or putative function, the organism of origin, the chromosomal position, and the GI number. The MBGD database displays both nucleotide and protein sequences and can link to other major databases including NCBI, Uniprot, KEGG, TIGR, GIB, and GTOP. The strain used for all *T. pallidum* sequence information was Nichols

(Fraser *et al*., 1998). The strain used for all *E. coli* sequence information was K12 (Blattner *et al.*, 1997).

Identifying Treponema pallidum BAM complex proteins

Putative *T. pallidum* BAM complex proteins were predicted by initially performing BLAST analysis using *E. coli* K12 BAM complex proteins. The NCBI database was used for BLASTp analysis (Altschul *et al.*, 1990) using the *T. pallidum* genome with the following *E. coli* proteins: BamA, BamB, BamC, BamD, BamE, Skp, DegP, and SurA. We used a cutoff of 20% sequence identity to classify a significant hit. The sequence identity cutoff is due to the inherent differences and vast evolutionary distance between *E. coli* and *T. pallidum.*

Analysis of protein sequences

Putative *T. pallidum* BAM complex homologues were further analyzed using various proteomic tools available from the Expert Protein Analysis System (EXPASY) proteomics server (Gasteiger *et al.*, 2003). EXPASY is a proteomics server of the Swiss Institute of Bioinformatics (SIB) and functions in collaboration with the European Institute of Bioinformatics. The theoretical isoelectric point (pI) and molecular weight (MW) were analyzed using the UniProt database (UniProt, 2010). The Translate tool was used from the EXPASY server in order to translate a nucleotide sequence to a protein sequence (Gasteiger *et al*., 2003). The NCBI BLAST tool was used against all other organisms in order to identify putative functions (Altschul *et al.*, 1990). Signal peptide prediction was done using SignalP 3.0, which incorporates a prediction of cleavage sites and a signal peptide/non-signal peptide prediction based on a combination of several artificial neural networks and hidden Markov models (Emanuelsson *et al*., 2007). The

LipoP 1.0 server was used to produce predictions of lipoproteins and discriminate between lipoprotein signal peptides, other signal peptides and N-terminal membrane helices in Gram negative bacteria (Juncker *et al*., 2003). Alignments between *T. pallidum* and *E. coli* homologues were performed using the Clustal-W alignment tool (Larkin *et al*., 2007).

Protein structure prediction

The secondary structure of *T. pallidum* protein Tp0326 was predicted and analyzed using the PredictProtein server (Rost *et al*., 2004). PredictProtein retrieves similar sequences and predicts aspects of protein structure and function based on sequence similarity. The SWISS-MODEL server (Arnold *et al*., 2006), HHpred homology detection and structure prediction tool (Soding, 2005), and InterProScan signature sequence detection program (Zdobnov & Apweiler, 2001) were also used for protein structure prediction and analysis. Three dimensional modeling of the predicted tertiary structure of *T. pallidum* protein Tp0326 was done using the Phyre protein folding recognition server (Kelley & Sternberg, 2009). Bacterial protein localization prediction was performed using the PSORTb prediction tool (Yu *et al.*, 2010) and the original PSORT prediction tool (Nakai & Horton, 1999).

Analyzing Treponema pallidum hits from LambdaZAP assay

The *T. pallidum* proteins identified to interact with *T. pallidum* protein Tp0326 POTRA 1-5, in the LambdaZAP assay, were first identified by using the NCBI BLAST tool using the *T. pallidum* genome (Altschul *et al.*, 1990). The proteins were then compared to the previously identified putative *T. pallidum* BAM complex homologues. Proteins which had not yet been identified as putative *T. pallidum* BAM complex proteins

were then further analyzed for sequence and structure information as previously described. This information is contained within Chapter 3 which describes all *in silico* results.

2.4 LambdaZAP II assay

The LambdaZAP assay utilized in this thesis fuses together the LambdaZAP system, designed by Stratagene (Santa Clara, California), with a traditional Far-Western blot assay.

2.4.1 LambdaZAP library

Library construction

The *T. pallidum* LambdaZAP library was originally constructed by Dr. Caroline Cameron using the Lambda ZAP II Predigested *EcoR* I/CIAP-Treated Vector Kit from Stratagene (Santa Clara, California). *Treponema pallidum* subspecies *pallidum* was first cultured in New-Zealand White rabbits. The *T. pallidum* cells were then harvested and genomic DNA from the organisms was isolated and digested with Tsp509I (TTAA). The digested DNA was then size selected from 0.5 – 4kb using agarose gel electrophoresis. The appropriate size digested DNA was then extracted and purified from the agarose gel. The digested *T. pallidum* DNA was then ligated to predigested Lambda ZAP II vector arms and packaged into the phage according to manufacturer's directions.

Establishing bacterial stocks

Escherichia coli SOLR and XL1-Blue MRF' glycerol stocks were purchased from Stratagene. The SOLR strain was plated on an LB + tet agar plate and the XL1-Blue MRF' strain was plated on an LB + kan agar plate. The plates were incubated O/N at 37

ºC. Once colonies could be seen isolated on the plates they were stored at 4 ºC. The plates can be stored for up to 1 week before use.

Titering the library

Initially a bacterial culture of *E. coli* XL1-Blue MRF' cells was created in 10mL of LB + supplements (0.2% (w/v) maltose and 10mM $MgSO_4$) in a 50mL conical tube and grown O/N at 37 ºC. Following O/N incubation, 10mL of LB + supplements was inoculated with 0.5mL of *E. coli* XL1-Blue MRF' O/N culture in a 50mL conical tube and incubated with shaking at 37 ºC until OD_{600} of 0.8-1.0. The *T. pallidum* LambdaZAP library was thawed on ice and then diluted in 990µL of SM buffer (100mM NaCl, 8mM $MgSO_4$ $7H_2O$, 50mM Tris-HCl, 0.002% (w/v) gelatin) to make a 10^{-2} dilution. Serial dilution in 900µL SM buffer was then performed to generate 10^{-5}, 10^{-6}, and 10^{-7} dilutions in SM buffer. After diluting the library,1µl of each dilution was added separately in 1.5mL tubes to 200µL of *E. coli* XL1-Blue MRF' cells, grown to an OD_{600} of 0.8-1.0. The phage + cells mixture was then incubated for 15 minutes at 37 ºC (no shaking) to allow the phage to attach to the cells. Each phage/cell mixture was then separately added to 3.5mL-4.0mL of NZY top agar (0.7% (w/v) agarose, 1L NZY broth: 5g NaCl, 2g $MgSO_4$ $7H_2O$, 5g Bacto-yeast extract, 10g NZ-amine casein hydrolysate, pH 7.5), melted and cooled to ~48 ºC-50 ºC. The top agar was then immediately plated onto dry, prewarmed NZY agar plates (1L NZY broth, 15g Bacto-agar, pH 7.5). The plates were allowed to set for 10 minutes and were then inverted and incubated O/N at 37 ºC. Following O/N incubation the plates were analyzed and stored at 4 ºC. The phage titer was determined by using the following equation: Titer = [(pfu x dilution) / Volume plated (µL)] x 1000µL/mL.

2.4.2 Performing LambdaZAP plaque lifts

Growth of cells for plating phage

A bacterial culture was created by inoculating a single colony of *E. coli* XL1-Blue MRF' in 50mL LB + supplements and growing O/N with vigorous shaking at 30 ºC. Following O/N incubation the cells were centrifuged at 1000g for 10 mins and then gently resuspended in 10mL of 10mM $MgSO_4$. Immediately before the cells were used they were diluted to an $O.D_{600}$ of 0.5 with 10mM $MgSO_4$.

Performing plaque lifts

The equivalent of 1.25 x 10^4 pfu of the T. *pallidum* LambdaZAP library and 600µL of freshly prepared *E. coli* XL1-Blue MRF' cells were combined into a single tube. This is repeated 4 times, in order to generate a total of 5 x 10^4 pfu of the T. *pallidum* LambdaZAP library. The number of pfu screened in the assay was determined by accounting for screening the *T. pallidum* genome in triplicate while accounting for the smallest possible DNA insert size, accounting for all three reading frames, and both orientations within the phage. The mixtures were incubated at 37 ºC for 15 minutes to allow attachment of phage to cells. Following incubation the mixtures were added to 6.5mL-7mL NZY top agar (~48-50 ºC) and immediately poured onto dry, prewarmed 150-mm NZY agar plates, which were at least 2 days old. The plates were allowed to set for 10 minutes before they were inverted and incubated at 37 ºC for ~5.5 hours (or until plaques could be seen). Once plaques could be visualized on the plates, they were chilled for 1 hour at 4 ºC to prevent the top agar from sticking to the nitrocellulose membrane. The plates were overlayed with 10mM IPTG-impregnated NitroPure, supported, pure nitrocellulose membranes (137mm), and using a needle with waterproof ink, orientation

marks were created through the membrane and agar. The plates were inverted and incubated at 37 ºC O/N. Following O/N incubation, the plates were chilled at 4ºC for 1 hour before removing the membranes from the plates. The membranes were washed in TBS-T, pH 7.4 for 10 mins four times each. The NZY agar plates were placed at 4 ºC for storage and further analysis.

2.4.3 Screening the plaques

Primary screening of nitrocellulose membranes

After the nitrocellulose membranes were removed from the NZY agar plates they were washed 4 x 10 min in TBS-T. Following washing they were placed in blocking buffer (50% seablock®/TBS-T) and incubated at RT for 60 mins with shaking. The membranes were treated to a Far-Western blot assay as described in section 2.2.4. The bait protein being used was recombinant *T. pallidum* Tp0326 POTRA1-5 at a concentration of 150µg/mL. Expression and purification of this protein is described in section 4.3. The bait protein was diluted in 1X PBS and a total volume of 5mL was incubated with each membrane in a 10" x 12" heat sealable pouch. Membranes were incubated with the bait protein for 2 hours and then subjected to washing for 4 x 10 min in TBS-T. The primary antibody used for this screen was chicken IgY anti-Tp92 from Aves Labs Inc. (Tigard, Oregon). The antibody was diluted 1:200 with 20% seablock®/TBS-T. The secondary antibody was goat-anti-rabbit IgG-AP conjugated from Sigma (St. Louis, Misouri) diluted 1:1000 in 20% seablock/TBS-T. Following antibody incubation the membranes were developed via the AP-development method, air dried O/N in the dark, and then analyzed visually by eye.

Coring out reactive plaques

The primary screen nitrocellulose membranes were initially analyzed for plaques which show positive immunoreactivity. The positive reacting plaques were cored out from the original NZY agar plate by lining up the orientation marks from the nitrocellulose membrane with those of the NZY agar plate. Cores were taken in a sterile manner, using the large end of a 200µL pipette tip. The cores, which contain the plaque of interest and surrounding plaques, were transferred to sterile microcentrifuge tubes containing 500µL of SM buffer and 20µL of chloroform. The tubes were then vortexed to release the phage particles into the SM buffer and incubated for 1-2 hours at RT or O/N at 4 ºC. This phage stock is stable for up to 6 months at 4 ºC. The tubes now contain new phage stocks which possess the immunoreactive plaques. These tubes were then re-titered, and screened with less pfu/plate to obtain isolated immunoreactive plaques.

Secondary screen

The secondary screen was performed using the cores taken from the original NZY plates from the primary screen. Each core was treated separately in a separate experiment. The secondary screen follows the same protocol as the primary screen with the exception of the amount of phage initially added to the prepared *E. coli* XL1-Blue MRF' cells. In the primary screen, 1.25×10^4 pfu/plate was used; the secondary screen uses 250 pfu/plate, with a total of 500 pfu screened. The diluted pfu/plate yields well isolated plaques on the NZY agar plates. The plaques were screened in the same manner as the primary screen. At this stage, isolated immunoreactive plaques were cored out and

stored in 500µL of SM buffer and 20µL of chloroform for *in vivo* excision and DNA sequencing.

2.4.4 Sequencing the plaques

In Vivo Excision of the pBluescript SK(-) Phagemid

Separate 50mL bacterial cultures of *E. coli* XL1-Blue MRF' and *E. coli* SOLR cells were grown O/N in LB + supplements at 30 °C with shaking. The cells were then centrifuged at 1000 x *g* for 10 minutes. The pellets were resuspended gently in 25mL of 10mM $MgSO_4$. The concentration of the cells was adjusted to OD_{600} of 1.0 (8.0×10^8 cells/mL) in 10mM $MgSO_4$. The following components were then added to a BD Falcon polypropylene tube from BD Biosciences (Mississauga, Ontario): 200µL of XL1-Blue MRF' cells at an OD_{600} of 1.0, 250µL of phage stock (>1×10^5 phage particles), and 1µL of ExAssist helper phage (>1×10^6 pfu/µL). The mixture was then incubated in the tube at 37 °C for 15 minutes to allow attachment of phage to cells. Following the incubation, 3mL of LB + supplements was added and the tubes were incubated for 2.5-3 hours at 37 °C with shaking. The tubes were then heated at 65-70 °C for 20 minutes to lyse the lambda phage particles and the cells. Following lysis, the tubes were centrifuged at 1000 x *g* for 15 minutes to pellet cell debris. The supernatant was decanted into a clean falcon tube. This stock contains the excised pBluescript phagemid packaged as filamentous phage particles. This stock can be stored at 4 °C for 1-2 months.

Plating the excised phagemids

The freshly prepared *E. coli* SOLR cells, at OD_{600} 1.0, were added to two 1.5mL tubes in a 200µL volume. The excised pBluescript phagemid particles were then added to the tubes, 5µL to one tube and 10µL to the other tube. The tubes were then incubated

at 37 °C for 15 minutes. Following incubation, 50µL of the cell mixture from each tube was plated on LB-amp agar plates and incubated O/N at 37 °C. An isolated colony was then used to generate O/N bacterial cultures in LB-amp media. The O/N culture was used to prepare a bacterial glycerol stock, which was stored at -80 °C. The O/N culture was also used to recover plasmid DNA and perform DNA sequencing.

DNA sequencing

Once pBluescript phagemid DNA had been recovered it was analyzed using a Nanodrop ND-1000 Spectrophotometer from Thermo-Fisher (Waltham, Massachusetts) for DNA concentration. The phagemids were then sent for DNA sequencing to the DNA sequencing facility at MWG Biotech Inc. Operon (Huntsville, Alabama) using 15pmol T7 forward primer and 15pmol M13 reverse primer and 1200ng phagemid DNA. The DNA sequences were then analyzed by NCBI BLAST against the *T. pallidum* genome (Altschul *et al*., 1990).

Chapter 3: Characterizing the putative *Treponema pallidum* BAM complex using bioinformatics

3.1 Contributions to the data

My contributions to the data presented in this chapter include: secondary structure analysis of *T. pallidum* protein Tp0326, delineation of *T. pallidum* Tp0326 POTRA domains, and identification of putative *T. pallidum* BAM complex proteins. This work was done in collaboration with Dr. Justin Radolf and Dr. Dan Desrosiers at the University of Connecticut Health Center, USA.

3.2 Introduction

Bioinformatics is the application of statistics and computer science to the field of molecular biology. In the last decade, bioinformatics has been used for research in the field of molecular biology in many ways: computationally analyzing wet-lab data, genome sequencing, identification of protein coding segments (Azad & Borodovsky, 2004), genome comparison to identify gene function (Altschul *et al*., 1990; Bansal, 1999), development of genomic and proteomics databases (Bairoch, 1991), and inference of phenotypes from genotypes (Bansal, 2005).

Three basic approaches have been used in bioinformatics: (i) use of computational search and alignment techniques to compare a new genome against a set of known genes in order to annotate the function of the genes in the newly sequenced genome (Altschul *et al*., 1990; Altschul *et al*., 1997), (ii) the use of mathematical modeling techniques such as data mining, statistical analysis, neural networks, genetic algorithm, and graph matching techniques to identify common patterns, features and high level functions (Bansal, 2005),

and (iii) an integrated approach that integrates search techniques with mathematical modeling (Bansal, 2005). The use of bioinformatics in molecular biology research includes sequence alignment, gene finding, genome assembly, drug design, drug discovery, protein structure alignment, protein structure prediction, prediction of gene expression and protein-protein interactions, genome-wide association studies and the modeling of evolution (Bansal, 2005).

The finding that *T. pallidum* possesses a putative BamA homologue as well as a putative *E. coli* Skp homologue suggests that an OMP translocation complex similar to that found in *E. coli* could potentially exist within *T. pallidum.* Bioinformatic analysis was initially used in order to investigate this hypothesis.

3.3 *Treponema pallidum* protein Tp0326 (BamA)

Treponema pallidum gene *tp0326* (GI # 15639317) is located at position 344276 - 346837 on the chromosome (Benson *et al*., 2003). The gene product, Tp0326, is an 837 a.a. protein with a predicted pI of 8.78 and MW of 96127.90 Da (EXPASY compute pI/Mw tool). Both SignalP3.0 and PSORTb predict a cleavable N-terminal signal peptide, with cleavage between a.a. 21 and 22, predicting a non-cytoplasmic protein. PSORTb localization prediction tool predicts that Tp0326 is located in the outer membrane with a localization probability of 99.3%. In addition, the C-terminus of Tp0326 contains hydrophobic residues at positions 1, 5, and 7 from the C-terminus, which loosely conforms to the consensus hydrophobicity pattern predicted for bacterial outer membrane proteins of hydrophobic residues at positions 1, 3, 5, 7, and 9 from the C-terminus (Struyve *et al.*, 1991).

NCBI BLAST of the Tp0326 protein sequence against the protein database gives hits of OMPs, OMP assembly complex proteins, and D15 major surface antigens from a long list of organisms including *B. burgdorferi, H. pylori, N. meningitidis, and N. gonorrhoeae*, which are listed in Table 1. It shares 23% sequence identity and 41% sequence similarity with *E. coli* YaeT (BamA). The observed sequence similarity is evenly distributed throughout the coding sequence of Tp0326, with the exception of a serine-rich region at the C-terminal end of Tp0326, which appears to be unique to *T. pallidum* (Cameron *et al*., 2000). *Treponema pallidum* protein Tp0326 also shares strong predicted structural similarities to *E. coli* BamA.

Table 1. BLAST results for *Treponema pallidum* protein Tp0326

Bacterium	Protein	Description	Sequence identity (%)
Escherichia coli	YaeT	OMP	23
Borrelia burgdorferi	BB0795	OMP	28
Borrelia hermsii	BH0795	OMP	29
Rhizobium etli	Omp1	OM lipoprotein	27
Geobacter uraniumreducens	Gura_3237	Surface antigen (D15) precursor	27
Helicobacter pylori	HP0655	Protective surface antigen (D15)	24
Neiserria meningitidis	Omp85	OMP	24
Neiserria gonorrhoeae	NGO1801	Hypothetical protein	24
Leptospira interrogans	LIC11623	OMP	24
Burkholderia cenocepacia	Bcen_6066	Surface antigen (D15)	25

The secondary structure of *T. pallidum* protein Tp0326 was analyzed by PredictProtein. The N-terminal region is predicted to possess 5 independent regions which conform to the consensus for a POTRA domain; that is β-sheet-α-helix-α-helix-β-sheet-β-sheet motif rather than the originally mis-identified β-sheet-β-sheet-α-helix-α-helix-β-sheet motif (Sanchez-Pulido *et al*., 2003). Three dimensional modeling of the Tp0326 POTRA domains using PyMOL confirmed the existence of the inherent POTRA motif. The predicted tertiary structure of Tp0326-POTRA2 is shown in Figure 8.

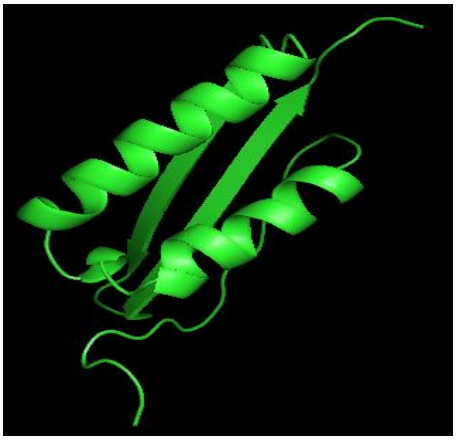


Figure 8. Three dimensional model of Tp0326-POTRA-2
Predicted 3D model of the Tp0326 POTRA-2 domain, using the PyMOL modeling program. The Tp0326 POTRA domain is predicted to be made up of the consensus β-sheet-α-helix-α-helix-β-sheet-β-sheet motif (The PyMOL Molecular Graphics System, Version 1.2.3, Schrödinger, LLC).

A β-barrel is a large beta-sheet that twists and coils to form a closed structure in which the first strand is hydrogen bonded to the last. The β-strands in β-barrels are typically arranged in an antiparallel fashion (Reardon & Farber, 1995). The C-terminal region of Tp0326 is predicted to be composed of alternating β-sheet-loop-β-sheet regions and is predicted to contain transmembrane regions by TMPred. These predictions are consistent with β-barrel regions. The predicted N-terminal POTRA domains and C-

terminal β-barrel regions conforms to the structure of BamA proteins. A schematic representation of the predicted structure of Tp0326 is shown in Figure 9. ProteinPredict was used to delineate the POTRA domains in Tp0326. The a.a regions of *T. pallidum* Tp0326 corresponding to each POTRA domain can be seen in Table 2. The overlapping regions of adjacent POTRA domains are predicted loop regions which lack secondary structure.

Table 2. Delineation of *Treponema pallidum* Tp0326 POTRA domains

Tp0326 POTRA Domain	Tp0326 amino acids	Amino acid total
POTRA 1	a.a 22-106	85
POTRA 2	a.a 105-183	79
POTRA 3	a.a 181-275	95
POTRA 4	a.a 272-355	84
POTRA 5	a.a 353-434	82

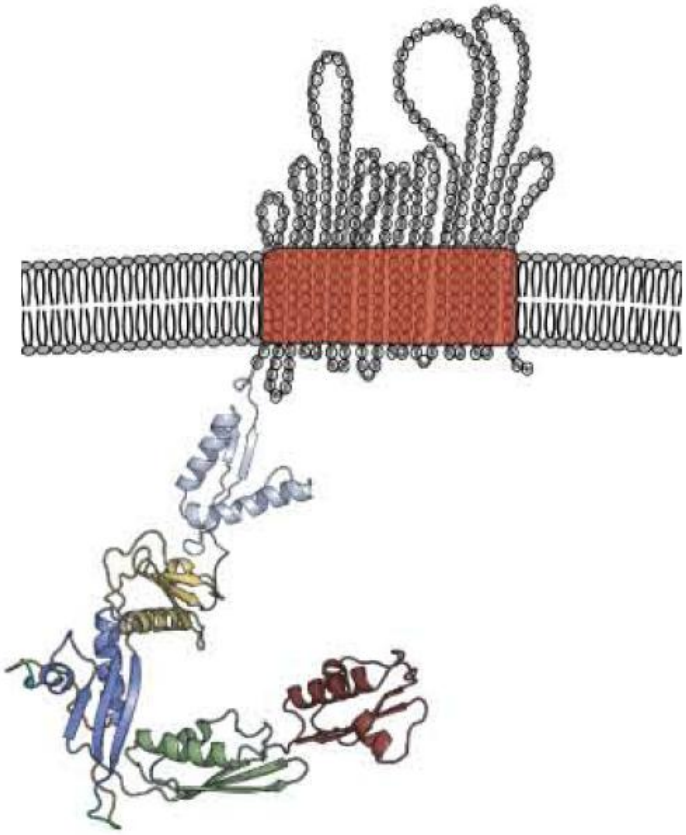


Figure 9. Schematic represenation of the predicted structure of Tp0326
Treponema pallidum protein Tp0326 is predicted to possess an N-terminal periplasmic region, consisting of 5 independently folded POTRA domains, and a C-terminal transmembrane β-barrel region with exposed loops. Adapted from Cox *et al*. (Cox *et al.*, 2010).

3.4 *Treponema pallidum* BAM complex lipoproteins

In *E. coli* there are four accessory lipoproteins which form a tight complex with BamA, these are Bam B-E (Sklar *et al.*, 2007; Wu *et al.*, 2005). The only lipoprotein which has been shown to be essential for cell viability is BamD (Malinverni *et al*., 2006). BamB is also highly conserved among many Gram-negative bacteria, however absent from some genomes such as *N. meningitidis* and *N. gonorrhoeae* (Knowles *et al*., 2009).

In order to identify BAM accessory lipoproteins in *T. pallidum*, I performed BLASTp searches, using the protein sequences from the *E.coli* lipoproteins, against the *T. pallidum* protein database.

The *E. coli* lipoprotein sequences for BamB-E were downloaded from GenBank at NCBI; sequence information (GI number) can be seen in Table 3. The protein sequences were used in BLASTp searches against the *T. pallidum* protein database in order to identify putative *T. pallidum* BAM complex lipoproteins. The results from the BLASTp searches are summarized in Table 3.

Table 3. *Treponema pallidum* BAM lipoprotein BLASTp results

E. coli lipoprotein	GI number	BLASTp result	E – value*	Sequence identity
BamB (YfgL)	26248876	Tp0133	3.5	32/105 (31 %)
BamC (NlpB)	157267368	-	-	-
BamD (YfiO)	26248958	Tp0622	0.046	17/40 (43 %)
BamE (SmpA)	90111468	Tp0690	5.2	13/49 (27 %)

* E-value refers to the entire sequence used in the BLASTp search.

The initial *E. coli* lipoprotein BLASTp results against the *T. pallidum* database did not yield any obvious *T. pallidum* homologues. Only three of the four *E. coli* proteins identified any hits in *T. pallidum*; no hit was identified for *E. coli* lipoprotein BamC. All three of the *T. pallidum* proteins identified are annotated as hypothetical proteins and all three BLASTp results had non-convincing E-value scores. The *T. pallidum* hit for BamE, Tp0690, had extremely high E-value score of 5.2. This hit is not considered significant and is not reliable. We also performed further BLASTp searches using smaller fragments

of the respective *E. coli* protein sequences which resulted in similar findings. The *T. pallidum* hits for BamB and BamD, Tp0133 and Tp0622, had high E-value scores, 3.5 and 0.046 respectively. These are generally not considered significant E-value scores, however, there is a 40 a.a stretch in Tp0622 which shares 43 % sequence identity with *E. coli* BamD and there is a 105 a.a stretch in Tp0133 which shares 31 % sequence identity with *E. coli* BamB.

In order to further investigate the findings from the BLASTp experiments, we analyzed the predicted secondary structures for the *T. pallidum* hypothetical proteins Tp0622 and Tp0133. *Treponema pallidum* protein Tp0690 was not further analyzed due to low sequence identity and similarity to *E. coli* BamE.

3.4.1 Predicted structure of *Treponema pallidum* hypothetical protein Tp0622

We first analyzed the potential for Tp0622 to be an outer membrane associated lipoprotein. PSORT localization prediction tool predicted a cleavable N-terminal signal sequence and a 93.1 % probability of localization in the periplasmic space. In contrast the SignalP 3.0 signal peptide detection program did not detect a cleavable signal peptide. LipoP lipoprotein detection program does not predict Tp0622 to be a lipoprotein. The ProteinPredict topology and secondary structure prediction program predicts Tp0622 to be composed almost entirely of α-helical secondary structure. InterProScan signature recognition program predicts Tp0622 to possess three tetratricopeptide repeat (TPR) motif domains from a.a 201 - 234, 409 - 442, and 544 - 577. *E. coli* protein BamD is predicted to contain up to six TPR motifs (Blatch & Lassle, 1999; D'Andrea & Regan, 2003).

3.4.2 Predicted structure of *Treponema pallidum* hypothetical protein Tp0133

We first analyzed the potential for Tp0133 to be an outer membrane associated lipoprotein. PSORTb and PSORT localization prediction tools do not predict any localization for Tp0133 and do not predict an N-terminal cleavable signal sequence. SignalP 3.0 signal peptide detection program predicts an N-terminal cleavable signal sequence with 74.1 % probability at a.a 30 using the neural-networks model but no cleavable signal sequence is detected using the hidden markov model. LipoP lipoprotein detection program does not predict Tp0133 to be a lipoprotein. The ProteinPredict topology and secondary structure prediction program predicts Tp0133 to be composed almost entirely of β-sheet secondary structure. The Phyre 3D modelling program predicts Tp0133 to form a 7-bladed β-sheet propeller structure, which is the same structure formed by *E. coli* lipoprotein BamB (Vuong *et al*., 2008). As well, *T. pallidum* protein Tp0133 and *E. coli* lipoprotein BamB are similar in size, 410 a.a and 392 a.a respectively.

3.5 *Treponema pallidum* BAM complex chaperones

In *E. coli* there are three periplasmic chaperones which have been identified to be the major factors involved with ushering OMPs to the BAM complex in the outer membrane, namely Skp, DegP, and SurA (Ruiz *et al*., 2006). It has been suggested that, in *E. coli*, SurA acts in a primary pathway responsible for the assembly of most OMPs, whereas Skp and DegP act in a secondary pathway (Sklar *et al.*, 2007).

In order to identify BAM accessory periplasmic chaperones in *T. pallidum*, we performed BLASTp searches, using the protein sequences from the *E.coli* periplasmic chaperones, against the *T. pallidum* protein database.

The *E. coli* periplasmic chaperone sequences for Skp, SurA, and DegP were downloaded from GenBank at NCBI; sequence information (GI number) can be seen in Table 4. The protein sequences were used in BLASTp searches against the *T. pallidum* protein database in order to identify putative *T. pallidum* periplasmic chaperones that may have a role in ushering OMPs to a putative BAM complex within *T. pallidum*. The results from the BLASTp searches are summarized in Table 4.

Initial BLASTp searches using the full length sequence for *E. coli* SurA did not identify any potential homologues in *T. pallidum.* Subsequent BLASTp searches used smaller portions of the *E. coli* SurA sequence to potentially identify a homologue in *T. pallidum*. Using amino acids 351-428 from *E. coli* SurA for the BLASTp search identified *T. pallidum* protein Tp0862, with an E-value of 3.1, which is not statistically significant. There is a stretch of 43 a.a that share 31 % sequence identity. The most prominent hit for *E. coli* DegP was *T. pallidum* protein Tp0773, with an E-value of 4 x 10^{-45}, which is extremely significant.

It was previously believed that *T. pallidum* protein Tp0327 was a homologue of the *E. coli* protein Skp due to the fact that directly downstream of the *tp0326* gene lays *tp0327*, and this genetic organization is conserved in bacteria that possess both a BamA and Skp protein (Voulhoux & Tommassen, 2004). Initial BLASTp searches using the sequence for the *E. coli* protein Skp against the *T. pallidum* protein database did not identify Tp0327 as a significant hit. For this reason Tp0327 is not listed as a putative homologue to *E. coli* protein Skp in Table 4. Subsequent BLASTp searches using the sequence for *T. pallidum* protein Tp0327 against the protein database identified several

Skp protein homologues in various other organisms such as *Spirochaeta smaragdinae*, several *Bacteroides* species, several *Borrelia* species, and *Victivallis vadensis*.

Initial BLASTp analyses identified *T. pallidum* proteins Tp0327, Tp0862, and Tp0773 as the most prominent homologues for *E. coli* Skp, SurA, and DegP, respectively. These three putative *T. pallidum* chaperone proteins were further analyzed using bioinformatic analyses.

Table 4. *Treponema pallidum* BAM accessory chaperone BLASTp results

E. coli protein	GI number	BLASTp result	E-value*	Sequence identity
Skp	944861	-	-	-
SurA	26245979	Tp0862	3.1	13/43 (31%)
DegP	16128154	Tp0773	4×10^{-45}	109/293 (38%)

* E-value refers to the entire sequence used in the BLASTp search

3.5.1 Investigating *Treponema pallidum* protein Tp0327

The *T. pallidum tp0327* gene is located directly downstream of the *tp0326* gene. This genetic organization is seen in the majority of bacteria which possess a BamA and Skp homologue (Voulhoux & Tommassen, 2004), and both are regulated by the σ^E envelope stress response (Rhodius *et al*., 2006). The sizes of both proteins are very similar, 161 and 172 a.a for Skp and Tp0327 respectively. SignalP 3.0 predicts that both proteins possess a cleavable N-terminal signal sequence, with a probability of 81.8 % for Tp0327. PSORTb localization prediction tool also predicts an N-terminal cleavable signal sequence for Tp0327; however, it predicts the protein could have multiple localization sties. PSORTb predicts cytoplasmic membrane localization with a 49 %

probability and periplasmic localization with a 25 % probability. Clustal-W alignment of Skp and Tp0327, which can be seen in figure 10, shows that the proteins align considerably well, with a high level of similarity at 58 %. The overall sequence identity between the two proteins is low, at 18 %, which could be why Tp0327 was not identified in the BLASTp search using the *E. coli* Skp protein sequence.

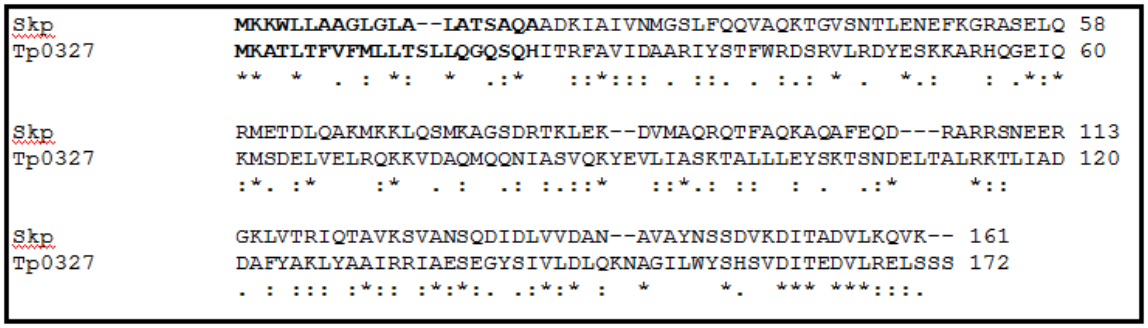


Figure 10. Clustal-W multiple sequence alignment of Skp and Tp0327
Clustal-W multiple sequence alignment of *E. coli* protein Skp and *T. pallidum* protein Tp0327. The SignalP 3.0 predicted signal peptides are shown in bold. Identical residues are designated with a (*), conserved substitutions are designated with a (:), and semi-conserved substitutions are designated with a (.) (Larkin *et al*., 2007).

3.5.2 Investigating *Treponema pallidum* protein Tp0773

The most significant hit from the BLASTp experiment for *E. coli* DegP was *T. pallidum* protein Tp0773 (GI 2611678), with an E-value of 4×10^{-45}, which is extremely significant. The two proteins share 38 % sequence identity and 58 % sequence similarity over a stretch of 293 a.a. SignalP 3.0 predicts that *T. pallidum* protein Tp0773 possesses an N-terminal cleavable signal sequence with a probability of 91.7 %. PSORTb localization prediction tool predicts a 94.4 % probability that Tp0773 is localized in the periplasm. The periplasmic chaperone DegP has been shown to possess both protease and chaperone activity which is regulated in a temperature-dependent manner (Lipinska

et al., 1990; Spiess *et al.*, 1999). The MEROPS protease prediction server (Rawlings, Barrett, & Bateman, 2010) predicts with 100% probability that Tp0773 is part of the subfamily S1B peptidases, which are catalytic serine endopeptidases (Rawlings & Barrett, 1993). The fact that Tp0773 is predicted to possess serine protease activity gives further evidence that it is a putative homologue of *E. coli* DegP. One finding that does not support the theory that Tp0773 is a periplasmic chaperone is that the LipoP lipoprotein prediction tool predicts Tp0773 to be a lipoprotein. LipoP predicts Tp0773 to possess a predicted SpII cleavage site at a.a 17.

3.5.3 Investigating *Treponema pallidum* protein Tp0862

The most prominent hit from the BLASTp experiment for *E. coli* SurA was *T. pallidum* protein Tp0862, with an E-value of 3.1, which is not statistically significant. Further diminishing this result is the fact that the BLASTp was performed using a truncated portion of the *E. coli* SurA protein. Full length *E. coli* SurA did not yield any hits from the BLASTp studies. There were some *T. pallidum* hypothetical protein hits from the BLASTp searches, however, none possessed predicted qualities of a periplasmic chaperone protein.

Escherichia coli protein SurA is annotated as a parvulin-like peptidyl-prolyl *cis-trans* isomerase (PPIases), whereas *T. pallidum* protein Tp0862 is annotated as a FKBP-type PPIase. Although both are peptidyl-prolyl cis-trans isomerases, which catalyze the interconversion of peptidyl-prolyl imide bonds in peptide and protein substrates, the two different types act on different substrates (Gothel & Marahiel, 1999). The Tp0862 protein sequence was used in a BLASTp search against the protein database and several different FKBP-type PPIases from various organisms were identified.

Further analyzing the *T. pallidum* protein Tp0862 by means of Signal P 3.0 signal peptide prediction predicts a cleavable N-terminal signal sequence with a 98.8 % probability. PSORT localization prediction tool also predicts Tp0862 to posses a cleavable N-terminal signal sequence and predicts a 92.6 % probability that Tp0862 is localized in the periplasmic space and a 17.8 % probability of outer membrane localization. In strong contrast, PSORTb predicts with 92.6 % probability that Tp0862 is localized in the cytoplasm.

3.6 Conclusions

Bioinformatic analysis was used in order to investigate the hypothesis that an OMP translocation complex similar to that found in *E. coli* exists within *T. pallidum.* Bioinformatic analysis was also used to analyze the predicted structure of *T. pallidum* protein Tp0326, a putative *E. coli* BamA homologue.

Using bioinformatic analysis, *T. pallidum* protein Tp0326 was determined to be a homologue of the *E. coli* protein BamA. PSORTb localization prediction tool predicted that Tp0326 is localized in the outer membrane. PredictProtein and Phyre structure prediction tools predicted an N-terminal region which possessed 5 independent regions, which conform to the consensus for a POTRA domain, and a C-terminal region composed of alternating β-sheet-loop-β-sheet regions. TMPred predicted the C-terminal region to contain transmembrane regions, which is indicative of a β-barrel region. The predicted N-terminal POTRA domains and C-terminal β-barrel regions conforms to the structure of BamA proteins. As BamA proteins are highly conserved throughout Gram-negative bacteria it is not surprising that *T. pallidum* possesses a similar protein. In other Gram-negative bacteria, BamA is the central protein in a multi-protein complex

composed of accessory lipoproteins and periplasmic chaperones. It is therefore quite likely that *T. pallidum* possesses some degree of BamA complexing proteins.

The search for putative *T. pallidum* homologues to the *E. coli* BAM lipoproteins was not very successful. Only low E-value hits were observed for BamB and BamD and no significant hits were found for BamC and BamE. Structural prediction tools predicted *T. pallidum* protein Tp0133 to possess similar secondary structure to *E. coli* BamB (β-propeller fold with seven or eight blades) and *T. pallidum* protein Tp0622 to possess similar secondary structure to *E. coli* BamD (multiple TPR motifs). The predicted structures of Tp0133 and Tp0622 give good evidence that these are indeed homologues to the *E. coli* lipoproteins BamB and BamD. The absence of BamC and BamE homologues in *T. pallidum* could be attributed to the inherently smaller genome of *T. pallidum* in comparison to *E. coli*. As well, BamD is the only lipoprotein in *E. coli* shown to be essential for the process of OMP folding and deposition into the outer memberane (Malinverni *et al*., 2006). The fact that *T. pallidum* was identified to possess a homologue to the only essential lipoprotein in the *E. coli* BAM complex strengthens the finding.

The search for putative *T. pallidum* homologues to the *E. coli* BAM associated periplasmic chaperones was slightly more successful than for the BAM lipoproteins. Putative *T. pallidum* homologues for all three of the *E. coli* periplasmic chaperones were identified: Tp0327 (Skp), Tp0773 (DegP), and Tp0862 (SurA). There was some ambiguity surrounding the bioinformatic analyses of Tp0773 and Tp0862.

Most of the bioinformatic evidence suggests that *T. pallidum* protein Tp0773 is a homologue of the *E. coli* periplasmic chaperone DegP; however it was also predicted to

be a lipoprotein by the LipoP prediction program. As the LipoP prediction tool is purely just prediction, it is possible that Tp0773 is not a lipoprotein. As well, due to the fact that *T. pallidum* is evolutionarily very distant from *E. coli*, and possesses a much smaller genome, it is possible that in *T. pallidum* this protein could serve a dual function. It is possible that even as a lipoprotein, Tp0773 could serve the same function as DegP.

Similarily to Tp0773, most of the bioinformatic evidence suggests that *T. pallidum* protein Tp0862 is a homologue of the *E. coli* periplasmic chaperone SurA. Complicating the finding is that the *E. coli* chaperone protein SurA and *T. pallidum* protein Tp0862 are believed to interact with different types of substrates. Parvulin-like PPIases, such as SurA, have been shown to interact with non-native periplasmic outer membrane protein folding intermediates (Hennecke *et al*., 2005), whereas FKBP-type PPIases are normally cytoplasmic (Gothel & Marahiel, 1999). As well, there was some discrepancy between PSORT and PSORTb as to predicted location of Tp0862. PSORT predicted Tp0862 to be periplasmic whereas PSORTb predicted it to be cytoplasmic. If the protein is indeed a periplasmic chaperone, it is possible given the vast differences between *T. pallidum* and *E. coli* that Tp0862 indeed functions as a BAM accessory chaperone. Arguing against *T. pallidum* protein Tp0862 as a periplasmic chaperone, is the notion that the there is functional redundancy amongst the *E. coli* periplasmic chaperones (Sklar *et al*., 2007). The *T. pallidum* genome is much smaller than the *E. coli* genome and it is likely that efficiency is of a greater concern in the *T. pallidum* genome. It could then be plausible that *T. pallidum* would remove this functional redundancy and not possess homologues of all three *E. coli* periplasmic chaperones.

The bioinformatic findings suggest that *T. pallidum* possesses proteins similar to those found in the *E. coli* BAM complex. The overall make up of the *T. pallidum* BAM complex and the number of proteins involved remains unclear from these studies. It appears from the bioinformatic analyses that the putative *T. pallidum* BAM complex is simpler than the *E. coli* BAM complex in terms of the number of proteins involved in the respective complexes.

Chapter 4: Characterizing the putative *Treponema pallidum* BAM complex using protein-protein interaction studies

4.1 Contributions to the data

My contributions to the data presented in this chapter include: cloning and construct creation for *T. pallidum tp0327*-pET28a, *tp0622*-pET28a, and *tp0773*-pET28a, soluble protein expression and purification for *T. pallidum* Tp0326-POTRA1-5, Tp0327, Tp0622, and Tp0773, *T. pallidum* lambdaZAP assay development, *T. pallidum* lambdaZAP assay, and Far-Western blot assays. The *T. pallidum tp0326-POTRA1-5*-pET28a construct was provided by Dr. Justin Radolf and Dr. Dan Desrosiers at the University of Connecticut Health Center, USA. The identification of the optimal expression conditions for the *tp0326-POTRA1-5*-pET28a construct was included in my contributions to the data.

4.2 Introduction

The interactions between proteins within the cell are important for the majority of biological functions. Specific examples of protein-protein interactions (PPIs) which are critical to properly functioning cellular processes include extracellular receptor/sensor PPIs and structural PPIs which are responsible for cell shape maintenance (Nooren & Thornton, 2003). PPIs between scaffold proteins represent a common mechanism used to control spatial and temporal aspects of cellular signalling (Iqbal *et al.*, 2010). For example, epidermal growth factor (EGF) induced Ras activation results in different outcomes dpending on the scaffold proteins used to induce Ras activation (Kholodenko *et al.*, 2010).

Protein-protein interactions are of central importance for virtually every process in a living cell, ranging from metabolism to structure (Miernyk and Thelen, 2008). Elucidating both individual protein associations and complex protein interaction networks, while challenging, is an essential goal of functional genomics. For example, discovering interacting partners for a 'protein of unknown function' can provide insight into actual function far beyond what is possible with sequence-based predictions (Miernyk and Thelen, 2008). If a previously unknown protein is observed to interact with other proteins or a complex of known function, then one has gained insight into the function of the unknown (Oliver, 2000). As well, identifying interactions between proteins involved in common cellular functions is a way to get a broader view of how they work cooperatively in a cell (Nooren & Thornton, 2003).

A protein generally resides in a crowded environment with many potential binding partners with different surface properties (Nooren & Thornton, 2003). Proteins have direct physical interactions at several different levels, including; the multi-subunit enzymic nanomachines (Mitra and Frank, 2006), cohort and client protein binding by molecular chaperones (Chapman *et al*., 2006), client protein binding by protein kinases (Chevalier and Walker, 2005), and multiple protein interactions in metabolic networks (Dhar-Chowdhury *et al*., 2005). Most proteins are very specific in their choice of partner, although some are multi-specific, having multiple binding partners on coinciding or overlapping interfaces which make up complexes (Nooren & Thornton, 2003). Discovering the PPIs that make up these complexes is an important technique for studying protein function.

Various approaches allow identification of physical PPIs in complexes (Chen & Xu, 2003; Phizicky & Fields, 1995). High-throughput techniques often used for screening for new unidentified PPIs include the yeast two-hybrid technique (Fields & Song, 1989), arrays of immobilized proteins (Bertone & Snyder, 2005), and phage display (Sidhu *et al*., 2003). More direct techniques are used to verify known or putative PPIs such as Co-immunoprecipitation and Far-Western blotting. Mass spectrometry based measurements of protein-protein interactions have also been used across the proteome in *Saccharomyces cerevisiae* (Gavin *et al.*, 2002) and in *Escherichia coli* (Arifuzzaman *et al.*, 2006).

Screening and direct PPI techniques were used in the research involved in this thesis in order to determine the possible presence of a *T. pallidum* BAM complex and to gain insight into OMP biogenesis in *T. pallidum*. A *T. pallidum* lambda genomic expression library was initially screened against recombinant *T. pallidum* protein Tp0326-POTRA1-5 in order to identify interacting *T. pallidum* proteins. Identified proteins in this screen have the potential to be interacting *T. pallidum* BAM complex proteins or *T. pallidum* OMPs, which are ushered to the OM through this complex. Far-Western blot analysis was used as a direct PPI technique in order to verify the interactions found in the expression library screen.

These PPI studies aim to support the hypothesis that an OMP translocation complex similar to that found in *E. coli* exists within *T. pallidum* and that this complex would be responsible for ushering *T. pallidum* OMPs to the bacterial surface.

4.3 Cloning and protein expression for Tp0326-POTRA1-5

The *T. pallidum tp0326-POTRA1-5*-pET28a construct was provided by collaborators Dr. Justin Radolf and Dr. Dan Desrosiers at the University of Connecticut Health Center, USA. This construct consists of Tp0326 a.a residues 22 – 434, which corresponds specifically to POTRA domains 1 through 5 without the N-terminal signal peptide. The sequence was cloned into the pET28a vector using engineered restriction sites XhoI and NdeI, which insert the sequence into pET28a directly downstream of an N-terminal 6x His-tag. The pET28a vector possesses a lac operator sequence upstream of the inserted sequence, which is inducible by IPTG, as well as a kan resistance cassette. The pET28a vector fuses the following N-terminal a.a residues to the expressed protein product, MGSSHHHHHHSSGLVPRGSHM, which result in an increased MW of 2312.53 Da. The expected MW of the protein product expressed using this construct is 49285.87 Da.

The *T. pallidum tp0326-POTRA1-5*-pET28a construct was transformed into *E. coli* BL21 Star^{TM} DE3 cells and tested for small scale soluble protein expression. It was determined from small scale expression that Tp0326-POTRA1-5 expressed best at 16 ºC, inducing with 0.4mM IPTG at OD_{600} 1.6, and growth O/N. The construct was then expressed in large scale and purified on a HisTrap column using an AKTAprime FPLC followed by gel-filtration chromatography. The gel-filtration chromatograph depicting the Tp0326-POTRA1-5 elutions off the gel-filtration column can be seen in Figure 11. Elutions from gel-filtration were analyzed via SDS-PAGE, which can be seen in Figure 12.

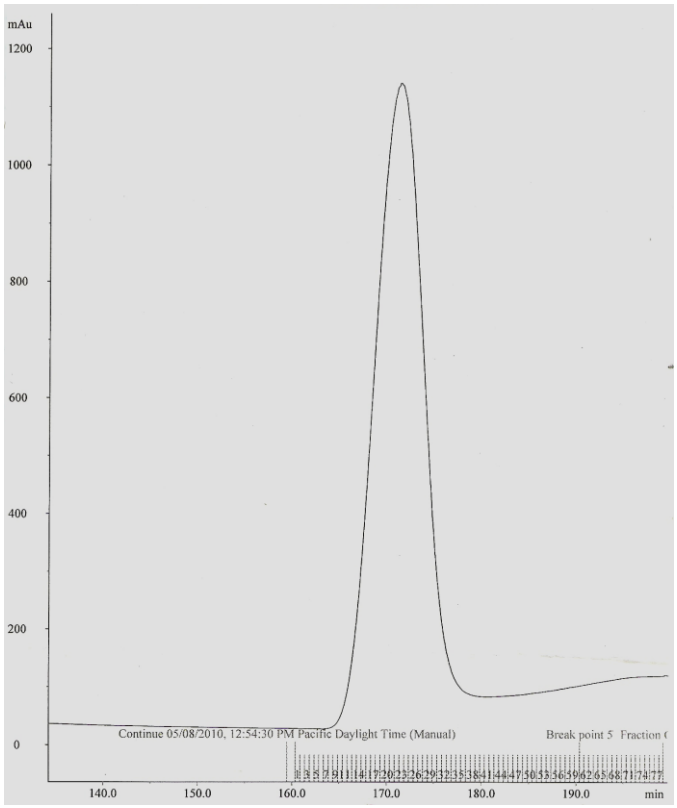


Figure 11. Gel filtration chromatograph for Tp0326-POTRA1-5
AKTAprime gel-filtration chromatograph for *T. pallidum* protein Tp0326-POTRA1-5. The y-axis shows the intensity of the protein eluted in mAu and the x-axis shows the elution time in mins. Tp0326-POTRA1-5 eluted over ~15 mins with a max intensity of ~1150 mAu.

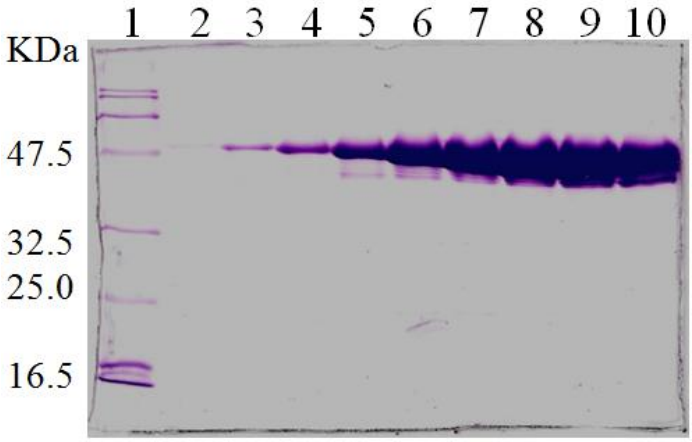


Figure 12. SDS-PAGE analysis of Tp0326-POTRA1-5
Analysis of gel-filtration elutions of *T. pallidum* protein Tp0326-POTRA1-5 using 15% SDS-PAGE. The apparent MW of Tp0326-POTRA1-5 is ~48-50 KDa. All elutions were pooled for futher use. Lane 1 – unstained broad range MW ladder, 2-10 Tp0326-POTRA1-5 elutions.

The Tp0326-POTRA1-5 protein was pooled and then the cryo-protectant trehalose was added to a final concentration of 7.5 % (De Carlo *et al*., 1999). Protein concentration was determined to be 1.15 mg/mL using the Pierce BCA protein concentration kit. Protein was then aliquoted and flash-frozen in liquid before being stored at -20 ºC for further use.

4.4 Lambda ZAP assay using Tp0326-POTRA1-5

4.4.1 Introduction

The Lambda ZAP system combines the high efficiency of lambda phage packaging, to create high titer libraries, with the convenience of a plasmid system, making it easier to analyze your clones. The Lambda ZAP system allows for *in vivo* rapid excision of the pBluescript phagemid, which contains the desired DNA insert (Short *et al.*, 1988). The Lambda ZAP vector can be seen in Figure 13. Once excised and isolated, the pBlueSK phagemid has 6 different primer sites for DNA sequencing, including T7 and M13 sites. The phagemid, which is seen in Figure 14, also has the bacteriophage f1 origin of replication, allowing rescue of single-stranded DNA, which can be used for DNA sequencing or site-directed mutagenesis (Stratagene Lambda ZAP II manual).

An extremely important feature of this system is that the Lambda ZAP vector allows for protein expression of your inserted coding sequence. The coding sequence is expressed from the lacZ promoter as a fusion protein with the β-galactosidase α-fragment using IPTG as an inducer (Stratagene Lambda ZAP II manual).

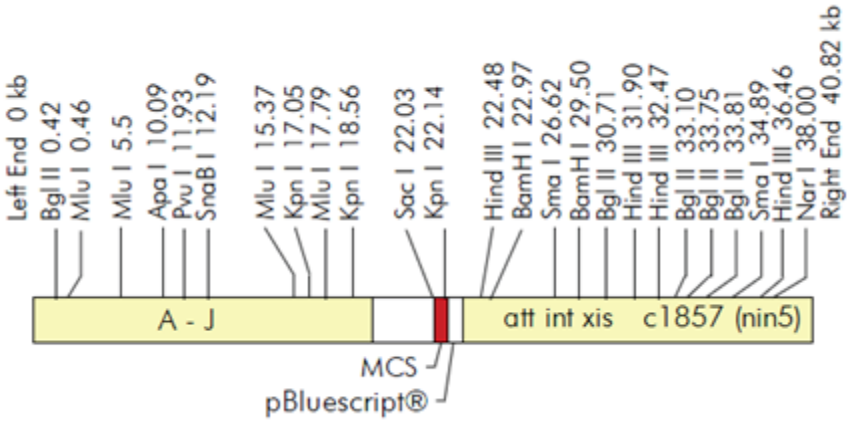


Figure 13. Map of the Lambda ZAP insertion vector
Map of the Lambda ZAP insertion vector depecting the multiple cloning site (MCS) and the region pertaining to the excisable pBluescript phagemid. The vector also possesses the *att*, *int*, and *xis* genes necessary for excision of the phagemid (Stratagene Lambda ZAP II manual).

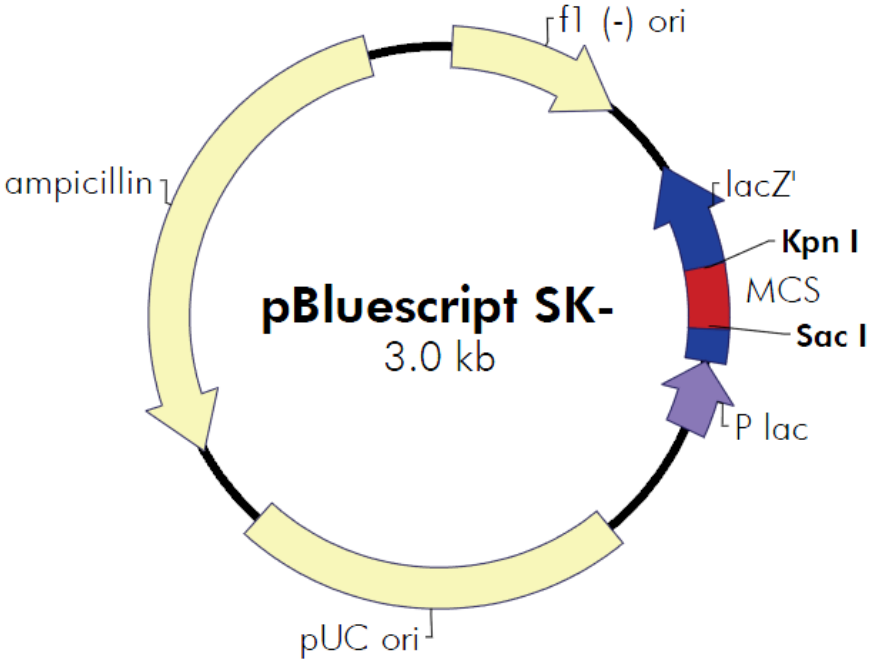


Figure 14. Lambda ZAP pBluescript SK(-) phagemid map
Map of the Lambda ZAP pBluescript SK (-) phagemid. The phagemid contains an f1 (-) origin of ss-DNA replication, pUC origin of replication, ampicillin resistance cassette, *lac* promoter, and a β-galactosidase α-fragment coding sequence which is where the MCS is inserted (Stratagene Lambda ZAP II manual).

This system uses *E.coli* XL1 Blue MRF' as the host strain because it contains a *lac* repressor which blocks transcription from the lacZ promoter in absence of IPTG (Short *et al.*, 1988). This host strain also possesses the genes forming the F´ pili found on the surface of the bacteria which is required for filamentous helper phage infection and *in vivo* excision of the phagemid (Stratagene Lambda ZAP II manual).

Excision of the pBluescript SK phagemid requires co-infection of the *E.coli* host strain with both the Lambda phage and the ExAssist helper phage (Stratagene Lambda ZAP II manual). The helper phage expresses proteins which recognize and nick the *int* DNA sequence within the Lambda ZAP vector. Another helper phage protein then synthesizes new DNA along the vector until it reaches the termination site, which is directly downstream of the pBluescript SK phagemid sequence. A final helper phage protein then circularizes the single stranded DNA molecule forming the pBluescript SK phagemid (Short *et al*., 1988). The pBluescript SK phagemid is isolated by transforming the *in vivo* excision reaction products into *E. coli* SOLR cells. Only the pBluescript phagemid can grow in *E. coli* SOLR cells because the Lambda and helper phage have an amber mutation that prevent replication in the nonsuppressing *E. coli* SOLR cells (Stratagene Lambda ZAP II manual). After successful transformation the pBluescript phagemids can be isolated from the *E. coli* SOLR cells and the DNA can be sequenced.

4.4.2 Screening the Lambda ZAP library

The *T. pallidum* Lambda ZAP library was originally created by Dr. Cameron (Cameron *et al*., 1998). The library was titered and found to be 4.5×10^9 pfu/mL. In order to adequately screen the *T. pallidum* genome we need to plate a high density of

plaques. Using the knowledge that the *T. pallidum* genome is roughly 1.14 million bp and the smallest possible Lambda ZAP DNA insert size is 500 bp, accounting for the different reading frames, orientation within the phage, and triplicate coverage would mean 4.1 x 10^4 pfu would need to be screened in this assay (1.14 x 10^6 (bp in the genome) / 500 (bp in DNA insert) x 3 (reading frames) x 2 (phage orientation) x 3 (triplicate) = 4.1 x 10^4). In order to screen this number of plaques we utilized a dual screen approach. A primary screen was first used with a high density of plaques, where the objective was to discover plaques that demonstrate interaction with our bait protein. The positive reacting plaque along with the surrounding plaques on the agar plate were excised and subjected to a secondary screen. The secondary screen used a greater dilution of the isolated primary plaque in order to achieve plaque isolation, where the objective was to discover plaques that interact specifically with the bait protein.

The primary screen of the *T. pallidum* Lambda ZAP library was conducted using 5 x 10^4 pfu spread out over four 150mm NZY agar plates. Plates were incubated for 5 hours at 37 ºC before being overlayed with nitrocellulose membranes impregnated with IPTG. After O/N incubation at 37 ºC with the nitrocellulose membranes, the membranes were subjected to a Far-Western blot like assay using recombinant Tp0326-POTRA1-5 as the bait protein and antibodies reactive to the bait protein. The membranes were developed using AP development and analyzed visually. A primary screen membrane can be seen in Figure 15. From the 4 membranes analyzed there were a total of 6 reactive plaques. The negative control membrane which was incubated with 1X PBS, instead of Tp0326-POTRA1-5 protein, did not show any positive reacting plaques. The 6 reactive

plaques were then individually cored out from the NZY agar plates and subjected to a secondary screen.

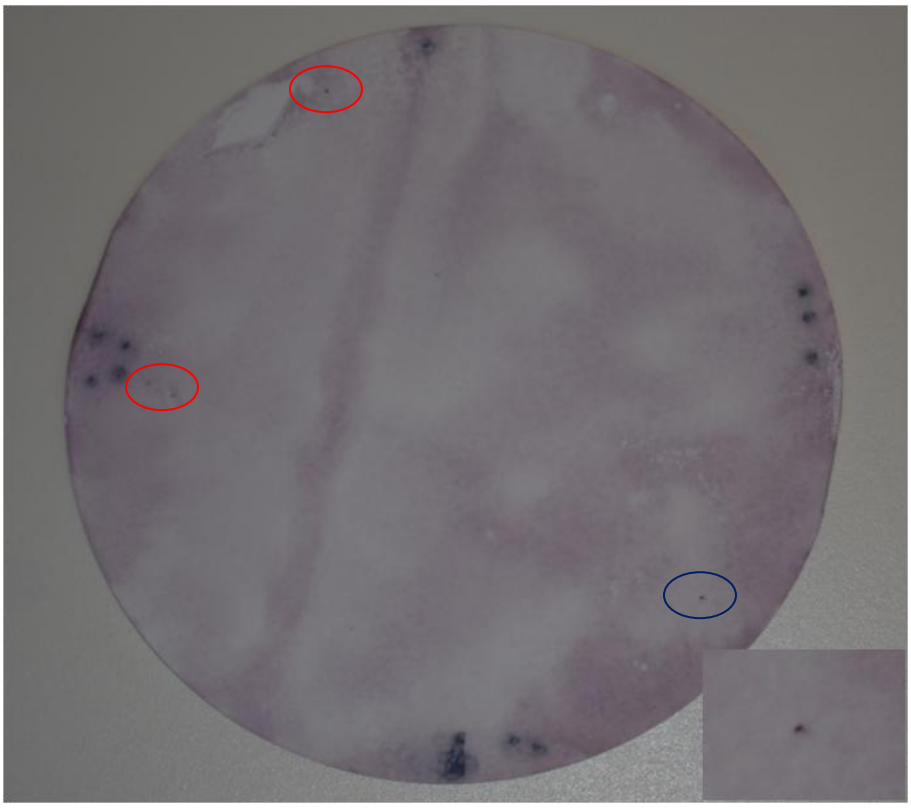


Figure 15. AP developed Lambda ZAP nitrocellulose membrane
Nitrocellulose membrane incubated with *T. pallidum* Lambda ZAP library O/N and subsequently subjected to Far-Western blot like assay using AP development. The bait protein used was recombinant Tp0326-POTRA1-5 (150µg/mL), 1º Ab – chicken-anti-Tp92, 2º Ab – goat-anti-chicken-AP. Black markings are waterproof ink orientation marks. Red and blue circles surround 4 positive reacting plaques. Plaque circled in blue is shown as an inset on the bottom right corner (with magnification).

The secondary screen used the phage which were cored out from the primary screen plates rather than the phage from the original Lambda ZAP library. In total 6 cores were taken from the original primary screen plates but only 4 were subjected to the secondary screen due to time constraints; the 4 plaques which showed the greatest intensity during AP-development were chosen. The phage contained in the 4 cores were titered and each core was subjected to an individual secondary screen, which screened

500 pfu spread out over two 150mm NZY agar plates. The secondary screen was performed in exactly the same manner as the primary screen except for the fact that each NZY plate contains 250 pfu compared to 12,500 pfu in the primary screen. Plaques were well isolated on the secondary screen plates. Each core taken from the primary screen contained one positive reacting plaque and several non-reacting plaques which surrounded it on the NZY plate. The nitrocellulose membranes from the secondary screen were treated to the same Far-Western blot like assay using recombinant Tp0326-POTRA1-5 as the bait protein. Positive reacting plaques which were isolated on the NZY plates were then cored out and subjected to *in vivo* excision and DNA sequencing.

4.4.3 Lambda ZAP DNA sequencing results

The DNA sequences obtained refer to that of the insert in the Lambda ZAP vector, whose expressed protein product interacted with recombinant Tp0326-POTRA1-5. The DNA sequences were identified using NCBI BLAST against the *T. pallidum* genome. The results of the DNA sequencing and subsequent NCBI BLAST experiments can be seen in Table 5. Sequence coverage refers to the portion of the *T. pallidum* genome that was identified in the Lambda insert DNA. Coverage refers to the portion of the matched *T. pallidum* gene that was identified in the Lambda insert DNA and a description of the protein function is also listed.

Table 5. Lambda ZAP DNA sequencing and NCBI BLAST results

Core	Sequence coverage	ORF match	Coverage (%)	Description
1	925652-928087 (2435bp)	*tp0854*	53	Hypothetical protein
2	269094-269994 (900bp)	*tp0257/tp0258*	16, 94	GPD*, hypothetical protein
3	815394-819816 (4422bp)	*tp0750-tp0754*	42, 100, 100, 100, 100	ECM adhesin, ECM adhesin, hypothetical proteins
4	675723-676235 (512bp)	*tp0622*	29	Hypothetical protein (putative BamD homologue)

* Glycerophosphodiester phosphodiesterase

The first core taken from the primary screen was identified to be a hypothetical protein, Tp0854; the sequence coverage of this gene found in the Lambda DNA insert was 53 %. There was some ambiguity amongst *in silico* prediction tools as to putative localization of this protein in *T. pallidum*. SignalP 3.0 signal peptide prediction tool predicts an 85 % probability that Tp0854 has a cleavable N-terminal signal peptide using neural networks models, however does not predict a signal peptide using hidden Markov models. PSORTb localization tool also predicts a cleavable N-terminal signal sequence, whereas PSORT does not predict a signal sequence. PSORTb predicts a 100 % probability that Tp0854 is localized in the cytoplasmic membrane as it predicts Tp0854 to possess 4 transmembrane α-helices. PSORT also predicts localization in the cytoplasmic membrane with a 61.8 % probability. It seems likely from these analyses

that Tp0854 is localized in the cytoplasmic membrane of *T. pallidum*; it is unclear whether it is orientated to the cytoplasm or periplasmic space.

The second core was identified to be a stretch of the *T. pallidum* genome which corresponded to both *tp0257* and *tp0258*; only 16 % of the *tp0257* gene was covered in the Lambda insert DNA compared to 94 % for *tp0258*. *Treponema pallidum* protein Tp0257 is annotated as a glycerophosphodiester phosphodiesterase (GPD) and Tp0258 is annotated as a hypothetical protein. Tp0257 was previously identified from a serologic screen as being reactive with syphilis patient sera, and was predicted by bioinformatic analyses to be a putative outer membrane protein (Cameron *et al*., 1998; Stebeck *et al.*, 1997). In a later study Tp0257 was shown to be a lipoprotein associated with the cytoplasmic membrane of *T. pallidum* (Shevchenko *et al*., 1999). The second core was comprised mainly of DNA from *tp0258*. PSORTb predicts an 89.6 % probability that Tp0258 is localized in the cytoplasm and PSORT predicts the same localization, with a 34 % probability. As well, neither localization prediction program nor SignalP 3.0, predict a cleavable N-terminal signal sequence. It seems likely based on these predictions that Tp0258 is a cytoplasmic protein.

The third core was identified to be a large stretch of the *T. pallidum* genome which corresponded to the genes *tp0750 – tp0754*. Only 42 % of the *tp0750* gene was covered in this Lambda DNA insert whereas 100 % of the genes *tp0751 – tp0754* were covered. Tp0752, Tp0753, and Tp0754 are all hypothetical proteins. Tp0752 is predicted to possess an N-terminal cleavable signal sequence by SignalP 3.0 with a probability of 64.1 %; PSORTb and PSORT also predict a signal sequence. Localization of Tp0752 is unclear as PSORTb predicts a 49 % probability that it is localized in the

cytoplasmic membrane, whereas PSORT only predicts a 17.4 % probability of this localization. Tp0753 is predicted to possess an N-terminal cleavable signal sequence by SignalP 3.0 with a probability of 57 % using neural network models. PSORTb predicts Tp0753 to be localized in the cytoplasmic membrane with 98.2 % probability and does not predict a signal sequence. In strong contrast, PSORT predicts an N-terminal cleavable signal sequence and predicts Tp0753 to be localized in the periplasm with 94.1 % probability. Tp0754 is not predicted to possess an N-terminal cleavable signal sequence by any of the prediction tools used. PSORTb predicts with 99.7 % probability that Tp0754 is a cytoplasmic protein. Tp0750 (Cameron, unpublished findings) and Tp0751 have both been shown to possess the ability to bind to host ECM components (Cameron, 2003; Cameron *et al*., 2005). As well, heterologous expression of Tp0751 in the spirochete *Treponema phagedenis* showed that Tp0751 is surface exposed (Cameron *et al*., 2008). Although there is no evidence that Tp0750 is surface exposed, SignalP 3.0 predicts a 97.7 % probability that it possesses an N-terminal cleavable signal sequence. As well, PSORT predicts a 91.7 % probability of localization in the periplasm and 8.3 % probability of localization in the outer membrane.

The fourth core was identified to be *T. pallidum* gene *tp0622*. This was a relatively small DNA insert at 512bp and covered only 29 % of the *tp0622* gene. *Treponema pallidum* protein Tp0622 was originally identified to be a putative *E. coli* BamD homologue by bioinformatic analyses. There is a 40 a.a stretch in Tp0622 which shares 42 % sequence identity with *E. coli* BamD. This 40 a.a stretch is contained within the 170 a.a stretch of Tp0622 which was identified in the Lambda ZAP screen. It is likely that this region could constitute a conserved BamA binding region.

4.5 Recombinant soluble protein expression and purification

The putative *T. pallidum* OMPs Tp0750 and Tp0751, putative *T. pallidum* BamD homologue Tp0622, and putative *T. pallidum* BAM associated periplasmic chaperone Tp0327 (Skp), were all cloned, solubly expressed, and purified for use in further PPI experiments in order to verify the findings from the Lambda ZAP screen and bioinformatic analyses. The putative *T. pallidum* BAM associated periplasmic chaperone Tp0773 (DegP) was cloned and tested for soluble protein expression. Putative *T. pallidum* OMPs Tp0750 and Tp0751 were kindly donated by Teresa Brooks and Rebecca Hof from the University of Victoria (Victoria, British Columbia).

4.5.1 Expression of *Treponema pallidum* protein Tp0327

Treponema pallidum gene *tp0327* was amplified by PCR, with an annealing temperature of 58 ºC and the following primers: CATGACCATATGATCACGCGCTTT GCCGTC (5') and GTCAGCTCGAGTCACGAGCTGCTCAGCTC (3'). The gene was cloned into pET28a without the predicted N-terminal signal sequence, using engineered restriction enzyme sites NdeI and XhoI. The a.a residues which are contained in the final expressed protein product are 23 – 172. The expected MW of the expressed protein is 19551.14 Da, including the N-terminal 6x His-tag. The optimal soluble protein expression conditions, as determined by small-scale expression, were growth at 23 ºC O/N and inducing with 1mM IPTG at OD_{600} 0.8. The construct was then expressed in large scale and purified on a HisTrap column using an AKTAprime FPLC followed by gel-filtration chromatography. Elutions from gel-filtration were analyzed via SDS-PAGE, Figure 16, and protein concentration was determined to be 910 µg/mL using the

Pierce BCA protein concentration kit. Protein was then aliquoted and flash-frozen in liquid nitrogen before being stored at -20 ºC for further use.

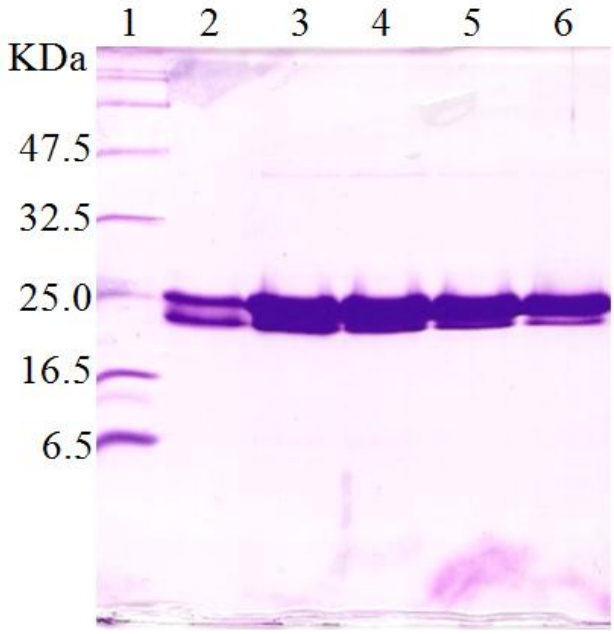


Figure 16. SDS-PAGE analysis of recombinant Tp0327
Analysis of gel-filtration elutions of *T. pallidum* protein Tp0327 using 15% SDS-PAGE. The apparent MW of Tp0327 is ~20-25 KDa. All elutions were pooled for futher use. Lane 1 – unstained broad range MW ladder, 2 - 6 Tp0327 elutions.

4.5.2 Expression of *Treponema pallidum* protein Tp0773

Treponema pallidum gene *tp0773* was amplified by PCR, with an annealing temperature of 58 ºC and the following primers: CTAGACCATATGGTGGGCTTCTTT CTAGGAAG (5') and GTCAGCTCGAGCTACTCATCTGAGCGTTCTG (3'). The gene was cloned into pET28a without the predicted N-terminal signal sequence, using engineered restriction enzyme sites NdeI and XhoI. The a.a residues which are contained in the final expressed protein product are 20 – 398. The expected MW of the expressed protein is 42607.18 Da, including the N-terminal 6x His-tag. From the small scale

expression trials it was determined that this construct did not express well under any condition. The *T. pallidum* protein Tp0773 was further analyzed using bioformatics, and was found to possess a predicted N-terminal transmembrane α-helix up to a.a residue 29. Transmembrane regions often do not express well so a new pET28a construct was created which incorporated a.a residues 30-398, thus removing the predicted transmembrane region. The new construct was amplified by PCR, with an annealing temperature of 55 ºC and the following primers: CTAGACCATATGTTCTCTGCTAGG TCCTCGG (5') and GTCAGCTCGAGCTACTCATCTGAGCGTTCTG (3'). The gene was cloned into pET28a without the predicted N-terminal signal sequence, using engineered restriction enzyme sites NdeI and XhoI. The protein product of the second construct, Tp-0773.30 – 398, had an expected MW of 41363.75 Da.

The second construct did not express well under any condition during small-scale expression trials. Large scale expression and soluble protein purification was still attempted, using growth conditions of 16 ºC O/N and protein induction with 0.4mM IPTG at OD_{600} 1.3. Protein was purified on a HisTrap column using an AKTAprime FPLC and analyzed using AKTAprime Primeview software. It was determined that the *tp0773.30-398*-pET28a construct did not express soluble protein. Small amounts of insoluble protein could be visualized on the SDS-PAGE gels from small scale protein expression.

4.5.3 Expression of *Treponema pallidum* protein Tp0622

Treponema pallidum gene *tp0622* was was amplified by PCR, with an annealing temperature of 55 ºC and the following primers: CTAGACCATATGCAAATTCA CAGCGCGCG (5') and GTCAGCTCGAGTTAGTACGCGCGAATAAGTAATTC (3').

The gene was cloned into pET28a using engineered restriction enzyme sites NdeI and XhoI. The portion of the *tp0622* gene that was cloned into pET28a for expression is the region of *tp0622* which was identified in the Lambda ZAP screen, which corresponds to a.a residues 277 – 452. This region also contains the 40 a.a stretch which shows ~50 % sequence identity to *E. coli* BamD. The expected MW of the expressed protein is 22437.19 Da, which includes the N-terminal 6x His-tag. The optimal soluble protein expression conditions, as determined by small-scale expression, were growth at 16 ºC O/N and inducing with 0.4mM IPTG at OD_{600} 1.6. The construct was then expressed in large scale and purified on a HisTrap column using an AKTAprime FPLC. It was determined that the construct did not express soluble protein well, however, minimal soluble expression was observed. Large scale expression was repeated and the soluble protein obtained from HisTrap column purification was pooled with that from the original purification, and together concentrated for gel-filtration chromatography. Elutions from gel-filtration were analyzed via SDS-PAGE, which can be seen in Figure 17. The Tp0622 protein was pooled and the cryo-protectant trehalose was added to a final concentration of 7.5 % (De Carlo *et al*., 1999). Protein concentration was determined to be 140 µg/mL using the Pierce BCA protein concentration kit. Protein was then aliquoted and flash-frozen in liquid nitrogen before being stored at -20 ºC for further use.

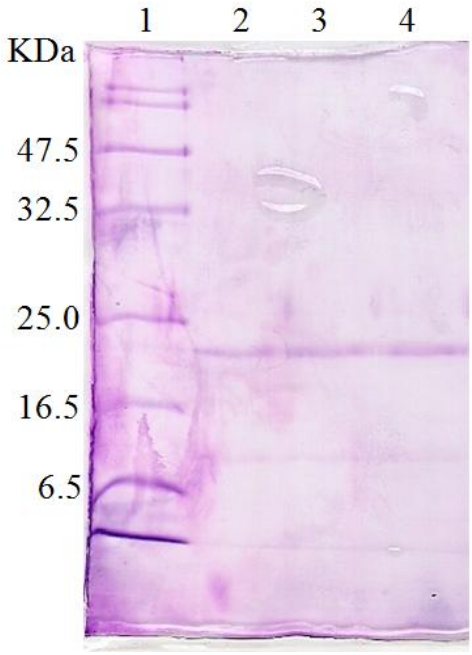


Figure 17. SDS-PAGE analysis of recombinant Tp0622
Analysis of gel-filtration elutions of *T. pallidum* protein Tp0622 using 15% SDS-PAGE. The apparent MW of Tp0622 is ~20-22 KDa. All elutions were pooled for futher use. Lane 1 – unstained broad range MW ladder, 2 - 3 Tp0622 elutions.

4.6 Protein-protein interactions through Far-Western blot analysis

4.6.1 Introduction

Screening and direct PPI techniques were used in this project in order to gain a better understanding of the *T. pallidum* BAM complex and OMP biogenesis in *T. pallidum*.

A *T. pallidum* Lambda genomic expression library was initially screened against recombinant *T. pallidum* protein Tp0326-POTRA1-5 in order to identify interacting *T. pallidum* proteins. Identified proteins in this screen have the potential to be interacting *T. pallidum* BAM complex proteins or *T. pallidum* OMPs, which are ushered to the OM through this complex. The *T. pallidum* Lambda ZAP screen identified several

hypothetical proteins (Tp0258, Tp0752, Tp0753, Tp0754, and Tp0854), ECM binding adhesins (Tp0750 and Tp0751), a Gpd protein (Tp0257), and a putative BamD protein (Tp0622).

Far-Western blot analysis was used as a direct PPI technique in order to verify the interactions found in the expression library screen, which included Tp0750, Tp0751, and Tp0622. The combined background information available and bioinformatic analyses best supported the putative interactions between Tp0750, Tp0751, Tp0622 and Tp0326-POTRA1-5. Due to time constraints only these proteins which were identified in the Lambda ZAP screen were further analyzed via Far-Western blotting. Far-Western blotting was also used to investigate a potential interaction between Tp0326-POTRA1-5 and the putative *E. coli* Skp homologue, *T. pallidum* protein Tp0327. These Far-Western blot studies aimed to support the hypothesis that an OMP translocation complex similar to that found in *E. coli* exists within *T. pallidum* and that this complex would be responsible for ushering *T. pallidum* OMPs to the bacterial surface.

4.6.2 Far-Western blot analysis

Recombinant *T. pallidum* proteins and BSA (negative control protein) were run on 15% SDS-PAGE gels, 2µg of each protein, and transferred to nitrocellulose membranes using a semi-dry immunoblotter; duplicate 15% SDS-PAGE gels were also run and stained with Coomassie-blue dye for total protein visualization (Figure 18). The putative *T. pallidum* OMP Tp0453 was also included as a negative control protein.

Treponema pallidum protein Tp0453 has been shown to be an integral OMP with characteristics that are unprecedented among bacterial OMPs of known structure. TP0453 lacks extensive β-sheet structure and does not traverse the OM to become surface

exposed (Hazlett *et al*., 2005). Lipid modification provides an OM tether for this partially amphiphilic polypeptide that contains multiple membrane inserting, amphipathic α-helices (Hazlett *et al.*, 2005). Due to these inherent differences between Tp0453 and β-barrel outer membrane proteins, it is unlikely that Tp0453 is ushered to the outer membrane through the putative *T. pallidum* BAM complex. While it is still possible that Tp0453 interacts with Tp0326, it poses as a good negative control for *T. pallidum* recombinant OMPs.

Information on the *T. pallidum* proteins tested for a PPI with Tp0326-POTRA1-5 can be found in Table 6. Following transfer, nitrocellulose membranes were initially treated to a Western blot assay and subsequently to a Far-Western blot assay as outlined below.

Western blotting using AP development was used in order to optimize the dilution of the primary antibody to be used in the Far-Western blots. The optimal dilution for the primary antibody was determined to be 1:2000.

Far-Western blot assays were conducted using Tp0326-POTRA1-5 as the bait protein, at a concentration of 250µg/mL, and 1X PBS as a negative control. Tp0326-POTRA1-5 protein or 1X PBS was incubated with separate nitrocellulose membranes for 2 hours at RT and then incubated with the primary antibody chicken-anti-Tp0326 for 1 hour. The blots were then incubated with the secondary antibody rabbit-anti-chicken IgG IRDye800® conjugated from Rockland Immunochemicals (Gilbertsville, Pennsylvania) and developed using the Odyssey® infrared imaging system from Li-Cor Biosystems (Lincoln, Nebraska). The Western and Far-Western blots as well as the coomassie stained gels can be seen in Figure 18.

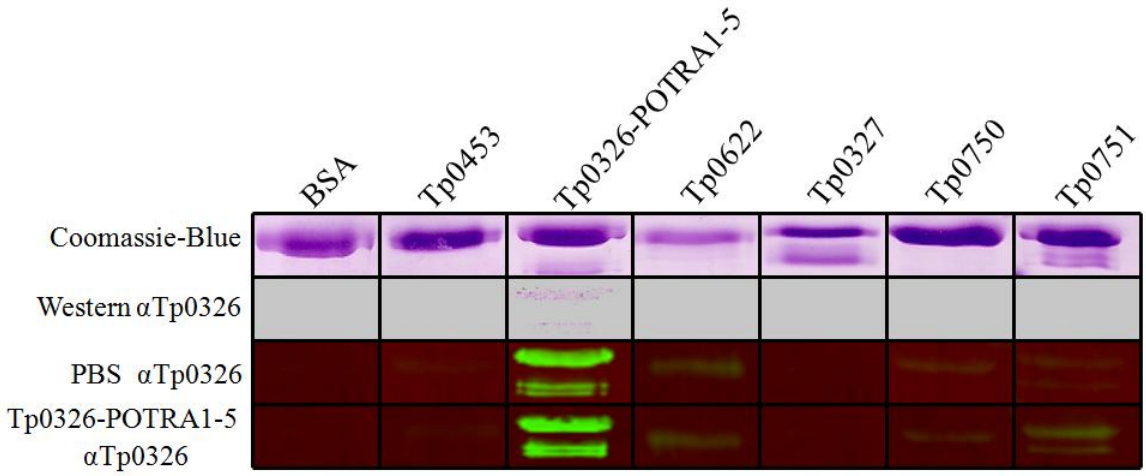


Figure 18. Far-Western blot analysis of Tp0326-POTRA1-5
Far-Western blot analysis of protein-protein interactions between Tp0326-POTRA1-5 and prey proteins BSA, Tp0453, Tp0326-POTRA1-5, Tp0622, Tp0327, Tp0750, and Tp0751. Coomassie-Blue - 2µg of each protein was run on a 15% SDS-PAGE and stained for 1 hour with coomassie-blue dye, this is the total protein. Western αTp0326 – all proteins were subjected to a Western blot assay which used chicken-anti-Tp0326 as the 1º Ab at a 1:2000 dilution and AP development. PBS αTp0326 – Far-Western blot negative control, nitrocellulose membranes were incubated with PBS in place of Tp0326-POTRA1-5 protein. Tp0326-POTRA1-5 αTp0326 – Far-Western blot assay which incubated nitrocellulose membranes with Tp0326-POTRA1-5 protein at a concentration of 250µg/mL. In both Far-Western blot assays the 1ºAb was chicken-anti-Tp0326, 1:2000 dilution, and the 2º Ab was rabbit-anti-chicken IRDye800® conjugated, 1:10000 dilution and blots were developed using the Odyssey® infrared imaging system.

Table 6. Information on proteins used in Far-Western blots

Protein	Sequence (a.a)	MW (KDa)	State
BSA	-	66.00	Native
Tp0453	32 - 288	30.82	Soluble rec.
Tp0326-POTRA1-5	22 - 434	49.29	Soluble rec.
Tp0622	277 - 452	22.43	Soluble rec.
Tp0327	23 - 172	19.55	Soluble rec.
Tp0750	23 - 223	25.79	Insoluble rec.
Tp0751	25 - 237	25.47	Soluble rec.

4.6.3 Analyzing protein-protein interactions

The level of protein binding in the Far-Western blots was quantified for fluorescence using the Odyssey® infrared imaging system software. The level of binding between Tp0326-POTRA1-5 and each individual prey protein was determined by normalizing all values against BSA and subtracting the fluorescence value in the negative control (PBS) from the value obtained from the Tp0326-POTRA1-5 incubated blots. This calculation gave the level of protein binding between the individual prey protein and Tp0326-POTRA1-5. The Far-Western blot experiment was repeated in triplicate and the level of protein binding between each individual prey protein and Tp0326-POTRA1-5 was averaged from the three experiments. The averaged level of protein binding over the three Far-Western blot experiments can be seen in Figure 19.

The negative control protein BSA was used to normalize all of the fluorescence values obtained from each individual Far-Western blot. This was in order to account for differences in amount of protein loaded from well to well and experiment to experiment. *T. pallidum* putative OMP Tp0453 is a recombinantly expressed soluble protein not believed to interact with the *T. pallidum* BAM complex and was thus used as a negative control for the experiment. *T. pallidum* protein Tp0326-POTRA1-5 was used a positive control for the primary antibody used in the experiment. We do not have a positive control protein for the experiment as there have not been any *T. pallidum* proteins shown definitively to bind to Tp0326.

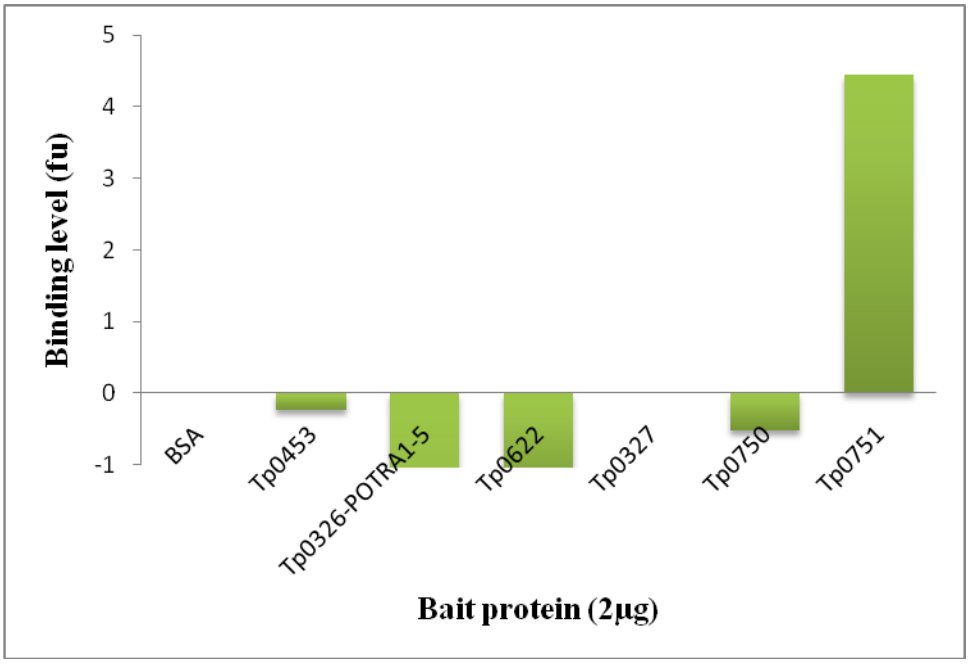


Figure 19. Level of protein binding from Far-Western blot assay
Average level of protein binding between prey proteins and Tp0326-POTRA1-5 observed from three identical Far-Western blot experiments. Level of fluorescence for each individual prey protein was quantified using Odyssey® infrared imaging system software for both the negative control PBS Far-Western blot and the Tp0326-POTRA1-5 Far-Western blot. Fluorescence level was then normalized against BSA and the level observed from the negative control PBS blot was subtracted from the Tp0326-POTRA1-5 blot to obtain the level of binding. The following binding levels were observed: BSA, 0.00; Tp0453, -0.24; Tp0326-POTRA1-5, -4.07; Tp0622, -1.60; Tp0327, 0.00; Tp0750, -0.53; Tp0751, 4.44. fu – fluorescence units.

Treponema pallidum proteins Tp0453, Tp0326-POTRA1-5, Tp0622, and Tp0750 all showed decreased binding of Tp0326-POTRA1-5 compared to PBS. These proteins all showed reactivity to the primary antibody when incubated with PBS rather than Tp0326-POTRA1-5, indicating an apparent level of background binding to the primary antibody. *Treponema pallidum* protein Tp0327 could not be visualized on any of the Far-Western blots, even with overexposure, and was thus assigned a value of zero for protein binding. The only *T. pallidum* protein tested which showed an increased level of binding to Tp0326-POTRA1-5 compared to PBS was Tp0751 (+4.44 fu).

4.7 Conclusions

Screening and direct PPI techniques were used in this thesis in order to gain a better understanding of the *T. pallidum* BAM complex and OMP biogenesis in *T. pallidum*. A *T. pallidum* lambda genomic expression library was initially screened against recombinant *T. pallidum* protein Tp0326-POTRA1-5 in order to identify interacting *T. pallidum* proteins. Interactions identified in the screen were then verified using recombinantly expressed *T. pallidum* proteins and a Far-Western blotting direct PPI approach.

The *T. pallidum* Lambda ZAP screen identified several hypothetical proteins (Tp0258, Tp0752, Tp0753, Tp0754, and Tp0854), two ECM binding adhesins (Tp0750 (Cameron laboratory, unpublished observations) and Tp0751 (Cameron, 2003; Cameron et al., 2005)), a Gpd protein (Tp0257), and a putative BamD protein (Tp0622). The Lambda ZAP screen identified two putative *T. pallidum* OMPs, Tp0751 (Cameron *et al*., 2008) and Tp0750 (Cameron laboratory, unpublished findings). There has been no direct evidence shown for Tp0750 being an OMP, however, it has been shown to bind to host ECM components (Cameron laboratory, unpublished observations), so it is likely that it resides on the surface of the bacterium where it can interact with host components.

It has recently been shown by cryo-electron tomography (CET) that the peptidoglycan layer in *T. pallidum* divides the periplasmic space into two distinct regions; above the cytoplasmic membrane and below the outer membrane (Izard *et al*., 2009). It was also shown in this study using CET that the the mean diameter of the periplasm is approximately 40nm (Izard *et al*., 2009). Given that the average diameter of a protein in *E. coli* is approximated to be 2-6nm (Lukatsky & Shakhnovich, 2008), it

seems rather unlikely that a *T. pallidum* protein which resides in the cytoplasmic membrane and is orientated to the periplasm could traverse the periplasm to interact with Tp0326. It therefore seems unlikely that the interactions observed in the Lambda ZAP screen between Tp0326-POTRA1-5 and *T. pallidum* proteins which are localized or predicted to be localized to the cytoplasmic membrane are real protein-protein interactions *in vivo*. These proteins include Tp0257, Tp0258, Tp0752, Tp0754, and Tp0854. The localization predictions for the hypothetical *T. pallidum* protein Tp0753 were ambiguous towards either cytoplasmic membrane or periplasmic localization.

The Lambda ZAP assay is a screen for putative protein-protein interactions. A protein-protein interaction observed in the assay is simply an interaction which has the potential to occur in the organism *in vivo*. It is therefore possible to have a real protein-protein interaction occur in the assay which is not actually a real protein-protein interaction in *T. pallidum*.

The only *T. pallidum* protein which was identified to bind Tp0326-POTRA1-5 through Far-Western blotting was putative *T. pallidum* OMP Tp0751. Binding could not be verified for Tp0750 or Tp0622. This does not necessarily mean that these proteins do not interact with Tp0326-POTRA1-5. It could mean that the conditions of the Far-Western blots need to be optimized for those specific interactions or that Far-Western blotting is not an appropriate method for detecting these particular interactions. The fact that *T. pallidum* protein Tp0751 was identified to bind to Tp0326-POTRA1-5 in both the LambdaZAP protein-protein interaction screen and Far-Western blotting gives good evidence that this is a real protein-protein interaction.

It remains unclear from these experiments whether Tp0751 resides in the outer membrane in isolation, in which case it would be strictly transported into the outer membrane through *T. pallidum* protein Tp0326, or if Tp0751 actually forms a stable interaction with Tp0326. Both the Lambda ZAP screen and Far-Western blotting are more adept at detecting stable protein-protein interactions compared to those that are transient in nature. It is therefore possible that the putative *T. pallidum* OMP Tp0751 forms a stable interaction with the putative *T. pallidum* BamA protein, Tp0326.

Together the findings from the PPI studies support the hypothesis that *T. pallidum* possesses a BAM complex and that this complex would be responsible for ushering *T. pallidum* OMPs to the bacterial surface. The precise number and identity of all *T. pallidum* proteins involved with the putative BAM complex remain unknown.

Chapter 5: Overview of the putative *Treponema pallidum* BAM complex

The outer membranes of Gram-negative bacteria possess OMPs that are associated with basic physiological functions, virulence, and drug resistance and therefore play a crucial part in the pathogenesis of the organisms and in maintaining cell viability (Bos *et al*., 2007). Understanding how OMPs are targeted and folded into the outer membrane of *T. pallidum* will likely be invaluable to understanding the pathogenesis of the pathogen and could potentionaly yield medical benefits such as vaccine production.

The core bacterial complex responsible for trafficking and folding proteins into the outer membrane of Gram-negative bacteria is the BAM complex (Knowles *et al*., 2009). In *E. coli* the core BAM complex is comprised of five proteins: BamA, an integral membrane protein, and four accessory lipoproteins, BamB - BamE, which localize to the inner leaflet of the outer membrane (Onufryk *et al*., 2005; Sklar *et al*., 2007; Wu *et al*., 2005). As well three periplasmic chaperones, SurA, Skp, and DegP are believed to be involved in transporting nascent OMPs across the periplasmic space to the BAM complex in the outer membrane of *E. coli* (Rizzitello *et al*., 2001; Sklar *et al*., 2007).

A homologue to BamA has been identified in *T. pallidum*, Tp0326 (Cameron *et al*., 2000), as has a homologue to *E. coli* Skp, *T. pallidum* protein Tp0327 (Cameron, unpublished observations). Other than these findings, there have previously been no reports on a *T. pallidum* BAM complex.

5.1 Connecting bioinformatic and protein-protein interaction data

In order to identify and characterize the *T. pallidum* BAM complex we have utilized a two-pronged approach: bioinformatic analysis and protein-protein interaction (PPI) studies. Bioinformatic analyses were used in order to identify potential *T. pallidum* BAM complex proteins. A *T. pallidum* lambda genomic expression library was screened against recombinant *T. pallidum* protein Tp0326 POTRA1-5, in order to identify putative *T. pallidum* BAM complex proteins as well as putative *T. pallidum* OMPs. The findings from the two approaches were compared, and putative *T. pallidum* BAM complex proteins were verified for their interaction with Tp0326 POTRA1-5 by means of Far-Western blot analysis.

5.1.1 Putative *Treponema pallidum* BAM associated lipoproteins

Bioinformatic analysis using BLASTp identified *T. pallidum* hits for *E. coli* BamB and BamD, Tp0133 and Tp0622 respectively. Although the respective E-values in the BLASTp searches were not significant, there is a 40 a.a stretch in Tp0622 which shares 42 % sequence identity with *E. coli* BamD and there is a 105 a.a stretch in Tp0133 which shares 30 % sequence identity with *E. coli* BamB. As well, structure recognition programs predict Tp0622 to possess three tetratricopeptide repeat (TPR) motif domains and Tp0133 to form a 7-bladed β-sheet propeller structure. These predicted structures are the same structures found in their putative *E. coli* homologues, BamD (Blatch & Lassle, 1999; D'Andrea & Regan, 2003) and BamB (Vuong *et al.*, 2008) respectively. Tetratricopeptide repeat motifs have been shown to be involved with protein-protein interactions and believed to be involved in forming large protein complexes (D'Andrea & Regan, 2003). Multiple bladed β-propellar structures have been shown to have a very

diverse set of functions, one of which could be acting as a scaffold for multiple ligands or substrates (Pons *et al*., 2003). Multiple bladed β-propellar structures have been proposed to interact with the BamA POTRA domains and/or nascent OMPs through β-augmentation (Gatsos *et al.*, 2008). It is possible given the predicted structure of *T. pallidum* proteins Tp0133 and Tp0622, that they could both form protein-protein interactions in a multi protein complex.

Supporting the bioinformatic data, is the finding that Tp0622 was identified to interact with Tp0326-POTRA1-5 in the Lambda ZAP PPI assay. Only a small portion of Tp0622 was identified in the screen, 29 %; however, the identified region overlaps with the 40 a.a stretch of the protein that shows high sequence identity to *E. coli* BamD. It is likely that this region of the protein is involved with binding to the POTRA domains of BamA. Futher supporting the notion that *T. pallidum* would possess a BamD homologue is the fact that BamD is the only lipoprotein which has been shown to be essential for cell viability and it is the only BAM lipoprotein which is ubiquitous in Gram-negative bacteria (Malinverni *et al.*, 2006). As well, *E. coli* BamD has been shown to interact with BamA directly in an interaction which requires the fifth POTRA domain (Malinverni *et al*., 2006; Sklar *et al*., 2007). In order to verify this interaction, the region of Tp0622 which was identified in the Lambda ZAP screen was cloned and expressed and tested for a direct PPI with Tp0326-POTRA1-5 by Far-Western blot analysis. We were not able to show an interaction between these two proteins using this assay. That is not to say that this is not a real interaction, nor does it conclusively identify Tp0622 as a BamD homologue. It could have been that we were not able to optimize the assay for that particular interaction to occur, or that this is not the appropriate methodology for

detecting such an interaction. Further study is needed in order to definitively label Tp0622 a BamD homologue.

The Lambda ZAP PPI screen did not identify any other putative *T. pallidum* BAM lipoproteins, such as Tp0133. However, we only screened 4 out of the 6 original cores which were taken from the primary screen. There is a possibility that the other 2 cores could contain putative *T. pallidum* BAM lipoproteins, chaperones, or OMPs. Lipoprotein BamB is very highly conserved among many Gram-negative bacteria (Knowles *et al.*, 2009) so it is likely that *T. pallidum* possesses a homologue of this protein. The bioinformatic data supports the notion that Tp0133 is a *T. pallidum* BamB homologue; however, further evidence is needed in order to qualify the bioinformatic finding.

There are a couple of explanations as to the lack of identification of any putative *T. pallidum* BamC or BamE homologues. BamC is not ubiquitous throughout Gram-negative bacteria (Knowles *et al*., 2009), nor is BamE. BamC does possess the ability to bind BamA directly; it requires the C-terminus of BamD (Malinverni *et al.*, 2006). BamE has also not been shown to possess the ability to bind to BamA; however, it has been shown to stabilize the binding of BamD to BamA (Sklar *et al.*, 2007).

Treponema pallidum has a much smaller genome than that of *E. coli*, and it is therefore likely that *T. pallidum* only possesses the necessary BAM lipoproteins for survival, which have been shown to be BamB and BamD. It could be that *T. pallidum* does not have any proteins that resemble BamC and BamE and that its BAM complex is much smaller than that of *E. coli*. If *T. pallidum* does actually possess BamC and BamE homologues, it could be that they were not identified in our screen because they can-not directly bind to BamA (Tp0326) on their own. A different experimental approach could

potentially identify these proteins in *T. pallidum*, such as using the putative *T. pallidum* BamD protein, Tp0622, as the bait protein in the Lambda ZAP screen, as BamC and BamE have both been shown to bind directly to BamD in *E. coli*. However, the bioinformatic data does not suggest that *T. pallidum* possesses BamC or BamE homologues and given the smaller, more efficient genome in *T. pallidum*, it is likely that it does not possess these proteins.

5.1.2 Putative *Treponema pallidum* BAM associated chaperones

Bioinformatic analyses identified putative *T. pallidum* homologues for all three of the *E. coli* periplasmic chaperones: Tp0327 (Skp), Tp0773 (DegP), and Tp0862 (SurA). There was some ambiguity surrounding the bioinformatic analyses of Tp0773 and Tp0862.

Most of the bioinformatic evidence presented suggests that *T. pallidum* protein Tp0773 is a homologue of the *E. coli* periplasmic chaperone DegP; however it was also predicted to be a lipoprotein by the LipoP prediction program. As the LipoP prediction tool is purely just prediction, it is possible that Tp0773 is not actually a lipoprotein. As well, due to the fact that *T. pallidum* is evolutionarily very distant from *E. coli*, and possesses a much smaller genome, it is possible that in *T. pallidum* this protein could serve a dual function. It is possible that even as a lipoprotein, Tp0773 could serve the same function as DegP.

Similarily to Tp0773, most of the bioinformatic evidence suggests that *T. pallidum* protein Tp0862 is a homologue of the *E. coli* periplasmic chaperone SurA. Complicating the finding is that the *E. coli* chaperone protein SurA (parvulin-type PPIase) and *T. pallidum* protein Tp0862 (FKBP-type PPIase) are believed to interact with

different types of substrates (Gothel & Marahiel, 1999). *Treponema pallidum* protein Tp0862 also shows high sequence identity to other FKBP-type PPIases from other organisms, when analyzed by BLASTp against the protein database. As well, FKBP-type PPIases are normally localized in the cytoplasm (Gothel & Marahiel, 1999). There was major discrepancy between PSORT and PSORTb as to the predicted location of Tp0862 within *T. pallidum*. PSORT predicted Tp0862 to be periplasmic whereas PSORTb predicted it to be cytoplasmic. Typically PSORT seems to give more accurate predictions for known *T. pallidum* proteins in comparison with PSORTb (Cameron, unpublished observations). If the protein is indeed a periplasmic chaperone, it is possible that Tp0862 functions as a BAM accessory chaperone. Arguing against *T. pallidum* protein Tp0862 as a periplasmic chaperone, is the notion that the there is functional redundancy amongst the *E. coli* periplasmic chaperones (Sklar *et al.*, 2007). The *T. pallidum* genome is much smaller than the *E. coli* genome and it is likely that efficiency is of a greater concern in the *T. pallidum* genome. It could then be plausible that *T. pallidum* would remove this functional redundancy and not possess homologues of all three *E. coli* periplasmic chaperones. Given the fact that other FKBP-type PPIases are typically cytoplasmic and do not catalyze reactions involving β-barrel substrates (Gothel & Marahiel, 1999), coupled with the above point related to removing functional redundancy in the *T. pallidum* genome, it seems more likely that *T. pallidum* does not possess a SurA homologue and instead relies on a two chaperone system.

The Lambda ZAP PPI screen did not identify any of the three putative *T. pallidum* periplasmic chaperones to interact with Tp0326-POTRA1-5. Soluble recombinant

protein Tp0327 was also tested for a direct PPI with Tp0326-POTRA1-5 using Far-Western blot analysis and no interaction could be identified.

The absence of any observed PPI between these putative *T. pallidum* periplasmic chaperones and BamA could be due to several reasons. This could be because it is likely that interactions between chaperones and the BAM complex are transient. Transient interactions, those that are not stable, are very difficult to pick-up in these types of PPI studies. It is possible that interactions could be observed between these putative *T. pallidum* periplasmic chaperones and BamA using different PPI studies which better identify transient interactions, such as cross-linking methodologies. Another possibility is that the periplasmic chaperones never actually come into direct contact with the putative *T. pallidum* BAM complex, but instead interact indirectly through the OMP which is being ushered to the complex. This is supported by the finding that only SurA in *E. coli* has been shown to actually interact with the BamA POTRA domains and it is unclear whether the interaction is direct or through a substrate intermediate (Sklar *et al*., 2007). If *T. pallidum* does not actually possess a SurA homologue, then it could be expected that no *T. pallidum* periplasmic chaperones actually interact with Tp0326. Another possibility as to the absence of an observed interaction in the Lambda ZAP screen between Tp0326 and any of the putative *T. pallidum* periplasmic chaperones is that the interaction has not been identified yet. As stated previously, only 4 of the original 6 cores taken from the primary screen in the Lambda ZAP assay have been further analyzed. It is possible that these remaining cores could correspond to one of the putative *T. pallidum* BAM associated chaperones.

5.1.3 Putative *Treponema pallidum* BAM interacting membrane proteins

The *T. pallidum* Lambda ZAP screen identified several hypothetical proteins (Tp0258, Tp0752, Tp0753, Tp0754, and Tp0854), two ECM binding adhesins (Tp0750 (Cameron laboratory, unpublished observations) and Tp0751 (Cameron, 2003; Cameron *et al*., 2005)), and a Gpd protein (Tp0257). Two of these identified proteins are putative *T. pallidum* OMPs, Tp0751 (Cameron *et al*., 2008) and Tp0750 (Cameron laboratory, unpublished findings). There has been no direct evidence shown for Tp0750 being an OMP, however, it has been shown to bind to host ECM components (Cameron laboratory, unpublished observations), so it is likely that it resides on the surface of the bacterium where it can interact with host components.

Several of the observed PPIs in the Lambda ZAP assay are not believed to be real PPIs in *T. pallidum*. This was discussed in section 4.7; these proteins include Tp0257, Tp0258, Tp0752, Tp0754, and Tp0854.

Bioinformatic analyses surrounding hypothetical *T. pallidum* protein Tp0753 were ambiguous. Different prediction programs predicted Tp0753 to either be localized in the cytoplasmic membrane or the periplasmic space. Because of time constraints and lack of evidence for Tp0753 to be involved in a putative *T. pallidum* BAM complex, only Tp0750 and Tp0751 were tested for a direct PPI with Tp0326-POTRA1-5 using Far-Western blot analysis. Tp0751, but not Tp0750, was shown to bind directly to Tp0326-POTRA1-5, verifying the interaction observed in the Lambda ZAP screen. The fact that *T. pallidum* protein Tp0751 was identified to bind to Tp0326-POTRA1-5 in both the LambdaZAP protein-protein interaction screen and Far-Western blotting gives good evidence that this is a real protein-protein interaction. The PPI between Tp0750 and

Tp0326-POTRA1-5 could still be a real interaction. The absence of a PPI could mean that the conditions of the Far-Western blots need to be optimized for that specific interaction or that Far-Western blotting is not an appropriate method for detecting that particular interaction. In contrast, because both *T. pallidum* proteins Tp0750 and Tp0751 were identified in the same Lambda ZAP screen core and the Tp0751 PPI with Tp0326-POTRA1-5 was verified, it could mean that only Tp0751 was interacting with Tp0326-POTRA1-5 in the Lambda ZAP assay.

It remains unclear from these experiments whether Tp0751 resides in the outer membrane in isolation, in which case it would be strictly transported into the outer membrane through *T. pallidum* protein Tp0326, or if Tp0751 actually forms a stable interaction with Tp0326. Both the Lambda ZAP screen and Far-Western blotting are more adept at detecting stable protein-protein interactions compared to those that are transient in nature. It is therefore much more likely that the putative *T. pallidum* OMP Tp0751 forms a stable interaction with the putative *T. pallidum* BamA protein Tp0326. The fact that no other putative *T. pallidum* OMPs or putative *T. pallidum* periplasmic chaperones were identified in this screen supports the notion that transient interactions are not likely to be picked up in this assay. It could however be that other putative *T. pallidum* OMPs have yet to be discovered in the Lambda ZAP assay using the Tp0326-POTRA1-5 bait protein.

Interestingly, Tp0751 has recently been predicted to be a lipoprotein (Houston *et al.*, submitted for publication). There is a certain degree of controversy as to whether *T. pallidum* possesses lipoproteins in its outer membrane. Recent research surrounding the ultrastructure of *T. pallidum* shows an electron dense layer associated with the outer

leaflet of the inner membrane and an absence of such a layer at the outer membrane, suggesting that the bacterium may not possess lipoproteins in the outer membrane (Izard *et al*., 2009). As well, *T. pallidum* does not possess a homologue of the LolB protein, which is critical in transporting lipoproteins to the outer membrane in Gram-negative bacteria (Tokuda, 2009). *Treponema pallidum* does, however, possess the rest of the Lol machinery (Fraser *et al*., 1998). The research presented in this thesis suggests that at least one *T. pallidum* lipoprotein associates with Tp0326 in the outer membrane. If this is true, then it is possible that Tp0326 is involved in ushering *T. pallidum* lipoproteins to the outer membrane.

5.2 Constructing the putative *Treponema pallidum* BAM complex

Compiling all the bioinformatic and PPI findings together, we begin to paint a preliminary picture as to what the make up of the putative *T. pallidum* BAM complex could look like. The central protein embedded in the outer membrane which makes up the BAM complex is believed to be Tp0326 (BamA).

Bioinformatic analyses suggest that *T. pallidum* possesses putative homologues to *E. coli* lipoproteins BamD and BamB, those are Tp0622 and Tp0133 respectively. It appears through *in silico* analyses that *T. pallidum* does not possess homologues to the non-essential *E. coli* BAM associated lipoproteins, BamC and BamE. Only Tp0622 was identified to interact with Tp0326 in the Lambda ZAP screen and the interaction could not be verified by Far-Western blot analysis. Although further experimentation is needed, it seems plausible that Tp0622 and Tp0133 could be involved in a putative *T. pallidum* BAM complex.

Bioinformatic analyses suggest that *T. pallidum* possesses putative homologues to *E. coli* periplasmic chaperones Skp and DegP, those are Tp0327 and Tp0773 respectively. Based on bioinformatic analyses it appears that *T. pallidum* does not possess a homologue for *E. coli* periplasmic chaperone SurA and instead relies on a two chaperone system. We were unable to show any kind of interaction experimentally between these putative *T. pallidum* chaperone proteins and Tp0326. This most likely could have been due to the probability that these interactions are transient in nature and the methodologies used in this project were unable to identify such interactions.

Protein-protein interaction studies identified that the putative *T. pallidum* OMP Tp0751 interacts with the POTRA domains of Tp0326. It is likely that the interaction between Tp0751 and Tp0326 is stable. Taking everything into account, the predicted model of the *T. pallidum* BAM complex is shown schematically in Figure 20.

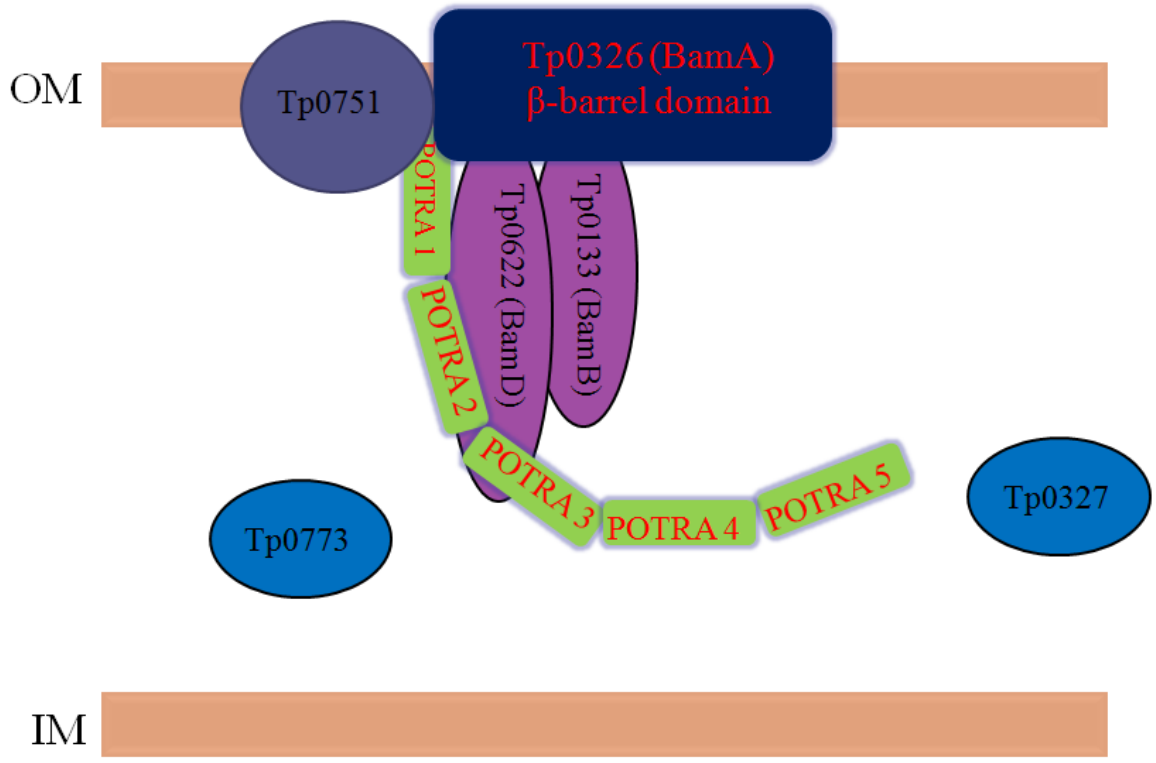


Figure 20. Schematic representation of the Treponema pallidum BAM complex
Schematic representation of the proposed *T. pallidum* BAM complex. The complex was created by compiling data obtained from bioinformatic analyses and protein-protein interaction studies using recombinant *T. pallidum* protein Tp0326-POTRA1-5. Tp0326 (BamA) is embedded in the outer membrane (OM) with its POTRA domains hanging free in the periplasmic space. Putative *T. pallidum* lipoproteins Tp0622 (BamD), and Tp0133 (BamB), form stable interactions with the Tp0326 POTRA domains in the periplasm and are anchored in the inner leaflet of the outer membrane. Putative *T. pallidum* BAM associated chaperone proteins Tp0773 and Tp0327 reside in the periplasm. Putative *T. pallidum* OMP Tp0751 is ushered into the OM through the BAM complex where it forms a stable interaction with Tp0326. IM – inner membrane.

5.3 Conclusions

The bioinformatic analyses and protein-protein interaction data presented in this research supports the hypotheses that *T. pallidum* possesses a BAM complex. With the exception of Tp0751 it remains to be seen if this complex is responsible for the ushering and folding of *T. pallidum* OMPs into the outer membrane.

The putative *T. pallidum* BAM complex appears to be much simpler, in terms of the number of involved proteins, in comparison to the *E. coli* BAM complex. This could likely be due to the fact that *T. pallidum* is quite distant from *E. coli* from an evolutionary point of view. Another factor involved in this finding is that the genome of *T. pallidum* is much smaller than that of *E. coli* and therefore the need for efficiency within the genome is much greater in *T. pallidum*. It is likely that proteins of redundant or non-essential function have been removed from the *T. pallidum* proteome and this is evident from bioinformatic analyses presented in this thesis. Another possibility is that the *T. pallidum* BAM complex is much more complex in comparison to what we have presented and that other proteins involved in the complex have not yet been identified. However, it seems more likely that the putative *T. pallidum* BAM complex is simpler than the *E. coli* BAM complex in terms of the number of proteins involved, given the smaller genome of *T. pallidum.*

The identification of a putative *T. pallidum* BAM complex contributes to the puzzle of lipoprotein sorting in the organism. There is a great deal of controversy as to whether *T. pallidum* possesses lipoproteins in its outer membrane. The research presented in this thesis suggests that *T. pallidum* may possess lipoproteins in its outer membrane. As *T. pallidum* does not possess a LolB homologue, which is involved in ushering lipoproteins across the periplasm to the outer membrane in Gram-negative bacteria, it is possible that Tp0326 is somehow involved in this process in *T. pallidum*. This type of lipoprotein sorting would be novel to all bacteria. Given the unusual ultrastructure of *T. pallidum* in comparison to other bacteria it would not be surprising for it to possess such a novel mechanism.

An important aspect of the research presented in this thesis is the success of the newly designed Lambda ZAP protein-protein interaction screen. This assay now gives us the ability to screen for putative protein-protein interactions in *T. pallidum* using a desired bait protein. This is something that has proven to be very difficult to research in *T. pallidum*.

The identification of a putative *T. pallidum* BAM complex has the potential to open a promising door to research surrounding this pathogen. Research surrounding the putative *T. pallidum* BAM complex yields the possibility of identifying *T. pallidum* OMPs which have yet to be discovered and should give us a better understanding of the pathogenesis of the organsim. This could obviously have great medical benefits in terms of vaccine production, something that has yet to be successfully accomplished and is of dire need.

5.4 Future directions

Preliminary characterization of the putative BAM complex within *T. pallidum* has been initiated through the studies contained within this thesis, although additional studies need to be conducted to fully elucidate the identities and functions of proteins that contribute to this complex within this bacterium. Possibly the most important area for future research is the continuation of the Lambda ZAP assay. Only 4 of the original 6 cores taken from the primary Lambda ZAP screen were processed for secondary screening and DNA sequencing. Processing of the remaining 2 cores could possibly identify further members of the putative *T. pallidum* BAM complex or could identify putative *T. pallidum* OMPs. The Lambda ZAP assay could also be utilized with different bait proteins. The putative BamD homologue Tp0622 could be used as a bait protein in

the screen, which could potentially identify other putative *T. pallidum* BAM complex proteins which do not specifically interact with the POTRA domains of Tp0326. For example the *E. coli* lipoprotein BamC does not interact directly with BamA but rather interacts with the C-terminus of BamD (Malinverni *et al*., 2006). The Lambda ZAP assay has the potential to be a very powerful research tool, especially for research surrounding the putative *T. pallidum* BAM complex and other protein-protein interaction complexes within *T. pallidum*.

Further research should be conducted in the area of protein expression and purification for the putative *T. pallidum* BAM complex protein Tp0133 and putative *T. pallidum* periplasmic chaperone Tp0773. If successfuly expressed and purified, these proteins could be tested for interactions with Tp0326 through Far-Western blotting or other protein-protein interaction studies such as co-immunoprecipitation. As well, the *T. pallidum* proteins which were shown to interact with Tp0326 in the Lambda ZAP assay but were not verified via Far-Western blotting (i.e Tp0622) should be retested for PPIs with Tp0326 using other PPI studies, such as co-immunoprecipitation. Cross-linking studies could also prove to be invaluable to this research as they are more prone to identifying transient interactions in comparison to Far-Western blotting.

Heterologous expression of the putative *T. pallidum* BAM complex proteins in model organisms, such as *Treponema phagedenis*, could provide invaluable functional information. Palmitoylation profiling in *T. phagedenis* could yield insights as to whether the putative *T. pallidum* BAM complex proteins are actually lipoproteins.

An area of future research for those *T. pallidum* proteins which are verified to be involved in the potential *T. pallidum* BAM complex would be protein crystallization.

Protein crystallization would provide alot of information about the *T. pallidum* BAM complex and potentially OMP biogenesis in *T. pallidum*. Crystallization of the putative *T. pallidum* BAM complex proteins would provide greater insight into how this complex functions. This should be a very feasible next step for research surrounding the *T. pallidum* BAM complex.

Bibliography

Alderete, J. F., & Baseman, J. B. (1980). Surface characterization of virulent *Treponema pallidum*. *Infect Immun, 30*(3), 814-823.

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.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., *et al*. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res, 25*(17), 3389-3402.

Antal, G. M., Lukehart, S. A., & Meheus, A. Z. (2002). The endemic treponematoses. *Microbes Infect, 4*(1), 83-94.

Arifuzzaman, M., Maeda, M., Itoh, A., Nishikata, K., Takita, C., Saito, R., *et al*. (2006). Large-scale identification of protein-protein interaction of *Escherichia coli* K-12. *Genome Res, 16*(5), 686-691.

Arnold, K., Bordoli, L., Kopp, J., & Schwede, T. (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics, 22*(2), 195-201.

Azad, R. K., & Borodovsky, M. (2004). Probabilistic methods of identifying genes in prokaryotic genomes: connections to the HMM theory. *Brief Bioinform, 5*(2), 118-130.

Bairoch, A. (1991). PROSITE: a dictionary of sites and patterns in proteins. *Nucleic Acids Res, 19 Suppl*, 2241-2245.

Baker-Zander, S. A., & Lukehart, S. A. (1992). Macrophage-mediated killing of opsonized *Treponema pallidum*. *J Infect Dis, 165*(1), 69-74.

Baker-Zander, S. A., Shaffer, J. M., & Lukehart, S. A. (1993). Characterization of the serum requirement for macrophage-mediated killing of *Treponema pallidum* ssp. *pallidum*: relationship to the development of opsonizing antibodies. *FEMS Immunol Med Microbiol, 6*(4), 273-279.

Bansal, A. K. (1999). An automated comparative analysis of 17 complete microbial genomes. *Bioinformatics, 15*(11), 900-908.

Bansal, A. K. (2005). Bioinformatics in microbial biotechnology--a mini review. *Microb Cell Fact, 4*(1), 19.

British Columbia Center for Disease Control. (2009). Annual surveillance report: HIV and Sexually Transmitted Infections.

British Columbia Center for Disease Control. (2010). Syphilis overview. http://www.bccdc.ca/dis-cond/a-z/_s/Syphilis/overview/default.htm. Retrieved August 21, 2010.

Beachey, E. H. (1981). Bacterial adherence: adhesin-receptor interactions mediating the attachment of bacteria to mucosal surface. *J Infect Dis, 143*(3), 325-345.

Behrens, S., Maier, R., de Cock, H., Schmid, F. X., & Gross, C. A. (2001). The SurA periplasmic PPIase lacking its parvulin domains functions *in vivo* and has chaperone activity. *Embo J, 20*(1-2), 285-294.

Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Wheeler, D. L. (2003). GenBank. *Nucleic Acids Res, 31*(1), 23-27.

Bertone, P., & Snyder, M. (2005). Advances in functional protein microarray technology. *Febs J, 272*(21), 5400-5411.

Bertot, G. M., Restelli, M. A., Galanternik, L., Aranibar Urey, R. C., Valvano, M. A., & Grinstein, S. (2007). Nasal immunization with *Burkholderia multivorans* outer membrane proteins and the mucosal adjuvant adamantylamide dipeptide confers efficient protection against experimental lung infections with *B. multivorans* and *B. cenocepacia*. *Infect Immun, 75*(6), 2740-2752.

Blanco, D. R., Giladi, M., Champion, C. I., Haake, D. A., Chikami, G. K., Miller, J. N., *et al*. (1991). Identification of *Treponema pallidum* subspecies *pallidum* genes encoding signal peptides and membrane-spanning sequences using a novel alkaline phosphatase expression vector. *Mol Microbiol, 5*(10), 2405-2415.

Blanco, D. R., Reimann, K., Skare, J., Champion, C. I., Foley, D., Exner, M. M., *et al.* (1994). Isolation of the outer membranes from *Treponema pallidum* and *Treponema vincentii*. *J Bacteriol, 176*(19), 6088-6099.

Blatch, G. L., & Lassle, M. (1999). The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. *Bioessays, 21*(11), 932-939.

Blattner, F. R., Plunkett, G., 3rd, Bloch, C. A., Perna, N. T., Burland, V., Riley, M., *et al*. (1997). The complete genome sequence of *Escherichia coli* K-12. *Science, 277*(5331), 1453-1462.

Bos, M. P., Robert, V., & Tommassen, J. (2007). Biogenesis of the gram-negative bacterial outer membrane. *Annu Rev Microbiol, 61*, 191-214.

Bos, M. P., Robert, V., & Tommassen, J. (2007). Functioning of outer membrane protein assembly factor Omp85 requires a single POTRA domain. *EMBO Rep, 8*(12), 1149-1154.

Brady, J. G., Schutze, G.E., Seibert, R., Horn, H.V., Marks, B., Parham, D.M. (1998). Detection of mycobacterial infections using the Dieterle stain. *Pediatr Dev Pathol 1*(4), 309-313.

Brinkman, M. B., McGill, M. A., Pettersson, J., Rogers, A., Matejkova, P., Smajs, D., *et al.* (2008). A novel *Treponema pallidum* antigen, TP0136, is an outer membrane protein that binds human fibronectin. *Infect Immun, 76*(5), 1848-1857.

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

Cameron, C. E. (2006). The *T. pallidum* outer membrane and outer membrane proteins. . In S. A. Lukehart, Radolf, J.D. (Ed.), *Pathogenic Treponema.* Norfolk, UK.: Horizon Scientific Press.

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

Cameron, C. E., Brown, E. L., Kuroiwa, J. M., Schnapp, L. M., & Brouwer, N. L. (2004). *Treponema pallidum* fibronectin-binding proteins. *J Bacteriol, 186*(20), 7019-7022.

Cameron, C. E., Castro, C., Lukehart, S. A., & Van Voorhis, W. C. (1998). Function and protective capacity of *Treponema pallidum subsp. pallidum* glycerophosphodiester phosphodiesterase. *Infect Immun, 66*(12), 5763-5770.

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., Castro, C., Molini, B., Godornes, C., & Van Voorhis, W. C. (2000). Opsonic potential, protective capacity, and sequence conservation of the *Treponema pallidum* subspecies *pallidum* Tp92. *J Infect Dis, 181*(4), 1401-1413.

Centurion-Lara, A., Castro, C., Barrett, L., Cameron, C., Mostowfi, M., Van Voorhis, W. C., *et al*. (1999). *Treponema pallidum* major sheath protein homologue Tpr K is a target of opsonic antibody and the protective immune response. *J Exp Med, 189*(4), 647-656.

Chapman, E., Farr, G.W., Usaite, R. *et al.* (2006). Global aggregation of newly translated proteins in an *Escherichia coli* strain deficient of the chaperonin GroEL. *Proc. Natl Acad. Sci. USA*, 103, 15800–15805.

Chen, R., & Henning, U. (1996). A periplasmic protein (Skp) of *Escherichia coli* selectively binds a class of outer membrane proteins. *Mol Microbiol, 19*(6), 1287-1294.

Chen, Y., & Xu, D. (2003). Computational analyses of high-throughput protein-protein interaction data. *Curr Protein Pept Sci, 4*(3), 159-181.

Chevalier, D. and Walker, J.C. (2005). Functional genomics of protein kinases in plants. *Brief. Funct. Genomic. Proteomic.* 3, 362–371.

Chi, B., Chauhan, S., & Kuramitsu, H. (1999). Development of a system for expressing heterologous genes in the oral spirochete *Treponema denticola* and its use in expression of the *Treponema pallidum* flaA gene. *Infect Immun, 67*(7), 3653-3656.

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*. *Infect Immun, 60*(3), 1076-1083.

Cox, D. L., Luthra, A., Dunham-Ems, S., Desrosiers, D. C., Salazar, J. C., Caimano, M. J. (2010). Surface immunolabeling and consensus computational framework to identify candidate rare outer membrane proteins of *Treponema pallidum*. *Infect Immun*. Ahead of press.

Cullen, P. A., & Cameron, C. E. (2006). Progress towards an effective syphilis vaccine: the past, present and future. *Expert Rev Vaccines, 5*(1), 67-80.

Cumberland, M. C., & Turner, T. B. (1949). The rate of multiplication of *Treponema pallidum* in normal and immune rabbits. *Am J Syph Gonorrhea Vener Dis, 33*(3), 201-212.

Cunningham, T. M., Walker, E. M., Miller, J. N., & Lovett, M. A. (1988). Selective release of the *Treponema pallidum* outer membrane and associated polypeptides with Triton X-114. *J Bacteriol, 170*(12), 5789-5796.

D'Andrea, L. D., & Regan, L. (2003). TPR proteins: the versatile helix. *Trends Biochem Sci, 28*(12), 655-662.

De Carlo, S., Adrian, M., Kalin, P., Mayer, J. M., & Dubochet, J. (1999). Unexpected property of trehalose as observed by cryo-electron microscopy. *J Microsc, 196*(1), 40-45.

de Cock, H., van Blokland, S., & Tommassen, J. (1996). In vitro insertion and assembly of outer membrane protein PhoE of *Escherichia coli* K-12 into the outer membrane. Role of Triton X-100. *J Biol Chem, 271*(22), 12885-12890.

De Mot, R., & Vanderleyden, J. (1994). The C-terminal sequence conservation between OmpA-related outer membrane proteins and MotB suggests a common function in both gram-positive and gram-negative bacteria, possibly in the interaction of these domains with peptidoglycan. *Mol Microbiol, 12*(2), 333-334.

Deacon, W. E., Falcone, V. H., & Harris, A. (1957). A fluorescent test for treponemal antibodies. *Proc Soc Exp Biol Med, 96*(2), 477-480.

Dhar-Chowdhury, P., Harrell, M.D., Han, S.Y. *et al.* (2005). The glycolytic enzymes, glyceraldehyde-3-phosphate dehydrogenase, triose- phosphate isomerase, and pyruvate kinase are components of the K(ATP) channel macromolecular complex and regulate its function. *J. Biol. Chem.* 280, 38464–38470.

Driessen, A. J., & Nouwen, N. (2008). Protein translocation across the bacterial cytoplasmic membrane. *Annu Rev Biochem, 77*, 643-667.

Drusin, L. M., Rouiller, G.C., Chapman, G.B. (1969). Electron Microscopy of *Treponema pallidum* Occurring in a Human Primary Lesion. *Journal of Bacteriology, 97*(2), 951-955.

Duguay, A. R., & Silhavy, T. J. (2004). Quality control in the bacterial periplasm. *Biochim Biophys Acta, 1694*(1-3), 121-134.

Emanuelsson, O., Brunak, S., von Heijne, G., & Nielsen, H. (2007). Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc, 2*(4), 953-971.

Fields, S., & Song, O. (1989). A novel genetic system to detect protein-protein interactions. *Nature, 340*(6230), 245-246.

Fieldsteel, A. H., Cox, D. L., & Moeckli, R. A. (1981). Cultivation of virulent *Treponema pallidum* in tissue culture. *Infect Immun, 32*(2), 908-915.

Fitzgerald, T. J. (1983). Attachment of treponemes to cell surfaces. In R. F. Schnell, Musher, D.M. (Ed.), *Pathogenesis and immunology of treponemal infections*. New York, NY: Marcel Dekker.

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, G. 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.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., & Bairoch, A. (2003). ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res, 31*(13), 3784-3788.

Gatsos, X., Perry, A. J., Anwari, K., Dolezal, P., Wolynec, P. P., Likic, V. A., *et al*. (2008). Protein secretion and outer membrane assembly in *Alphaproteobacteria*. *FEMS Microbiol Rev, 32*(6), 995-1009.

Gatzeva-Topalova, P. Z., Walton, T. A., & Sousa, M. C. (2008). Crystal structure of YaeT: conformational flexibility and substrate recognition. *Structure, 16*(12), 1873-1881.

Gavin, A. C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., *et al*. (2002). Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature, 415*(6868), 141-147.

Giacani, L., Molini, B. J., Kim, E. Y., Godornes, B. C., Leader, B. T., Tantalo, L. C., *et al.* (2010). Antigenic variation in *Treponema pallidum*: TprK sequence diversity accumulates in response to immune pressure during experimental syphilis. *J Immunol, 184*(7), 3822-3829.

Gjestland, T. (1955). The Oslo study of untreated syphilis; an epidemiologic investigation of the natural course of the syphilitic infection based upon a re-study of the Boeck-Bruusgaard material. *Acta Derm Venereol Suppl (Stockh), 35*(Suppl 34), 3-368; Annex I-LVI.

Gothel, S. F., & Marahiel, M. A. (1999). Peptidyl-prolyl cis-trans isomerases, a superfamily of ubiquitous folding catalysts. *Cell Mol Life Sci, 55*(3), 423-436.

Harms, N., Koningstein, G., Dontje, W., Muller, M., Oudega, B., Luirink, J., *et al.* (2001). The early interaction of the outer membrane protein phoe with the periplasmic chaperone Skp occurs at the cytoplasmic membrane. *J Biol Chem, 276*(22), 18804-18811.

Hayashi, S., & Wu, H. C. (1990). Lipoproteins in bacteria. *J Bioenerg Biomembr, 22*(3), 451-471.

Hazlett, K. R., Cox, D. L., Decaffmeyer, M., Bennett, M. P., Desrosiers, D. C., La Vake, C. J., et al. (2005). TP0453, a concealed outer membrane protein of *Treponema pallidum*, enhances membrane permeability. *J Bacteriol, 187*(18), 6499-6508.

Hennecke, G., Nolte, J., Volkmer-Engert, R., Schneider-Mergener, J., & Behrens, S. (2005). The periplasmic chaperone SurA exploits two features characteristic of integral outer membrane proteins for selective substrate recognition. *J Biol Chem, 280*(25), 23540-23548.

Holt, S. C. (1978). Anatomy and chemistry of spirochetes. *Microbiol Rev, 42*(1), 114-160.

Hsu, P. L., Chamberlain, N. R., Orth, K., Moomaw, C. R., Zhang, L. Q., Slaughter, C. A., *et al*. (1989). Sequence analysis of the 47-kilodalton major integral membrane immunogen of *Treponema pallidum*. *Infect Immun, 57*(1), 196-203.

Iqbal, J., Zaidi, M., Narayan, G., Avadhani. (2010). Cell signaling. *Mol. Int. Phys. Musc. Sys*., 1211, 3-8.

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.

Jepsen, O. B., Hougen, K. H., & Birch-Andersen, A. (1968). Electron microscopy of *Treponema pallidum* Nichols. *Acta Pathol Microbiol Scand, 74*(2), 241-258.

Juncker, A. S., Willenbrock, H., Von Heijne, G., Brunak, S., Nielsen, H., & Krogh, A. (2003). Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein Sci, 12*(8), 1652-1662.

Keenan, J., Day, T., Neal, S., Cook, B., Perez-Perez, G., Allardyce, R., *et al*. (2000). A role for the bacterial outer membrane in the pathogenesis of *Helicobacter pylori* infection. *FEMS Microbiol Lett, 182*(2), 259-264.

Kelley, L. A., & Sternberg, M. J. (2009). Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc, 4*(3), 363-371.

Kholodenko, B.N., Hancock, J.F., Kolch, W. (2010). Signalling ballet in space and time. *Nat. Rev. Mol. Cell Biol*., 11, 1051-1053.

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.

Knowles, T. J., Jeeves, M., Bobat, S., Dancea, F., McClelland, D., Palmer, T., *et al*. (2008). Fold and function of polypeptide transport-associated domains responsible for delivering unfolded proteins to membranes. *Mol Microbiol, 68*(5), 1216-1227.

Knowles, T. J., Scott-Tucker, A., Overduin, M., & Henderson, I. R. (2009). Membrane protein architects: the role of the BAM complex in outer membrane protein assembly. *Nat Rev Microbiol, 7*(3), 206-214.

Koebnik, R., Locher, K. P., & Van Gelder, P. (2000). Structure and function of bacterial outer membrane proteins: barrels in a nutshell. *Mol Microbiol, 37*(2), 239-253.

Krojer, T., Sawa, J., Schafer, E., Saibil, H. R., Ehrmann, M., & Clausen, T. (2008). Structural basis for the regulated protease and chaperone function of DegP. *Nature, 453*(7197), 885-890.

Kubitschek, H. E. (1990). Cell volume increase in *Escherichia coli* after shifts to richer media. *J Bacteriol, 172*(1), 94-101.

Lafond, R. E., & Lukehart, S. A. (2006). Biological basis for syphilis. *Clin Microbiol Rev, 19*(1), 29-49.

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

Lee, J. H., Choi, H. J., Jung, J., Lee, M. G., Lee, J. B., & Lee, K. H. (2003). Receptors for *Treponema pallidum* attachment to the surface and matrix proteins of cultured human dermal microvascular endothelial cells. *Yonsei Med J, 44*(3), 371-378.

Lipinska, B., Zylicz, M., & Georgopoulos, C. (1990). The HtrA (DegP) protein, essential for *Escherichia coli* survival at high temperatures, is an endopeptidase. *J Bacteriol, 172*(4), 1791-1797.

Lukatsky, D. B., & Shakhnovich, E. I. (2008). Statistically enhanced promiscuity of structurally correlated patterns. *Phys Rev E Stat Nonlin Soft Matter Phys, 77*(2 Pt 1), 020901.

Lukehart, S. A., Baker-Zander, S. A., Lloyd, R. M., & Sell, S. (1980). Characterization of lymphocyte responsiveness in early experimental syphilis. II. Nature of cellular infiltration and *Treponema pallidum* distribution in testicular lesions. *J Immunol, 124*(1), 461-467.

Lukehart, S. A., & Miller, J. N. (1978). Demonstration of the *in vitro* phagocytosis of *Treponema pallidum* by rabbit peritoneal macrophages. *J Immunol, 121*(5), 2014-2024.

Malinverni, J. C., Werner, J., Kim, S., Sklar, J. G., Kahne, D., Misra, R., *et al.* (2006). YfiO stabilizes the YaeT complex and is essential for outer membrane protein assembly in *Escherichia coli*. *Mol Microbiol, 61*(1), 151-164.

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

Mecsas, J., Rouviere, P. E., Erickson, J. W., Donohue, T. J., & Gross, C. A. (1993). The activity of sigma E, an *Escherichia coli* heat-inducible sigma-factor, is modulated by expression of outer membrane proteins. *Genes Dev, 7*(12B), 2618-2628.

Miernyk, J.A., Thelen, J.J. (2008). Biochemical approaches for discovering protein-protein interactions. *The Plant Journal,* 53, 597-609.

Mitra, K., Frank, J. (2006). Ribosome dynamics: insights from atomic structure modeling into cryo-electron microscopy maps. *Annu. Rev. Biophys. Biomol. Struct.,* 35, 299-317.

Mori, H., & Ito, K. (2001). The Sec protein-translocation pathway. *Trends Microbiol, 9*(10), 494-500.

Mulay, V., Caimano, M. J., Liveris, D., Desrosiers, D. C., Radolf, J. D., & Schwartz, I. (2007). *Borrelia burgdorferi* BBA74, a periplasmic protein associated with the outer membrane, lacks porin-like properties. *J Bacteriol, 189*(5), 2063-2068.

Nakai, K., & Horton, P. (1999). PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends Biochem Sci, 24*(1), 34-36.

Nakamoto, H., & Bardwell, J. C. (2004). Catalysis of disulfide bond formation and isomerization in the *Escherichia coli* periplasm. *Biochim Biophys Acta, 1694*(1-3), 111-119.

Nikaido, H. (1996). Outer membrane. In *Escherichia coli and Salmonella: cellular and molecular biology.* (2nd ed., pp. 29-47). Washington, DC: American Society for Microbiology.

Nooren, I. M., & Thornton, J. M. (2003). Diversity of protein-protein interactions. *Embo J, 22*(14), 3486-3492.

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.

Oliver, S. (2000). Guilt-by-association goes global. *Nature, 403*(6770), 601-603.

Onufryk, C., Crouch, M. L., Fang, F. C., & Gross, C. A. (2005). Characterization of six lipoproteins in the sigmaE regulon. *J Bacteriol, 187*(13), 4552-4561.

Papanikou, E., Karamanou, S., & Economou, A. (2007). Bacterial protein secretion through the translocase nanomachine. *Nat Rev Microbiol, 5*(11), 839-851.

Parsons, L. M., Lin, F., & Orban, J. (2006). Peptidoglycan recognition by Pal, an outer membrane lipoprotein. *Biochemistry, 45*(7), 2122-2128.

Patti, J. M., Allen, B. L., McGavin, M. J., & Hook, M. (1994). MSCRAMM-mediated adherence of microorganisms to host tissues. *Annu Rev Microbiol, 48*, 585-617.

Penn, C. W., Cockayne, A., & Bailey, M. J. (1985). The outer membrane of *Treponema pallidum*: biological significance and biochemical properties. *J Gen Microbiol, 131*(9), 2349-2357.

Phizicky, E. M., & Fields, S. (1995). Protein-protein interactions: methods for detection and analysis. *Microbiol Rev, 59*(1), 94-123.

Pickering, L. K. (2006). *Syphilis, Red Book.* Elk Grove Village, IL: American Academy of Pediatrics. pp. 631-644.

Pike, R. M. (1976). Laboratory-associated infections: summary and analysis of 3921 cases. *Health Lab Sci, 13*(2), 105-114.

Pons, T., Gomez, R., Chinea, G., & Valencia, A. (2003). Beta-propellers: associated functions and their role in human diseases. *Curr Med Chem, 10*(6), 505-524.

Public Health Agency of Canada. Website: Syphilis (2008). Retrieved August 30, 2010

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. *Infect Immun, 56*(7), 1825-1828.

Radolf, J. D., Norgard, M. V., & Schulz, W. W. (1989). Outer membrane ultrastructure explains the limited antigenicity of virulent *Treponema pallidum*. *Proc Natl Acad Sci U S A, 86*(6), 2051-2055.

Radolf, J. D., Robinson, E. J., Bourell, K. W., Akins, D. R., Porcella, S. F., Weigel, L. M., *et al*. (1995). Characterization of outer membranes isolated from *Treponema pallidum*, the syphilis spirochete. *Infect Immun, 63*(11), 4244-4252.

Raivio, T. L., & Silhavy, T. J. (2001). Periplasmic stress and ECF sigma factors. *Annu Rev Microbiol, 55*, 591-624.

Rawlings, N. D., & Barrett, A. J. (1993). Evolutionary families of peptidases. *Biochem J, 290 (Pt 1)*, 205-218.

Rawlings, N. D., Barrett, A. J., & Bateman, A. (2010). MEROPS: the peptidase database. *Nucleic Acids Res, 38*(Database issue), D227-233.

Reardon, D., & Farber, G. K. (1995). The structure and evolution of alpha/beta barrel proteins. *Faseb J, 9*(7), 497-503.

Remaut, H., & Waksman, G. (2006). Protein-protein interaction through beta-strand addition. *Trends Biochem Sci, 31*(8), 436-444.

Rhodius, V. A., Suh, W. C., Nonaka, G., West, J., & Gross, C. A. (2006). Conserved and variable functions of the sigmaE stress response in related genomes. *PLoS Biol, 4*(1), e2.

Riviere, G. R., Thomas, D. D., & Cobb, C. M. (1989). In vitro model of *Treponema pallidum* invasiveness. *Infect Immun, 57*(8), 2267-2271.

Rizzitello, A. E., Harper, J. R., & Silhavy, T. J. (2001). Genetic evidence for parallel pathways of chaperone activity in the periplasm of *Escherichia coli*. *J Bacteriol, 183*(23), 6794-6800.

Robert, V., Volokhina, E. B., Senf, F., Bos, M. P., Van Gelder, P., & Tommassen, J. (2006). Assembly factor Omp85 recognizes its outer membrane protein substrates by a species-specific C-terminal motif. *PLoS Biol, 4*(11), e377.

Rost, B., Yachdav, G., & Liu, J. (2004). The PredictProtein server. *Nucleic Acids Res, 32*(Web Server issue), W321-326.

Rouviere, P. E., & Gross, C. A. (1996). SurA, a periplasmic protein with peptidyl-prolyl isomerase activity, participates in the assembly of outer membrane porins. *Genes Dev, 10*(24), 3170-3182.

Ruiz, N., Kahne, D., & Silhavy, T. J. (2006). Advances in understanding bacterial outer-membrane biogenesis. *Nat Rev Microbiol, 4*(1), 57-66.

Rutten, L., Mannie, J. P., Stead, C. M., Raetz, C. R., Reynolds, C. M., Bonvin, A. M., *et al.* (2009). Active-site architecture and catalytic mechanism of the lipid A deacylase LpxR of *Salmonella typhimurium*. *Proc Natl Acad Sci U S A, 106*(6), 1960-1964.

Salazar, J. C., Hazlett, K. R., & Radolf, J. D. (2002). The immune response to infection with *Treponema pallidum*, the stealth pathogen. *Microbes Infect, 4*(11), 1133-1140.

Sanchez-Pulido, L., Devos, D., Genevrois, S., Vicente, M., & Valencia, A. (2003). POTRA: a conserved domain in the FtsQ family and a class of beta-barrel outer membrane proteins. *Trends Biochem Sci, 28*(10), 523-526.

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

Sellati, T. J., Bouis, D. A., Caimano, M. J., Feulner, J. A., Ayers, C., Lien, E., *et al*. (1999). Activation of human monocytic cells by *Borrelia burgdorferi* and *Treponema pallidum* is facilitated by CD14 and correlates with surface exposure of spirochetal lipoproteins. *J Immunol, 163*(4), 2049-2056.

Shevchenko, D. V., Akins, D. R., Robinson, E. J., Li, M., Shevchenko, O. V., & Radolf, J. D. (1997). Identification of homologs for thioredoxin, peptidyl prolyl cis-trans isomerase, and glycerophosphodiester phosphodiesterase in outer membrane fractions from *Treponema pallidum*, the syphilis spirochete. *Infect Immun, 65*(10), 4179-4189.

Shevchenko, D. V., Sellati, T. J., Cox, D. L., Shevchenko, O. V., Robinson, E. J., & Radolf, J. D. (1999). Membrane topology and cellular location of the *Treponema pallidum* glycerophosphodiester phosphodiesterase (GlpQ) ortholog. *Infect Immun, 67*(5), 2266-2276.

Short, J. M., Fernandez, J. M., Sorge, J. A., & Huse, W. D. (1988). Lambda ZAP: a bacteriophage lambda expression vector with in vivo excision properties. *Nucleic Acids Res, 16*(15), 7583-7600.

Sidhu, S. S., Fairbrother, W. J., & Deshayes, K. (2003). Exploring protein-protein interactions with phage display. *Chembiochem, 4*(1), 14-25.

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*. *Proc Natl Acad Sci U S A, 104*(15), 6400-6405.

Sklar, J. G., Wu, T., Kahne, D., & Silhavy, T. J. (2007). Defining the roles of the periplasmic chaperones SurA, Skp, and DegP in *Escherichia coli*. *Genes Dev, 21*(19), 2473-2484.

Soding, J. (2005). Protein homology detection by HMM-HMM comparison. *Bioinformatics, 21*(7), 951-960.

Spiess, C., Beil, A., & Ehrmann, M. (1999). A temperature-dependent switch from chaperone to protease in a widely conserved heat shock protein. *Cell, 97*(3), 339-347.

Stamm, L. V., & Bassford, P. J., Jr. (1985). Cellular and extracellular protein antigens of *Treponema pallidum* synthesized during in vitro incubation of freshly extracted organisms. *Infect Immun, 47*(3), 799-807.

Stebeck, C. E., Shaffer, J. M., Arroll, T. W., Lukehart, S. A., & Van Voorhis, W. C. (1997). Identification of the *Treponema pallidum* subsp. *pallidum* glycerophosphodiester phosphodiesterase homologue. *FEMS Microbiol Lett, 154*(2), 303-310.

Struyve, M., Moons, M., & Tommassen, J. (1991). Carboxy-terminal phenylalanine is essential for the correct assembly of a bacterial outer membrane protein. *J Mol Biol, 218*(1), 141-148.

Tokuda, H. (2009). Biogenesis of outer membranes in Gram-negative bacteria. *Biosci Biotechnol Biochem, 73*(3), 465-473.

Tokuda, H., Matsuyama, S., Tanaka-Masuda, K. (2007). Structure, function and transport of lipoproteins in *Escherichia coli.* In M. Ehrmann (Ed.), *The Periplasm* (pp. 67-79). Washington DC: ASM Press.

Tommassen, J. (2010). Assembly of outer-membrane proteins in bacteria and mitochondria. *Microbiology, 156*(Pt 9), 2587-2596.

Tsukazaki, T., Mori, H., Fukai, S., Ishitani, R., Mori, T., Dohmae, N., *et al.* (2008). Conformational transition of Sec machinery inferred from bacterial SecYE structures. *Nature, 455*(7215), 988-991.

Turner, T. B., & Hollander, D. H. (1957). Biology of the treponematoses based on studies carried out at the International Treponematosis Laboratory Center of the Johns Hopkins University under the auspices of the World Health Organization. *Monogr Ser World Health Organ*(35), 3-266.

Uchiyama, I. (2003). MBGD: microbial genome database for comparative analysis. *Nucleic Acids Res, 31*(1), 58-62.

UniProt. (2010). The Universal Protein Resource (UniProt) in 2010. *Nucleic Acids Res, 38*(Database issue), D142-148.

Ureta, A. R., Endres, R. G., Wingreen, N. S., & Silhavy, T. J. (2007). Kinetic analysis of the assembly of the outer membrane protein LamB in *Escherichia coli* mutants each lacking a secretion or targeting factor in a different cellular compartment. *J Bacteriol, 189*(2), 446-454.

Van Voorhis, W. C., Barrett, L. K., Lukehart, S. A., Schmidt, B., Schriefer, M., & Cameron, C. E. (2003). Serodiagnosis of syphilis: antibodies to recombinant Tp0453, Tp92, and Gpd proteins are sensitive and specific indicators of infection by *Treponema pallidum*. *J Clin Microbiol, 41*(8), 3668-3674.

von Specht, B., Knapp, B., Hungerer, K., Lucking, C., Schmitt, A., & Domdey, H. (1996). Outer membrane proteins of *Pseudomonas aeruginosa* as vaccine candidates. *J Biotechnol, 44*(1-3), 145-153.

Voulhoux, R., Bos, M. P., Geurtsen, J., Mols, M., & Tommassen, J. (2003). Role of a highly conserved bacterial protein in outer membrane protein assembly. *Science, 299*(5604), 262-265.

Voulhoux, R., & Tommassen, J. (2004). Omp85, an evolutionarily conserved bacterial protein involved in outer-membrane-protein assembly. *Res Microbiol, 155*(3), 129-135.

Vuong, P., Bennion, D., Mantei, J., Frost, D., & Misra, R. (2008). Analysis of YfgL and YaeT interactions through bioinformatics, mutagenesis, and biochemistry. *J Bacteriol, 190*(5), 1507-1517.

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

Walsh, N. P., Alba, B. M., Bose, B., Gross, C. A., & Sauer, R. T. (2003). OMP peptide signals initiate the envelope-stress response by activating DegS protease via relief of inhibition mediated by its PDZ domain. *Cell, 113*(1), 61-71.

Weiser, J. N., & Gotschlich, E. C. (1991). Outer membrane protein A (OmpA) contributes to serum resistance and pathogenicity of *Escherichia coli* K-1. *Infect Immun, 59*(7), 2252-2258.

World Health Organization. (2003). Primary and secondary syphilis--United States, 2002. *MMWR Morb Mortal Wkly Rep, 52*(46), 1117-1120.

World Health Organization. Website: Disease watch 2008: Syphilis. Retrieved August 21, 2010.

Wilson, M., McNab, R., Henderson, B. (2002). Bacterial protein secretion systems. In *Bacterial Disease Mechanisms* (pp. 61-71). Eastman Dental Institute, University College London: Cambridge University Press.

Workowski, K. A., & Berman, S. M. (2006). Sexually transmitted diseases treatment guidelines, 2006. *MMWR Recomm Rep, 55*(RR-11), 1-94.

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.

Yu, N. Y., Wagner, J. R., Laird, M. R., Melli, G., Rey, S., Lo, R., *et al*. (2010). PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics, 26*(13), 1608-1615.

Zdobnov, E. M., & Apweiler, R. (2001). InterProScan--an integration platform for the signature-recognition methods in InterPro. *Bioinformatics, 17*(9), 847-848.