


12-2018

# ASSEMBLY AND DISPLAY OF SURFACE PROTEINS IN ACTINOMYCES ORIS

Sara Siegel

Follow this and additional works at: [https://digitalcommons.library.tmc.edu/utgsbs\\_dissertations](https://digitalcommons.library.tmc.edu/utgsbs_dissertations)

 Part of the [Bacteriology Commons](#), [Medicine and Health Sciences Commons](#), and the [Microbial Physiology Commons](#)

---

## Recommended Citation

Siegel, Sara, "ASSEMBLY AND DISPLAY OF SURFACE PROTEINS IN ACTINOMYCES ORIS" (2018). *UT GSBS Dissertations and Theses (Open Access)*. 900.

[https://digitalcommons.library.tmc.edu/utgsbs\\_dissertations/900](https://digitalcommons.library.tmc.edu/utgsbs_dissertations/900)

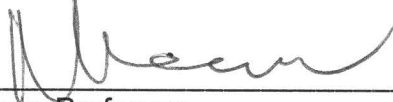
This Dissertation (PhD) is brought to you for free and open access by the Graduate School of Biomedical Sciences at DigitalCommons@TMC. It has been accepted for inclusion in UT GSBS Dissertations and Theses (Open Access) by an authorized administrator of DigitalCommons@TMC. For more information, please contact [laurel.sanders@library.tmc.edu](mailto:laurel.sanders@library.tmc.edu).

ASSEMBLY AND DISPLAY OF SURFACE PROTEINS  
IN ACTINOMYCES ORIS

By


Sara Danielle Siegel, B.S.

APPROVED:

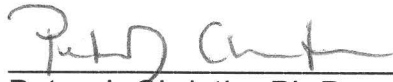


Advisory Professor

Hung Ton-That, Ph.D.



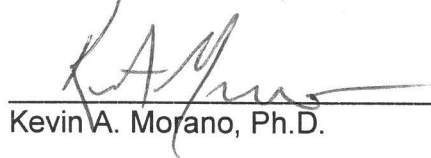
Jeffrey K. Actor, Ph.D.



Peter J. Christie, Ph.D.



Nicholas R. De Lay, Ph.D.



Kevin A. Morano, Ph.D.

APPROVED:

\_\_\_\_\_  
Dean, The University of Texas

MD Anderson Cancer Center UTHHealth Graduate School of Biomedical Sciences

ASSEMBLY AND DISPLAY OF SURFACE PROTEINS  
IN ACTINOMYCES ORIS

A

DISSERTATION

Presented to the Faculty of

The University of Texas

MD Anderson Cancer Center UTHealth

Graduate School of Biomedical Science

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

By

Sara Danielle Siegel, B.S.

Houston, Texas

December, 2018

## ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Hung Ton-That for his guidance through which I have learned patience, independence and resilience. He taught me to pursue the questions that intrigue me and how to tell a scientific story, complete with illustrations. Hung is always in my corner, and I have benefited greatly from working with him over the last five years.

I would like to thank all of the members of my advisory committee Dr. Kevin Morano, Dr. Peter Christie, Dr. Nicholas De Lay and Dr. Jeffrey Actor for all of your great advice to move my projects forward and your kind, encouraging words. I would also like to thank my exam chair, Dr. Danielle Garsin, for preparing me to take the intimidating candidacy exam. As well as Lyz Culpepper who simplified so many of the moving parts during my graduate education.

I am so thankful for the excellent lab mates that I have had over the years. Everything would have been more difficult without the advice, encouragement, and help that I received every day from Julie. I was also fortunate to have Truc as a friend, she brought so much life into the lab, and I continue to seek her advice. My bench mate Belkys has made my time so enjoyable, and I will miss turning around to talk to her about science and life. Amar was always a positive and helpful lab mate and friend. I'm so lucky to have worked with Alexis and Cara who were always cheerful and were a reminder that science is pretty cool. My source of *Actinomyces* know-how was always Wu, and his help made it possible for me to accomplish a lot of my goals. Luis, Melissa, Ju Huck, Bo, Nguyen, and Cuong have all contributed valuable technical expertise and suggestions for my projects. I look forward to seeing all of your future successes!

It has not been an easy road, but it was definitely easier with the support from my friends going through the same thing. My classmates Chris and Elisa have been

endlessly supportive and offered a reprieve when everything felt overwhelming. I appreciate that Chris has always offered a different perspective. And that Elisa has been a positive influence in all things. I am excited to see where we will end up! I have been encouraged by Robert and Sarah to climb to new heights and try new things. Kara has never failed to make me laugh an embarrassing amount and brighten my day. Alex was an amazing source of knowledge and encouragement. Makayla made life feel a little bit less stressful, and I am going to sorely miss board game nights. And finally, all of the people in the MMG department have been amazing and I was fortunate to unwittingly join a department as excited about Halloween costumes as I am.

My life would not be as colorful without my husband, Andy. Even though he was catapulted into the very strange world of graduate school, he never seemed to mind. I'm thrilled that he decided to marry me despite my terrible health habits, long hours, and bizarre stress laughing. I look forward to all of our future adventures together. My Lila, who is always excited when I come home no matter how grumpy I am, I feel so much better when you give me an adorable puppy hug. I love you both with all of myself.

My dad was always my voice of reason and supported me even when he thought what I wanted to do was weird. If I needed a boost of enthusiasm I could also go to my mom – also, thanks for reading to me, even in utero. My step dad always believed in me. And my little sister was always on my side. I am so lucky to have had all of your love and support.

ASSEMBLY AND DISPLAY OF SURFACE PROTEINS  
IN ACTINOMYCES ORIS

Sara Danielle Siegel, B.S.

Advisory Professor: Hung Ton-That, Ph.D.

Bacteria are an integral part of human health and disease. In the human host, dental plaques form as a result of up to 700 individual bacterial species colonizing oral surfaces and forming a multispecies biofilm. These biofilms are the cause of prevalent human diseases such as dental caries, gingivitis, and periodontitis. The microbes present in the oral biofilm are highly spatially and temporally structured and require a primary colonizing species to adhere to host tissue. As an important primary colonizer of the oral biofilm, the actinobacterium *Actinomyces oris* utilizes cell wall anchored proteins and glycoconjugates to initiate adherence to host surfaces, recruit additional bacterial species that could not bind otherwise, and maintain the structural integrity of the oral biofilm. In this thesis, I reveal mechanisms involved in the assembly and display of surface proteins that are central to these processes in *A. oris*.

Cell wall-anchored proteins contain a signal peptide to direct their secretion to the exoplasmic side of the membrane, where they are liberated from the secretion machine by signal peptidases. Cell wall-anchored proteins also contain a cell wall sorting signal, which is required for their covalent attachment to peptidoglycan by transpeptidase enzymes called sortases. Furthermore, a subset of cell wall-anchored proteins are polymerized to form pili prior to being anchored. I found that pilin proteins

require a distinct signal peptidase for their maturation and function and uncovered residues required for adherence in a minor pilin protein. In certain cases after translocation, proteins are modified by the addition of glycopolymers, and I characterized a phosphotransferase enzyme with a novel role in protein glycosylation. These studies contribute to the understanding of the role of *A. oris* as a primary colonizer in the oral biofilm. Additionally, using *A. oris* as a model for general processes has led to findings which are applicable to principles of biofilm formation, interspecies interactions, glycoconjugate formation, and bacterial pathogenesis.

## TABLE OF CONTENTS

APPROVAL SHEET	i
TITLE PAGE	ii
ACKNOWLEDGMENTS	iii
ABSTRACT	v
TABLE OF CONTENTS	vii
LIST OF FIGURES	ix
LIST OF TABLES	xi
<b>CHAPTER I: Introduction</b>	<b>1</b>
<i>Actinomyces</i> and the Oral Biofilm	3
Glycoconjugates of the Gram-positive Cell Envelope	7
Cell Wall Anchored Proteins	10
Significance of these Studies	19
<b>CHAPTER II: Materials and Methods</b>	<b>21</b>
<b>CHAPTER III: A Type I Signal Peptidase is Required for Pilus Assembly in the Gram-positive, Biofilm-forming Bacterium <i>Actinomyces oris</i></b>	<b>48</b>
Introduction	49
Results	51
Discussion	72
<b>CHAPTER IV: A Phosphotransferase LCP Enzyme Mediates Glycosylation of a Gram-positive Cell Wall Anchored Protein</b>	<b>75</b>
Introduction	76
Results	78
Discussion	98



<b>CHAPTER V: Mapping the Molecular Domains in CafA Responsible for Interkingdom Adherence</b>	102
Introduction	103
Results	106
Discussion	116
<b>CHAPTER VI: Exploring the Role of the Twin-Arginine Residues in the CafA Signal Peptide</b>	118
Introduction	119
Results	121
Discussion	130
<b>CHAPTER VII: Future Implications and Perspective</b>	132
A Model for Bacterial Type I Signal Peptidases	133
A New Paradigm for an LCP Enzyme	137
Interkingdom Adhesin with a Pivotal Role in Oral Biofilm Formation	142
Essentiality of the Twin-Arginine Translocon	145
Final Remarks	147
<b>BIBLIOGRAPHY</b>	148
<b>VITA</b>	167

## LIST OF FIGURES

<b>Figure 1-1</b>	Known cell surface protein dependent processes in <i>Actinomyces oris</i>	6
<b>Figure 1-2</b>	Sortase anchors monomeric and polymeric substrates to the cell wall	14
<b>Figure 1-3</b>	Type 2 pili are polymerized by the pilus-specific sortase SrtC2	18
<b>Figure 3-1</b>	Signal peptidase LepB2 is a genetic suppressor of SrtA essentiality	52
<b>Figure 3-2</b>	<i>lepB1</i> and <i>lepB2</i> are independently expressed	54
<b>Figure 3-3</b>	Involvement of SPases in sortase-associated GspA glycosylation	56
<b>Figure 3-4</b>	Mapping the signal peptide cleavage site by Edman degradation	58
<b>Figure 3-5</b>	The signal peptidase LepB2 is required for pilus assembly	65
<b>Figure 3-6</b>	Requirement of LepB2 for surface expression of tip pilins	66
<b>Figure 3-7</b>	The Ser-Lys catalytic dyad of LepB2 is necessary for pilus assembly	68
<b>Figure 3-8</b>	Requirement of LepB2 for polymicrobial interactions and biofilm formation	71
<b>Figure 4-1</b>	LcpA is solely responsible for GspA glycosylation	80
<b>Figure 4-2</b>	Crystal structure of <i>A. oris</i> LcpA and structural requirements for glycosylation activity	84
<b>Figure 4-3</b>	The disulfide bond C179-C365 is required for LcpA stability	88
<b>Figure 4-4</b>	LcpA exhibits pyrophosphatase activity	92

<b>Figure 4-5</b>	LcpA interacts with GspA in solution and catalyzes phosphotransfer	96
<b>Figure 4-6</b>	Proposed model for LcpA-mediated glycosylation of the cell wall-anchored protein GspA	101
<b>Figure 5-1</b>	CafA adheres to receptor polysaccharides present on HGF-1 and oral streptococci	107
<b>Figure 5-2</b>	Characterization of type 2 pili and pilus-related functions in <i>Actinomyces oris</i> strain N11A12	112
<b>Figure 5-3</b>	Analysis of nonsynonymous mutations between CafA from MG1 and N11A12 identified residues necessary for coaggregation	115
<b>Figure 6-1</b>	Mutation of Arg 11 and Arg 12 to Ala abrogates CafA secretion and destabilizes CafA	122
<b>Figure 6-2</b>	Evaluation of Tat-dependent reporter protein fusions	125
<b>Figure 6-3</b>	Requirement of bulky hydrophobic residues for protein stability and examination of GFP expressing strains by fluorescence microscopy	127
<b>Figure 6-4</b>	Analysis of the <i>tatC</i> conditional mutant by electron microscopy and SDS sensitivity	129
<b>Figure 7-1</b>	Sequence alignment of <i>A. oris</i> LepB1 and LepB2	135
<b>Figure 7-2</b>	Phylogenetic analysis of LCP proteins	140
<b>Figure 7-3</b>	Purification of CafA from <i>A. oris</i> supernatant	144

## LIST OF TABLES

<b>Table 2-1</b>	Bacterial strains and cell lines used in these studies	24
<b>Table 2-2</b>	Plasmids used in these studies	32
<b>Table 2-3</b>	Primers used in these studies	35
<b>Table 3-1</b>	Protein sequencing of GspA proteins purified from four <i>A. oris</i> strains by Edman Degradation	61
<b>Table 3-2</b>	Protein sequencing of FimA proteins purified from four <i>A. oris</i> strains by Edman Degradation	63
<b>Table 5-1</b>	Coaggregation and CafA polymerization phenotypes of <i>Actinomyces</i> clinical isolates	109
<b>Table 7-1</b>	Predicted Tat-dependent proteins in the <i>A. oris</i> genome	146

**CHAPTER I**  
**Introduction**

*This work is, in part, adapted or reprinted from works published in 2016: Siegel SD, Liu J, Ton-That H. (2016) Biogenesis of the Gram-positive bacterial cell envelope. Curr Opin Microbiol. doi: 10.1016/j.mib.2016.07.015 and Siegel SD, Reardon ME, Ton-That H. (2016) Anchoring of LPXTG-Like Proteins to the Gram-Positive Cell Wall Envelope. Curr Top Microbiol Immunol. doi: 10.1128/EC.00083-15. I am the first author of these publications and was responsible for preparing the original manuscript. I have permission to reproduce any and all of this manuscript, in print or electronically, for the purpose of my thesis in accordance with Elsevier (publisher of Current Opinion in Microbiology) "Journal Author Rights." <https://www.elsevier.com/about/policies/copyright> and have permission to reprint Figure 1 from Springer Nature Terms and Conditions for RightsLink Permissions Springer Customer Service Centre GmbH : Springer Nature. Chapter "Anchoring of LPXTG-Like Proteins to the Gram-Positive Cell Wall Envelope" in Protein and Sugar Export and Assembly in Gram-positive Bacteria by Sara D. Siegel, Melissa E. Reardon and Hung Ton-That. © Springer International Publishing Switzerland 2016 (2016).*

The bacterial cell envelope serves as a platform for molecular interactions. The outermost layer of the Gram-positive envelope, the cell wall, is decorated with proteins and glycoconjugates that mediate interactions with the local environment. In this thesis, I reveal molecular mechanisms underlying the assembly and display of proteins on the surface of the Gram-positive actinobacterium *Actinomyces oris*. My work expands the understanding of how these proteins are elaborated and their role in interaction with host and microbes that make *A. oris* a major contributor to the formation of the multispecies oral biofilm.

## ***Actinomyces oris* and the Oral Biofilm**

Dental plaques are multispecies biofilms that form on oral surfaces and consist of up to 700 different bacterial species (1). The oral biofilm is more easily accessible than other human-associated microbial communities and can serve as a model for biofilms that form on non-shedding surfaces (2). Early realizations of the complexity of this biofilm sparked a foray into understanding the species present in the biofilm, what drove these interactions, and whether they proceeded at random or with high specificity (3). These studies demonstrated that pairwise interactions between two species, known as coaggregation, is a regulated process mediated by specific receptor and adhesin pairs (4,5). Expansion of genome based research also provided extensive knowledge about the players involved at different stages of biofilm formation (6).

From these initial works multispecies biofilm development is proposed to occur through a stepwise process (1,2). First, primary colonizers bind a conditioned substratum then expand and secrete exopolysaccharide (EPS) eventually forming a microcolony. In the oral biofilm, salivary proteins deposited on the tooth surface serve as the conditioned substratum. These initial microcolonies modify the local environment attracting secondary colonizers to the biofilm. Growth and secretion of EPS by secondary colonizers culminates in the formation of a mature multispecies biofilm (1,2,7). This regulated, hierarchical accretion of bacterial species has been confirmed by modern micron-scale biogeographical studies that show oral biofilms exhibit a high degree of spatial structuring (8,9). A shared matrix and a variety of molecular interactions resulting from physical attachment and soluble molecule signals maintain the biofilm through harsh conditions and inconstant nutrient availability (10,11).

Although oral biofilms do form in healthy hosts, problems arise from overgrowth of these complex communities and subsequent recruitment of key pathogenic species that have been demonstrated to subvert host immunity and modulate local and systemic

inflammatory responses (12). Pathogenic microbes in the oral biofilm can cause a number of infectious diseases such as dental caries, gingivitis, and periodontitis (6). Dental diseases caused by poor oral hygiene are considered by the World Health Organization as among the most important global oral health burdens due to their high prevalence and impact on human health (13). There is also extensive evidence that suggests oral bacteria contribute to systemic diseases including cardiovascular disease and preterm birth via a systemic inflammatory response to the presence of microbial species disseminated through the bloodstream of those with poor oral health and hygiene (14,15).

One bacterial genus that has emerged as a key primary colonizer and a major contributor to oral biofilm formation is *Actinomyces*. *Actinomyces* demonstrate the unique ability to interact with the tooth, gingiva and bacterial species (16-18), are frequently isolated from oral biofilms (19-21), and remain abundant at the basal layer as the biofilm expands (9,22-24). Molecular analyses have uncovered drivers of *Actinomyces* ability to facilitate this structural and temporal hierarchy (18,25,26). This means that *Actinomyces* can maintain the oral biofilm adherence to the tooth as it expands below the gum line. Therefore, *Actinomyces* acts as a support to attract bacteria that directly release harmful molecules to damage tissue and persist in the periodontal pocket where they would otherwise be unable to adhere and persist (9,27).

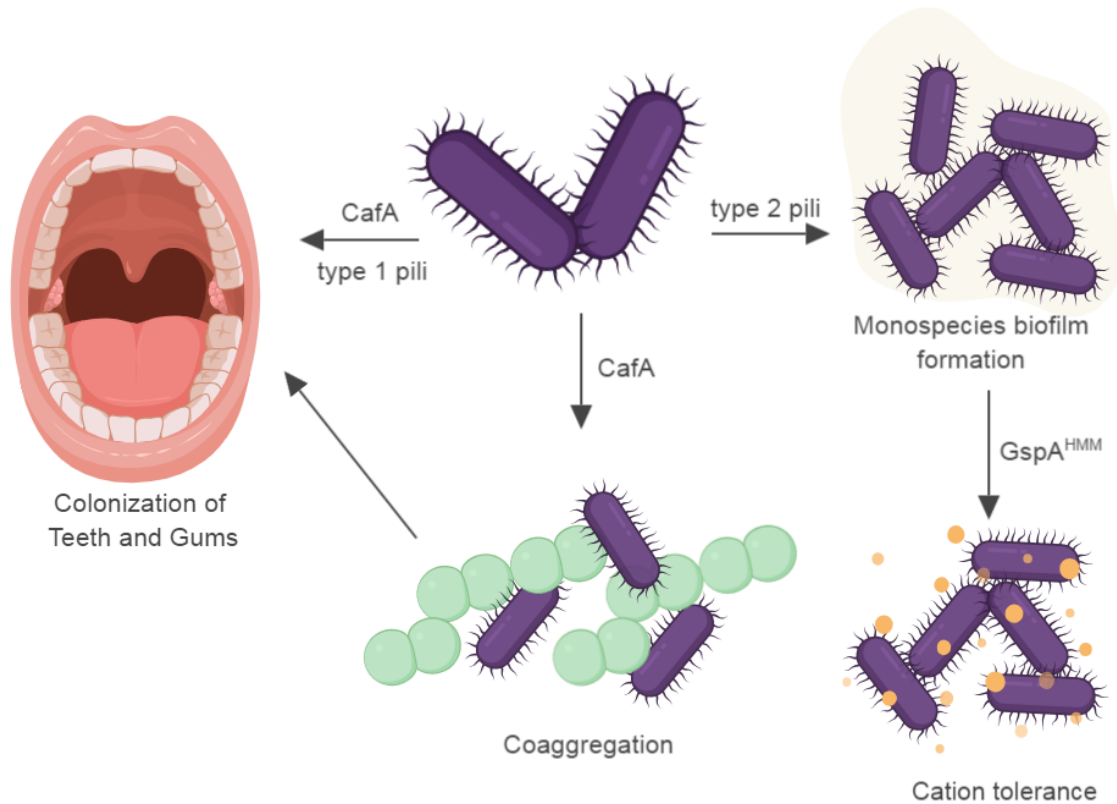
There are several species of *Actinomyces* associated with the oral biofilm, and of these *Actinomyces oris* has been developed as a primary model (3,21). The niche of *A. oris* as a primary colonizer is heavily dependent on proteins anchored to the cell wall. Therefore, *A. oris* serves as an outstanding model for basic principles of cell wall anchoring, pilus polymerization, and molecular interactions necessary for multi- and mono-species biofilm formation. We also have defined morphological characteristics



(Fig. 1-1), established molecular techniques, and an arsenal of approaches to address the physiological impact of genetic manipulations in *A. oris*.

The role of *A. oris* as a primary colonizer of the oral biofilm is associated with the expression of two distinct pilus types (21). *Actinomyces* pili are among the earliest imaged bacterial macrostructures by electron microscopy (28) and were the first sortase-catalyzed pilus operons identified by sequencing (29). Early studies in *A. oris* demonstrated that these pili are accessible to immune cells, and inhibition of either type 1 or type 2 pili leads to loss of colonization (30-34). However, molecular details of the biogenesis of these pili remained a mystery until sortase-catalyzed pili were described in *Corynebacterium diphtheriae* (35), a finding that uncovered conserved mechanisms for several Gram-positive pili including those from *Actinomyces* (36).

These features have established *A. oris* as a model to study actinobacterial protein secretion, cell envelope biogenesis, cell wall sorting, and molecular interactions that occur at the cell surface. Cell surface display of macromolecules begins after proteins are translocated across the cytoplasmic membrane and may require oxidative protein folding. Once outside of the cell, proteins can be further modified by a suite of membrane-associated enzymes. These modifications can include cell wall anchoring by sortase enzymes or the addition of glycopolymers. Proteins displayed on the cell surface are accessible to the environment and contribute to modulation of cellular physiology and interactions required for pathogenicity. And although pili are the most well described contributors to biofilm formation, we still do not have a complete understanding of either the players mediating additional interactions or the role of the suite of other cell wall associated proteins.



**Figure 1-1: Known cell surface protein dependent processes in *Actinomyces oris*.**

*A. oris* are represented as purple rod-shaped cells undergoing snapping division. Each arrow points to a physiological property that we can assess and the associated cell surface protein(s). Specifically, these proteins are covalently attached and/or polymerized by sortases. The colonization of teeth and gums depends on type 1 pili and CafA-containing type 2 pili, respectively. Additionally, CafA is required to mediate coaggregation with oral streptococci, shown here in light green. Monospecies biofilm formation is mediated by type 2 pili, specifically FimA polymers, and EPS is represented in the background. Maintenance of the mono- and multi-species biofilm in the presence of salts is mediated by GspA<sup>HMM</sup> (unpublished data from Abu Amar Al Mamun, used with permission), Na<sup>+</sup> and K<sup>+</sup> ions are represented with orange dots. Figure was generated using © BioRender.

## **Glycoconjugates of the Gram-positive Cell Envelope**

The Gram-positive cell envelope is a broadly defined structure encompassing the exoplasmic face of the cytoplasmic membrane and the characteristically thick layered meshwork of peptidoglycan called the cell wall as well as all of the components associated with these megastructures (37). The cytoplasmic membrane allows for nutrient acquisition, embeds transporter proteins, and maintains an electrochemical gradient important for energy production and protein transport. A thick layer of peptidoglycan surrounds the cytoplasmic membrane. Although we tend to regard the peptidoglycan as a static structure, new evidence suggests that there is a stunning variety of wall-associated glycoconjugates that are variable in presence and composition (38,39). In addition to glycoconjugates, there are also proteins covalently attached to the cell wall (40). The combination of these cell envelope features underlies bacterial viability, environmental interactions, and niche selection (39). Recent investigations of the Gram-positive cell envelope have focused on the individuality of molecular composition and how these variations in the cell envelope are derived and translated into distinct functions and phenotypes.

A major source of variation in the cell envelope arises from envelope associated glycoconjugates (38,39). In addition to the notably conserved glycoconjugate peptidoglycan, there are several additional glycoconjugates that contribute to the individuality of the bacterial cell envelope. These include the capsular polysaccharide, exopolysaccharide, glycan-modified pili and flagella, wall and lipo-teichoic acids, and glycoproteins.

### ***Peptidoglycan***

Peptidoglycan (PG) is the rigid bacterial glycoconjugate that provides structure for cell morphology and protects cells from turgor pressure resulting from external

osmotic changes. It is the most highly conserved glycoconjugate in bacteria consisting of N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc) sugar moieties and a pentapeptide attached to the MurNAc subunit. It is present in all bacteria except *Mycoplasma* (41). The GlcNAc-MurNAc-pentapeptide precursors are synthesized in the cytoplasm, attached to the universal lipid carrier undecaprenol phosphate (UndP), and flipped across the membrane. In the exoplasmic space, the precursors are accessible to transglycosylases that link MurNAc to GlcNAc moieties to form the glycan chains. Transpeptidase enzymes then covalently link these repeating MurNAc-GlcNAc-pentapeptide units together forming a peptide crossbridge. This process is repeated until eventually a durable mesh structure is formed (36). PG glycan chain length differs between species, but variations of the peptidoglycan structure are introduced primarily to the pentapeptide composition (41).

### ***Wall teichoic acids***

Wall teichoic acids (WTA) are long anionic glycopolymers that are attached to and threaded through the PG layer. These molecules extend beyond the cell wall and account for a significant portion of the cell mass (37,42). WTA polymers consist of repeating polyribitol phosphate or polyglycerol phosphate subunits attached by a phosphodiester bond to the MurNAc subunit on peptidoglycan. WTAs are negatively charged, and their display has been implicated in several different and species specific functions (37,39). The negative charge of the polymers gives teichoic acids the capacity to bind cations (42). Accordingly, one of the earliest functions attributed to WTAs was a role in cation homeostasis (43). WTAs have since been found to have a range of functions including protection against cationic antimicrobial peptides, reduced biofilm formation on abiotic surfaces, and resistance to lysozyme (39).

The attachment of teichoic acids to the cell wall is mediated by a family of enzymes containing LytR-CspA-PsrA domains or LCP enzymes (44). The activity of the

LCP enzymes extends beyond just WTA synthesis and have been described to catalyze the attachment of other glycopolymers to the bacterial cell wall such as arabinogalactan, glucose-rhamnose-polymer, and capsular polysaccharides (45-48). LCP enzymes have recently been found to protect *Mycobacterium tuberculosis* from LC3-associated phagocytosis without affecting arabinogalactan display or cell wall integrity suggesting an entirely different role for this particular LCP (49). In *A. oris*, the function of one LCP enzyme, LcpA, is the transfer of an unknown glycopolymer to a cell wall anchored protein called GspA (50). Combined, these data suggest that LCP enzymes recognize a spectrum of substrate and acceptor molecules. The specific activity of *A. oris* LcpA that is required to glycosylate GspA is investigated in chapter IV.

### ***Glycoproteins***

The function of LcpA in the glycosylation of a protein substrate came as a surprise because it is typically accepted that bacterial proteins are either N- or O-glycosylated meaning that the glycopolymers are attached to asparagine or serine/threonine residues, respectively. Protein glycosylation mechanisms are distinguished by whether they proceed via sequential or *en bloc* glycosylation (38). Substrates that are sequentially glycosylated are directly targeted by specific glycosyl transferase (GTase) enzymes that transfer specific sugar molecules one at a time. In contrast, if the glycan is fully synthesized in the cytoplasm and transferred across the membrane on the UndP lipid carrier prior to transfer, this is *en bloc* glycosylation. Oligosaccharyltransferase (OSTase) enzymes are responsible for *en bloc* protein glycosylation in bacteria. Determination of all glycopolymer composition is notably challenging because glycosylation occurs through a non-templated process and often results in complex structure (38,39).

## **Cell Wall Anchored Proteins**

The surface proteins that decorate the *A. oris* cell are part of the larger cell envelope structure and are another major source of cell envelope variation (37,51). Proteins destined for secretion contain an N-terminal signal peptide and are addressed for cell wall anchoring via a C-terminal cell wall sorting signal. Cell wall anchored proteins mediate many of the specific functions that require direct interactions such as adherence.

### ***General protein translocation machinery***

The assembly of proteins on the exoplasmic side of the membrane requires the proteins to first be secreted beyond the cytoplasmic membrane. Secreted proteins are specifically addressed to translocation machines after being synthesized in the cytoplasm. These addresses, or signal peptides, specify the protein for either general secretion (Sec) or twin-arginine translocation (Tat) transport. Recognition of the signal peptide by the secretion machine components is the first step of the translocation event. Signal peptides are located at the N-terminus of a polypeptide and are short, unstructured primary amino acid sequences. In Gram-positive bacteria signal peptides are on average 30 amino acids in length (52). Signal peptides can be divided into three distinct regions, the positively charged n-region followed by a hydrophobic domain (h-region) and a polar cleavage (c-region) containing an AXA motif, but overall have little sequence similarity (53).

The subsequent translocation step is the most energy intensive step in the process. The Sec and Tat pathways are the two major secretion systems in bacteria. A major distinction between the secretion machines is the folding state required for successful substrate translocation. The Sec protein conducting channel is narrow allowing for only one or two polypeptide chains at a time. In contrast, the Tat system

secretetes folded proteins through a variably sized channel appropriate for individual substrates. These mechanism also differ in energetics. Although both require high energy expenditure, Sec consumes NTP to translocate proteins, while Tat utilizes the proton motive force (PMF) (54).

Proteins targeted to the Sec pathway proceed in an unfolded state. Because the substrates are unfolded, Sec translocation can occur co- or post-translationally. Co-translational targeting is mediated by a protein called signal recognition particle, and post-translational polypeptides are maintained in an unfolded state by cytoplasmic chaperones that interact with SecA at the secretion channel. The protein conducting channel of SecYEG is a heterotrimer, which interacts with the soluble ATPase SecA to power translocation (55).

The twin-arginine translocon signal peptide, as the name suggests, contains two consecutive arginine residues with the conserved consensus sequence being (S/T)RRxFLK (55). Gram-positive Tat machines have two core components, TatA and TatC, and can include the accessory components TatB and TatE depending on the organism (56,57). *A. oris* encodes *tatA* and *tatC* in a gene cluster and *tatB* elsewhere in the genome. TatA is a membrane bound monomer with a cytoplasmic C-terminal extension. TatC has six transmembrane helices with the N-terminus and C-terminus facing into the cytoplasm (56). TatB and TatE are TatA-like proteins, and TatB has been shown to be constitutively associated with TatC to form the TAT signal peptide recognition complex whereas TatE has been shown to be functionally redundant to TatA. Following signal peptide binding to the TatBC recognition complex, TatA moves laterally through the membrane to the site of translocation and oligomerizes in a PMF dependent manner to form a protein conducting channel matching the size of the substrate (56,58).

The terminal step of translocation regardless of whether proteins are secreted by Sec or Tat is removal of the signal peptide so that the protein can be freed from the

translocation machine to fulfill its subsequent function. Non-lipoprotein signal peptide cleavage is catalyzed by type I signal peptidases (SPase). SPases are serine endopeptidases that recognize an AXA motif present in the signal peptide c-region and cleave immediately after this motif leaving a mature, secreted protein (59,60). Integral membrane proteins that do not contain a cleavage site are laterally inserted from the secretion machine by YidC (40,61). A type I SPase from *A. oris* with novel substrate specificity is analyzed in chapter III.

### ***Oxidative protein folding***

Proteins secreted via the Sec translocon often contain cysteine residues that participate in a disulfide bond. Due to the lack of a traditional periplasm, disulfide bond (DSB) formation in Gram-positive bacteria was considered to be only a specialized phenomenon with a paired substrate-enzyme. However, recent *in vivo* analysis of a disulfide bond forming pathways revealed the physiological importance of DSB formation in the biogenesis of the actinobacterial cell envelope (62,63).

The major complex required for oxidative protein folding in *A. oris* consists of two proteins, the main oxidoreductase MdbA and its reoxidizing partner protein VKOR (62,64). As secreted proteins exit the translocation machine, MdbA catalyzes the formation of consecutive disulfide bonds. In the absence of MdbA or if participating cysteines are mutated so that they can no longer form a disulfide bond, many proteins including actinobacterial pili and toxins, lose stability and become degraded or remain nonfunctional (25,62,65). Additionally, in the absence of the reoxidizing factor VKOR, protein stability is compromised, although not to the extent of the MdbA deletion (64).

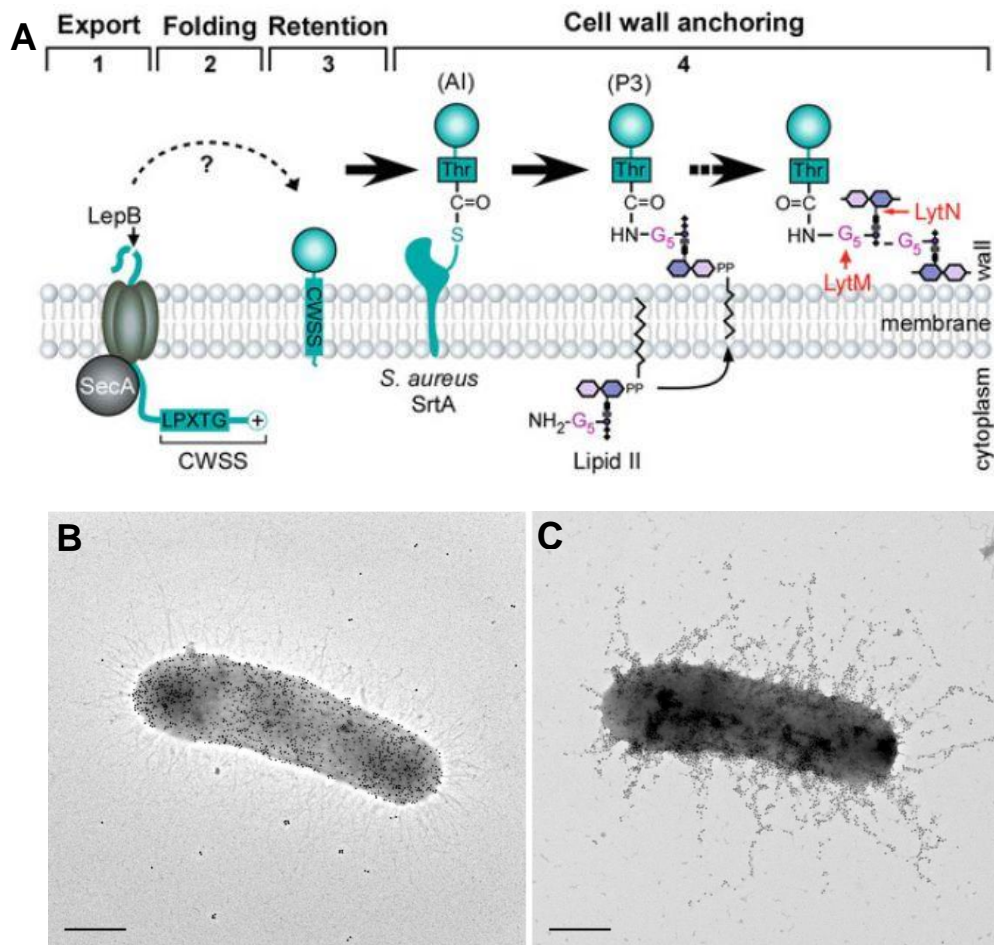
### ***Cell wall sorting pathway***

Proteins which are secreted and contain a cell wall sorting signal (CWS) are covalently attached to the cell wall by a transpeptidase enzyme called sortase. Thus far,



all known cell wall anchored proteins proceed through the Sec pathway (40). The CWS is a tripartite signal containing an LPXTG motif followed by a hydrophobic domain and a positively charged tail (51). After secretion, CWS containing proteins are retained in the membrane by the hydrophobic domain of the CWS. Within the CWS is a conserved motif of five residues referred to as the LPXTG motif (66). Sortase recognizes the CWS to anchor the protein to the cell wall (40). During the anchoring reaction, the LPXTG residues interact with the sortase catalytic pocket, which is necessary for the formation of a sortase-substrate acyl intermediate (67). This intermediate is then resolved by a nucleophilic attack of the free amine group from the lipid II precursor, then incorporated into the cell wall (Fig. 1-2A). Cell wall sorting allows for the attachment of both monomeric and polymeric proteins to the cell wall (Fig. 1-2B and C). Using immunogold labeling and electron microscopy (IEM) monomeric and polymeric cell wall proteins can be visualized and have a distinctive pattern. Here, monomeric cell wall anchored proteins are represented by GspA (Fig. 1-2B) and polymeric proteins are represented by the type 2 pilin backbone FimA (Fig. 1-2C).

All Gram-positive bacteria sequenced thus far encode at least one sortase (68). Based on primary amino acid sequence homology and substrate preference, six sortase classes have been proposed, known as class A – F (66). Sortases can have a broad substrate range as is the case of housekeeping sortases, or in the case of pilin-specific sortases recognize single or few substrates with high specificity. Housekeeping sortases from firmicutes exemplify the Class A sortases that recognize the canonical LPXTG motif and include SrtA from *S. aureus*. Actinobacterial housekeeping sortases fall into Class E and recognize an LAXTG motif (66). Class C sortases are those that are uniquely responsible for pilus polymerization (35,66). Class B sortases have a role in anchoring iron-regulated proteins, class D sortases are related to sporulation and class F sortases have an unknown function (66).



**Figure 1-2. Sortase anchors monomeric and polymeric substrates to the cell wall.**

(A) Schematic depicting cell wall anchoring pathway for the paradigm housekeeping sortase SrtA from *S. aureus*. Cell wall anchored proteins are first exported to the exoplasmic side of the membrane by their N-terminal signal peptide (1). Then, a signal peptidase cleaves off the signal peptide so that the protein can fold (2). The cell wall anchored protein precursor is retained in the membrane by the hydrophobic domain of the C-terminal cell wall sorting signal (CWSS) (3). Sortase recognizes the CWSS and cleaves between the threonine and glycine residues of the LPXTG motif forming an acyl intermediate with the substrate (AI). The sortase-substrate intermediate is resolved by the free amino group of the glycine crossbridge on the lipid II precursor (P3) and is incorporated into the cell wall via normal cell wall building processes (4). *Permission to reprint Figure 1 from Springer Nature Terms and Conditions for RightsLink Permissions*

*Springer Customer Service Centre GmbH : Springer Nature. Chapter "Anchoring of LPXTG-Like Proteins to the Gram-Positive Cell Wall Envelope" in Protein and Sugar Export and Assembly in Gram-positive Bacteria by Sara D. Siegel, Melissa E. Reardon and Hung Ton-That. © Springer International Publishing Switzerland 2016 (2016). (B) Monomeric cell wall anchored protein, GspA. (C) Polymeric cell wall anchored protein, FimA. For B and C, the proteins were reacted with a specific primary antibody and then treated with a secondary antibody conjugated to a 12-nm gold particle and viewed by electron microscopy. Scale bar represents 0.2 nm.*

### ***Sortase-Catalyzed Pilus Assembly***

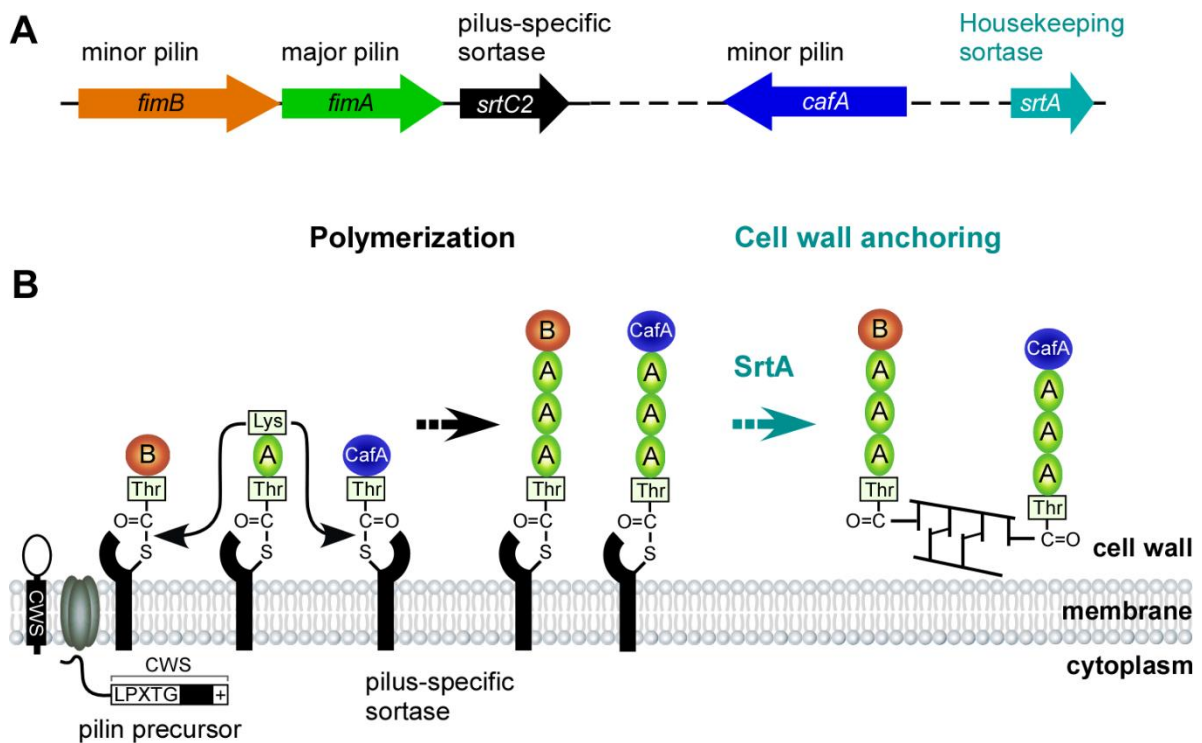
Utilizing Class C sortases, Gram-positive bacteria elaborate covalently linked polymeric proteins called pili or fimbriae (Fig. 1-2C). These structures are strongly associated with pathogenicity and play a direct role in bacterial adherence (35,69-74). Pilin subunits are covalently linked by pilin specific sortases and to the cell wall by the housekeeping sortase enzyme. Individual pilin subunits have an isopeptide bond in the Cna domain that contribute to overall stability of the subunits and polymer (25,75,76). The Cna domain superfamily was first defined in an *S. aureus* collagen binding protein, it forms a stalk-like fold that present the ligand binding domains in a variety of cell and pilus-associated adhesin proteins (77). By measuring the amount of mechanical force that could be applied to a pilus by atomic force microscopy, *A. oris* type 2 pili have been demonstrated to withstand forces of approximately 690 pN, owing largely to the isopeptide bond (76,78). These forces are the strongest reported for globular proteins and likely contribute to the ability to withstand large mechanical disturbances that exist in the *A. oris* niche (78).

*A. oris* pili are heterodimeric polymers made of a repeating major pilin subunits, FimP or FimA, and a single minor pilin tip protein. Pilin-specific sortases are typically genetically clustered with the pilin subunits they polymerize and have high specificity for their substrates. For example, *A. oris* encodes SrtC1 and SrtC2, which polymerize type 1 and type 2 pili, respectively. Type 1 pili are encoded by the *fimQ-fimP-srtC1* cluster (18). Type 2 pili are encoded by *fimB-fimA-srtC2* (29,79). An interesting exception is the alternative tip protein CafA, which is encoded outside of the type 2 gene cluster (Fig. 1-3A), but is polymerized by SrtC2 with the major pilin subunit FimA (Fig. 1-3B) (25). CafA is a pilus-associated adhesin that plays a major role the ability of *A. oris* to colonize the oral biofilm. CafA mediates both coaggregation with oral streptococci, and adherence to

host tissues. Studies of the molecular requirements for CafA binding are described in chapter V.

Cognate pilin-specific sortase recognizes the LPXTG motif in the membrane retained pilin precursors and cleaves between the threonine (T) and glycine (G) residues to form a sortase-pilin acyl intermediate. Polymerization occurs when the intermediate is resolved by a nucleophilic attack from a lysine (K) group of the major pilin subunit (35). Pili are anchored to the cell wall by a free amine group from the lipid II cell wall precursor (Fig. 1-3B) (40).

In *A. oris*, SrtC1 and SrtC2 require a tip protein to initiate polymerization. Thus, substrate entry and polymerization steps seem to be regulated (18,25). However, the regulation and molecular requirements for initiation and specificity are not fully understood. A feature that is conserved among the pilus tip proteins is the presence of a putative Tat signal peptide. The possible involvement of the Tat translocation system in CafA export is explored in chapter VI.



**Figure 1-3: Type 2 pili are polymerized by the pilus-specific sortase SrtC2.** (A) The type 2 pilus gene cluster encoding *fimB-fimA-srtC2*. *cafA* and *srtA* are positioned outside of the gene cluster, but are also involved in assembly of type 2 pili. (B) Graphic representation of the type 2 pilus polymerization pathway in *A. oris*. Pilin precursors are secreted via a signal peptide then retained in the membrane by the cell wall sorting signal (CWS). The type 2 pilus-specific sortase SrtC2 forms an acyl intermediate with the pilin precursors, as described above. A minor tip pilin, either FimB or CafA, initiates polymerization of the major pilin subunit FimA. Once a minor-major pilin heterodimer is formed the major pilin subunit is polymerized successively until the housekeeping sortase SrtA catalyzes cell wall anchoring.

## Significance of these Studies

Oral biofilm-associated diseases are significant global oral health burdens, however, there is no effective therapy to control oral biofilms (13). *A. oris* plays a critical role in development of the oral biofilm as a primary colonizer enriched at the base of the biofilm and facilitates the adherence of species that cannot bind otherwise (8,9). The studies in this thesis investigate protein and glycan assembly and display on the *A. oris* cell envelope and their roles in oral biofilm formation.

Proteins displayed on the surface must be translocated before surface attachment. The final step of translocation is cleavage of the signal peptide as the protein emerges from the secretion machinery. I have shown that a type SPase LepB2 is required for pilus assembly, but dispensable for other cell wall anchored proteins. The deletion of LepB2 also decreases biofilm formation and coaggregation, which are physiologically relevant phenotypes (80). Importantly, the second signal peptidase encoded by *A. oris* LepB1 is not involved in these processes. While the presence of multiple signal peptidases is not uncommon in Gram-positive bacteria, the specific role and target specificity of these homologous enzymes is often unknown (55). This system provides a new model containing only two non-redundant SPase I enzymes with known substrate specificity.

After translocation, proteins on the surface can be modified by *en bloc* glycosylation. The cell wall anchored protein GspA is glycosylated in an LcpA-dependent manner (50). This represents an entirely new mode of protein glycosylation (81), because LCP enzymes have been shown to catalyze the transfer of glycopolymers to the cell wall, but had not been previously demonstrated to modify protein substrates. I have characterized the residues and activities of LcpA that are required for this novel functionality. My work revealed that the *A. oris* LcpA has phosphotransferase activity similar to other LCP enzymes, and our results are applicable for a broad range of LCP

enzymes. This system can be used to address how LCP enzymes select acceptor substrates.

Once proteins on the cell surface are fully translocated, modified, and anchored, they can contribute to adherence, a critical factor contributing to bacterial pathogenesis. I have investigated how a pilus-associated adhesin, CafA, contributes to oral biofilm formation in two distinct ways by adhering to both oral streptococci and human gingival cells. This study also addressed the prevalence and conservation of CafA among different *A. oris* strains. These studies can be extended to other bacterial adhesins, and enhance our overall understanding of adhesins in formation of multispecies biofilms.

I have also examined the role of a putative twin arginine signal peptide in CafA secretion. This work has shown that there are likely additional factors contributing to the secretion of proteins through either Sec or Tat. Bioinformatics predicted the presence of a Tat signal peptide in CafA. However, our mutagenesis studies suggested that the twin arginine residues were required for stability, but could not mediate secretion of two Tat-dependent reporter proteins.

Overall, my studies have explored how cell surface proteins are successfully translocated, glycosylated, and contribute to biofilm formation in the oral microbe *Actinomyces oris*. My work has established new possibilities for investigating signal peptidase specificity, LCP-mediated protein glycosylation, protein adherence function, and tools for the Tat pathway.



**CHAPTER II**  
**Materials and Methods**

*This chapter is based, in part, upon work published in 2016: Siegel SD, Wu C, Ton-That H "A type I signal peptidase is required for pilus assembly in the Gram-positive biofilm-forming bacterium Actinomyces oris." J Bact. 198(15):2064-73. doi: 10.1128/JB.00353-16. I am the first author of this publication and was responsible for preparing the original manuscript and conducting the experiments described in this paper. I have permission to reproduce any and all of this manuscript, in print or electronically, for the purpose of my thesis in accordance with the American Society for Microbiology (publisher of Journal of Bacteriology) "Journals Statements of Authors' Rights."  
[http://journals.asm.org/site/misc/ASM\\_Author\\_Statement.xhtml](http://journals.asm.org/site/misc/ASM_Author_Statement.xhtml).*

**Bacterial strains and media.** *Actinomyces* strains were grown in heart infusion broth (HIB) and on heart infusion agar (HIA) supplemented with 0.5% glucose at 37°C with 5% CO<sub>2</sub>. *Streptococcus* strains were grown in HIB and on HIA supplemented with 1.0% glucose at 37°C in anaerobic conditions. *E. coli* strains were grown in Luria broth (LB) or Luria agar (LA) at 37°C in aerobic conditions. When appropriate, 50 µg mL<sup>-1</sup> kanamycin (Kan) or streptomycin (Sm) or 100 ug mL<sup>-1</sup> of ampicillin (Amp) was added to the media. *E. coli* DH5α was used for cloning experiments. Restriction enzymes, T4 polynucleotide kinase, LIC-qualified T4 DNA polymerase, T4 DNA ligase and Phusion DNA polymerase were purchased from New England Biolabs. A list of strains can be found in Table 2-1.

**Human cell culture.** Human gingival fibroblast (HGF-1) cells (ATCC® CRL-2014™) were grown in Dulbecco's Modified Eagle's Medium (DMEM) with 10% heat-inactivated fetal bovine serum and 1% penicillin/streptomycin at 37°C with 5% CO<sub>2</sub>, with medium renewal every three days. To subculture, the medium was removed and treated with 1-mL 0.25% trypsin, 0.53 mM EDTA (TE) at 37°C for 10 minutes, until the cells detached from the flask. Cells in suspension were pelleted by centrifugation, and the supernatant

containing TE was aspirated. The cells were sub-cultured at a 1:4 ratio and new media was added.

***E. coli* competent cells and heat shock transformation.** *E. coli* cells were subcultured from overnight cultures in LB and grown until  $OD_{A600} = 0.6$ . The cells were chilled for 10 minutes (min), then centrifuged in a pre-chilled 4°C rotor at 6,000 rpm for 5 min. The supernatant was discarded and the pellet was suspended in 20 mL TfbI (30 mM  $KCH_3CO_2$ , 100 mM RbCl, 10 mM  $CaCl_2$ , 50 mM MnCl, pH 5.8) then chilled on ice for 5 mins before pelleting by centrifugation. The cell pellets were then suspended in 2 mL TfbII (10 mM PIPES, 10 mM RbCl, 75 mM CaCl, 15% v/v glycerol, pH 6.5) and chilled on ice for 15 mins. Then, 100  $\mu$ L cells were then aliquoted into sterile, pre-chilled 1.5 mL tubes and snap frozen in a dry ice ethanol bath. The aliquots were stored at -80°C. To transform the competent cells, an aliquot of competent cells was thawed on ice for 10 mins. Then, 100 ng of plasmid was added to the cells for 1 hour followed by heat shock at 42°C for 40 seconds. Cells were recovered in LB for 20 mins at 37°C with shaking, then plated on LA with Amp.

***A. oris* competent cells and electroporation.** Overnight *A. oris* cells were subcultured into 25 mL HIB, and grown at 37°C to an  $OD_{A600} = 0.6$ . 10% glycine in HIB was added and the cells were incubated for an additional hour. Cells were chilled on ice for 20 mins then centrifuged at 6,000 rpm for 10 min at 4°C. The cells were washed twice with cold 15% glycerol and finally suspended in 1.5 mL 15% glycerol before aliquoting 100  $\mu$ L into 1.5 mL tubes, snap frozen (as above) and stored at -80°C. To transform the *A. oris* cells, plasmid was added to a thawed competent cell aliquot, swirled to mix and transferred to a 2 mm gap electroporation cuvette (Sigma). The cells were electroporated at 2,500 V with 25  $\mu$ F capacitance and 400  $\Omega$  of resistance (82). HIB was added to the cuvette immediately following electroporation and recovered for 2 hours at 37°C before plating onto HIA containing either Kan or Sm.

**Table 2-1.** Bacterial strains and cell lines used in this studies

Strain	Description	Reference
<i>E. coli</i> DH5 $\alpha$	Plasmid propagation and cloning strain	(83)
<i>A. oris</i> MG1	Parental strain	(84)
<i>A. oris</i> CW1	$\Delta galk$ ; an isogenic derivative of MG-1	(79)
<i>S. oralis</i> So34	RPS positive	(85)
<i>S. oralis</i> OC1	RPS-negative isogenic mutant of So34	(85)
<i>A. oris</i> AR4	$\Delta fimA$ ; an isogenic derivative of CW1	(79)
<i>A. oris</i> WU36	Conditional <i>srtA</i> deletion mutant ( $\Delta srtA$ ), containing pTetR- $\Omega$ -SrtA	(50)
<i>A. oris</i> WU51	$\Delta gspA$ ; an isogenic derivative of CW1	(50)
<i>A. oris</i> WU73	$\Delta gspA/\Delta srtA$	(50)
<i>A. oris</i> WU42	$\Delta lepB2$ ; an isogenic derivative of CW1	(80)
<i>A. oris</i> WU50	$\Delta lepB1$ ; an isogenic derivative of CW1	(80)
<i>A. oris</i> WU80	$\Delta lepB1/2$ ; an isogenic derivative of CW1	(80)
<i>A. oris</i> WU81	$\Delta lepB2/\Delta srtA$	(80)
<i>A. oris</i> WU42c1	WU42 containing pLepB2	(80)
<i>A. oris</i> WU42c2	WU42 containing pLepB2(S101A)	(80)
<i>A. oris</i> WU42c3	WU42 containing pLepB2(K169A)	(80)
<i>A. oris</i> WU72	$\Delta lcpA$ ; an isogenic derivative of CW1	(50)
<i>A. oris</i> D0299	$\Delta lcpB$ ; an isogenic derivative of CW1	This study
<i>A. oris</i> D1578	$\Delta lcpD$ ; an isogenic derivative of CW1	This study
<i>A. oris</i> <i>lcp</i> $\Delta$ 3	$\Delta lcpA\Delta lcpB\Delta lcpD$	This study
<i>A. oris</i> MR108	$\Delta vkor$ ; an isogenic derivative of CW1	(62)
<i>E. coli</i> C3209/ Shuffle $\text{\textcircled{R}}$	Derived from <i>E. coli</i> K12 parental strains, engineered to form disulfide bonds in the cytoplasm; expresses T7 RNAP.	(86)
<i>A. oris</i> AR5	$\Delta cafA$ ; an isogenic derivative of CW1	(25)
<i>A. oris</i> <i>tetR</i> - $\Omega$ - <i>tatC</i>	Conditional <i>srtA</i> deletion mutant, inducible by addition of tetracycline and theophylline	This study
<i>A. naeslundii</i> N28B15	Clinical isolate	J. Cisar collection
<i>A. naeslundii</i> N34A24	Clinical isolate	J. Cisar collection
<i>A. naeslundii</i> N35B3	Clinical isolate	J. Cisar collection
<i>A. oris</i> N11A16	Clinical isolate	J. Cisar collection
<i>A. oris</i> N12A2B	Clinical isolate	J. Cisar collection
<i>A. oris</i> ATCC $\text{\textcircled{R}}$ 49339 $\text{\textsuperscript{TM}}$	Human abscess isolate	J. Cisar collection
<i>A. oris</i> N11A12	Clinical isolate	J. Cisar collection
<i>A. oris</i> ATCC $\text{\textcircled{R}}$ 27044 $\text{\textsuperscript{TM}}$	Human sputum isolate	J. Cisar collection

<i>A. oris</i> N33A2B	Clinical isolate	J. Cisar collection
<i>A. oris</i> N37B13	Clinical isolate	J. Cisar collection
<i>A. oris</i> N34A23	Clinical isolate	J. Cisar collection
<i>A. oris</i> N28B1	Clinical isolate	J. Cisar collection
<i>A. oris</i> N29A27	Clinical isolate	J. Cisar collection
<i>A. oris</i> N32A8	Clinical isolate	J. Cisar collection
<i>A. oris</i> N37B9	Clinical isolate	J. Cisar collection
Actinomyces n/s <sup>a</sup> N38B10	Clinical isolate	J. Cisar collection
Actinomyces n/s N33A3	Clinical isolate	J. Cisar collection
Actinomyces n/s N34A14	Clinical isolate	J. Cisar collection
<i>A. johnsonii</i> PK1259	Clinical isolate	J. Cisar collection
<i>A. johnsonii</i> ATCC® 49338™	Human gingival crevice isolate	J. Cisar collection
<i>Streptomyces</i> <i>violaceoruber</i> M145	Prototrophic derivative of <i>S. coelicolor</i> A3(2)	ATCC® BAA-471™
HGF-1	Human gingival fibroblast cell line (ATCC® CRL- 2014™)	ATCC

---

<sup>a</sup> non-serotypable

## Plasmid construction

A list of plasmids can be found in Table 2-2 and list of primers in Table 2-3.

pCWU10 – To generate a RSF1010 derivative that is functional in *A. oris*, the ampicillin resistance gene of pCVD047, a cyanobacterial broad-host-range vector (87), was replaced by the kanamycin (kan) cassette from pJRD215 (88). Briefly, using the primer pair pCVD047-noAmp-F/R for an inverse PCR reaction, an amplicon containing pCVD047 sequence without the ampicillin resistance gene was generated. Next, the fragment encompassing the kan resistance gene and a multiple cloning site from pJRD215 was amplified with the primer pair 215Kan<sub>MCS</sub>-F/R. Both amplicons were digested with *SacI* and *HindIII*. The isolated products were ligated to generate pCWU10.

pGspA<sub>Δcws</sub> – Using pAcaC<sub>Δcws</sub> (50) as a template, primers com-AcaC-F and GspA<sub>Δcws</sub>-His6-R were used to PCR-amplify a fragment encompassing the promoter region of *gspA* and its open reading frame (ORF) lacking the cell wall sorting signal (CWS) while appending a hexa-histidine tag to the GspA C-terminus. The *gspA* amplicon was digested with *NdeI* and *EcoRI* and ligated into pCWU10 digested with the same enzymes to generate pGspA<sub>Δcws</sub>.

pFimA<sub>Δcws</sub> – The primer sets prFimB-BamHI-F and prFimB-R and FimA-F and FimA<sub>Δcws</sub>-His6-EcoRI-R were used in PCR reactions with MG1 genomic DNA as a template to amplify the promoter region of *fimB* and the *fimA* ORF lacking the CWSS while appending a hexa-histidine tag to the FimA C-terminus, respectively. The *fimB* promoter product was digested with *BamHI* whereas the *fimA* coding region was digested with *EcoRI* and then treated with T4 polynucleotide kinase to phosphorylate the 5' end. The digested DNA fragments were ligated into pCWU10 digested with *BamHI* and *EcoRI* to generate pFimA<sub>Δcws</sub>.

pCWu2- $\Delta$ *lepB1* – 1 kb upstream and downstream the *lepB1* (*ana\_1188*) open reading frame was amplified using primer pairs 1188up-F and 1188up-R and 1188dn-F and 1188dn-R. 1188up fragment was engineered with a 5' EcoRI and 3' XbaI restriction site and the 1188dn fragment was engineered with 5' XbaI and 3' KpnI restriction sites to facilitate ligation of the fragments to the pCWu2 vector digested with EcoRI and KpnI, in the presence of T4 DNA ligase.

pCWu2- $\Delta$ *lepB2* – 1 kb upstream and downstream the *lepB2* (*ana\_1190*) open reading frame was amplified using primer pairs 1190up-F and 1190up-R and 1190dn-F and 1190dn-R. 1190up fragment was engineered with a 5' XbaI and 3' HindIII restriction site, and the 1190dn fragment was engineered with 5' HindIII and 3' EcoRI restriction sites to facilitate ligation of the fragments to the pCWu2 vector digested with XbaI and EcoRI, in the presence of T4 DNA ligase.

pLepB2 – The *lepB2* coding sequence and *rimM* promoter region were amplified from pLepB2 with the primers Pro-1192-F and com-*lepB2*-R. The PCR product was digested by EcoRI and HindIII and ligated into pJRD215 (84) digested with the same enzymes.

prGspA – *gspA* (*ana\_1291*) sequence containing only residues 31 and 256 (lacking signal peptide and cell wall sorting signal) was amplified from the *A. oris* MG1 genomic DNA template with the primers rGspA-F and rGspA-R, containing adapter sequences for subsequent ligation independent cloning (LIC) into pMCSG7. The resultant amplicons were treated with LIC-qualified T4 DNA polymerase and dCTP. In parallel, the SspI digested vector was treated with T4 DNA polymerase and dGTP. The insert and vector fragments were ligated in a step-down annealing reaction. The vector was transformed into chemically competent *E. coli* DH5 $\alpha$ .

pLcpA – LcpA (*ana\_1292*) complement strain was cloned by amplifying the native promoter, 152 bp upstream of the start codon until the native stop codon using primers pLCP-F and pLCP-R engineered with a BamHI and EcoRI cut site respectively and using *A. oris* MG1 genomic DNA as a template. The resultant amplicon was gel purified. Simultaneous restriction digest with BamHI-HF and EcoRI-HF of the amplicon and parental plasmid (pCWu10) was performed, and the products were again gel purified. The restricted products were ligated with T4 DNA ligase at 16°C for 18 h. The ligation mixture was transformed into chemically competent *E. coli* DH5 $\alpha$ , and colony PCR was used to identify plasmids with the correct insert. The plasmid was introduced to *A. oris*  $\Delta$ *lcpA* cells via electroporation (88).

pCWu2- $\Delta$ *lcpB* – 1 kb upstream and downstream the *lcpB* (*ana\_0299*) open reading frame was amplified using primer pairs 0299up-F and 0299up-R and 0299dn-F and 0299dn-R. 0299up fragment was engineered with a 5' KpnI and 3' XbaI restriction site, and the 0299dn fragment was engineered with 5' XbaI and 3' EcoRI restriction site to facilitate ligation of the fragments to the pCWu2 vector digested with KpnI and EcoRI in the presence of T4 DNA ligase.

pCWu2- $\Delta$ *lcpD* – 1 kb upstream and downstream the *lcpD* (*ana\_1578*) open reading frame was amplified using primer pairs 1578up-F and 1578up-R and 1578dn-F and 1578dn-R. 1578up fragment was engineered with a 5' KpnI and 3' XbaI restriction site, and the 1578dn fragment was engineered with a 5' XbaI and 3' EcoRI restriction site to facilitate ligation of the fragments to the pCWu2 vector digested with KpnI and EcoRI in the presence of T4 DNA ligase.

rLcpA – *lcpA* (*ana\_1292*) sequence containing residues 78 and 370 (lacking the transmembrane region) was amplified from the *A. oris* MG1 genomic DNA template with



the primers rLcpA-F and LCP<sup>Δ</sup>™-R containing adapter sequences for subsequent ligation independent cloning (LIC) into pMCSG7 (89).

pCafA – The *A. oris* MG1 CafA (ana\_2235) complement strain was cloned by amplifying the native promoter, 407 bp upstream of the start codon until the native stop codon, using primers CafA-pro-F and CafA\_comp-R engineered with a KpnI and BamHI cut site respectively and using *A. oris* MG1 genomic DNA (gDNA) as a template. The resultant amplicon was gel purified. The amplicon and parental plasmid (pCWu10) were simultaneously digested with the restriction enzymes KpnI-HF and BamHI-HF, and the products were purified. The restricted products were ligated with T4 DNA ligase at 16°C for 18 hours.

pCafA<sub>ΔCWS</sub>-H6 – The CafA (ana\_2235) native purification strain was cloned by amplifying the native promoter, 407 bp upstream of the start codon until the CWS, using primers CafA\_pro-F and CafA<sub>ΔCWS</sub>-H6-R engineered with a KpnI and hexa-histidine-BamHI cut site respectively and using *A. oris* MG1 pCafA as a template. The resultant amplicon was gel purified. The amplicon and parental plasmid (pCWu10) were simultaneously digested with the restriction enzymes KpnI-HF and BamHI-HF (NEB), and the products were purified. The restricted products were ligated with T4 DNA ligase at 16°C for 18 hours.

pGFP – Green fluorescent protein (GFP) was amplified from pBsk-GFP using primers GFP-F and GFP-R. To drive expression, the EF-Tu (ana\_0022) promoter was cloned upstream of GFP using the primers prom-EF-tu-F and prom-EF-tu-R. To facilitate ligation, prom-EF-tu-F was designed with a KpnI restriction site and GFP-R was designed with a BamHI restriction site. The GFP fragment was treated with T4 polynucleotide kinase and ATP for intramolecular ligation. The two amplified restricted

fragments were combined with pCWu10, treated with KpnI and BamHI in the presence of T4 DNA ligase.

pDagA<sub>SV</sub> – The agarase gene (*dagA*) was amplified from *Streptomyces violaceoruber* M145 gDNA using primers dagA<sub>SV</sub>-F and dagA<sub>SV</sub>-R. To drive expression, the EF-Tu (*ana\_0022*) promoter was cloned upstream of *dagA* using the primers prom-EF-tu-F and prom-EF-tu-R. To facilitate ligation, prom-EF-tu-F was designed with a KpnI restriction site and GFP-R was designed with a BamHI restriction site. The GFP fragment was treated with T4 polynucleotide kinase and ATP to facilitate intramolecular ligation. The two amplified restricted fragments were combined with pCWu10 treated with KpnI and BamHI in the presence of T4 DNA ligase.

DagA and GFP reporter plasmids – Inverse PCR was used to insert signal peptides between the EF-tu promoter and *dagA* or *gfp* sequence, the signal peptide fragment contained a 5' phosphate and 3' restriction site that was complementary to the restriction site on the reporter gene.

pCWu2- $\Delta$ *tatC* – 1 kb upstream and downstream the *tatC* (*ana\_0769*) open reading frame was amplified using primer pairs tatCup-F and tatCup-R and tatCdn-F and tatCdn-R. tatCup fragment was engineered with a 5' KpnI and 3' HindIII restriction site, and the tatCdn fragment was engineered with a 5' HindIII and 3' EcoRI restriction site to facilitate ligation of the fragments as well as the pCWu2 vector digested with KpnI and EcoRI in the presence of T4 DNA ligase.

ptetR- $\Omega$ -TatC – The tetR- $\Omega$  inducible promoter was amplified from the ptetR- $\Omega$ -*srtA* plasmid using tetR- $\Omega$ -F engineered with a KpnI restriction site and tetR- $\Omega$ -R primers. The *tatC* open reading frame was amplified using primers tatC-F and tatC-R, and the reverse primer was designed with an EcoRI restriction site. The digested tatC orf

fragment was treated with T4 polynucleotide kinase and ATP for subsequent intramolecular ligation. The vector pCWu9 was digested with KpnI and EcoRI and ligated with the tetR- $\Omega$  and tatC fragments.

**Table 2-2. Plasmids used in these studies**

Plasmid	Description	Reference
pJRD215	<i>Actinomyces/E. coli</i> shuttle vector, Km <sup>R</sup> , Sm <sup>R</sup>	(82)
pRMC2	a tetracycline-inducible expression vector	(90)
pCVD047	Broad host range vector	(87)
pCWU10	<i>Actinomyces/E.coli</i> shuttle vector, Km <sup>R</sup> . A derivative of pCVD, a generous gift from Dr. Golden	(80)
pTetR-Ω-SrtA	A derivative of pTetR-SrtA with a riboswitch element incorporated upstream of the start codon ATG	(50)
pGspA	pJRD215 expressing GspA	(50)
pGspA <sub>Δcws</sub>	pJRD215 expressing GspA lacking the cell wall sorting signal (CWS)	(80)
pFimA <sub>Δcws</sub>	pCWU10 expressing <i>fimA</i> under the <i>srtC2</i> promoter, lacking the cell wall sorting signal (CWS)	(80)
pLepB2	pJRD215 expressing <i>lepB2</i>	(80)
pLepB2(S101A)	pJRD215 expressing <i>lepB2</i> with the mutation S101A	(80)
pLepB2(K169A)	pJRD215 expressing <i>lepB2</i> with the mutation K169A	(80)
pCWU2	Derivative of pHTT177, expressing GalK under the control of the <i>rpsJ</i> promoter	(79)
pCWU2-Δ <i>lepB1</i>	An allelic replacement vector of <i>lepB1</i> using pCWU2	(80)
pCWU2-Δ <i>lepB2</i>	An allelic replacement vector of <i>lepB2</i> using pCWU2	(80)
pCafA <sub>Δcws-H6</sub>	pCWU10 expressing CafA lacking the cell wall sorting signal (CWS) containing a 6xHis tag	This study
pMCSG7	Expression vector for protein purification	(89)
pKO-0299	pCWU2 derived KO plasmid Δ <i>lcpB</i>	This study
pKO-1578	pCWU2 derived KO plasmid Δ <i>lcpD</i>	This study
rLcpA	pMCSG7 expressing LcpA(78-370)	This study
rLcpA(R149A)	rLcpA with the mutation R149A	This study
rLcpA(Δs-s)	rLcpA with the mutation C179A/C365A	This study
pLcpA	pCWu10 expressing LCP from native promoter	This study
pLcpA(R128A)	pLcpA with the mutation R128A	This study
pLcpA(R149A)	pLcpA with the mutation R149A	This study
pLcpA(R266A)	pLcpA with the mutation R266A	This study
pLcpA(C365A)	pLcpA with the mutation C365A	This study
pLcpA(C179/365A)	pLcpA with the mutation C179/365A	This study
pVKOR-2HA	pCWU10 expressing VKOR with double C-terminal influenza hemagglutinin (HA) tag	(64)
pGspA <sub>Δcws-H6</sub>	pCWU10 expressing GspA lacking the cell wall sorting signal (CWS) with His <sub>6</sub> tag	(80)
rGspA	pMCSG7 expressing GspA(31-286)	This study

pCafA	pCWu10 expressing <i>cafA</i> from native promoter	This study
pCafA(K111M)	pCafA with K111M mutation	This study
pCafA(R123H)	pCafA with R123H mutation	This study
pCafA(Y145R)	pCafA with Y145R mutation	This study
pCafA <sub>Δcws</sub> -H6	pCWu10 expressing CafA variant lacking the cell wall sorting signal and containing a Hisx6 tag from CafA promoter	This study
pCafA-RR1	pCafA <sub>Δcws</sub> -H6 with CafA(R3,4A) mutation	This study
pCafA-RR2	pCafA <sub>Δcws</sub> -H6 with CafA(R11,12A) mutation	This study
pCafA-RR3	pCafA <sub>Δcws</sub> -H6 with CafA(R25,26A) mutation	This study
pBsk-GFP	pBluescript SK expressing GFP	Addgene plasmid #29459 (91)
pGFP	pCWu10 expressing GFP from the EF-Tu promoter	This study
pDagA <sub>Sv</sub>	pCWu10 expressing the DagA ORF from <i>S. violaceoruber</i> from the EF-Tu promoter	This study
pFimA <sup>SP</sup> -DagA	pCWu10 expressing FimA-SP fused to DagA without signal peptide from the EF-Tu promoter	This study
pFimA <sup>SP</sup> -GFP	pCWu10 expressing FimA-SP fused to GFP from the EF-Tu promoter	This study
pDagA <sup>SP</sup> -GFP	pCWu10 expressing DagA-SP fused to GFP from the EF-Tu promoter	This study
pCafA <sup>SP</sup> -DagA	pCWu10 expressing CafA-SP fused to DagA without signal peptide from the EF-Tu promoter	This study
pCafA <sup>SP</sup> -GFP	pCWu10 expressing CafA-SP fused to GFP from the EF-Tu promoter	This study
pTatC-KO	pCWu2 expressing ~1.8kb fragments upstream and downstream of TatC ORFs for allelic exchange and counter-selection of mutants	This study
ptetR-Ω-tatC	pCWu9 expressing <i>tatC</i> ORFs under transcriptional control from a tetR inducible promoter and post-transcriptional control by a theophylline responsive riboswitch	This study

---

**Site-directed mutagenesis.** For site-directed mutagenesis of LepB2, pLepB2, rLcpA, pLcpA and pCafA<sub>Δcws</sub>-H6 mutations were incorporated into the 5' end of synthesized primers (Table 2-3) for inverse PCR as previously described (50). The PCR products were purified by gel extraction and phosphorylated to facilitate circularization of the amplicons. After transformation into *E. coli*, the genes were sequenced to ensure the mutation was present and in-frame.

**Generation of nonpolar, in-frame deletion mutants in *A. oris*.** Nonpolar, in-frame deletion mutants used in this study were generated according to a previously published protocol (79). Briefly, 1.0 kb DNA fragments upstream and downstream of a target gene were PCR-amplified using appropriate primers. The two fragments were treated with restriction enzymes and linked together by a single-step ligation. Subsequently, the 2 kb fragment was cloned into the deletion vector pCWU2 (50). The generated plasmids were introduced into the *A. oris*  $\Delta galK$  strain by electroporation. Selection of corresponding in-frame deletion mutants was performed using GalK as a counter-selectable marker. Generated mutants were characterized by PCR and/or immunoblotting.

**Table 2-3.** Primers used in this study

Primer	Sequence <sup>(a)</sup>	Application
1190upF	GGCGTCTAGATCCGGACAAACCGTTCCATG CCCC GA	pCWU2- $\Delta$ lepB2
1190upR	GGCGAAGCTTTGCGCTGCTCATAGGCCTTC TCCTG	pCWU2- $\Delta$ lepB2
1190dnF	GGCGAAGCTTGACTGATCGCCCCGAAAGCG TGCTG	pCWU2- $\Delta$ lepB2
1190dnR	GGCGGAATTCACCGACCTCGTCCAGGCCGC CGACGT	pCWU2- $\Delta$ lepB2
Pro-1190F	GGCGGAATTC <del>C</del> CCCTCGGCCGAGTCATCGGCC GCTCG	pLepB2
pro-1190R	GGCGGGTACCCTCCTGGGATCGGGGCATGGAACGG	pLepB2
com-1190F	GGCGGGTACCCTGATGAGCAGCGCACCCGACCA GAGC	pLepB2
com-1190R	GGCGAAGCTTTCAGTGGTGGTGGTGGTGGT GGTCCCCG GAGCCCGCCA GCCTG	pLepB2
1188upF	GGCGGAATTCACTGATCGCCCCGAAAGCGT GCTGG	pCWU2- $\Delta$ lepB1
1188upR	GGCGTCTAGAACGGGGAAGTGCAGGCCGGT GTG	pCWU2- $\Delta$ lepB1
1188dnF	GGCGTCTAGAGGCCCGGACACTTACGGCGG CATGG	pCWU2- $\Delta$ lepB1
1188dnR	GGCGGGTACCCTAGCCCCATGACGCATCCA CCG	pCWU2- $\Delta$ lepB1
1190S101A-F	GCGATGGAGAACACCCTCAACGAGGGCG	pLepB2(S101A)
1190S101A-R	CCCCGAGATGGTGAAGCTGCTCTGGATG	pLepB2(S101A)
1190K169A-F	GCGCGGGTCATCGGAATGCCCGGTGACCACG	pLepB2(K169A)
1190K169A-R	GATGAGGTGGTGACCGCGTTCTGGGG	pLepB2(K169A)
pCVD047- noAmp-F	GGCGGAGTCATGATTTAGAAAAATAAACAATAGGG G	pCVD047 $\Delta$ Amp
pCVD047- noAmp-R	GGCGAAGCTTCTGTCAGACCAAGTTACTCATATATA	pCVD047 $\Delta$ Amp
215Kan <sub>MCS</sub> -F	GGCGGAGCTCTCAGAAGAAGTTCGTCAAGAA GGCGA	pCWU10
215Kan <sub>MCS</sub> -R	GGCGAAGCTTATCG ATGATAAGCTGTCAA	pCWU10
GspA $\Delta$ cws-His6- R	GGCGGAATTCTCAGTGGTGGTGGTGGTGGTGGGGC TTGCCGGAGGTGGAGGCCGCC	pGspA $\Delta$ cwss
prFimB-BamHI-F	AAAAAGGATCCGACGTCACCGGTGTCATCACCTCC	pFimA $\Delta$ cwss
prFimB-R	GGGACCGCCTTCTCTTAGGCGTCCG	pFimA $\Delta$ cwss
FimA-F	GTGACGCCGTCGGACAAGACGGAG	pFimA $\Delta$ cwss

FimA <sub>Δcws</sub> -His6- EcoRI-R	AAAAA <u>GAA</u> TTCTCAGTGGTGGTGGTGGTGGTGAAC CGA CTG CTT GGT GTT CTC AAC GG	pFimA <sub>Δcws</sub>
RT-16s-R	GGTGTTGCCGACTTTTCATG	RT- PCR (16s rRNA)
RT-16s-F	GTCGCTAGTAATCGCAGATCAG	RT- PCR (16s rRNA)
RT-1291F	GACGGCACCTACAAGATCAC	RT-PCR ( <i>gspA</i> )
RT-1291R	AGGAGTCGGTCTTGCTGA	RT-PCR ( <i>gspA</i> )
P1	CCTCCAGGTCC CGATCACAC	RT and qRT-PCR
P2	GCCTGCGGGGTTGGATAGAGG	RT and qRT-PCR
P3	CCCGGCCGGTCAGCCTCCGAGGTCCG	RT and qRT-PCR
P4	GCTGGACCAGTGGGAGTATGGCCAC	RT and qRT-PCR
P5	GTGTGATCGCGGACCTGGAGG	RT- PCR
P6	GGCCATACTCCCACTGGTCCAG	RT- PCR
D0299up-F	GGCGGGTACC <u>GTG</u> ACGAGCAGCGCCGCTGCGCT	<i>ΔlcpB</i>
D0299up-R	GGCGT <u>CTAG</u> AGCGAGGTCGTGTTCGGCCCCTGACGA G	<i>ΔlcpB</i>
D0299dn-F	GGCGT <u>CTAG</u> AGTCACGCTCGACGCCGACGCGGACA C	<i>ΔlcpB</i>
D0299dn-R	GGCGGAAT <u>TCAG</u> CTCCTCAACCGCCTCGGGCAC	<i>ΔlcpB</i>
D1578up-F	GGCGGGTACC <u>TCT</u> CCTACGTCCTGGAGAAGACGA	<i>ΔlcpD</i>
D1578up-R	GGCGT <u>CTAG</u> AGCGATCATAGGGAACGAGACTGCTA	<i>ΔlcpD</i>
D1578dn-F	GGCGT <u>CTAG</u> AGGCCAACATCGGCGAGACGGTACTG	<i>ΔlcpD</i>
D1578dn-R	GGCGGAAT <u>TCG</u> CGCAGATGTTGCGCACCCCTCACGT	<i>ΔlcpD</i>
rLcpA-F	TACTTCCAATCCAATGCAGCTCATCGCCCTGCACGC G	rLcpA
rLcpA-R	TTATCCACTTCCAATGTTACTAGGCGCCCGCTGGCC	rLcpA
LcpA-F	<u>GGATCCT</u> CGCCTCCTTCCAGTCTGACTGG	pLcpA
LcpA-R	<u>GAATTC</u> CCTCGGGGTCTCTCCGGCGAGTG	pLcpA
LcpA(R128A)-F	GCCGCCGATGTCATCGCCCTGGTACGC	SDM <sup>b</sup> LcpA
LcpA(R128A)-R	GGAACCCTCCACCTCCTGAGTG	SDM LcpA
LcpA(R149A)-F	GCGGACCTGACCATCAACAGCAAGG	SDM LcpA
LcpA(R149A)-R	GGGCAGGTTGATGATGGTACTC	SDM LcpA
LcpA(R266A)-F	GCCAGCCAGTCCACGGCCACCGTG	SDM LcpA
LcpA(R266A)-R	GCGTTGGGCGCCATCGGCC	SDM LcpA
LcpA(C179A)-F	GCCACCGGGCTCGGAATCCCCAC	SDM LcpA



LcpA(C179A)-R	CAGGGCGTTGACCGTGTTCTGAGG	SDM LcpA
LcpA(C365A)-F	CGCACGGCCAGCGGGCGCCTAG	SDM LcpA
LcpA(C365A)-R	GTTTTGGGACTGTACCCATAGCGGGC	SDM LcpA
RT-lcpA-F	CGGCAGATGGGTGACCATGAGC	RT-PCR <sup>c</sup>
RT-lcpA-R	CAGTGCGGCCAGGTCGCTGAG	RT-PCR
rGspA-F	TACTTCCAATCCAATGCATCCCTCGCCTTCAAGATCG CCG	rGspA
rGspA-R	TTATCCACTTCCAATGTTACTTGCCGGAGGTGGAGG CCGC	rGspA
L/CafA(K111M)	ATCGCCGGAGATCCGTGAGGG	SDM CafA
R/CafA(K111M)	ATGTGGCGAACCAGTCGATGCACC	SDM CafA
L/CafA(R123H)	CACATTGGATAGGGTGCATCGACTGG	SDM CafA
R/CafA(R123H)	CACACCACTGCCGTGGGAAAGAATGAG	SDM CafA
L/CafA(Y145R)	GTATCCGACACTCAGGCCCCCG	SDM CafA
R/CafA(Y145R)	CGTCCGGGTAAGTGGAGGGGAGACG	SDM CafA
CafA-pro-F	<u>GGATCCCAAGAAGCGCGTCGTAGATCTCCCAG</u>	pCafA, pCafA <sub>Δcws</sub> -H6
CafA <sub>Δcws</sub> -H6-R	<u>GGTACCTCAGTGGTGGTGGTGGTGGTGCACCGCCG</u> ACTTGCGGTTG	pCafA <sub>Δcws</sub> -H6
CafA-RR1-F	GCCGCATTCTTTGTCCGCTCACACCGGCGG	pCafA-RR1
CafA-RR1-R	AAGCATGAAGCGCTACCTCAGTTTTAGGGTC	pCafA-RR1
CafA-RR2-F	GCGGCGGAAGTACGTCATCACGTCAACCCTCA	pCafA-RR2
CafA-RR2-R	GTGTGAGCGGACAAAGAATCGGCGAA	pCafA-RR2
CafA-RR3-F	GCTGCGAGACTGAGGAGCGGCGCCGCCATCTCG	pCafA-RR3
CafA-RR3-R	AGACTGGCGTGAGGGTTGACGTGATGACGTC	pCafA-RR3
PromEF-Tu-F	<u>GGTACCCTGCCTCCGGGGTCCGCACC</u>	GFP and DagA reporters
PromEF-Tu-R	TGGTGTCTCCTGGGACTCGGGTAG	GFP and DagA reporters
DagA <sub>Sv</sub> -F	GTGGTCAACCGACGTGATCTCATCAAG	pDagA <sub>Sv</sub>
DagA <sub>Sv</sub> -R	<u>GGATCCCTACACGGCCTGATACGTCCTGAC</u>	pDagA <sub>Sv</sub> and derivatives
GFP-F	ATGTCTAAAGGTGAAGAACTGTTACCG	pGFP, pDagA <sup>SP</sup> - GFP, pFimA <sup>SP</sup> -GFP, pCafA <sup>SP</sup> -GFP
GFP-R	<u>GGATCCCTATTTGTAGAGCTCATCCATGCCGTG</u>	pGFP, pDagA <sup>SP</sup> - GFP, pFimA <sup>SP</sup> -GFP, pCafA <sup>SP</sup> -GFP
DagA <sup>SP</sup> -R	GAGGTCTGCGGCATGAGCGGC	pDagA <sup>SP</sup> -GFP
FimA <sup>SP</sup> -F	ATGAAGCACAACGCCAGCACGCTG	pFimA <sup>SP</sup> -GFP, pFimA <sup>SP</sup> -DagA

FimA <sup>SP</sup> -R	CGTTTCCGTGGCCACGGCC	pFimA <sup>SP</sup> -GFP, pFimA <sup>SP</sup> -DagA
CafA <sup>SP</sup> -R	GGCCTCAGCCGGGGGAGCCG	pCafA <sup>SP</sup> -GFP, pCafA <sup>SP</sup> -DagA
TatC-up-F	<u>GGTACC</u> CTGGCCTGCTTGACGGCCTTAAGCAG	pTatAC-KO
TatC-up-R	<u>AAGCTT</u> CTTCACCCCGTCCCTGCCTGACTTG	pTatAC-KO
TatC-dn-F	<u>AAGCTT</u> GATGGTTCTCCCCAAGGCGGACGG	pTatAC-KO
TatC-dn-R	<u>GAATTC</u> GACTACACTTTGCCCGCGCTGCAC	pTatAC-KO
tetR-Ω-F	<u>GGTACC</u> TAAAGACCCACTTTTCACATTTAAGTTG	pTetR-Ω-tatC
tetR-Ω-R	CTTGTTGCCTCCTTAGCAGGGTG	ptetR-Ω-tatC
TatC-F	GTGAACCTCTTTAAGCCGTCGCAC	pTetR-Ω-tatC
TatC-R	<u>GAATTC</u> CCTACTCGGCCAGGGCGGCCTCGAG	pTetR-Ω-tatC
RT-cafA-F	GCGCGGTAAGACCGCCTCAG	cafA, qRT-PCR
RT-cafA-R	CGGCCGACTTGGGAGACGATG	cafA, qRT-PCR
RT-16s-F	GTCGCTAGTAATCGCAGATCAG	16S rRNA, qRT-PCR
RT-16s-R	GGTGTTGCCGACTTTCATG	16S rRNA qRT-PCR

<sup>a</sup> Engineered restriction sites are underlined

<sup>b</sup> SDM – primers used for site directed mutagenesis by inverse PCR

<sup>c</sup> RT-PCR – reverse transcription PCR

**Cell growth assays.** *A. oris* growth was assessed by a plate assay and optical density (OD<sub>600</sub>) according to a previous protocol (50). For the plate assay, overnight cultures of various strains were harvested and normalized to the same OD<sub>600</sub> in HIB. Equivalent cell suspensions were subjected to 10-fold serial dilution and spotted on HIA plates supplemented with or without 100 ng mL<sup>-1</sup> AHT and 2 mM theophylline. Cell growth at 37°C was recorded after 3 days. For growth in liquid broth, strains were sub-cultured to OD<sub>600</sub> of ~ 0.02. Cell growth was monitored by OD<sub>600</sub> over 19 h. The OD values were presented as an average of three independent experiments performed in duplicate.

**Reverse transcriptase and quantitative real-time polymerase chain reactions.**

Overnight cultures were used to inoculate fresh cultures of various *A. oris* strains, which were grown until OD<sub>600</sub> ~ 0.25. Cells were harvested by centrifugation, and cell pellets were suspended in Trizol (Ambion) and lysed by mechanical disruption using 0.1 mm silica spheres (MP Bio). Total RNA was extracted using a Direct-zol RNA MiniPrep kit (Zymo Research). Complementary DNA (cDNA) was synthesized using SuperScript III reverse transcriptase (Invitrogen). For quantitative real-time PCR, cDNA was mixed with iTAQ SYBR green supermix (Bio-Rad), along with appropriate primer sets (Table 2-3). Threshold concentration (C<sub>t</sub>) values were determined and the relative expression level was calculated using the 2<sup>-ΔΔC<sub>t</sub></sup> method (92) with the 16s rRNA gene serving as an internal control. The data were obtained from three independent experiments performed in triplicate.

**Cell fractionation and immunoblotting.** Overnight cultures of *A. oris* strains were diluted 40-fold in HIB and grown at 37°C until OD<sub>600</sub> ~ 0.4 - 0.5. Normalized aliquots were subject to cell fractionation as described previously (93). Protein samples obtained from supernatant (S), cell wall (W), membrane (M) and cytoplasm (C) fractions were analyzed by 15% or 3-20% polyacrylamide gels, separated by SDS-PAGE, and immunoblotted with specific antibodies (1:1,333 dilution for α-GspA; 1:10,000 for α-

FimP; 1:8,000 for  $\alpha$ -FimA and 1:4,000 for  $\alpha$ -MdbA; 1:10,000 dilution for  $\alpha$ -GspA; 1:1,000, affinity purified  $\alpha$ -LcpA; and 1:8,000,  $\alpha$ -SrtA; 1:8,000 for  $\alpha$ -CafA). The proteins were detected by chemiluminescence.

LcpA signal from four independent blots was normalized against a Coomassie R250 stained loading control band from the same blot and quantified using ImageJ, <https://imagej.nih.gov/>. The obtained intensity values were normalized to those of the wild-type strain, which were set to 1. The results were presented as average of four independent experiments.

**Protein purification and analysis.** Recombinant plasmids pGspA $_{\Delta$ cwss and pFimA $_{\Delta$ cwss generated above were introduced to *A. oris* strains containing combinations of the signal peptidases *lepB1* and *lepB2* by electroporation. Each strain was inoculated into 200 mL HIB supplemented with 50  $\mu$ g mL $^{-1}$  kan and grown overnight. The cell-free cultures were obtained by centrifugation and filtration using 0.45  $\mu$ m-pore size filters. 2 ml of nickel resin, washed twice with EQ buffer (150 mM NaCl, 50 mM Tris.HCl, pH 7.5) were added to the supernatants and gently agitated at 4°C overnight. The suspensions were decanted onto a protein purification column, which was then washed with 20 ml EQ buffer. Bound proteins were eluted in 3-ml fractions with 1 M imidazole. The eluted proteins were desalted using a desalting column (BioRad), and concentrated using an Amicon Ultra 0.5 mL centrifugal unit (Millipore) with 10K (GspA) or 30K (FimA) cutoff.

GspA<sup>LMM</sup> was purified as above for pGspA $_{\Delta$ cwss from the  $\Delta$ /*lcpA* background.

rLcpA and alanine substitution mutants – generated as described above, were introduced into *E. coli* SHuffle for protein purification. The strains were inoculated into 500 mL of LB supplemented with 100 Amp  $\mu$ g mL $^{-1}$  and grown to OD<sub>600</sub> = 0.8 at 30°C. Protein expression was induced using 0.1 mM IPTG and the culture was transferred to 16°C overnight. Cells were pelleted by centrifugation and washed by suspension in EQ

buffer (150 mM NaCl, 50 mM Tris HCl, pH 7.4). Cells were treated with 1X protease inhibitor cocktail (GenDEPOT). The treated cells were chilled and then lysed with a French press. Lysates were centrifuged to remove cell debris and nucleic acids. The remaining soluble fraction was decanted onto an equilibrated NiNTA column (Qiagen). The bound proteins were washed with up to 100 mM imidazole to remove nonspecific proteins, and eluted with 500 mM imidazole in wash buffer (1X EQ and 10% glycerol). Imidazole was removed from the eluates with a desalting column (Bio-Rad) exchanged with wash buffer.

Purified proteins were analyzed by SDS-PAGE and visualized with Coomassie Blue and Periodic acid-Schiff staining as previously described (50).

For Edman sequencing, the proteins were transferred to a PVDF membrane. The membrane was rinsed with ultrapure water and stained with 0.02% Coomassie Brilliant blue for 30 seconds followed by de-staining in 40% methanol, 5% acetic acid solution for 1 minute and rinsed with ultrapure water 3 times. The membrane was completely air-dried, and protein bands were excised for Edman sequencing at the TUCF core facility (Tufts University).

CafA was purified from the supernatant of *A. oris* cells transformed with the pCafA<sub>ΔCWS</sub>-H6 plasmid. One liter of cells was grown to OD = 0.8, and then the cells were centrifuged twice to collect the cell-free supernatant. To remove remaining cells, the supernatant was filtered through a vacuum filtration device (Olympus) with a 0.45 μM pore. The supernatant was then treated with 1X protease inhibitor cocktail (GenDepot-Xpert) and decanted onto a Capturem™ His-Tagged Purification Large Volume unit twice. The column was washed twice with EQ buffer, and the proteins were eluted with 30 mL of 0.5 M imidazole in EQ buffer. The eluate was desalted using a size exclusion

column (Bio-Rad) exchanged with EQ buffer, and concentrated using a 15 mL centrifugation filter with a 50K cutoff according to manufacturer instructions (Millipore).

**Coaggregation and biofilm assays.** Coaggregation between *A. oris* and *Streptococcus oralis* was assayed according to a previously published protocol (93). Briefly, stationary cultures of *A. oris* and *S. oralis* were collected by centrifugation, washed, and suspended in coaggregation buffer (93). To quantify coaggregation, optical density (OD<sub>600</sub>) of individual *A. oris* and *S. oralis* cell suspensions in coaggregation buffer was measured. Equal cell volumes (0.5 ml) were mixed and allowed to form aggregates. The resulting aggregates were photographed by an Alpha-Imager or removed by centrifugation at 100 x g, and OD<sub>600</sub> of the remaining supernatants were recorded. Relative coaggregation was determined as previously described (94).

*A. oris* biofilms were cultivated *in vitro* as reported before (79). Briefly, cells were inoculated in HIB supplemented with 1% sucrose in a 24-well polystyrene plate and incubated at 37°C with 5% CO<sub>2</sub> for 48 h. After incubation, the media was removed, and each well was gently washed with PBS three times and dried over night at room temperature. Biofilms were stained with 1% crystal violet and washed with water to remove the unbound dye. To quantify biofilm production, the biofilms were destained with 95% ethanol and the released dye was quantified by measuring absorbance at 580 nm using a Tecan Infinite M1000 microplate reader. The assays were performed three times in triplicate.

**Whole-cell ELISA.** Surface expression of *A. oris* tip proteins was quantified by whole-cell ELISA as previously described with some modifications (95,96). Briefly, mid-log phase cells of *A. oris* MG1 and its isogenic mutant strains were harvested by centrifugation and washed twice with phosphate-buffered saline (PBS). The cells were suspended in carbonate-bicarbonate buffer (15 mM sodium carbonate, 35 mM sodium

bicarbonate, pH 9.6) and normalized to an OD<sub>600</sub> of 1.0. 100- $\mu$ L aliquots of cell suspensions were dispensed into 96-well high-binding ELISA plates (Corning Costar EIA/RIA plate) to allow cell binding to wells at 4°C overnight. Unbound cells were removed by washing with PBS containing 0.05% Tween 20 (PBST). Adhered cells were blocked with 2% bovine serum albumin (BSA) in PBST prior to incubating with specific antibodies against CafA and FimB (1:10,000) for 2 h at 25°C, followed by washing with PBST, and staining with anti-rabbit IgG HRP-linked antibodies (Cell Signaling; 1:20,000) for 1 h at 25°C. Washed cells were treated with 100  $\mu$ L of the chromogenic 3,3',5,5'-Tetramethylbenzidine (TMB) substrate (Affymatrix eBioscience) prior to absorbance measurement at 450 nm using a Tecan M1000 plate reader. The absorbance for MG1,  $\Delta$ *lepB1*, and  $\Delta$ *lepB2* strains as compared with the  $\Delta$ *cafA* and  $\Delta$ *fimB* mutants used as background was determined from at least two independent experiments performed in triplicate.

**Immunogold-labeling and electron microscopy.** Pili were visualized by electron microscopy (EM) as previously described (62). Briefly, *A. oris* cells grown in HIB were pelleted, washed once, and suspended in 0.1 M NaCl. For negative staining, a drop of bacterial suspension was placed onto carbon-coated nickel grids and stained with 1% uranyl acetate before viewing with a JEM1400 electron microscope. For immunogold electron microscopy (IEM), after immobilizing cells on the nickel grids, samples were stained with specific antibodies ( $\alpha$ -FimP, 1:100 dilution;  $\alpha$ -FimA, 1:100 dilution) followed by incubation with 12-nm gold-goat anti-rabbit IgG (Jackson ImmunoResearch). Finally, the samples were washed five times with water before staining with 1% uranyl acetate.

**Pyrophosphatase assay.** This assay uses a 24-hour endpoint reaction of rLcpA and its derivatives combined with farnesyl pyrophosphate (FPP) to detect the amount of phosphate released. LcpA concentration for all assays was 3  $\mu$ M, and FPP substrate

was titrated as indicated. The reactions were buffered in 20 mM Tris-HCl (pH 8.0), and incubated for 24 hours at 30°C (97). Controls for LcpA only and FPP only were included for each reaction to determine any nonspecific signal. To detect release of inorganic phosphate a Phosphate Fluorometric Assay Kit (MAK031, Sigma) was used according to manufacturer instructions. The presence of inorganic phosphate leads to the conversion of sucrose to glucose-1-phosphate via an enzymatic reaction. The glucose-1-phosphate is then oxidized and reacts with a probe, which results in the release of a fluorescent signal proportional to the amount of phosphate in the sample. The phosphate detection master mix was added to either a phosphate standard or the endpoint samples in a black 96-well plate and then incubated in the dark for 1 h. The fluorescence from each sample was measured by a Tecan microplate reader at  $\lambda_{\text{ex}} = 535/\lambda_{\text{em}} = 587$  nm. Results from the phosphate standards were used to generate a standard curve and included for each biological replicate. Each run also included a zero standard, which did not include any phosphate and this was used to determine background fluorescence to be subtracted from all readings. Linear regression analysis of the standard curve was used to determine the concentration of phosphate in the test samples. Non-linear regression and statistical analyses were performed in Prism GraphPad (version 5.04).

**Phosphotransferase assay and 2D protein electrophoresis.** 12  $\mu\text{M}$  GspA<sup>LMM</sup> was incubated with 4  $\mu\text{M}$  of LcpA (WT or R149A) and 50  $\mu\text{M}$  of FPP in 20 mM Tris-HCl (pH 8.0) for 72 h at 30°C. After incubation, the protein samples were treated with hydrofluoric acid (HF) following published protocols (48,98-100) or mock-treated with H<sub>2</sub>O. Briefly, protein samples were treated with 46% HF at 4°C for 18 h. After acid removal by vacuum evaporation, the protein samples were washed with 500  $\mu\text{L}$  of deionized water followed by vacuum evaporation.

To resolve GspA modified by a phosphate group, samples were solubilized for 30 mins at 25°C in rehydration buffer (Bio-Rad). To resolve GspA<sup>LMM</sup> species by isoelectric



point, we used an 11 cm IPG strip with a narrow linear range (pH 3 – 6) (Bio-Rad). The sample was added during the rehydration step and overlaid with mineral oil. The IPG strips were rehydrated actively (50 V) with the sample for 12 hours, then the voltage was increased rapidly to a max of 8,000 V for the isoelectric focusing step with a Protean IEF system (Bio-Rad). After IEF, the IPG strips were placed onto a 4 – 20% Criterion™ TGX™ IPG+1 gel (Bio-Rad) in overlay agarose (0.5% agarose in 1X Tris/Glycine/SDS (TGS) buffer and 0.003% bpb) alongside a precision plus protein™ dual color standard (Bio-Rad). After the overlay agarose was set, 1X TGS buffer was added and the samples were run at 200 V for 57 min. The proteins were then transferred to a PVDF membrane for subsequent immunoblotting as above with an  $\alpha$ -GspA antibody.

**HGF-1 bacterial adherence assay.** Human gingival fibroblast (HGF-1) cells (ATCC® CRL-2014™) were grown in DMEM with 10% fetal bovine serum (FBS) and penicillin/streptomycin (Thermofisher Scientific) (DMEM +/+) at 37°C in 5% CO<sub>2</sub> until 100% confluency. For the adherence assay 6-well plates were seeded at a density of  $6.0 \times 10^4$  cells/well, and allowed to reattach in DMEM +/+ for 48 hours at 37°C in 5% CO<sub>2</sub> (101). Bacterial cells were grown to midlog phase from overnight cultures, washed in Dulbecco's phosphate buffered saline (DPBS), and then normalized to  $1 \times 10^8$  cells. The washed cells were diluted 1:100 in DMEM with 10% fetal bovine serum (FBS) without antibiotics (DMEM+/-) and plated for the colony forming unit (cfu) input count. Each strain was plated in triplicate for each experiment. Meanwhile, the HGF-1 cells were treated with 2U neuraminidase/sialidase (*Clostridium perfringens*; Sigma) and 1X protease inhibitor cocktail (GenScript) for 2 hours at 37°C in 5% CO<sub>2</sub> in DMEM without FBS or antibiotics (DMEM+/-). After washing sialidase from HGF-1 cells in DMEM+/-, 200  $\mu$ L of  $1 \times 10^6$  *A. oris* cells were added for a final multiplicity of infection (MOI) of 200 bacterial cells/HGF-1 cell. The plates were centrifuged at 200 x g for 5 min and then incubated 2 hours at 37°C in 5% CO<sub>2</sub>. HGF-1 cells were lysed with cold water and vigorous pipetting.

Bacterial cells were enumerated by plating for cfu/mL on HIA in duplicates per well.

Adherence assay were all performed in technical triplicates with duplicate enumeration, and each assay was replicated three times independently.

**Genomic DNA sequencing.** Genomic DNA (gDNA) sequencing steps were performed at CARMiG by An Dinh. First, gDNA was isolated from 1 mL of cells from overnight cultures of *A. oris* MG1, N11A16, N32A8, and N11A12 according to the DNeasy Blood & Tissue Purification Kit (Qiagen) instructions for Gram-positive bacteria. To determine the concentration and quality of the prepared gDNA Qubit dsDNA High Sensitivity kit and gel electrophoresis were used. Samples were diluted to 0.2 ng/μL. A gDNA library was prepared using Nextera XT Library Prep kit (Illumina) according to manufacturer instructions. Libraries were assessed for yield (Qubit), and average insert size using Agilent D5000 High Sensitivity ScreenTape on Agilent 4200 TapeStation. Normalized sample libraries were pooled to achieve 80-fold genome coverage depth, and then denatured and diluted to 12 pM with a 1% PhiX spike-in control. Sequencing was performed on an Illumina MiSeq platform with MiSeq V2 reagent kit. Paired-end reads with onboard Illumina base-calling and adapted-trimming to yield \*.fastq files. CLC Genomics workbench was utilized to filter, trim, and *de novo* assemble the reads. The data was analyzed using the comparative analysis platform EDGAR.

**Agarase clearance assay.** To determine whether signal peptides were sufficient to export proteins via the Tat pathway, we utilized a modified version of the agarase assay originally developed for testing heterologous signal peptide fusions from *Streptomyces coelicolor* in *Streptomyces lividans* (102,103). *A. oris* cells expressing native DagA, DagA signal peptide fusions, or an empty vector control were grown to mid logarithmic phase in HIB with kanamycin, and then concentrated to OD<sub>600</sub> of 2.0. Three microliters of concentrated cells were spotted onto HIA and grown for 48 hours at 37°C with 5% CO<sub>2</sub>.

Lugol solution (Sigma) was added to the plates for 1 h, then decanted to image the zone of clearance.

**Fluorescence intensity measurements.** Fluorescence intensity of was measured from the supernatant of bacterial cells grown in HIB or from cells washed and suspended in PBS. A blank measurement was made for HIB or PBS alone, depending on the sample. Additionally, a gain reference was made using 40  $\mu\text{g}/\text{mL}$  purified GFP in either HIB or PBS. The mean of technical triplicates was measured and averaged. Overall percentages of fluorescent signal in the media compared to the whole cell were calculated from three independent experiments. The time-course fluorescence was only taken from the media fraction and assessed every hour for 8 h with a final reading at 18 h.

**Fluorescence microscopy.** 10-well multi-test slides (MP Biomedicals) were prepared by adding 2  $\mu\text{L}$  of 1.5% agarose in 1X PBS. Bacterial samples grown to mid-log phase were centrifuged, washed with 1X PBS, pipetted onto an individual agarose pads, and a cover slip was placed over the slide to secure the samples. Images were obtained on an Olympus IX81-ZDC inverted microscope using Slidebook imaging software. Fluorescent images were taken at 0.5 sec exposure with 255 gain and exported as TIFF files.

## **CHAPTER III**

### **A Type I Signal Peptidase is Required for Pilus Assembly in the Gram-positive, Biofilm-forming Bacterium**

*Actinomyces oris*

*The results and figures in this chapter are from work published in 2016: Siegel SD, Wu C, Ton-That H "A type I signal peptidase is required for pilus assembly in the Gram-positive biofilm-forming bacterium Actinomyces oris." J Bact. 198(15):2064-73. doi: 10.1128/JB.00353-16. I am the first author of this publication and was responsible for preparing the original manuscript and conducting the experiments described in this paper. I have permission to reproduce any and all of this manuscript, in print or electronically, for the purpose of my thesis in accordance with the American Society for Microbiology (publisher of Journal of Bacteriology) "Journals Statements of Authors' Rights." [http://journals.asm.org/site/misc/ASM\\_Author\\_Statement.xhtml](http://journals.asm.org/site/misc/ASM_Author_Statement.xhtml).*

## **Introduction**

The role of the housekeeping sortase in Gram-positive bacterial pathogenesis has been demonstrated in a diverse abundance of species. For example, in *S. aureus* deletion of housekeeping sortase results in attenuation of pathogenicity in several different models of infection (104). This is due to the reduction of anchored virulence factors on the cell surface, such as protein A, the paradigm sortase substrate. Similar virulence-associated phenotypes have also been implicated in other Gram-positive bacteria including *Listeria monocytogenes*, *Streptococcus pneumoniae*, and *Bacillus anthracis* (105-107).

In an effort to identify how deletion of the *A. oris* housekeeping sortase affects biofilm formation it was discovered that it is essential for cell viability, a feature unique to this species (50). While the sortase cannot be deleted or catalytically inactivated, it was possible to delete all predicted LPXTG containing proteins (25). This led to the hypothesis that a toxic protein may accumulate at the membrane in the absence of housekeeping sortase. Using a transposon screen for suppressors of  $\Delta srtA$  lethality, 13 suppressor mutants were identified, and two were characterized further. The two proteins, which were characterized after the transposon screen, were LCP (LytR-CpsA-

*Psr*) and *AcaC* (*Actinomyces cell wall associated protein C*). These genes are clustered in the genome and both had multiple transposon hits. *AcaC* (now *GspA*) was glycosylated in an LCP-dependent manner and required *SrtA* for anchoring. A model was proposed that accumulation of glycosylated *GspA* in the membrane in the absence of sortase leads to membrane glycol-stress and eventual cell death (50).

Additional mutations that could suppress  $\Delta srtA$  lethality were identified from the original transposon mutagenesis screen. We hypothesized that glycosylation or anchoring pathway of *GspA* was central to the essentiality of *srtA*. Thus, all of the mutations were specifically associated with one of the two pathways. One such mutation was found to disrupt the open reading frame of *ana\_1190*, annotated as a putative type I signal peptidase (SPase) and renamed *lepB2* (50).

Secreted proteins encode a signal peptide at the N-terminus. After secretion, the non-lipoprotein signal peptides are cleaved by a bacterial type I SPase. This constitutes the terminal step of the secretion pathway. Type I SPases are endopeptidases that utilize a serine-lysine catalytic dyad. SPases recognize an AxA motif that is a broadly conserved cleavage sequence among Gram-negative and Gram-positive bacteria (53). Gram-negative bacteria generally encode only a single essential type I SPase, however, Gram-positive bacteria often encode several SPases, some of which cleave only specific substrates (60,108,109).

In this work, we investigated the role of *LepB2* in the *GspA* glycosylation and anchoring pathway and determined that it is dispensable for *GspA* signal peptide cleavage. Rather, it modulated the glycosylation of *GspA*, and may have specifically impacted a yet unknown glycosylation enzyme. We further showed that *LepB2* does cleave the signal peptide of pilin proteins and that signal peptide cleavage is required for pilus polymerization. Improper cleavage of the signal peptide impacted bacterial physiology.

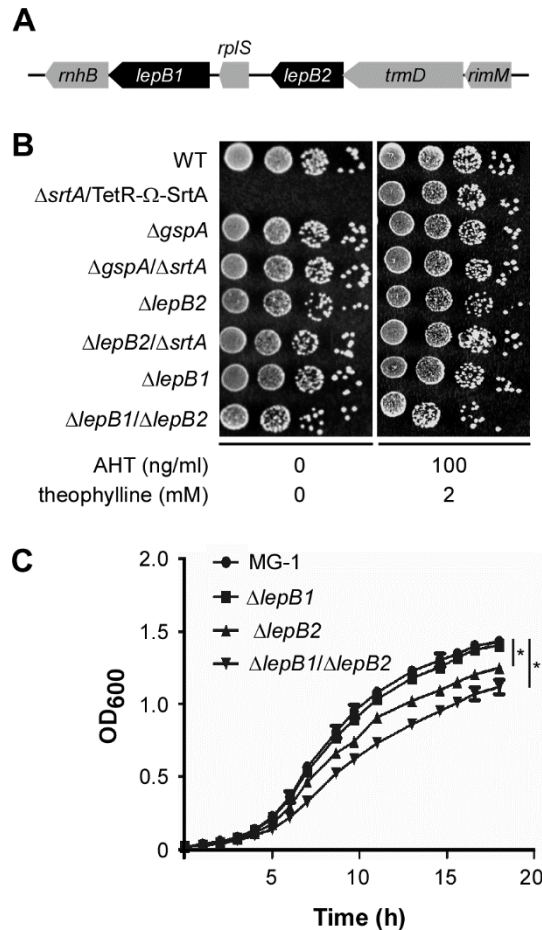
## Results

### **The type I signal peptidase LepB2 is a genetic suppressor of *srtA* essentiality –**

We have previously reported that *A. oris srtA* is an essential gene and identified the Tn5::*lepB2* mutation as one of the genetic suppressors of *srtA* essentiality (50). *lepB2* is genetically linked to *trmD* and *rimM*, which encode a tRNA (guanine(37)-N(1))-methyltransferase and a 16S rRNA processing protein, respectively (Fig. 3-1A).

According to the BioCyc databases, <http://biocyc.org/>, the three genes are predicted to be expressed as one transcription unit in *Actinomyces naeslundii* Howell 279 (110).

Farther upstream of *lepB2* is another signal peptidase-encoding gene annotated as *lepB1* (Fig. 3-1A). To exclude the possibility that the Tn5::*lepB2* has a polar effect on adjacent genes, we created in-frame, nonpolar deletion mutations by allelic exchange (79) including the *lepB2/srtA* double deletion mutation ( $\Delta lepB2/\Delta srtA$ ). At first, we generated a *lepB2* mutation which enabled further deletion of *srtA* from the bacterial chromosome without ectopic expression of SrtA. Of note, we were unable to delete *srtA* from a  $\Delta lepB1$  mutant strain. The  $\Delta lepB2/\Delta srtA$  strain and respective single mutations were assessed for their ability to grow in laboratory conditions using both plate and liquid broth assays (50). In the plate assay, normalized cell cultures were spotted in serial dilution on HIA plates and grown at 37°C. Compared to the conditional *srtA* deletion mutant, which failed to grow in the absence of two inducers anhydrous tetracycline (AHT) and theophylline (50), the  $\Delta lepB2/\Delta srtA$  mutant did not display any visible defects when grown in any conditions (Fig. 3-1B). Growth of  $\Delta gspA$ ,  $\Delta lepB2$ ,  $\Delta lepB1$  and  $\Delta lepB1/\Delta lepB2$  mutants was comparable to that of the parental strain MG1 (Fig. 3-1B). Nonetheless, when grown in liquid broth, the  $\Delta lepB2$  and  $\Delta lepB1/\Delta lepB2$  mutants displayed a slight growth defect as compared to the MG1 and  $\Delta lepB1$  strains (Fig. 3-1C).

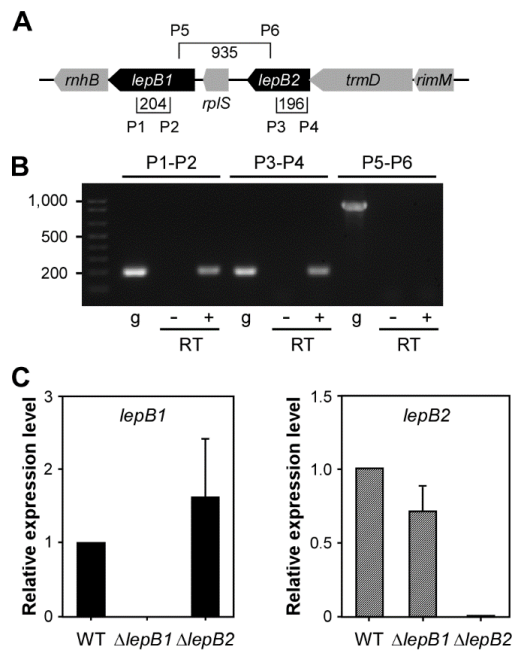


**Figure 3-1: Signal peptidase LepB2 is a genetic suppressor of SrtA essentiality.**

(A) Shown is a graphic representation of the *lepB2* gene locus from *A. oris* MG1. This locus encodes a ribonuclease H (*rnhB*), SPase I (*lepB1*), a ribosomal protein L19 (*rplS*), SPase I (*lepB2*), a tRNA (guanine-N1)-methyltransferases (*trmD*), and a 16S rRNA processing protein (*rimM*). (B) Growth of wild-type MG1, conditional *srtA* mutant, and non-polar deletion mutant strains on solid media in the presence or absence of inducers anhydrotetracycline (AHT) and theophylline. (C) Growth of *A. oris* MG1,  $\Delta lepB1$ ,  $\Delta lepB2$  and  $\Delta lepB1/\Delta lepB2$  was determined in liquid media. The optical density (OD<sub>600</sub>) values were presented as an average of three independent experiments done in duplicate. Symbols \* indicate the p values less than 0.05 that were determined using the paired, two tailed t-test with Prism GraphPad.

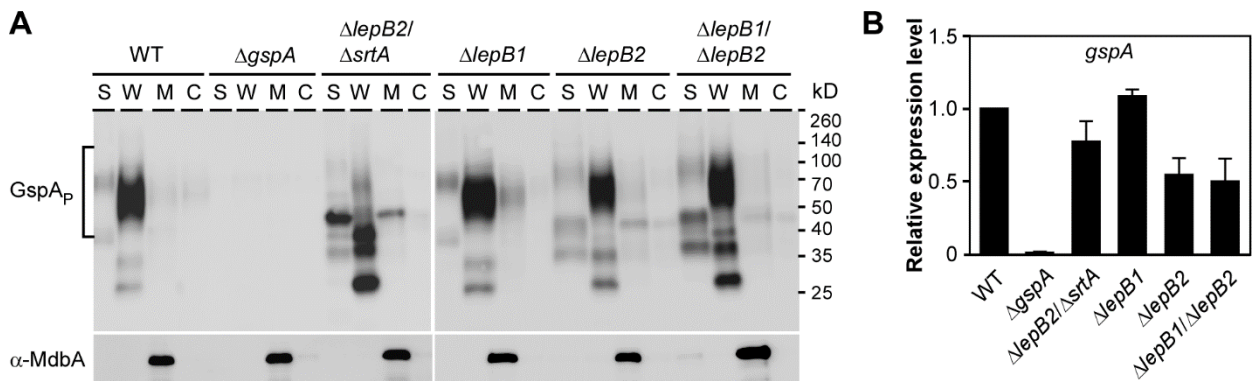


We next examined if *lepB1* and *lepB2* are part of a transcriptional unit using reverse transcription polymerase chain reaction (RT-PCR) as previously described (111). Total RNA isolated from MG1 was used in reverse transcription reactions to produce cDNA. As shown in Fig. 3-2A and 3-2B, RT-PCR detected the presence of *lepB1* or *lepB2* transcripts using probes specific for *lepB1* (P1 and P2 primers) or *lepB2* (P3 and P4), respectively. Neither transcript was observed in the absence of reverse transcriptase (RT), indicative of no gDNA contamination. No transcripts were detected using probes specific for the region encompassing *lepB1* and *lepB2* (P5 and P6), whereas the same probes enabled amplification of this region from gDNA (Fig. 3-2B, g lanes). To ascertain whether deletion of *lepB2* does not affect *lepB1* expression, we examined expression levels of both genes using quantitative real-time PCR (qRT-PCR). Compared to the wild-type strain, relative expression of *lepB1* or *lepB2* was not altered when *lepB2* or *lepB1* was deleted, respectively (Fig. 3-2C) suggesting that *lepB1* and *lepB2* are independently expressed. Altogether, the results in Figures 1 and 2 confirmed that LepB2 is a genetic suppressor of *srtA* deficiency and suggest that LepB2 is required for processing of substrates that are involved in the SrtA-mediated glycosylation pathway.



**Figure 3-2: *lepB1* and *lepB2* are independently expressed.** (A) Specific primers (P) were designed to detect *lepB1*, *lepB2* and a region encompassing both. Brackets and numbers specify the primer position and sizes of amplicons. (B) Using specific primers in (A), reverse transcription PCR (RT-PCR) was employed to determine if *lepB1* and *lepB2* are independently expressed. Plus and minus signs indicate RT-PCR reactions were performed in the presence and absence of reverse transcriptase (RT). PCR reactions with genomic DNA (g) were used as control. (C) Quantitative real-time PCR (qRT-PCR) was employed to determine the expression levels of *lepB1* and *lepB2* relative to the parental strain MG1, using the  $2^{-\Delta\Delta C_t}$  method, where 16s rRNA serves as the internal control. The data were presented averages of three independent experiments performed in triplicate; error bars represent standard deviation (SD).

**Involvement of SPases LepB1 and LepB2 in GspA glycosylation** – We previously showed that glycosylated GspA is a substrate of SrtA since depletion of *srtA* leads to the membrane accumulation of glycosylated GspA causing lethal phenotypes (50). Subsequent Tn5 mutagenesis identified viable mutants that lack both *srtA* and *gspA* or *srtA* and *lepB2* (50). In this work, we have confirmed that LepB2 is a genetic suppressor of *srtA* essentiality by generating a nonpolar deletion mutant devoid of *srtA* and *lepB2* (Fig. 3-1). Therefore, we hypothesized that LepB2 works on the same GspA glycosylation pathway by processing GspA before it is glycosylated and then anchored to the cell wall by sortase SrtA. To examine this possibility, we first examined the status of GspA glycosylation in the absence of SPases by analyzing GspA proteins isolated from the culture medium (S), cell wall (W), membrane (M), and cytoplasmic (C) fractions of wild-type and SPase mutant strains. Protein samples were immunoblotted with antibodies against GspA and MdbA, a membrane-bound protein as a control (62). In the wild-type MG1 strain, glycosylated GspA polymers (GspA<sub>P</sub>) were detected as smeared bands in the cell wall fraction, and no GspA signal was observed in the  $\Delta gspA$  mutant (Fig. 3-3A, first 8 lanes), as previously reported (50). While the *lepB1*, *lepB2*, and  $\Delta lepB1/\Delta lepB2$  mutants did not exhibit significant defects in GspA glycosylation, the  $\Delta lepB2/\Delta srtA$  double mutant failed to produce wild-type levels of GspA polymers. Instead, the accumulation of three low molecular weight (LMW) species was detected in the cell wall fraction, in addition to a 50-kDa species observed in the culture medium and membrane fractions (Fig. 3-3A, lanes  $\Delta lepB2/\Delta srtA$ ). To determine whether the observed defects might be due to low expression levels of *gspA*, we quantitated gene expression by qRT-PCR. Overall, no significant reduction in gene expression levels was observed in the mutant strain when compared with the wild-type strain (Fig. 3-3B).

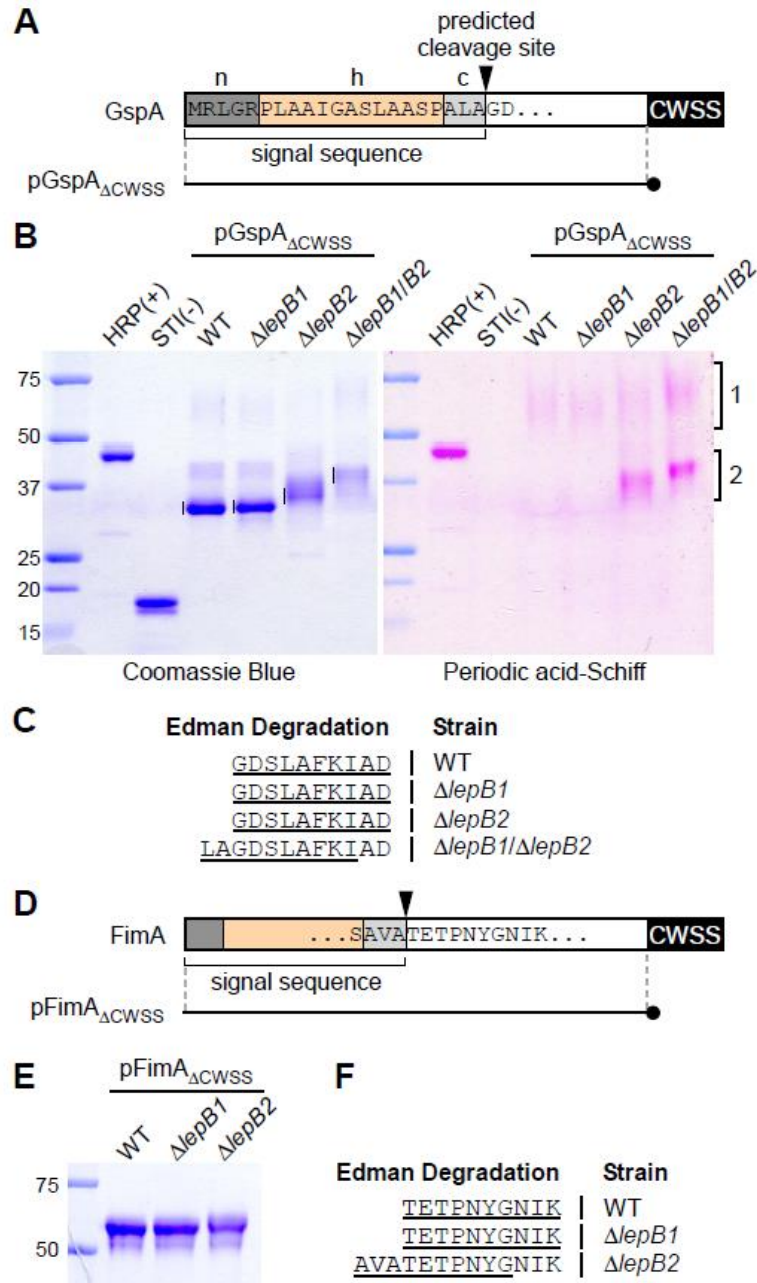


**Figure 3-3: Involvement of SPases in sortase-associated GspA glycosylation. (A)**

Cells were grown to mid-log phase and normalized by optical density. Culture supernatant (S), cell wall (W), membrane (M), and cytoplasmic (C) fractions were separated by SDS-PAGE and immunoblotted with antibodies against GspA ( $\alpha$ -GspA) and membrane protein MdbA ( $\alpha$ -MdbA), which served as loading and fractionation controls. Brackets designate glycosylated, high molecular weight (HMW) species of GspA polymers (GspA<sub>P</sub>). **(B)** The relative expression level of *gspA* in indicated strains was determined by qRT-PCR as described in Fig. 2. Error bars are the SD from two independent experiments performed in triplicate.

Altogether, the results suggest that either LepB1 or LepB2 is sufficient for cleavage of the cell wall-anchored glycoprotein GspA and that LepB2 may process factors involving GspA glycosylation.

**Specificity of *A. oris* SPases for processing of LPXTG-containing proteins** – GspA is one of 18 cell wall-anchored proteins with a CWSS and a putative signal peptide. A typical signal peptide is comprised of a net positive charged region (n-region), followed by a hydrophobic region (h-region) and a cleavage region (c-region) (53) (Fig. 3-4A). The c-region harbors a typical cleavage site, i.e. an AXA motif, which is recognized and cleaved by a type I SPase after the second Ala residue (53). To determine the cleavage site of GspA and the role of LepB1 and LepB2 in this process, we purified GspA by affinity chromatography from *A. oris* strains lacking either gene or both and determined the N-terminal sequences of the purified proteins by Edman degradation sequencing. Because GspA is heavily glycosylated and its glycosylation sites are not known (50), we expressed in these strains a GspA variant lacking its membrane-bound CWSS and containing a His-tag at the C-terminus (Fig. 3-4A; pGspA $\Delta$ CWSS). Using this approach, we could avoid the formation of high-molecular weight glycosylated GspA by forcing the protein to be released as a monomer to facilitate detection of small changes in pre-protein processing. After harvesting the cell-free culture supernatants of *A. oris* strains expressing this construct by centrifugation and filtration, His-tagged proteins were captured by Ni-NTA agarose and purified according to a published protocol (25). Purified proteins were analyzed by SDS-PAGE, followed by Coomassie Blue and Periodic acid-Schiff (PAS) staining to distinguish glycosylated proteins (see Materials and Methods). In the WT strain, the predominant GspA band and a weaker band migrated below and above the 37-kDa marker, respectively, in addition to the last one smeared between the 50- and 75-kDa markers by Coomassie Blue staining (Fig. 3-4B; left panel). As shown previously, this predominant GspA species was negative for PAS staining, while



**Figure 3-4: Mapping the signal peptide cleavage site by Edman degradation.** (A) Shown is a graphic representation of GspA. The N-terminal sequence of GspA contains a typical signal peptide comprised of a positively charged polar n-region, a hydrophobic central region (h-region), and a c-region with the signal peptidase recognition motif AxA. The arrowhead marks the predicted SPase I cleavage site. pGspA $\Delta$ CWSS denotes a

plasmid expressing a GspA variant, in which the CWSS is replaced by a six His-tag (black circle). **(B)** His-tagged GspA proteins were purified from the culture medium of indicated strains by affinity chromatography, separated by SDS-PAGE and visualized by Coomassie Blue (~ 1.25 µg of protein in all lanes) and Periodic acid-Schiff (PAS) (2.5 µg of protein) staining. Glycosylated horseradish peroxidase (HRP) and non-glycosylated soybean trypsin inhibitor (STI) were used as positive and negative controls. Brackets with numbers indicates glycosylated forms of GspA. **(C)** Indicated bands (vertical lines) in B were subjected to Edman degradation. The N-terminal sequence of GspA was deduced from the first 10 sequencing cycles (underline). Samples from the  $\Delta lepB1/\Delta lepB2$  mutant produced ragged N-terminal sequencing with the major sequence shown. **(D)** Like GspA, FimA harbors a signal peptide with the conserved cleavage site motif AXA. pFimA $\Delta_{CWSS}$  represents a plasmid expressing a FimA molecule, in which the CWSS is replaced by a six His-tag (black circle). **(E)** His-tagged FimA proteins were purified from the culture medium of indicated strains by affinity chromatography, separated by SDS-PAGE and stained by Coomassie Blue. **(F)** The purified proteins were subjected to N-terminal sequencing. Underlines indicate the sequence resulted from first 10 sequencing cycles. The *lepB2* mutant also produced ragged N-terminal sequencing with the major sequence shown.

some trace of PAS staining was detected with the smeary band (Fig. 3-4B; bracket 1). The same phenotypes were observed in the  $\Delta lepB1$  mutant; intriguingly, in the  $\Delta lepB2$  mutant a smeary band migrating around 37-kDa was positive for PAS staining (Fig. 3-4B; bracket 2), while the high molecular weight (HMW) band stained positive for PAS remained weakly visible (Fig. 3-4B; bracket 1). Finally, in the double  $lepB1/B2$  mutant a further-upshifted band with positive PAS staining was detected. In addition, the HMW species of GspA showed a slight increase in PAS staining (Fig. 3-4B; bracket 1, last lane). These results indicate that glycosylation of GspA still occurs, albeit weakly, in the absence of the CWSS. Together with the data presented in Fig. 3-3A, we conclude that neither the deletion of  $lepB1$  nor  $lepB2$  affects GspA glycosylation.

If either SPase is sufficient to process GspA, the cleaved products in each deletion mutant should contain the same N-terminal sequence. To determine if this is the case, we excised the major GspA bands indicated in Fig. 3-4B for Edman sequencing. As shown in Table 3-1, the first ten cycles of Edman degradation for GspA samples purified from the WT,  $\Delta lepB1$ , and  $\Delta lepB2$  strains revealed the matching sequence of GDSLAFKIAD (Fig. 3-4C), consistent with the predicted cleavage site between Ala (-1 position) and Gly (+1 position) (Fig. 3-4A). In contrast, the same analysis of GspA samples isolated from the  $\Delta lepB1/B2$  mutant produced a mixture of residues in the majority of the first 10 Edman sequencing cycles (Table 3-1) indicative of multiple polypeptides present in these samples; with the major polypeptide that has the sequence of LAGDSLAFKI (Fig. 3-4C). The results indicated that both LepB1 and LepB2 are capable of cleaving the GspA signal peptide. In the absence of both SPases, the GspA signal peptide might be proteolytically degraded at different residues by unidentified protease(s).



**Table 3-1: Protein sequencing of GspA proteins purified from four *A. oris* strains by Edman degradation**

Sequencing Cycle	Residue (pmol)			
	MG1	$\Delta lepB1$	$\Delta lepB2$	$\Delta lepB1/\Delta lepB2^a$
1	G (4.761)	G (13.250)	G (7.833)	L (3.150); A (3.665)
2	D (4.438)	D (13.830)	D (8.280)	A (3.887)
3	S (2.529)	S (5.692)	S (4.623)	G (2.938); P (1.390)
4	L (3.972)	L (10.650)	L (6.026)	D (2.962); A (2.836)
5	A (4.082)	A (11.050)	A (6.184)	S (1.774)
6	F (2.913)	F (8.309)	F (5.394)	L (3.162)
7	K (3.882)	K (10.280)	K (5.761)	A (4.124); G (2.501)
8	I (2.661)	I (8.125)	I (4.958)	F (2.520); D (2.494)
9	A (3.427)	A (9.922)	A (5.492)	K (2.520); S (1.729)
10	D (2.794)	D (7.996)	D (5.265)	I (1.875); L (2.144)

<sup>a</sup> Two major residues were detected in the majority of the first 10 sequencing cycles.

**LepB2 is required for pilus assembly** – As previously mentioned, in addition to GspA and 13 other cell wall anchored proteins (25), the *A. oris* MG1 strain expresses 4 pilus proteins FimP/Q and FimA/B, which constitute type 1 and type 2 fimbriae, respectively (18,79,84). We asked if the function of the two SPases is extended to these LPXTG-containing pilus proteins. Using the same approach described for GspA, whereby the CWSS of FimA was replaced by a His-tag (Fig. 3-4D), we purified secreted FimA proteins from the culture medium and analyzed them by SDS-PAGE and Edman degradation. While the FimA proteins isolated from the WT and  $\Delta lepB1$  strains migrated with the same mobility by SDS-PAGE, the FimA proteins isolated from the  $\Delta lepB2$  mutant migrated slightly slower as compared to the first two (Fig. 3-4E). Consistently, Edman degradation revealed the same sequence of TETPNYGNIK in the first two samples, supportive of the cleavage site between Ala (-1) and Thr (+1) (Fig. 3-4D), whereas the last one had ragged sequences, with the major species of AVATETPNYG (Fig. 3-4F and Table 3-2). The results indicate that LepB1 is not required for proper cleavage of the FimA signal peptide and that LepB2 is essential for FimA pre-protein processing.

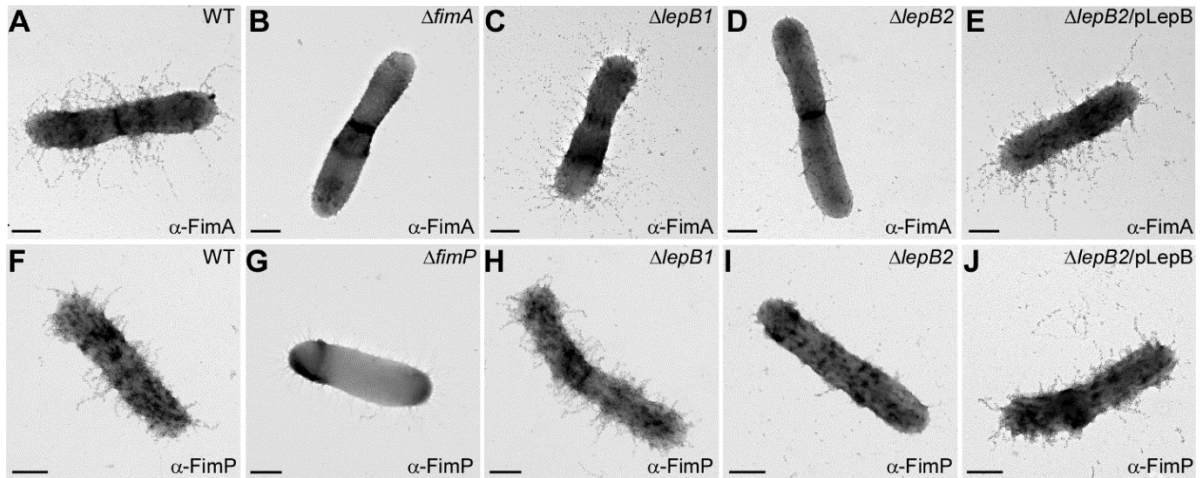
**Table 3-2: Protein sequencing of FimA proteins purified from four *A. oris* strains by Edman degradation**

Sequencing Cycle	Residue (pmol)		
	MG1	$\Delta lepB1$	$\Delta lepB2^a$
1	T (8.363)	T (2.341); G (1.394)	A (10.082); T (8.999); G (70.840);
2	E (6.034)	E (1.391)	V (6.512); E (6.271); L (6.009)
3	T (7.885)	T (2.137)	A (10.563); T (9.934)
4	P (6.338)	P (1.786)	T (10.321); P (8.828)
5	N (4.784)	N (1.467)	E (5.881); N (4.545)
6	Y (5.789)	Y (1.512)	T (11.027); A (9.149); Y (5.218)
7	G (5.246)	G (2.061)	P (7.555); G (11.852); V (6.863)
8	N (4.597)	N (1.412)	N (5.679); A (9.687)
9	I (4.777)	I (1.771)	I (6.770); Y (4.509)
10	K (5.922)	K (1.133)	K (6.659); G (12.054); E (5.618)

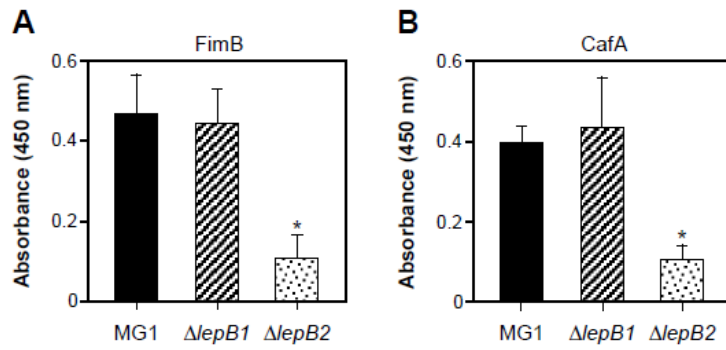
<sup>a</sup> Multiple residues were detected in each of the first 10 sequencing cycles.

We next investigated the impact of LepB2-mediated cleavage of FimA precursors by examining pilus assembly on the bacterial cell surface by IEM. *A. oris* cells were immobilized on nickel grids, washed and stained with antibodies against FimA ( $\alpha$ -FimA) and then reacted with secondary IgG antibodies conjugated to gold particles. Samples were then stained with 1% uranyl acetate and viewed by an electron microscope. As expected, FimA labeled pili were abundant on the surface of wild-type cells, but absent in a mutant lacking *fimA* (Fig. 3-5A & 3-5B). Note that visible unstained pili in the  $\Delta$ *fimA* mutant were FimP pili (79). While the deletion of *lepB1* did not affect FimA pilus assembly, deletion of *lepB2* severely affected assembly (compare Fig. 3-5C & 3-5D). Ectopic expression of LepB2 in the *lepB2* mutant rescued this assembly defect to the wild-type levels (Fig. 3-5E). To examine if LepB2 is required for the assembly process of tip fimbrial proteins, we quantified the protein level of FimB and CafA on the bacterial surface by whole-cell ELISA. Consistent with the above results, the  $\Delta$ *lepB2* mutant displayed drastic reductions in FimB and CafA signal the  $\Delta$ *lepB1* mutant exhibited wild-type levels of both pilus tip proteins (Fig. 3-6).

To determine if LepB2 also acts on type 1 fimbriae, we analyzed pilus assembly using IEM with antibodies against the fimbrial shaft protein FimP ( $\alpha$ -FimP). As shown in Fig. 3-5F-J, the FimP phenotypes mirrored those of FimA presented above. Altogether, the results indicated that LepB2 is specifically utilized for processing of pilus precursors in *A. oris*.

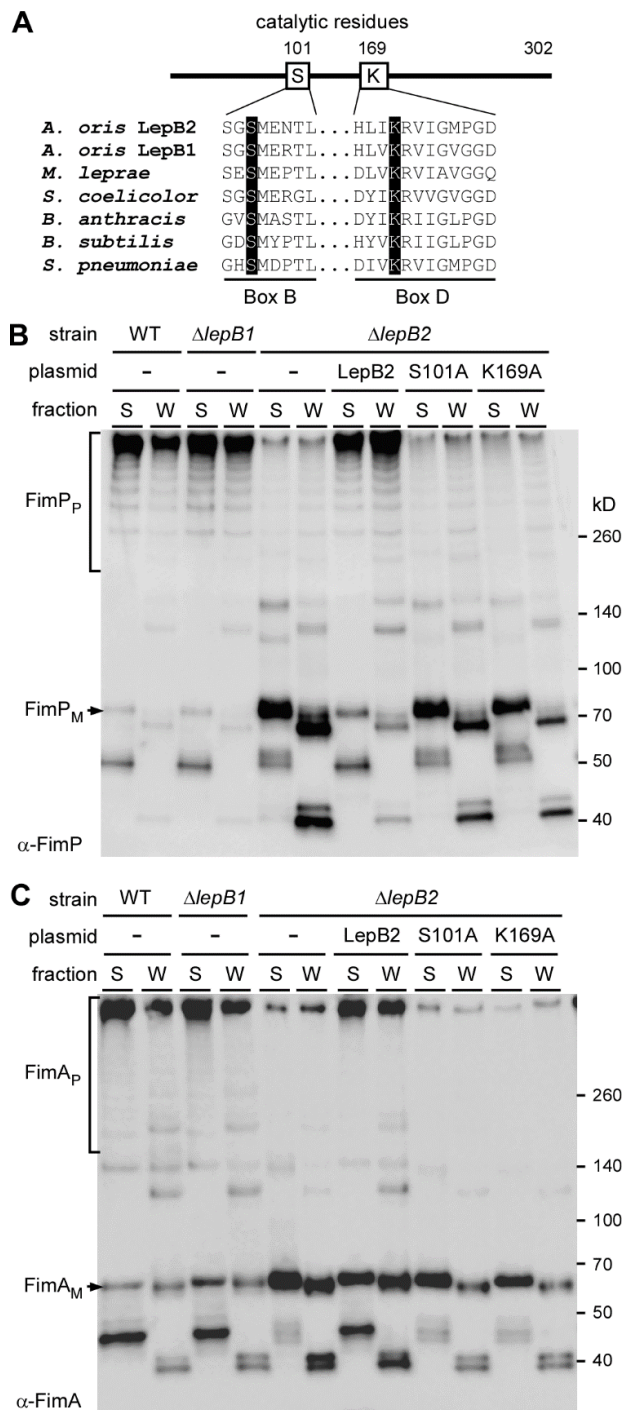


**Figure 3-5: The signal peptidase LepB2 is required for pilus assembly.** *A. oris* cells were immobilized on nickel grids, reacted with the specific antiserum against the type 2 major pilin subunit, FimA ( $\alpha$ -FimA) (A – E) or the type 1 major pilin subunit FimP ( $\alpha$ -FimP) (F – J) followed by goat anti-rabbit IgG conjugated to 12-nm gold particles, and stained with 1.0% uranyl acetate. Samples were viewed by transmission electron microscopy. Scale bars, 0.2  $\mu$ m.



**Figure 3-6: Requirement of LepB2 for surface expression of tip pilins.** The presence of FimB (**A**) and CafA (**B**) on the bacterial cell surface was analyzed by whole cell ELISA with specific antibodies to FimB and CafA, respectively. The absorbance values, as compared to those of corresponding mutants as background, were determined from at least two independent experiments performed in triplicate. Error bars represent standard deviations. Asterisks (\*) indicate P-values of 0.02 (**A**) and 0.03 (**B**); all were determined using the paired, two-tailed t-test with Prism GraphPad.

**Requirement of the conserved catalytic dyad for LepB2 activity** – Bacterial type 1 SPases utilize a conserved Ser-Lys catalytic dyad present within the conserved Box B and Box D, respectively, for their proteolytic activities (60). Our sequence alignment of LepB1 and LepB2 with many other Gram-positive signal peptidases revealed the conservation of this catalytic dyad (Fig. 3-7A). In LepB2, Ser is located at position 101 and Lys at 169. To determine if these residues are indeed required for LepB2 activity, we generated recombinant plasmids expressing LepB2 with alanine substitutions at S101 or K169. The resulting plasmids were introduced into the  $\Delta lepB2$  mutant and the effects of these mutations on pilus assembly were analyzed by immunoblotting and EM. By Western blot analysis, we observed high molecular weight species of FimP and FimA indicative of pilus polymers (denoted as P) in the culture medium (S) and cell wall (W) fractions of wild-type cells (Fig. 3-7B & 3-7C; first two lanes), as previously reported (84). While deletion of *lepB1* did not affect pilus assembly (Fig. 3-7B & 3-7C; next two lanes), deletion of *lepB2* greatly reduced pilus polymers and increased the accumulation of LMW products, presumably degradation products, in the membrane as well as their secretion into the culture medium (Fig. 3-7B & 3-7C; lanes  $\Delta lepB2$ ). The defects of this mutant were rescued by ectopic expression of *lepB2* (Fig. 3-7B & 3-7C; lanes  $\Delta lepB2/pLepB2$ ). As expected, alanine substitutions of S101 and K169 resulted in the same defects as deletion of *lepB2* (Fig. 3-7B & 3-7C; last 4 lanes). The effects of the catalytic dyad mutations on pilus assembly were also confirmed by negative-staining EM (data not shown). Thus, LepB2 contains a canonical catalytic dyad present in type I signal peptidase enzymes, and this dyad is critical for LepB2 activity.



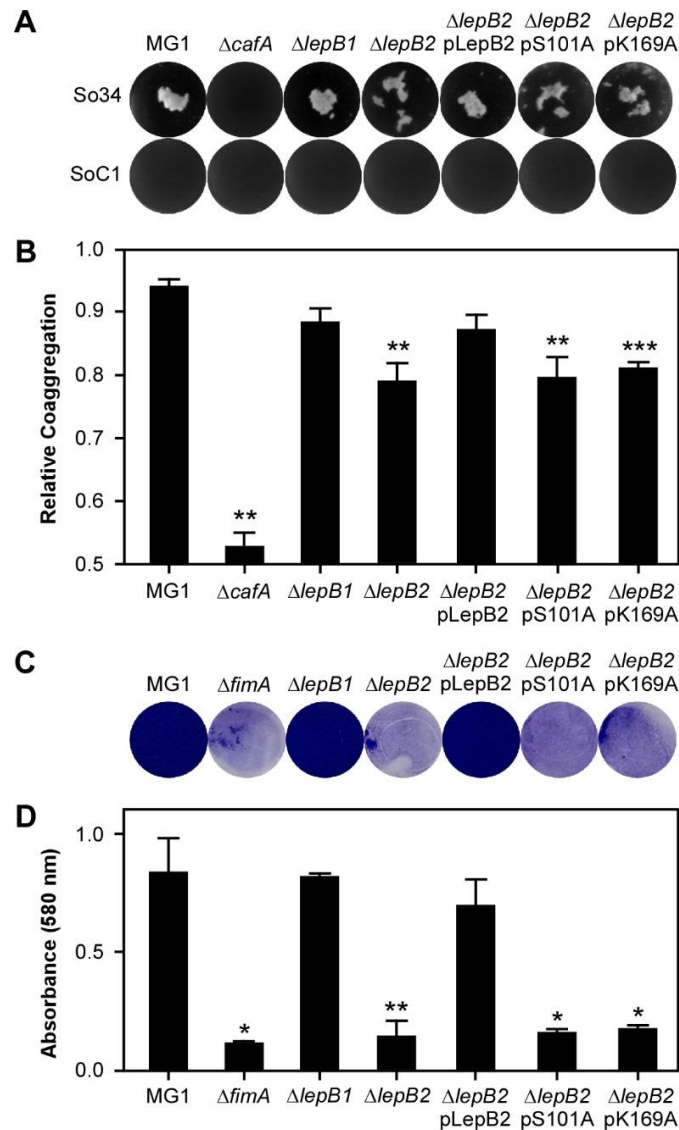
**Figure 3-7: The Ser-Lys catalytic dyad of LepB2 is necessary for pilus assembly.**

(A) Multiple sequence alignment of type 1 SPases from *A. oris* (LepB1 and LepB2), *Mycobacterium leprae* (LepB), *Streptomyces coelicolor* (Sip1), *Bacillus anthracis* (SipS), *Bacillus subtilis* (SipS), and *Streptococcus pneumoniae* (Spi) was performed using



CLUSTAL W (112). The conserved Box B and Box D, which contain the catalytic Ser and Lys residues (highlighted in black), respectively, are shown. Numbers indicate Ser and Lys positions in *A. oris* LepB2. **(B)** Supernatant (S) and cell wall (W) fractions were collected from MG1 and its isogenic derivatives. Equivalent protein samples were subjected to immunoblotting with  $\alpha$ -FimP. **(C)** The same samples in B were immunoblotted with  $\alpha$ -FimA. The positions of fimbrial monomer (M), HMW polymers (P), and molecular mass markers are indicated.

Because LepB2 is required for pilus assembly and type 2 fimbriae are involved in polymicrobial interactions (or coaggregation) and biofilm formation, we then examined if LepB2 is also important for these processes. To test for polymicrobial interactions, *Actinomyces* cells were mixed with *Streptococcus oralis* in equal numbers and coaggregation was determined both visually (113) and quantitatively as previously reported (93). As shown in Fig. 3-8A & 3-8B, *Actinomyces* coaggregation with *S. oralis* was dependent on CafA as deletion of *cafA* abrogated this interaction, consistent with our previous report (25). Compared to the parental MG1 strain, the *lepB1* mutant did not display any noticeable defect in coaggregation. In contrast, deletion of *lepB2* significantly reduced bacterial coaggregation; this defect was rescued by overexpressing wild-type LepB2 from a plasmid, but not from the catalytically inactive LepB2 variants, i.e. S101A or K169A. Finally, the ability of the LepB2 mutants to form biofilms was evaluated using an established protocol (93), whereby *Actinomyces* biofilms were cultivated in the presence of 1% sucrose and quantified by staining with crystal violet (see Methods). Unlike the *lepB1* mutant, which produced biofilms at the wild-type level, the *lepB2* mutant was unable to form biofilms. The biofilm defect of the *lepB2* mutant was restored when wild-type LepB2 was expressed ectopically, but not with the S101A or K169A mutant (Fig. 3-8C and 3-8D). Altogether, the results support the notion that LepB2 is the signal peptidase specific for LPXTG-containing pilus proteins in *A. oris*.



**Figure 3-8: Requirement of LepB2 for polymicrobial interactions and biofilm formation.** The parental *A. oris* MG1 strain and its variants were examined for their ability to interact with *Streptococcus oralis*. Coaggregation was scored visually (**A**) or quantitatively by optical density (**B**). To cultivate biofilms, *Actinomyces* cells were grown in microtiter plates in the presence of 1% sucrose at 37°C with 5% CO<sub>2</sub>. Generated biofilms were stained with crystal violet (**C**) and subsequently quantified by measuring absorbance at 580 nm (**D**). The values are expressed as averages of three independent experiments performed in triplicate. Symbols \*, \*\*, and \*\*\* indicate p values of 0.05, 0.01, and 0.005, respectively; all determined using the paired, two-tailed t-test with Prism GraphPad.

## Discussion

We recently reported that the signal peptidase LepB2 is a genetic suppressor of *srtA* essentiality in *A. oris*, a lethal phenotype associated with glyco-stress caused by accumulation of the glycosylated protein GspA in the cytoplasmic membrane when SrtA is disabled (50). The suppressor was found when we recovered a viable *srtA*-deleted mutant with the Tn5 transposon inserted into the *lepB2* gene (50). In this report, we confirmed the suppression phenotype of *lepB2* by deleting chromosomal *srtA* in a strain already devoid of *lepB2* (Fig. 3-1). It is important to note that deletion of *srtA* in the absence of *gspA* is also non-lethal (50) suggesting that LepB2 acts on the GspA glycosylation pathway. Indeed, in the  $\Delta lepB2/\Delta srtA$  double mutant, GspA glycosylation was severely defective (Fig. 3-3) unlike the phenotype of *srtA* depletion, which does not affect GspA glycosylation (50). *lepB2* is part of a gene locus that encodes another signal peptidase gene, *lepB1* (Fig. 3-1A). This raises a question of redundancy as multiple copies of SPases are typically present in a single species of Gram-positive bacteria (60).

To address this question, we mapped the cleavage site of the GspA signal peptide by Edman degradation. Because GspA is heavily glycosylated leading to a smeared migration pattern on SDS-PAGE (Fig. 3-3A), detecting the cleavage of the GspA signal peptide was challenging. To circumvent this problem, we constructed a GspA molecule lacking its CWSS (denoted as GspA $\Delta_{CWSS}$ ), which caused it to be secreted into the extracellular milieu with less chance of glycosylation. To our surprise, the GspA $\Delta_{CWSS}$  precursor was properly processed when expressed in the absence of LepB1 or LepB2 (Fig. 3-4B & 3-4C). Especially when considering that LepB2 is linked to GspA glycosylation. The results support the idea that both SPases are capable of cleaving the GspA signal peptide. This is consistent with the observation that in the absence of both *lepB1* and *lepB2* the GspA signal peptide was not properly processed

(Fig. 3-4C). It is intriguing, however, that the glycosylation level of GspA detected in the cell wall fraction of this double mutant was comparable to that of the parental and single deletion strains (Fig. 3-3A). This seems contradictory to a general belief that processing of the signal peptide is a prerequisite for protein maturation (40). We speculate that GspA glycosylation occurs immediately as the protein precursors emerge from the Sec translocon. This is in line with the “Pilusosome Hypothesis” that sortase and its substrates are in close proximity with a protein secretion machine for efficient assembly (114). It is also possible that in the absence of the two SPases aberrant cleavage of the protein precursors, probably by membrane-bound protease(s), is sufficient to release the polypeptides from the secretion machine for sortase and glycosyltransferases to be able to perform their functions.

While both LepB1 and LepB2 are capable of processing GspA, we demonstrated that only LepB2 is specific for cleavage of fimbrillin signal peptides. Using the major fimbrillin shaft FimA as an experimental model, we showed by N-terminal sequencing that cleavage of the FimA signal peptide depends on LepB2 (Fig. 3-4 D-F). In the absence of the cognate signal peptidase LepB2, the FimA signal peptide was proteolytically cleaved producing ragged polypeptides (Table 3-2). Unlike GspA, this failure of signal peptide cleavage severely affected fimbrial assembly (Fig. 3-5). Consequently, the ability of this mutant strain to interact with oral streptococci and to form biofilm is significantly hindered (Fig. 3-8). LepB2 function is not limited to type 2 fimbriae, as the *lepB2* mutant also fails to assemble type 1 fimbriae with FimP as the fimbrial shaft; in contrast, LepB1 is dispensable for these processes (Fig. 3-5, 3-6 and 3-7).

The notion that a signal peptidase is involved in pilus formation has been previously reported with the signal peptidase-like protein SipA in the Gram-positive

pathogen *Streptococcus pyogenes* (115). SipA is expressed by the gene locus that encodes the T3 pilin components in many strains of *S. pyogenes* (115). Although it has the same core fold as the type I SPase of *Escherichia coli* (116), it lacks the catalytic Ser and Lys residues typical of type I SPases (117) (Fig. 3-7A). In place of Ser and Lys are Asp and Gly, but they do not play any role in signal peptidase activity nor cleavage of T3 pre-pilins (117). Unlike SipA, LepB2 contains a canonical catalytic dyad with the conserved S101 and K169 (Fig 3-7A). Alanine substitution of these residues abrogates pilus assembly as well as biofilm formation (Fig. 3-7 and 3-8). Altogether, it is clear that LepB2 is a type I SPase, which may serve as the prototype of type I SPases specific for pilus assembly in Gram-positive bacteria.

**CHAPTER IV**  
**A Phosphotransferase LCP Enzyme Mediates Glycosylation of a**  
**Gram-positive Cell Wall Anchored Protein**

## Introduction

Glycopolymers such as wall teichoic acids (WTAs) displayed on the cell envelope of Gram-positive bacteria play critical roles in cell physiology by modulating immunogenicity, host and bacterial surface interactions, protein stability, cell division and affinity for charged molecules including antimicrobial peptides and cations (38,39,118,119). WTAs are polyol repeats ending in a disaccharide linkage unit attached to the C6 hydroxyl group of N-acetyl muramic acid (MurNAc) of bacterial peptidoglycan via a phosphodiester bond (120). Attachment of WTAs to the anchor molecules of the cell envelope requires LytR-CpsA-Psr (LCP) family enzymes widespread in Gram-positive bacteria (121). The first crystal structure of an LCP enzyme was CpsA2 from *Streptococcus pneumoniae*. The enzyme contains a hydrophobic tunnel capped with surface exposed catalytic arginine residues, and of these both features that are essential for functionality (44). Serendipitously, the CpsA2 structure co-crystallized with octaprenylpyrophosphate (oprPP), where the isoprenyl tail is nestled within the hydrophobic pocket and the pyrophosphate head group interacted with the surface exposed arginine residues. In the case of WTAs, LCP proteins catalyze the formation of a phosphodiester bond to link the glycan to the MurNAc of the cell wall. A pyrophosphatase reaction removes the glycan from the lipid donor molecule (122). It has been previously demonstrated that *S. pneumoniae* CpsA2 and *Corynebacterium glutamicum* LcpA possess *in vitro* pyrophosphatase activity and this is likely a characteristic of most LCP enzymes that mediate phosphotransfer (44,47,97).

An LCP protein has also been identified in the Gram-positive actinobacterium *Actinomyces oris*, a key colonizer of the oral cavity that plays an important role in the development of oral biofilms or dental plaque (2). The identification of *A. oris* LcpA that is linked to the glycosylation of the cell wall anchored protein GspA was revealed by an



unbiased transposon mutagenesis screen. Deletions of both *lcpA* and *gspA* were shown to suppress the essentiality of the housekeeping sortase *srtA*. Additionally, the adjacent encoding of *lcpA* and *gspA* genes in *A. oris* indicates that their protein products are functionally linked (50). GspA harbors a typical C-terminal cell wall sorting signal (CWS), which is recognized by sortase enzymes for covalent attachment to peptidoglycan via an LPXTG motif within the CWSS (51). Biochemical evidence indicates that GspA is highly glycosylated and this glycosylation requires LcpA. A mutant strain lacking *lcpA* no longer produces high molecular mass glycopolymers of GspA and concomitantly accumulated intermediate forms (50). A model for *srtA* essentiality involving both GspA and LcpA has been proposed: as GspA is translocated across the cytoplasmic membrane by the Sec machine, it is glycosylated by LcpA with the glycan chain synthesized by a separate pathway and subsequently anchored to the cell wall by the housekeeping sortase SrtA (50). In the absence of *srtA*, glycosylated GspA accumulates in the membrane leading to toxic glycol-stress. Consistent with this model, genetic disruption of *srtA* in the absence of *lcpA*, *gspA*, or a GspA mutant devoid of the membrane anchored CWSS results in viable cell types (50).

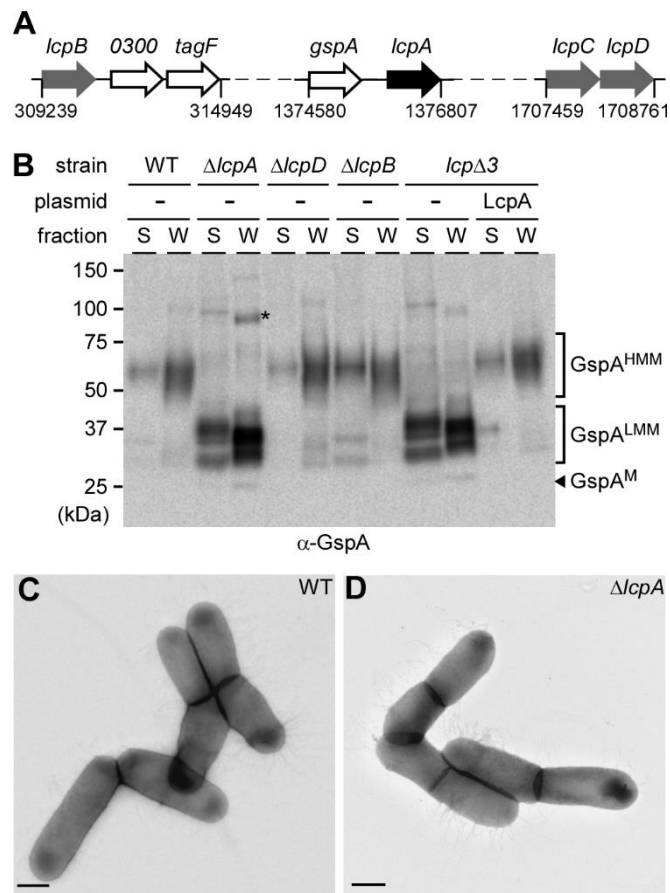
While the exact nature and composition of the GspA glycans remain to be biochemically determined, *A. oris* LcpA represents the first example of an LCP enzyme that modifies a cell wall anchored protein substrate. Here, we present a high resolution crystal structure of *A. oris* LcpA revealing conserved features of known LCP enzymes and unique characteristics that may be typical of actinobacterial LCP proteins. Further biochemical characterization provides evidence that not only does LcpA possess pyrophosphatase activity but it also functions as a phosphotransferase that catalyzes glycosylation of the cell wall anchored protein GspA.

## Results

**LcpA is required for GspA glycosylation.** As previously mentioned, a screen for sortase SrtA suppressors identified an LCP homolog (*ana\_1292*) hereafter named *lcpA* (50), which is located immediately downstream of *gspA* (Fig. 4-1A), which is another suppressor of *srtA* lethality (50). In addition to LcpA, *A. oris* MG1 encodes three additional proteins with LCP domains. *ana\_0299*, hereafter called *lcpB*, is adjacent to two conserved genes (Fig. 4-1A) coding for a UDP-N-acetyl-D-mannosaminuronic acid dehydrogenase (*ana\_0300*) and a homolog of glycosyl/glycerophosphate transferase TagF, which has previously been implicated in the wall teichoic acid (WTA) synthesis of *Staphylococcus epidermidis* (123). *ana\_1577* (*lcpC*) and *ana\_1578* (*lcpD*) appear to reside in the same transcriptional unit (Fig. 4-1A). Because LcpA has been linked to GspA glycosylation (123), we examined if genetic disruption of LcpB, LcpC, and LcpD affects this process, although all three were not identified from the original suppressor screen. We obtained mutations in *lcpB* and *lcpD*, but we were unable to generate a deletion mutation of the *lcpC* gene after several attempts suggesting *lcpC* is an essential gene. A triple mutation (*lcpΔ3*) of *lcpA*, *lcpB*, and *lcpD* was also obtained.

To analyze LcpA-mediated glycosylation, cell cultures of *A. oris* MG1 and these mutant strains were grown to mid-log phase, normalized by optical density and subjected to cell fractionation, as previous described (50). Protein samples from the culture medium (S) and cell wall (W) fractions were analyzed by western blotting with a specific antibody against GspA ( $\alpha$ -GspA). As reported before (123), the MG1 strain (WT) produced a high molecular-mass species of GspA with glycan polymers, i.e. GspA<sup>HMM</sup>, detected mostly in the cell wall fractions (Fig. 4-1B, lanes WT). Deletion of *lcpA* abrogated formation of GspA<sup>HMM</sup> resulting in accumulation of a low molecular weight species of GspA termed GspA<sup>LMM</sup>. GspA<sup>LMM</sup> migrates near the 37-kDa marker, although the GspA monomer (GspA<sup>M</sup>; arrowhead) migrates near the 25-kDa marker (Fig. 4-1B,

lanes  $\Delta lcpA$ ). The single mutant strains  $\Delta lcpB$  and  $\Delta lcpD$  displayed no significant defects in formation of GspA<sup>HMM</sup> (Fig. 4-1B, lanes  $\Delta lcpB$  and  $\Delta lcpD$ ), whereas the triple mutant  $lcp\Delta3$  failed to produce GspA<sup>HMM</sup> phenocopying the  $lcpA$  mutant; this defect was rescued by ectopic expression of  $lcpA$  in the  $lcp\Delta3$  mutant (Fig. 4-1B, last 4 lanes). To determine if deletion of  $lcpA$  affects cell morphology and pilus assembly, the parental and  $lcpA$  mutant strains were examined by electron microscopy, whereby bacterial cells, immobilized on carbon-coated nickel grids, were stained with 1% uranyl acetate prior to viewing with an electron microscope. As shown in Fig. 4-1C-D, both strains displayed similar cell morphology and pilus assembly phenotypes. Altogether, the results support that LcpA is necessary and sufficient for the production of GspA<sup>HMM</sup> and suggest that GspA<sup>LMM</sup> might represent an intermediate form of the glycoprotein GspA<sup>HMM</sup>.



**Figure 4-1: LcpA is solely responsible for GspA glycosylation.** (A) Presented are gene clusters that encode four LCP proteins (LcpA-D) with numbers indicating the nucleotide positions of *lcp* genes. (B) *A. oris* cells of indicated strains grown to early log phase were subjected to cell fractionation. Culture medium (S) and cell wall (W) fractions were analyzed by immunoblotting with specific antibodies against GspA. High molecular mass (HMM) and low molecular mass (LMM) species of GspA, GspA monomer (M), and molecular mass markers are indicated. (C & D) *A. oris* cells were immobilized on nickel grids, stained with 1% uranyl acetate prior to viewing with an electron microscope. Scale bar represents 0.5  $\mu$ m.

***(Note: Portions of the writing in the following section is credited to Brendan Amer, who performed the X-ray crystallography, and was used with permission.)***

**X-ray structure of LcpA from *Actinomyces oris*.** To obtain insight into how LcpA glycosylates GspA we first determined the molecular structure of the LcpA enzyme. An inspection of its primary sequence reveals a proposed tripartite structure: (i) residues 1-54 presumably reside in the cytoplasm and are predicated to adopt helical secondary structure, (ii) residues 55-77 are non-polar and likely form a single transmembrane helix (TM), and (iii) residues 78-370 presumably reside on the extracellular surface and share primary sequence homology to LCP-type enzymes (Pfam family PF03816). The structure of the extracellular LCP domain (rLcpA, residues 78-370) was solved at 2.5-Å resolution. Electron density was observed for residues 79-106 and 126-368, which form a single domain that adopts an  $\alpha$ - $\beta$ - $\alpha$  architecture. A seven-stranded anti-parallel  $\beta$ -sheet forms the core of the protein with a total of eight  $\alpha$ -helices flanking the  $\beta$ -sheet on both of its faces forming a hydrophobic tunnel (Fig. 4-2A). The tunnel is ~23 Å in length and is lined by residues located on the central  $\beta$ -sheet, helices H5, H6 and H7. The tunnel varies in width from ~6 to 14 Å and is widest in the core of the protein. The surface of the tunnel contains many non-polar residues consistent with it interacting with lipid substrates. Interestingly, during refinement, an additional electron density was observed near the exit point of the tunnel defined by helices H6 and H7 indicating that a ligand was bound. However, it was not possible to conclusively define the identity of this ligand using MALDI-TOF mass spectrometry and modeling the ligand as a phosphate-isoprenoid molecule or other membrane-associated lipid yielded poor refinement statistics. The best match to the data was obtained by modeling the ligand as a PEG-4000 molecule that was used as a precipitant during crystallization. This ligand is bound with 50% occupancy and defines the exit point for the tunnel distal to the active site. The presence of a hydrophobic tunnel leading into the active site suggests that LcpA

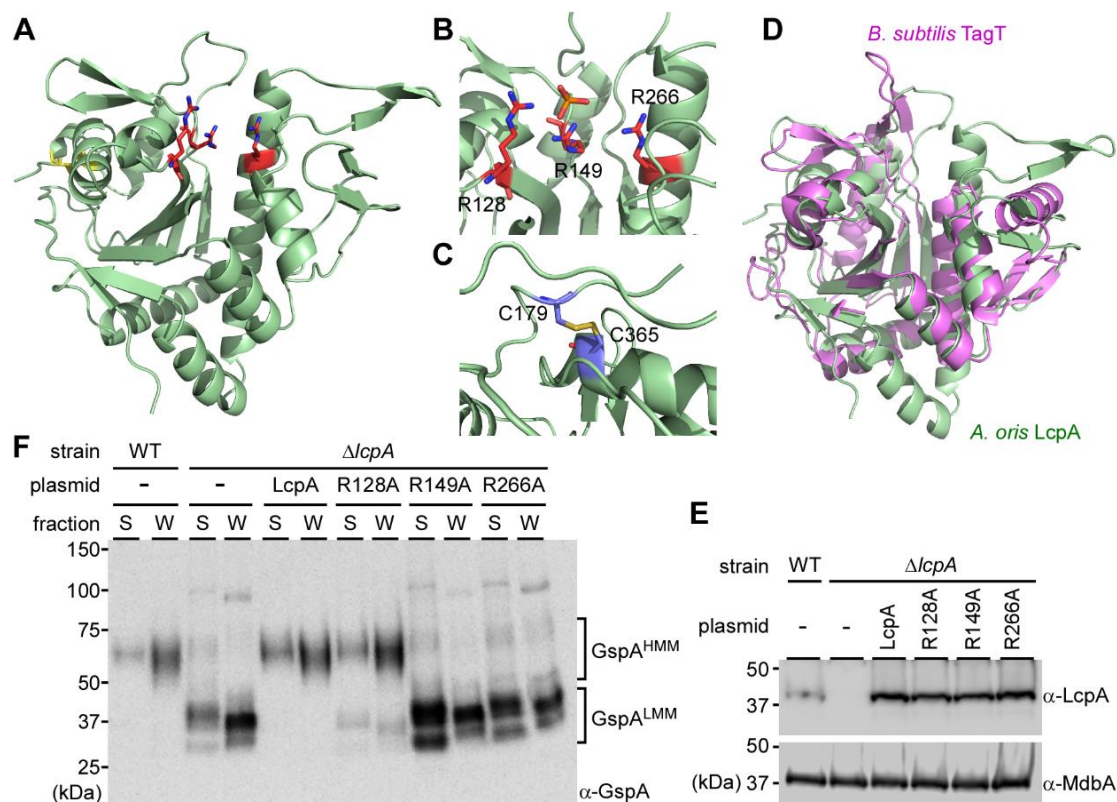
could bind a lipid-linked glycan donor substrate similar to other members of the LCP superfamily (44).

Members of the LCP superfamily contain conserved arginine residues, which are thought to mediate a phosphotransfer reaction that attaches glycopolymers to acceptors (122). In *A. oris* LcpA, R128, R149, and R266 are conserved residues that cluster together within a surface exposed pocket (Fig 4-2A and Fig. 4-2B, shown in red). One surface of this exposed active site is formed by residues in strand  $\beta$ 3 and helix H1, while the top and side of the pocket is formed by helix H4 and H5, respectively, packing against the core  $\beta$ -sheet. R128 and R149 in the pocket are positioned towards the surface and located in strands B  $\beta$ 3 and  $\beta$ 4/  $\beta$ 5 loop, respectively. Helix H5 spans the length of the protein and contains the third conserved active arginine (R266), which is located closer to the body of the enzyme where the pocket narrows. Electron density is observed between the guanidino sidechains of R128 and R149 and the modeled phosphate atom. The hydrophobic tunnel leads from this conserved site to the opposite face of the protein structure.

Intriguingly, unlike other LCP enzymes, LcpA contains a disulfide bond formed between residues C179 and C365 linking the C-terminus to  $\alpha$ -helix H2 (Fig. 4-2C). This disulfide is presumably stabilizing, since it persists despite the presence of a reducing agent in the protein buffer used in the final purification step. The cysteine residues are also conserved in other LCP homologs present in Actinobacteria.

The structure of rLcpA is similar to that of previously reported LCP enzymes that attach polymers to the cell wall and is most closely related to the YwtF (TagT) enzyme from *Bacillus subtilis* based on a DALI analysis with a Z-score of 21.8 (PDB: 4DE9 (46)); the backbone atoms can be superimposed with a root-mean-square deviation (RMSD) of 2.5Å (Fig. 4-2D). The structural conservation and presence of arginine residues in the

surface exposed pocket prompted us to investigate functional similarities to TagT related to GspA glycosylation.



**Figure 4-2: Crystal structure of *A. oris* LcpA and structural requirements for glycosylation activity.** (A) The structure of the extracellular LCP domain (residues 78-370) was determined to 2.5-Å resolution. The proposed catalytic Arg residues are shown as sticks and colored red, and the cysteine residues participating in the disulfide bond are shown in yellow. (B) Presented is a detailed view of the LcpA active site with the conserved catalytic arginine residues (R128, R149, and R266) shown in red. (C) Shown is a close-up view of the disulfide bond that links the C-terminus via C365 to the second  $\alpha$ -helix via C179 present in the LCP extracellular domain. (D) The LcpA structure (light green) is superimposed with the *Bacillus subtilis* YwtF (TagT) (PBD: 4DE9) (magenta). (E) Protein samples of indicated strains were prepared as described in Fig. 1B and analyzed by immunoblotting with anti-GspA. (F) Protein samples from the membrane fractions in (E) were immunoblotted with antibodies against LcpA. A membrane protein, MdbA, was used as a control. Molecular mass markers in kDa are shown. **Note:** *Crystallization, structural determination and modeling were performed by Brendan Amer and Jason Gosschalk and used with permission.*



**The conserved arginine residues in LcpA are required for glycosylation activity.**

As presented above, LcpA is required for glycosylation of GspA (Fig. 4-1B) and LcpA contains conserved arginine residues (R128, R149 and R266) (Fig. 4-2B). Conserved Arg residues have been implicated in LCP activity by interacting with the pyrophosphate of the lipid-linked glycan donor (44). To determine whether these Arg residues affect the glycosylation activity of *A. oris* LcpA, we generated alanine-substitution mutants of these arginine residues using pLcpA as a template (Fig. 4-2E). Plasmids expressing mutant proteins were introduced into the  $\Delta$ *lcpA* mutant and expression of LcpA proteins was determined by immunoblotting membrane lysates of various strains with specific antibodies against LcpA ( $\alpha$ -LcpA) or with  $\alpha$ -MdbA, to detect MdbA as a control for the membrane bound protein (62). As expected, LcpA was detected in the parental strain and absent from the  $\Delta$ *lcpA* mutant (Fig. 4-2E, first two lanes). Complementation of the  $\Delta$ *lcpA* mutant with a multi-copy plasmid enhanced LcpA production as compared to the WT strain (Fig. 4-2E, lane LcpA). Mutations of the three Arg residues did not affect the stability of mutant proteins as compared to ectopically expressed wild-type LcpA (Fig. 4-2E, last 3 lanes). We then examined GspA glycosylation by Western blotting the supernatant and cell wall fractions as described in Fig. 4-1B. Interestingly, the LcpA-R128A mutant was able to produce GspA<sup>HMM</sup> at the level comparable to that of the WT strain and the rescued strain  $\Delta$ *lcpA*/LcpA, whereas the LcpA-R149A and LcpA-R266A mutants were defective in glycosylation of GspA phenocopying the  $\Delta$ *lcpA* mutant (Fig. 4-2F). Altogether, the results support that the R149 and R266 residues are essential for the glycosylation activity of LcpA.

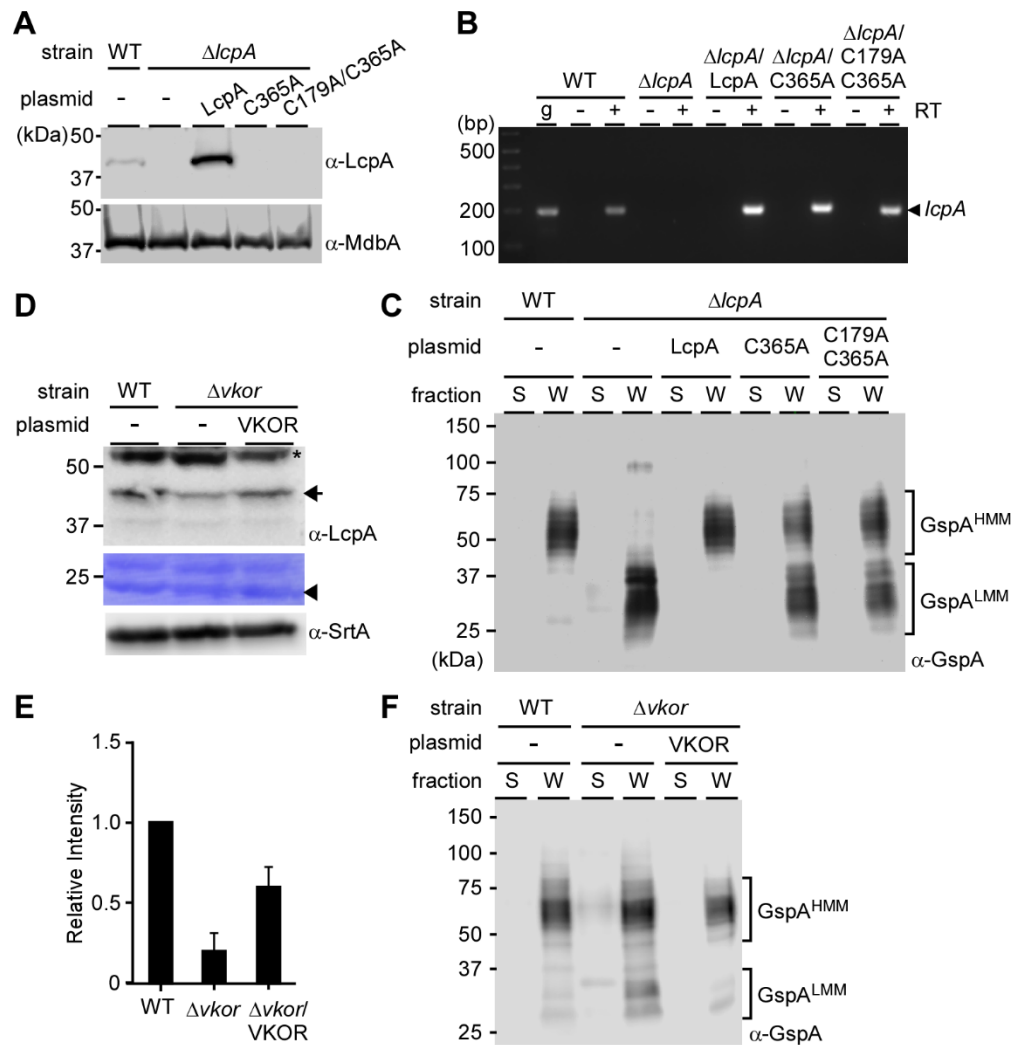
**The disulfide bond in *A. oris* LcpA is required for protein stability.** *A. oris* LcpA has a stable disulfide bond (Fig. 4-2C), and it appears that the disulfide linkage is a conserved feature in Actinobacterial LCP proteins. To determine the role of the disulfide bond in LcpA stability and glycosylation activity, we generated alanine-substitution

mutants of either one (C365) or both Cys residues (C179 and C365). Membrane fractions of the parental and mutant strains were immunoblotted with  $\alpha$ -LcpA as described previously. Enhanced signal of LcpA was observed in a strain expressing LcpA from a plasmid, as compared to the parental strain (Fig. 4-3A, first 3 lanes). However, no LcpA signal was detected in the membrane of strains expressing LcpA with C365A or C179A/C365A mutation (Fig. 4-3A, last 2 lanes) suggesting that the disulfide bond formed by C179 and C365 is required for stability of LcpA.

To ensure the protein production defect above was not due to the lack of *lcpA* transcription, we collected mRNA in these strains and used reverse transcription PCR (RT-PCR) to amplify a 196-bp region specific to the *lcpA* gene. In the WT strain, the *lcpA* transcript was only detected when reverse transcriptase (RT) was added, with *lcpA* amplified from genomic DNA (g) used as a control for the length and specificity of the amplicon (Fig. 4-3B, lanes WT). As expected, no *lcpA* transcript was detected in the *lcpA* mutant (Fig. 4-3B, lanes  $\Delta$ *lcpA*), while the transcript levels of *lcpA* expressed from these recombinant plasmids were comparable to the *lcpA* level in the WT strain (Fig. 4-3B, remaining lanes). Altogether, the results suggest that in the absence of the disulfide bond there is a defect in LcpA protein stability and is not due to lack of gene expression.

We next examined if mutations of these cysteine residues affect LcpA glycosylation activity by immunoblotting the culture medium and cell wall fractions of the same set of strains, according to the procedure described in Fig. 4-2F. Surprisingly, strains expressing LcpA with C365A or C179A/C365A mutation produced GspA<sup>HMM</sup>, albeit less abundant as compared to the WT and rescued strains with accumulation of the intermediate GspA<sup>LMM</sup> unlike the aforementioned strains (Fig. 4-3C). The data strongly indicate that the disulfide bond is necessary for full activity of LcpA.

We previously reported that disulfide bond formation in *A. oris* requires the activity of a membrane bound thiol-disulfide oxidoreductase named MdbA (62), and reactivation of MdbA involves another oxidoreductase called VKOR (124,125). A mutant strain of *mdbA* is not possible, so because *vkor* contributes to, but is not required for oxidative protein folding (62), we examined if LcpA stability is affected in the *vkor* mutant. To test this possibility, the parent, its isogenic  $\Delta vkor$  mutant, and rescued strains were subjected to cell fractionation. To determine if deletion of *vkor* affects LcpA expression, protoplast fractions were analyzed by Western blotting with  $\alpha$ -LcpA; protein levels were quantified by densitometry from four independent experiments with loading controls from the same blots stained by Coomassie. As compared to the WT and rescued strains, the  $\Delta vkor$  mutant produced significantly less LcpA (Fig. 4-3D-E). As a control, the protein level of the housekeeping sortase SrtA remained the same in three strains (Fig. 4-3D). When the culture medium and cell wall fractions were immunoblotted with anti-GspA there were no significant defects in GspA glycosylation. In the  $\Delta vkor$  mutant GspA<sup>LMM</sup> species accumulated in this strain as compared to the WT and complementing strains (Fig. 4-3F). Altogether, the results support that disulfide bond formation is critical for LcpA stability and this oxidative protein folding is mediated by the major oxidoreductase machinery MdbA and VKOR as previously reported (62).



**Figure 4-3: The disulfide bond C179-C365 is required for LcpA stability. (A)**

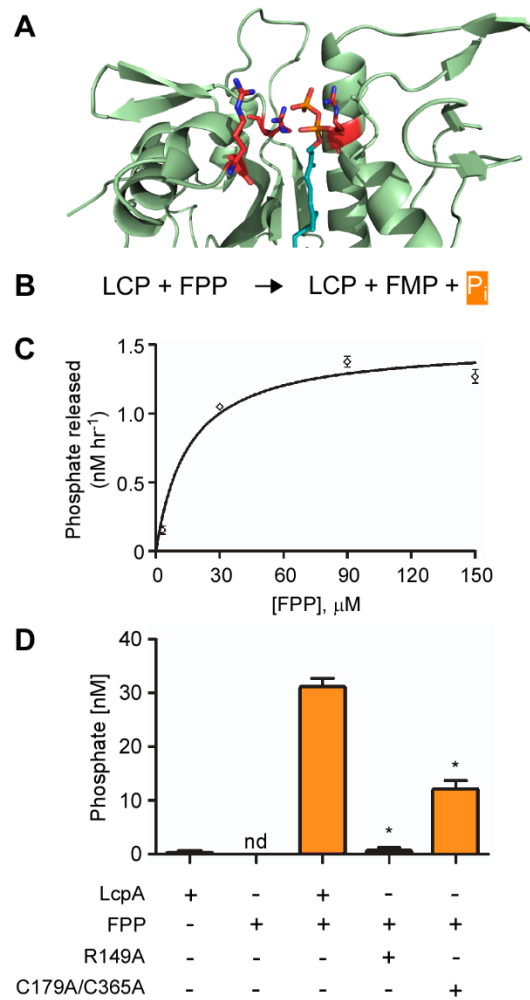
Immunoblotting of the membrane fractions of indicated strains were performed as described in Fig. 4-2F with anti-MdbA used for the control membrane protein MdbA. **(B)** Expression of *lcpA* in indicated strains was analyzed by RT-PCR using primers specific for a 196-bp region of *lcpA*. *A. oris* MG1 genomic DNA (gDNA) was used as controls for length and specificity. (+) and (-) indicate the presence or absence of reverse transcriptase (RT). **(C)** Protein samples from the indicated strains were prepared and analyzed by immunoblotting as described in Fig. 4-2F. **(D)** Shown is a representative Western blot of protoplast fractions of the MG1 strain (WT), the  $\Delta vkor$  mutant, and this mutant expressing VKOR from a plasmid. LcpA is marked with an arrow, whereas a non-

specific band is shown as an asterisk (\*); an arrowhead for a loading control band from the immunoblotted membrane stained with Coomassie blue. The membrane-bound protein SrtA serves as a control. **(E)** Relative steady state stability of LcpA was determined by comparing the relative intensity of the LcpA bands in **(D)**, which were normalized against the loading control band. The relative intensity of the wild-type LcpA bands was set to 1. Error bars represent standard deviation (SD) of four independent replicates. **(F)** The culture medium and cell wall fractions of the indicated strains were analyzed by immunoblotting as described in panel C.

**LcpA exhibits pyrophosphatase activity *in vitro*.** LCP enzymes studied to date possess pyrophosphatase activity, i.e., they catalyze the hydrolysis of pyrophosphate bonds. For example, the LCP enzyme TagT from *B. subtilis* showing that the enzyme exhibits pyrophosphatase activity *in vitro* as do LCP proteins from *Mycobacterium tuberculosis* and *Corynebacterium glutamicum* (44,47,97). In LCP enzymes interaction of the arginine residues with the pyrophosphate is necessary for pyrophosphatase activity. We modeled an Opr-PP molecule into the LcpA hydrophobic pocket. This model was created using electron density of the modeled phosphate ion to place the phosphate head groups of Opr-PP and the electron density used to model PEG4000 to model the lipid component of the Opr-PP polyprenyl (Fig. 4-4A). Indeed, R149 and R266 interact with the pyrophosphate head group according to the model. To test for pyrophosphatase activity of *A. oris* LcpA, we utilized an *in vitro* assay with a diphosphate mimetic substrate, farnesyl pyrophosphate (FPP) and rLcpA and its mutant derivatives purified from *E. coli*. Pyrophosphatase activity of LcpA proteins was determined by quantitatively measuring inorganic phosphate ( $P_i$ ) release from FPP (Fig. 4-4B). It was found that rLcpA was able to hydrolyze FPP exhibiting a  $V_{max}$  of  $1.509 \pm 0.077$  nM hr<sup>-1</sup> and  $K_m$  of  $15.16 \pm 3.656$   $\mu$ M (Fig. 4-4C). The saturating substrate concentration occurred at an enzyme-to-substrate ratio of approximately 1:30 (Fig. 4-4C).

We then examined if mutations of the catalytic residue R149 and disulfide bond C179/C365 affect the pyrophosphatase activity of LCP using the above assay with the saturating substrate concentration. As expected, the LcpA enzyme and FPP contained little to no background  $P_i$  (Fig. 4-4D, first 2 columns). Compared to the wild-type rLcpA enzyme, alanine-substitution of R149 in rLcpA abrogated the enzymatic pyrophosphatase activity (Fig. 4-4D, compare lane 4 to lane 3) further confirming the essential role of this catalytic residue. Consistent with the *in vivo* results above, the rLcpA protein lacking the disulfide bond C179-C365 exhibited significantly reduced

pyrophosphatase activity, approximately 3-fold less than the wild-type (Fig. 4-4D, last column). Altogether, the results indicate that LcpA possesses pyrophosphatase activity and that the disulfide bond C179-C365 plays an important role in maintaining the full activity of LcpA.



**Figure 4-4: LcpA exhibits pyrophosphatase activity.** (A) Octaprenyl-pyrophosphate (oprPP) bound to *A. oris* LcpA was modeled with the prenyl chain shown in turquoise and the pyrophosphate in orange. Potential interactions of Arg residues (red) with the pyrophosphate head group are presented. (B) Presented is the hydrolysis reaction of farnesyl pyrophosphate (FPP) by LCP enzymes resulting in formation of farnesyl monophosphate (FMP) and inorganic phosphate (P<sub>i</sub>, orange). (C) 3 μM of recombinant LcpA was incubated with increasing concentrations of FPP for 24 h at 30°C. Released P<sub>i</sub> was detected by a fluorescent method and quantified from three biological replicates; the V<sub>max</sub> and K<sub>m</sub> values were calculated using the Michaelis-Menten equation in Prism GraphPad, and error bars represent standard error of the mean (SEM) fit by nonlinear regression. (D) Pyrophosphatase activity at saturating substrate concentrations (1:30) of



recombinant LcpA and mutant derivatives, LcpA<sup>R149A</sup> and LcpA<sup>C179A-C365A</sup>, was determined as described in (C) with sole LcpA and FPP included as controls. The results were derived from three independent experiments performed in triplicate with phosphate standards performed in parallel. Error bars represent SEM, and statistical analysis with a one-tailed Mann-Whitney-Wilcoxon test was determined using Prism GraphPad. The asterisk (\*) indicates p-values of 0.0383 and 0.0500 for reactions with R149A and C179A-C365A enzymes, respectively; nd, not detected. ***Note: Modeling of LcpA with OprPP was performed by Jason Gosschalk and used with permission.***

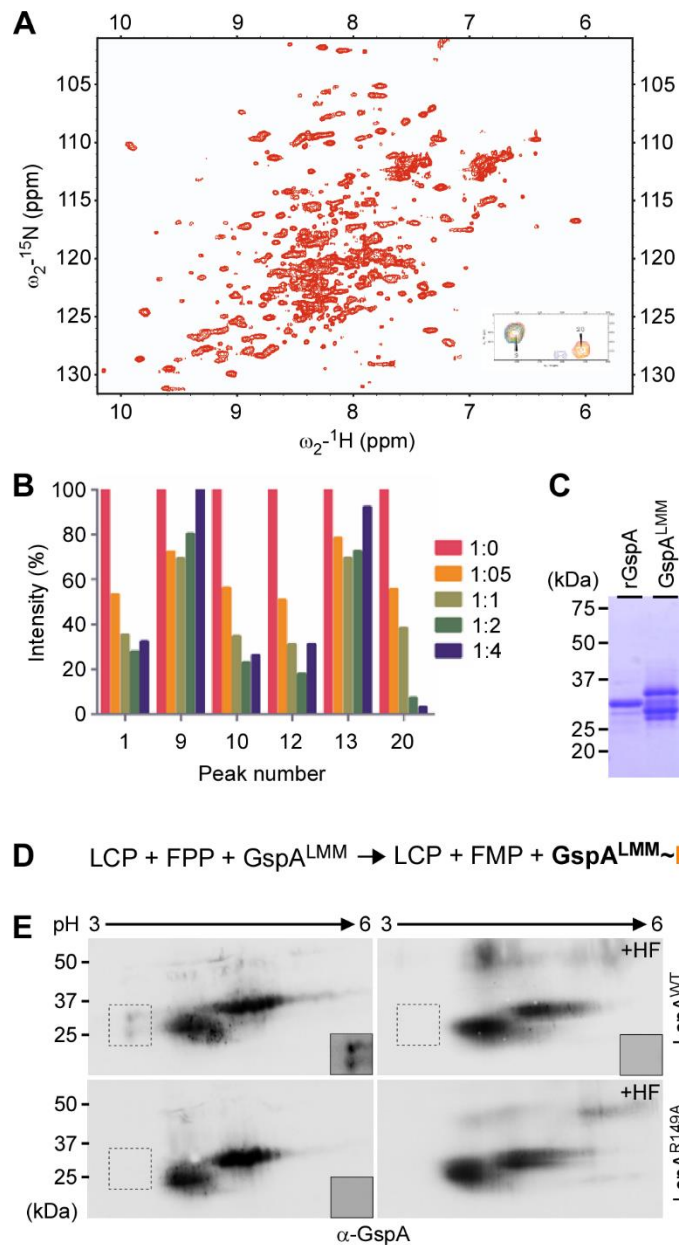
**(Note: A portion of the writing in the following section is credited to Brendan Amer, who performed the NMR experiments, and was used with permission.)**

**A. *oris* LcpA catalyzes phosphotransfer.** To further define the mechanism of surface protein glycosylation by LcpA, we investigated its interactions with its GspA substrate using solution NMR spectroscopy, which can detect transiently interacting proteins.  $^1\text{H}$ - $^{15}\text{N}$ -HSQC titration studies were performed with  $^{15}\text{N}$ -isotopically enriched rLcpA and  $^{14}\text{N}$ -rGspA, a truncation of GspA lacking its predicted N-terminal signal peptide and C-terminal transmembrane region. A series of  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR spectra of  $^{15}\text{N}$ -rLcpA with various amounts of the  $^{14}\text{N}$ -rGspA was acquired. The various spectra of the  $^{15}\text{N}$ -rLcpA (up to 1:4  $^{15}\text{N}$ -rLcpA-to- $^{14}\text{N}$ -rGspA ratio) titrations were partially resolved enabling for line-shape, specifically peak-height analysis (Fig. 4-5A). Spectra of  $^{15}\text{N}$ -rLcpA and  $^{14}\text{N}$ -rGspA at a 1:8 ratio respectively are completely broadened either due to sample dilution or more likely spin-diffusion caused by complex formation. Unfortunately, due to the low quality of the spectra, site-specific interactions or chemical-exchange equilibria could not be estimated. However, analysis of 43 resolved peaks revealed that 4 of these peaks with high signal-to-noise (approximately 20-fold over background), i.e., peaks 1, 10, 12, and 20 exhibited dose-dependent reduction in peak-height during the titration experiment (Fig. 4-5B). This suggests that rLcpA and rGspA interact weakly *in vitro*. Further refinement of this interaction will help define the LcpA-mediated mechanism of glycopolymer transfer, and this data supports further studies of these interactions.

The structural evidence, pyrophosphatase activity, *in vivo* glycosylation, and enzyme-substrate interaction above are consistent with the phosphotransfer activity of LCP enzymes that have previously been shown to mediate WTA synthesis (44); WTA is linked to the N-acetyl-muramic acid component of the cell wall via a phosphodiester linkage (39,126). To examine if *A. oris* LcpA possesses phosphotransfer activity, we

employed an *in vitro* phosphotransfer assay in which the recombinant enzyme rLcpA was mixed with FPP and GspA proteins. After 24 h incubation at 30°C, protein samples were analyzed by 2-D gel electrophoresis followed by immunoblotting with  $\alpha$ -GspA. Because deletion of *lcpA* results in the accumulation of several GspA<sup>LMM</sup> species (Fig. 4-5C), we surmised that the GspA<sup>LMM</sup> proteins are a substrate of LcpA. To facilitate purification of GspA<sup>LMM</sup> in *A. oris*, we engineered a GspA protein with its CWSS replaced by a His-tag and the recombinant protein was expressed in the  $\Delta$ *lcpA* mutant; GspA<sup>LMM</sup> proteins were purified from the culture medium by affinity chromatography. As compared to the recombinant protein rGspA, which was used in Fig. 4-5A, the GspA<sup>LMM</sup> proteins migrated between the 25-kDa and 37-kDa markers (Fig. 4-5C). The identity of these GspA proteins was also confirmed by mass spectrometry. If the GspA<sup>LMM</sup> proteins are substrates of LcpA, addition of LcpA and FPP should lead to phosphate modifications of GspA (Fig. 4-5D), hence increasing acidity due to the negatively charged phosphate group. As shown in Fig 5E, in the presence of the wild-type LcpA enzyme, two new spots migrating between the 25-kDa and 37-kDa markers and towards the acidic pI were detected as compared to samples treated with the inactive enzyme LcpA<sup>R149A</sup>.

To test if phosphate modification of GspA<sup>LMM</sup> occurs via a phosphodiester bond the LcpA + FPP + GspA<sup>LMM</sup> samples were treated with hydrofluoric acid (HF), which hydrolyzes phosphodiester bonds as previously demonstrated in *Staphylococcus aureus* with an LCP enzyme (48) prior to 2D-gel electrophoresis and immunoblotting. Indeed, HF treatment resulted in abrogation of phosphate modification (Fig. 4-5E; HF panels). Altogether the results support the notion that LcpA is a phosphotransferase and that GspA<sup>LMM</sup> is a bona fide substrate for LcpA catalyzed-glycosylation.



**Figure 4-5: LcpA interacts with GspA in solution and catalyzes phosphotransfer.**

(A) Presented is the full  $^1\text{H}$ - $^{15}\text{N}$ -HSQC of  $250\ \mu\text{M}$   $^{15}\text{N}$ -recombinant LcpA with inset showing representative data of overlaid  $^1\text{H}$ - $^{15}\text{N}$ -HSQC titration spectra displaying two isolated peaks with high signal-to-noise. Red spectra 1:0 molar equivalents of  $^{15}\text{N}$ -rLcpA to  $^{14}\text{N}$ -rGspA. Orange, Yellow, Green, and Blue represent 1:0.5, 1:1, 1:2, and 1:4 spectra, respectively. Peak 9 is an example, which does not exhibit a dose-dependent decrease in peak height upon adding rGspA, and Peak 20 is shown as an example,

which does exhibit drastic effects on peak height. **(B)** Normalized plot of peak intensity of selected residues from titration experiment with high signal-to-noise. Intensity data was normalized to 1:0 titration peak intensities. **(C)** Recombinant GspA (rGspA) and GspA<sup>LMM</sup> were purified affinity chromatography from *E. coli* and *A. oris* lysates, respectively, and analyzed by SDS-PAGE and Coomassie staining. **(D)** Presented is a simplified schematic for *in vitro* phosphotransfer. **(E)** The phosphotransfer reaction contained 12  $\mu\text{M}$  of GspA<sup>LMM</sup>, 4  $\mu\text{M}$  of LcpA (WT or R149A) and 50  $\mu\text{M}$  of FPP in 20 mM Tris-HCl (pH 8.0). After 72 h incubation at 30°C, protein samples were treated with hydrofluoric acid (HF) or mock-treated prior to 2D-electrophoresis followed by immunoblotting with anti-GspA antibodies. Insets with increased contrast were shown for boxed regions. ***NMR analysis was performed by Brendan Amer and used with permission.***

## Discussion

Members of the LCP protein family studied to date have been shown to attach glycopolymers to peptidoglycan (120,121) with many demonstrated to possess pyrophosphatase and phosphotransferase activities (44,47,97,127,128). LCP enzymes are characterized as the terminal enzyme, which catalyzes the linkage of glycopolymers to the muramic acid component of the peptidoglycan via a phosphodiester bond from a prenyl pyrophosphate glycan donor (122,129). None of these enzymes, however, are involved in glycosylation of cell wall anchored proteins. We present here experimental evidence that *A. oris* LcpA – capable of catalyzing hydrolysis of diphosphate bonds and phosphotransfer – glycosylates the cell wall anchored protein GspA prior to attachment to peptidoglycan, a process that is facilitated by the housekeeping sortase SrtA (50).

Crystallization studies authenticate LcpA as a member of the LCP protein family, revealing that LcpA is structurally related to *B. subtilis* TagT, a previously crystallized LCP that mediates the linkage of WTAs to peptidoglycan (46). Both enzymes have a similar hydrophobic tunnel lined with arginine residues (R149 and R266 in *A. oris* LcpA), a conserved feature of LCP enzymes that is necessary for interaction with glycan donor substrates. Consistent with this, alanine-substitution of the catalytic R149 residue abrogates pyrophosphatase and phosphotransferase activities and glycosylation of GspA. Unlike *B. subtilis* TagT, *A. oris* LcpA does not attach glycopolymers to peptidoglycan, as a GspA mutant lacking the CWS still contains glycans (50). This raises an intriguing question as to where glycopolymers are attached to GspA. While the biochemical nature of the glycans and locations of the glycosylation sites remain to be elucidated, the results presented in our previous publication (50) and Fig. 4-5 suggest that an intermediate form of GspA, GspA<sup>LMM</sup>, may serve as a substrate of LcpA. It is interesting to note that *A. oris* harbors four LCP homologs, but only LcpA is involved in GspA glycosylation (Fig. 4-1). Because *lcpC* is an essential gene and a conditional

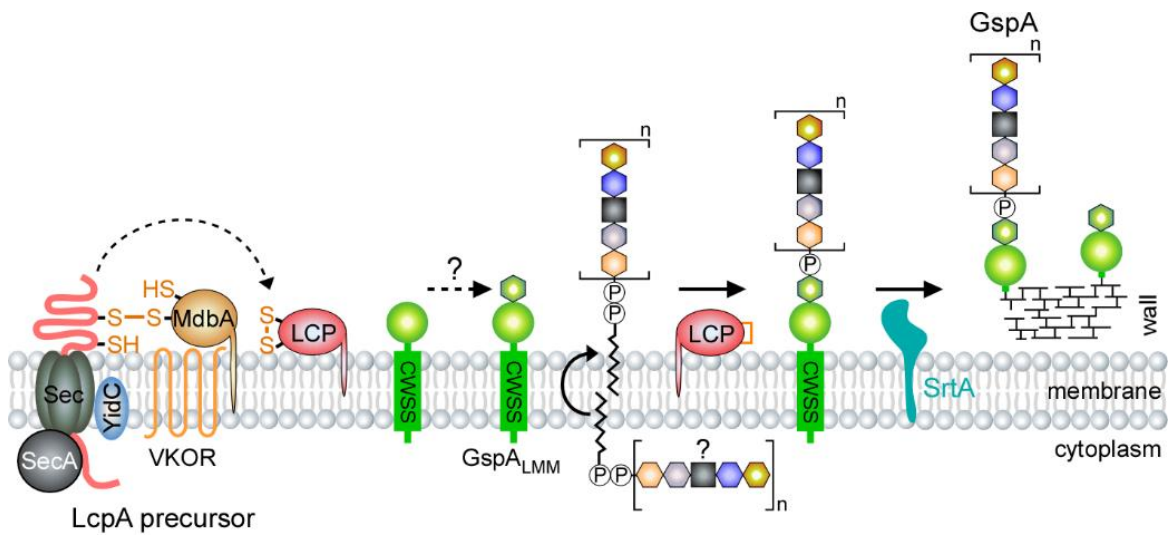
deletion mutant is not available, this does not exclude the possibility that LcpC may modify GspA, leading formation of GspA<sup>LMM</sup>. Future experiments will address this issue.

Intriguingly, the presence of a hydrophobic tunnel in the LcpA and TagT structures as mentioned above suggests that the enzymes use a pyrophosphate-lipid linked glycan donor. To gain insight into how this substrate bound, we used the structure of the TagT enzyme bound to all *cis* octaprenyl-pyrophosphate (opr-PP) (PDB 4DE9) to model the opr-PP:rLcpA complex. This was achieved by superimposing the protein coordinates as well as the coordinates of the phosphate proximal to the glycan strand in the structure of TagT and the active site phosphate atom present in the structure of rLcpA. The model suggests rLcpA catalyzes a phosphotransfer reaction in which the pyrophosphate linkage joining the lipid to the sugar molecule is broken, presumably as a result of nucleophilic attack by an oxygen or nitrogen atom present an amino acid sidechain within the GspA protein. As a result, the proximal phosphate and glycan are transferred to GspA. In this reaction, R149 may stabilize the phosphate leaving group, whereas R266 may favorable interact with the trigonal bipyramidal intermediate that likely forms during catalysis. The process is thermodynamically favorable, as breakage of the phosphoanhydride linkage in the substrate releases more free energy than is required to attach a sugar molecule to the protein (the Gibbs standard free energy for phosphoanhydride breakage in the substrate is  $\sim -7.3$  kcal/mol, whereas only  $\sim 3.3$  kcal/mol is required to form the phosphodiester bond that joins the sugar to the protein). The complexity of the glycans has prohibited our ability to determine the exact identity of our glycans species, although this is a subject of current work.

Unlike other LCP proteins studied to date, *A. oris* LcpA possesses a distinct feature, which appears to be commonly present in the actinobacterial LCP enzymes, i.e., a disulfide bond. Given disulfide bond formation is critical for oxidative folding of

exported proteins in Actinobacteria (130), a process that is catalyzed by a pair of thiol-disulfide oxidoreductase enzymes MdbA/VKOR in *A. oris* (62), we hypothesized that the disulfide bond formed between C179 and C365 is essential for post-translocational folding of LcpA. This is evident by the fact that mutations that abrogate the disulfide bond C179-C365 severely affects stability of LcpA, whereas deletion of VKOR significantly reduces LcpA stability (Fig. 4-3). Altogether, we propose that as the LcpA precursor emerges from the Sec machine, it is folded by the MdbA/VKOR enzymes and inserted into the membrane by the membrane protein insertase YidC. Separately, the membrane-bound GspA is also transported by the Sec and further modified by an unknown mechanism resulting in an intermediate form named GspA<sup>LMM</sup>. LcpA catalyzes the attachment of an unknown glycan chain to GspA<sup>LMM</sup>, which is then anchored to the bacterial peptidoglycan by the housekeeping sortase SrtA (Fig. 4-6). Given the conservation of LCP and GspA proteins, this glycosylation pathway may be conserved in Actinobacteria.





**Figure 4-6: Proposed model of LcpA-mediated glycosylation of the cell-anchored protein GspA.** *A. oris* LcpA is proposed to catalyze the linkage of unknown glycopolymers to GspA, which is then anchored to the bacterial cell wall by the housekeeping sortase SrtA. The oxidoreductase enzymes MdbA/VKOR is thought to catalyze oxidative folding of LcpA (see text for details).

**CHAPTER V**  
**Mapping the Molecular Domain in CafA Responsible**  
**for Interkingdom Adherence**

## Introduction

Bacterial adherence is an early and important step during infection and is mediated by adhesins that define tissue tropism and interspecies interactions (131). Many Gram-positive pathogens elaborate covalently linked sortase-catalyzed polymers with an adhesive tip pilin to mediate this step in the infection (18,25,71,73,84). Sortases are a broad class of transpeptidase enzymes that catalyze pilus polymerization and cell wall anchoring of surface proteins. *Actinomyces oris* strain MG1 has become a model organism to study the biogenesis of sortase-catalyzed pili (21). *A. oris* produces two distinct pilus structures called type 1 and type 2 pili. Both types of pili confer distinct adherence abilities to *A. oris* that are necessary to fulfill the role as a primary colonizer of the oral biofilm (21). The role of *A. oris* as a primary colonizer is exemplified by studies that show its early and specific colonization and localization at the biofilm base (9,16,132).

Type 1 pili mediate bacterial adherence to the conditioned enamel of the tooth via interactions with proline-rich proteins (18), whereas type 2 pili are involved in coaggregation with oral streptococci and adherence to host cells (79,113). Therefore, using pili, *A. oris* attaches to the tooth surface and enables further biofilm development via the recruitment of bacteria, which cannot otherwise adhere to the substratum (2,133). Accumulation of dental plaque can lead to dental diseases such as dental caries and periodontitis that are considerable health burdens on a global scale (13). Ultimately, accretion requires specific molecular interactions that serve as the basis for the spatiotemporal development of the oral biofilm and is dependent on primary colonizers providing the initial site of attachment for subsequent biofilm maturation (2,5). Following attachment and microcolony formation by primary colonizers, coaggregation between bacterial species can occur which serves to enhance bacterial attachment and mature the biofilm (2). *A. oris* strain MG1 and *S. oralis* strain 34 are well-defined coaggregation partners. Their interaction which can be monitored in by simply combining the two

species in a defined buffer and observing the supernatant clearing and aggregate formation (134). In this interaction, *S. oralis* provides a defined receptor polysaccharide (RPS) with a terminal D-Gal- $\beta$ -(1-3)-GalNac. *A. oris* provides the specific adhesin for this interaction. It has been known that this interaction required type 2 pili from *A. oris* (33,79). However, the specific adhesin remained unidentified until more recently when coaggregation factor A (*cafA*) was identified in *A. oris* as required for this specific interspecies interaction (25). The CafA adhesin is particularly interesting because unlike other pilus-associated proteins that genetically cluster with the cognate pilin specific sortase, *cafA* is located distant from SrtC2, which is required for its incorporation into type 2 pili. Type 2 pili were also shown to be important for sialidase-dependent binding of epithelial cells, however, CafA was not identified as the specific adhesin in this case (17). Sialic acid normally conceals the TF antigen, which has a terminal residue of D-Gal- $\beta$ -(1-3)-GalNac that bears similarity to the RPS from So34.

CafA hijacks the FimA backbone pilus to form a distinct CafA-FimA type 2 pilus (25). On average, pilin tip proteins are larger compared to the respective backbone pilins and have a multidomain structure including separable stalk and adherence domains. Pilus tip proteins often serve as the functional adhesin for pili (71,135). This property, combined with the fact that pili are accessible to immune cells and chemical modifiers, makes pili and particularly adhesins good targets for anti-adhesive strategies (135,136).

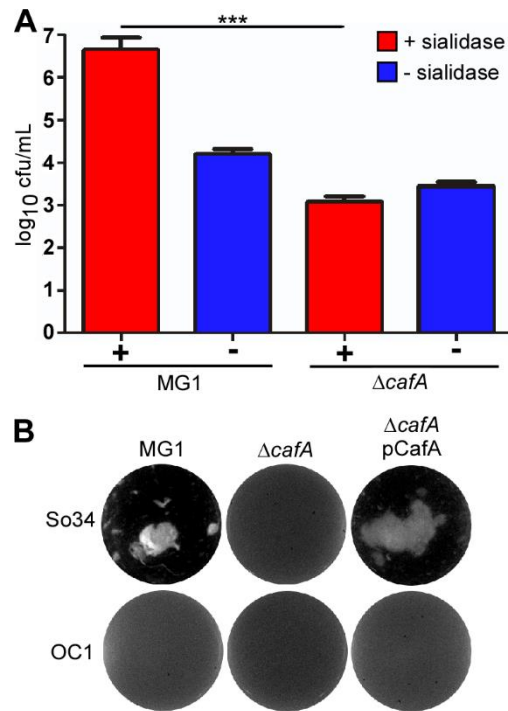
Our past characterization of pilus components and pilus-related functions utilized the lab strain *A. oris* MG1, however, a collection of clinical isolates taken from various mouth sites has previously been characterized by the Cisar laboratory with serological methods. *Actinomyces* have undergone major reclassification based on 16s rRNA and multi-locus sequence typing of several housekeeping genes, and we refer to the reclassified species designation rather than the serological description, and retain the original serological strain nomenclature (137,138).

Here I report the study of the molecular basis of CafA interkingdom adherence. We first find that *A. oris* MG1 has the ability to adhere to human gingival fibroblasts (HGF-1), and this interaction depends on CafA. We then utilized a clinical isolate screen for the presence of CafA polymers and concurrence with So34 coaggregation. From this screen we identified a single clinical isolate, *A. oris* strain N11A12, that displays CafA polymers, but does not coaggregate with So34. Based on characterization and sequencing of CafA from this isolate, we identified the specific amino acids responsible for coaggregation with So34.

## Results

**CafA binds a receptor that is conserved between oral streptococci and human fibroblasts.** Previous work in *Actinomyces* demonstrated the necessity of sialidase/neuraminidase treatment to enhance epithelial cell binding (17,139). To determine if *A. oris* MG1 binding to human cells was CafA dependent, we infected human gingival fibroblasts (HGF-1) with *A. oris*. We found that wild-type adherence increases about a hundred-fold (2-log difference) when the HGF-1 cells are treated with sialidase prior to infection. *A. oris*  $\Delta$ *cafA* binds HGF-1 at the level of MG1 without sialidase and does not exhibit increased adherence after sialidase pretreatment (Fig. 5-1A). This indicates that sialidase treatment of HGF-1 uncovers a receptor that is normally blocked by sialic acid, and that *A. oris* binding to the receptor is dependent upon CafA. In human cells, sialic acid often caps the TF antigen, which has a terminal Gal- $\beta$ -(1-3)-GalNac (135). *Actinomyces spp.* possess and express a sialidase gene, *nanH*, and a sialidase transporter, *nanT*, which are significantly upregulated in *Actinomyces naeslundii* isolated specifically from root caries (140).

Coaggregation is mediated by an adhesin from one species and the coordinating receptor of another. Moreover, prior work demonstrated the interaction between *S. oralis* strain 34 and *A. oris* MG1 was determined to be mediated by a Gal- $\beta$ -(1-3)-GalNac moiety on So34, also called coaggregation group 3 (141). CafA was recently identified as the specific adhesin necessary for binding So34 (25). We confirmed that the coaggregation phenotype relies on the presence of CafA in *A. oris* MG1 (Fig. 5-1B). Although, these results suggest that CafA is an interkingdom adhesin that utilizes D-Gal- $\beta$ -(1-3)-GalNac as a receptor on both oral streptococci and human cells.



**Figure 5-1: CafA adheres to receptor polysaccharides present on HGF-1 and oral streptococci. (A)** HGF-1 adherence by *A. oris* MG1 or the isogenic deletion strain  $\Delta$ cafA in the presence (+) or absence (-) of sialidase treatment by direct measurement of colony forming units per milliliter (cfu/mL). Student's two-tailed, unpaired T-test; \*\*\* P-value < 0.0001. Input values are comparable (MOI ~ 200) as determined by measurement of cfu/mL. **(B)** Coaggregation phenotype of *A. oris* MG1,  $\Delta$ cafA, and the CafA complement strain ( $\Delta$ cafA pCafA) with the RPS-positive *S. oralis* strain 34 (So34) and RPS-negative strain *S. oralis* OC1 (SoC1).

***Actinomyces* clinical isolates have differential coaggregation phenotypes and CafA expression.** We utilized coaggregation with So34 as a measure of the presence of binding to the RPS on oral streptococci which is recognized by the species adhesin CafA from *A. oris* MG1 (25). We wanted to determine whether *Actinomyces* clinical isolates had differential phenotypes and whether we could identify at least one of which displayed polymeric CafA, but did not coaggregate with So34. We first performed a coaggregation assay of 23 clinical isolates (Table 5-1). Then we examined the cell wall fractions of each isolate for the presence of polymeric CafA (Table 5-1). This screen identified different coaggregation groups as previously reported (142,143). The majority of the strains displaying polymeric CafA coaggregated with So34, as expected from our results with the lab strain MG1. We also identified some isolates which bind both So34 and SoOC1, but exhibited either monomeric CafA or absence of CafA. Therefore, this interaction may not necessarily be CafA dependent, because CafA requires polymerization by FimA to be functional (25). Additionally, we found a monomeric CafA with no So34 binding. Our most interesting finding, however, was that isolate 9, which corresponds to *A. oris* strain N11A12, does not coaggregate with So34, despite elaborating CafA polymers. This is considerably different from the other isolates that we screened where polymeric CafA is obligatorily associated with So34 coaggregation. We characterized this isolate to ensure that the coaggregation defect was not due to significant changes of pilus structure or expression.



**Table 5-1: Coaggregation and CafA polymerization phenotypes of *Actinomyces* clinical isolates.**

Number	Species	Strain	Coaggregation with <sup>(b)</sup>		CafA
			So34	OC1	
1	<i>A. oris</i>	MG1	+	-	p <sup>(c)</sup>
2	<i>A. oris</i>	$\Delta cafA$ <sup>(a)</sup>	-	-	- <sup>(d)</sup>
3	<i>A. naeslundii</i>	N28B15	+	+	-
4	<i>A. naeslundii</i>	N34A24	+	+	m <sup>(e)</sup>
5	<i>A. naeslundii</i>	N35B3	+	+	m
6	<i>A. oris</i>	N11A16	+	-	p
7	<i>A. oris</i>	N12A2B	+	-	p
8	<i>A. oris</i>	ATCC 49339	+	-	p
9	<i>A. oris</i>	N11A12	-	-	p
10	<i>A. oris</i>	ATCC 27044	+	-	p
11	<i>A. oris</i>	N33A2B	+	+	-
12	<i>A. oris</i>	N37B13	-	-	m
13	<i>A. oris</i>	N34A23	+	-	p
14	<i>A. oris</i>	N28B1	+	-	p
15	<i>A. oris</i>	N29A27	+	-	p
16	<i>A. oris</i>	N32A8	+	-	p
17	<i>A. oris</i>	N37B9	+	-	p
18	Non-serotypeable	N38B10	+	-	p
19	Non-serotypeable	N33A3	-	-	m
20	Non-serotypeable	N34A14	+	-	p
21	<i>A. johnsonii</i>	PK1259	+	-	p
22	<i>A. johnsonii</i>	ATCC 49338	+	-	p

<sup>(a)</sup>Isogenic derivative of MG1

<sup>(b)</sup>Positive and negative signs indicate presence and absence of coaggregation, respectively.

<sup>(c)</sup>CafA polymers

<sup>(d)</sup>No CafA signal

<sup>(e)</sup>CafA monomer

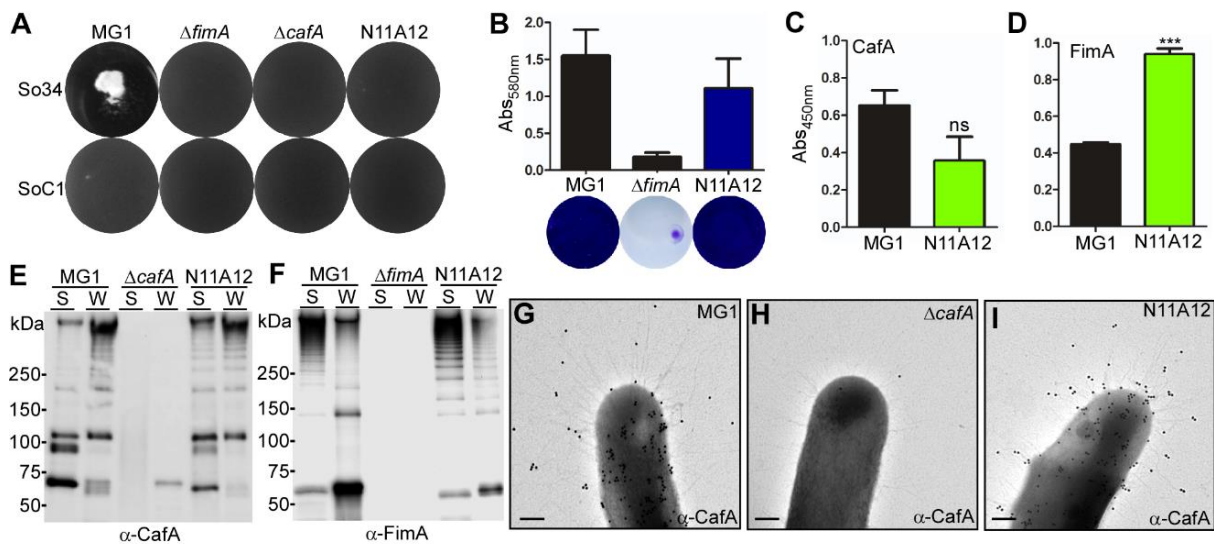
**Characterization the clinical isolate *A. oris* strain N11A12.** The coaggregation deficient phenotype and display of CafA polymers of *A. oris* N11A12 prompted us to further characterize this isolate. First, we repeated the coaggregation result outside of the screening context and found that indeed coaggregation of this isolate with So34 was defective (Fig. 5-2A).

We wanted to ensure that this isolate was not dissimilar to MG1 in monospecies biofilm formation, which is mediated by FimA, the major pilin subunit of the type 2 pili (84). To do this we utilized an *in vitro* biofilm assay and assessed them qualitatively and quantitatively. Our lab strain MG1 and N11A12 are stained robustly by the crystal violet and are not significantly different in their ability to form a monospecies biofilms. In contrast,  $\Delta fimA$  is defective in biofilm formation under the same conditions (Fig. 5-2B).

Next, we quantified the amount of FimA and CafA on the surface of MG1 and N11A12 using whole-cell ELISA. We found that CafA was not significantly different between the lab strain and the clinical isolate (Fig. 5-2C). We do note that FimA level is increased significantly in N11A12 compared to MG1 (Fig. 5-2D). We concluded from these data that lack of expression of type 2 pilus structural components was not the cause of the coaggregation defect. However, to be functional, CafA must be polymer associated (25). We previously detected the polymer through the screening process, but wanted to ensure the presence of CafA polymers in the media and wall fraction of N11A12 were similar to MG1 (Fig. 5-2E). We also detected FimA polymers in N11A12 which were similar to MG1 in the media and cell wall fraction, which we expected due to the necessity of FimA polymers for monospecies biofilm formation (Fig. 5-2F). Also of note, the size of the CafA and FimA monomer was very similar between MG1 and N11A12. Additionally, we tested whether we could visualize CafA at the pilus tip by immunogold labeling and electron microscopy, as we saw in our MG1 control, and was well above nonspecific binding from  $\Delta cafA$  (Fig. 5-2G & 5-2H). Indeed, in *A. oris*

N11A12, CafA was present on pili and extended away from the cells (Fig. 5-21).

Collectively, these data indicated that the coaggregation deficient phenotype of *A. oris* strain N11A12 was not due to reduced expression or polymerization of CafA containing pili.



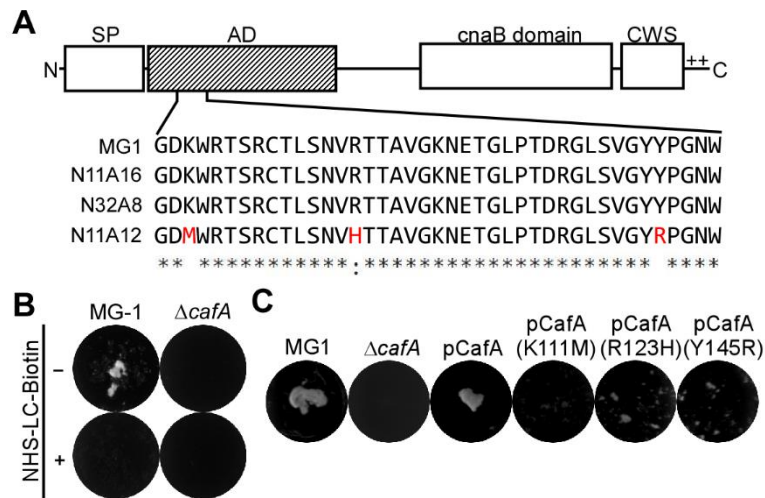
**Figure 5-2: Characterization of type 2 pili and pilus-related functions in**

**Actinomyces oris** strain N11A12. (A) Coaggregation assay with *A. oris* and *S. oralis* cells suspended in coaggregation buffer in a 1:1 ratio. (B) *In vitro* biofilm assay of *A. oris* MG1,  $\Delta fimA$ , and N11A12. Biofilms were grown statistically in HIB + 1% sucrose for 48 h. Planktonic cells were washed away and the remaining biofilm was detected with crystal violet. Crystal violet was released with EtOH and quantified at A<sub>580nm</sub>. The values shown are the average of three independent replicates that were each run in triplicate, and error bars represent standard error of the mean (SEM). Statistical analysis was performed using an unpaired, two-tailed Student's t-test. (C & D) Whole cell ELISA with of *A. oris* MG1 and *A. oris* N11A12 with (C)  $\alpha$ -CafA and (D)  $\alpha$ -FimA antibodies. Data shown from three independent replicates and all replicates were performed in quadruplicate. Error bars represent SEM, and statistical analysis using unpaired, two-tailed Student's t-test. Designation ns refers to  $p > 0.05$ , and \*\*\*  $p = 0.0001$ . **This data was generated by Alexis Bradford and used with permission.** (E & F) Western blot analysis of supernatant (S) and cell wall (W) fractions using (E)  $\alpha$ -CafA and (F)  $\alpha$ -FimA antibodies. (G - I) IEM performed with (G) MG1, (H)  $\Delta cafA$  (parental strain MG1) and (I) N11A12 using an  $\alpha$ -CafA primary antibody and 18-nm gold particles conjugated to the secondary antibody. Scale bar represents 0.2  $\mu$ M.

**Identification of specific residues that are critical for CafA adherence through comparison with the CafA sequence from *A. oris* N11A12.** We reasoned that because the coaggregation defect did not appear to be due to differences in type 2 pilus expression or structure, we could compare the primary amino acid sequence of CafA from MG1 to that of N11A12 to find potential residues necessary for adherence. CafA was present outside of the N11A12 cells and anchored to both pili and the cell wall, we reasoned that any changes occurring in the signal peptide or cell wall sorting signal domains were not significant for the adherence function. CafA from the lab strain MG1 and the clinical isolate N11A12 are 95.98% identical at the amino acid level, and we did not detect any major insertions or deletions. By multiple sequence alignment we identified 36 non-synonymous mutations. Eleven of the substitutions retained strong conservation (e.g. Ser to Thr), and eight substitutions maintained weakly similar properties, based on the Gonnet PAM 250 matrix. To begin to narrow down which nonsynonymous mutations were meaningful, we sequenced *cafA* from two additional *A. oris* isolates, which display the MG1 coaggregation and polymerization phenotypes, N11A16 and N32A8. We compared changes in MG1, N11A16, and N32A8 to those in N11A12 to identify changes in these strains that overlapped and thus could not be implicated in the loss of CafA binding function. We were able to eliminate all but three major changes found in the N-terminal domain. The remaining suspect substitutions were K111M, R123H and Y145R (Fig. 5-3A). These amino acid changes were present in the N-terminal portion of the protein that was previously purified and used to block coaggregation (25). Additionally, amino acid changes occurring in basic amino acids were of particular interest to us because previous reports of *E. coli* adhesins indicated that Lys and Arg residues are involved in glycan receptor binding (135,144). Tyr has also been implicated in glycan binding (135). To determine whether basic amino acids may be involved in CafA binding, we first reacted MG1 and  $\Delta$ *cafA* cells with sulfosuccinimidyl-6-[biotin-amido]hexanoate which reacts with free amine groups found in Lys and Arg

residues only on the cell surface. Sulfosuccinimidyl-6-[biotin-amido]hexanoate can effectively block interaction with glycan receptors (144). We utilized binary qualitative coaggregation to determine the effect of the chemical on coaggregation efficiency. We determined that exposing *A. oris* MG1 to sulfosuccinimidyl-6-[biotin-amido]hexanoate reduces coaggregation to  $\Delta cafA$  levels (Fig. 5-3B).

Due to the fact that sulfosuccinimidyl-6-[biotin-amido]hexanoate could abrogate coaggregation, we introduced the changes that occurred in K111 and R123 residues of N11A12 CafA into the MG1 CafA by site-directed mutagenesis. These three substitutions were introduced singly into pCafA<sup>MG1</sup> by site-directed mutagenesis. I also included the Y145 substitution, because it was substituted in N11A12 compared to the other strains, and Tyr has been demonstrated to bind glycan moieties previously. Then, I used these strains to test whether these residues were indeed necessary for coaggregation with So34. All three mutant strains were defective in coaggregation with RPS-positive, So34 (Fig. 5-3C).



**Figure 5-3: Analysis of nonsynonymous mutations between CafA from MG1 and N11A12 identifies residues necessary for coaggregation. (A)** Schematic with the variable sequences between *A. oris* strains MG1, N11A16, N32A8 and N11A12 for amino acids in the adherence domain. Sequencing was performed by Anh Dinh and used with permission. **(B)** Coaggregation assay of MG1 and  $\Delta cafA$  with So34 performed in the absence (-) or presence (+) of sulfo-NHS-LC-biotin. **(C)** Coaggregation assay carried out using So34 and *A. oris* and derived strains encoding CafA<sup>MG1</sup> site-directed mutants K111M, R123H and Y145R.

## Discussion

The oral biofilm is a polymicrobial community of bacterial species that relies on hierarchical interactions with host surfaces and bacterial species. These interactions are dictated by a sequential addition of new species to the biofilm based on their molecular properties and can be defined by their temporal association, such as primary, bridging and secondary colonizing species (2). They can also be defined by their pathogenic potential, being inflammophilic, accessory pathogens compatible with inflammation or completely incompatible with inflammation. These classifications have an underlying molecular mechanism dictating temporal, spatial and inflammatory compatibility. Here, we have defined molecular mechanisms and interactions associated with a primary colonizer and accessory pathogen compatible with inflammation (12).

*A. oris* is well defined as a primary colonizer (2,9). Using type 1 pili to bind the tooth surface and type 2 pili to mediate bacterial coaggregation and contact with host cells which is dependent on either sialidase or inflammation to uncover the host cell receptor (18,25). *A. oris* encodes a sialidase gene, *nanH*, and a transporter *nanT* which is upregulated in the caries biofilm (140). Upregulation of *nanH* and *nanT* could promote uncovering of the receptor for *A. oris* binding when the local concentration of the enzyme increases, if the bacterium is below the surface of the gum. Additionally, chronically inflamed cells often reveal the TF antigen due to dysregulation of sialic acid production (135). Bifunctional receptor binding may promote a pathogenic phenotype such that upon recruitment of inflammophilic pathobionts to the biofilm, the resulting inflammation would reveal the TF antigen. This would support a role for *A. oris* as a compatible accessory pathogen, where it could maintain a subgingival biofilm (12,135).

The *Actinomyces* clinical isolate screen revealed differential phenotypes for the presence and display of CafA. It also reinforced research that demonstrated *Actinomyces* species participate in different coaggregation groups (141). An interesting finding that



came from this work is the presence of monomeric CafA proteins in only certain strains, for example the only monomeric CafA in *A. oris* is from strain N37B13 (Table 5-1). This isolate displays a predictable negative coaggregation phenotype, as it was previously demonstrated that CafA needs to be polymerized by FimA to mediate coaggregation (25). Therefore, this strain may provide clues about how CafA initially evolved to hijack the type 2 pilus.

## **CHAPTER VI**

### **Exploring the Role of the Twin-Arginine Residues in the CafA Signal Peptide**

## Introduction

Pili are assembled on the exoplasmic membrane, and therefore the subunits must be translocated prior to their assembly. General protein secretion in bacteria is mediated either by the general secretion pathway (Sec) or twin-arginine translocon (Tat). The major difference between Sec and Tat is that substrates of Sec are translocated as unfolded polypeptides, whereas Tat transports only fully-folded proteins (54).

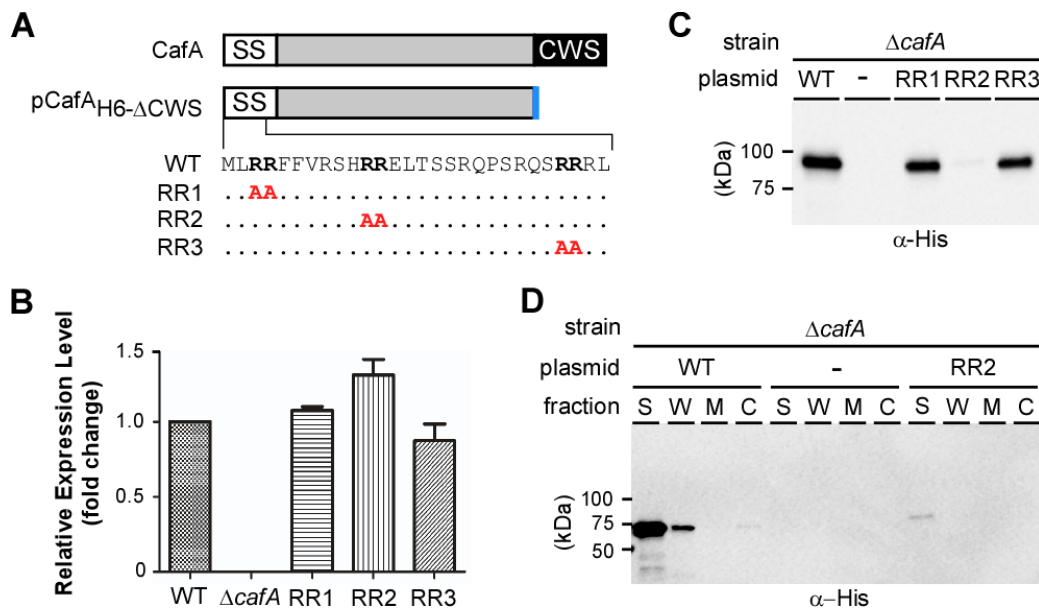
The current model for the process of polymerization begins when the pilin precursors are translocated through the general secretory pathway (Sec) by an N-terminal signal peptide. The precursor is folded after secretion and retained in the membrane by a C-terminal domain called the cell wall sorting signal (CWS) (40,62). The CWS is required to covalently attach the tip protein to the incoming backbone pilin and the backbone pilins to each other. In addition to the aforementioned CWS requirements for sortase interaction, an additional sequence element, the triple glycine (TG) motif GGxG or GGxA, is only found in proteins that interact with pilin-specific sortases. The TG motif overlaps with the LPxTG motif in the pilin subunits of many organisms and was shown to be critical for pilus assembly in *Lactobacillus rhamnosus* GG (145).

In some organisms the polymerization process requires the presence of a tip protein, but the molecular details surrounding this process have not been elucidated (18,25,146). In *A. oris*, the tip proteins FimQ, FimB and CafA are indeed required for initiating the polymerization of the major pilin subunits (18,25). Interestingly, all of the pilus initiation proteins, CafA, FimB and FimQ, have a putative Tat signal peptide, suggesting that the tip proteins may use the twin arginine translocase (Tat). Moreover, the conservation among them provides evidence that it may be important for initiating pilus assembly. However, to date, all sortase substrates are known to utilize the general secretion pathway for their secretion (40).

The consensus sequence used to discriminate Tat from Sec signal peptides is (S/T)RRxFLK. Primarily, this sequence information can be directly applied to identifying *E. coli* Tat substrates, but Tat signal peptides from other organisms have demonstrated tolerance for conservative mutations in this motif (103). The presence of a twin arginine motif in signal peptide of the pilin precursors suggested the use of the Tat pathway, although it was unclear if these proteins are secreted through the Tat pathway. In this study, I investigated whether the CafA signal peptide was sufficient to mediate Tat transport.

## Results

The CafA protein sequence contains three potential twin arginine residues within the signal peptide where a putative signal peptidase cleavage site is predicted to be between A56 and V57 (Fig. 6-1A). The presence of these residues suggested that CafA may be secreted by the Tat pathway (56). In addition to the twin arginine residues, two of the potential signals have a leucine residue at position +2, which has been shown to be important an important residue for Tat translocation (147). To determine whether any of these twin arginine motifs are important for CafA secretion, I created a variant that lacks the cell wall sorting signal (CWS) and included a hexa-histidine tag at the C-terminus so that the mutant protein is secreted into the supernatant rather than polymerized or cell wall anchored (Fig. 6-1A). To substitute the potential arginine residues to alanine I used site-directed mutagenesis (Fig. 6-1A). I confirmed that these constructs are expressed at levels comparable to *cafA* containing a wild-type signal peptide using qRT-PCR of *cafA* (Fig. 6-1B). The culture supernatants were concentrated and subjected to Western blotting with an antibody against the C-terminal His-tag to measure CafA secretion and levels. The RR1 and RR3 CafA signal peptide variants secrete CafA at levels similar to the wild-type signal peptide. Only the RR2 mutation resulted in the lack of CafA secretion (Fig. 6-1C). These results suggested that residues R11 and R12 are critical for CafA secretion, and secretion may indeed require the Tat pathway. Because the RR2 variant was not secreted, I expected the protein to accumulate in the cytoplasmic fractionation. However, I was unable to detect a significant signal in any fraction for the RR2 mutant, which may indicate that reduced secretion was due to low protein production or stability (Fig. 6-1D).



**Figure 6-1: Mutation of Arg 11 and Arg 12 to Ala abrogates CafA secretion and destabilizes CafA.** (A) Schematic of full-length CafA indicating the signal peptide (SP) and cell wall sorting signal (CWS), and substitutions made in CafA variants RR1, RR2, and RR3. The pCafA<sub>ΔCWS</sub>-H6 variant lacks the CWS and contains a hexa-histidine tag at the C-terminus (blue rectangle). The SP is wild-type (WT) or mutated to Ala as indicated in RR1, RR2 and RR3. (B) Western blot analysis of the supernatant fraction from *A. oris*  $\Delta$ cafA strain expressing WT CafA, no CafA (-), or variants of CafA RR1, RR2, or RR3 reacted with an  $\alpha$ -His monoclonal antibody conjugated to HRP. (C) Quantitative PCR data expressed as a fold change calculated using the  $2^{-\Delta\Delta C_T}$  method and normalized to the 16s rRNA. (D) Cell fractionation for supernatant (S), cell wall (W), membrane (M) and cytoplasmic (C) fractions subjected to Western blot with an  $\alpha$ -His monoclonal antibody conjugated to HRP.

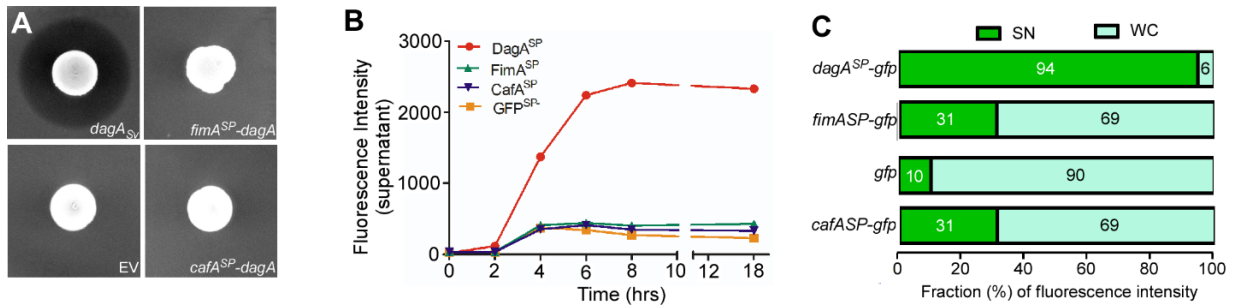
To overcome both the protein stability challenge and determine whether the signal peptide could indeed target the polypeptide for secretion through the Tat system, we utilized heterologous Tat-dependent protein fusions. I chose two established reporters to assess twin-arginine translocation, agarase and green fluorescent protein (GFP) (103,148).

As a proxy for Tat transport, I first used the agarase secretion reporter. Agarase is a natural actinobacterial Tat substrate in *Streptomyces violaceoruber*. Agarase activity is Tat-dependent and is easily measurable using an agar clearance assay (102). This assay has been previously used to assess whether heterologous signal peptides are directed for transport by Tat (103). We began by investigating whether agarase with the native *S. violaceoruber* signal peptide was functional when expressed in *A. oris*. The promoter from *A. oris* EF-Tu (*ana\_0022*) was used to drive expression of the agarase encoding gene *dagA* (pDagA<sub>SV</sub>). I grew *A. oris* pEV or *A. oris* pDagA<sub>SV</sub> over 48 h as a spot on a nutrient rich agar. The plates were then stained with Lugol's reagent to visually assess agar degradation. A distinct zone of clearance around the colonies demonstrated that agarase is produced, exported, and functional in *A. oris* under our conditions and that *A. oris* has no native agarase activity (Fig. 6-2A, *dagA*<sub>SV</sub> versus EV). As a control for Sec-mediated secretion, we fused the sequence encoding the FimA signal peptide to *dagA* (Fig. 6-2A, *fimA*<sup>SP</sup>-*dagA*) (62). To test whether CafA can secrete a functional agarase, I fused the sequence encoding the CafA WT signal peptide to the DagA sequence (Fig. 6-2A, *cafA*<sup>SP</sup>-*dagA*). The results of the agarase fusion suggested that CafA is not secreted by the Tat machine, as no zone of clearance is apparent in the strain expressing the *cafA*<sup>SP</sup>-*dagA* fusion. I wanted to confirm these results with a secondary reporter protein assay.

For the GFP fluorescence Tat exclusive reporter assay, I used fluorescence to measure secretion in a folded state. GFP can be targeted to the Sec pathway, and

secreted in an unfolded state, but does not maintain complete fluorescence after translocation. In contrast, when folded GFP is targeted to the Tat system, the secreted protein fluoresces (148,149). Therefore, GFP translocation and fluorescence can be assessed to determine the folding state during translocation. Additionally, I expected GFP to be more stable in the cytoplasm making it possible to assess the localization of a signal peptide fusion protein. There were no studies that identified Tat substrates in *A. oris*. Therefore, I utilized the signal peptide from *S. violaceoruber* DagA as a positive control, since I found that it is sufficient to mediate Tat transport when expressed in *A. oris* (Fig. 6-2A) (103). I first employed a kinetic assay to measure fluorescence in the supernatant over time. This assay determined that only *dagA<sup>SP</sup>-gfp* had a detectable signal above the *gfp<sup>SP-</sup>* background (Fig.6-2B). I then assayed the fraction of fluorescence in the supernatant versus the whole cell to ensure that the proteins were indeed being produced after 6 hours of growth. When the DagA signal peptide was fused to GFP 94% of the fluorescent signal was in the supernatant compared to the cytoplasmic GFP control in which fluorescence is primarily localized to the whole cell fraction (Fig. 6-2C). When the CafA signal peptide was fused to GFP this construct phenocopied the FimA signal peptide fusion, suggesting that Tat does not play a role in secreting CafA (Fig. 6-2C).



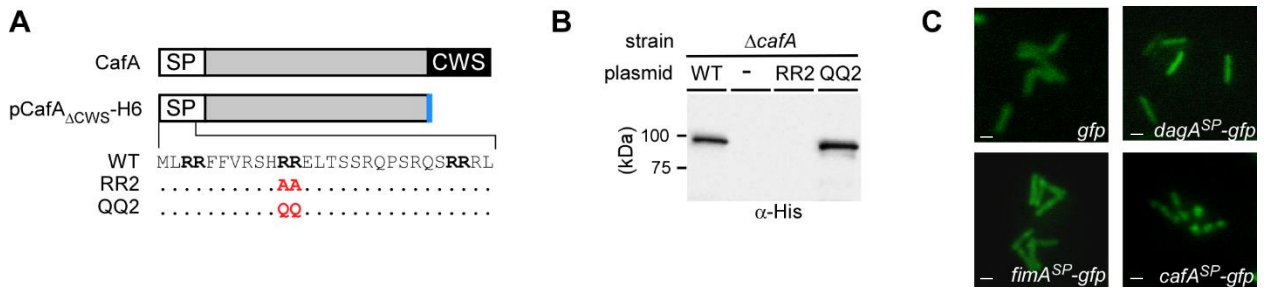


**Figure 6-2: Evaluation of Tat-dependent reporter protein fusions. (A)**

Representative images of *A. oris* CW1 colony grown on solid media overnight with pCWu10-empty vector control (EV), pCWu10 expressing the DagA ORF from *S. violaceoruber* (pDagA<sub>sv</sub>) as a Tat control, pCWu10 expressing the FimA signal peptide translationally fused to DagA lacking a signal peptide (pFimA<sup>SP</sup>-DagA) as a Sec control, or pCWu10 expressing the CafA signal peptide translationally fused to DagA lacking a signal peptide (pCafA<sup>SP</sup>-DagA). **(B)** Measurement of fluorescence intensity in the supernatant over time of strains containing pCWu10 expressing the GFP with no signal peptide (pGFP<sup>SP</sup>), pCWu10 expressing the FimA signal peptide translationally fused to GFP (pFimA<sup>SP</sup>) as a Sec control, pCWu10 expressing the DagA signal peptide translationally fused to GFP (pDagA<sup>SP</sup>) or pCWu10 expressing the CafA signal peptide translationally fused to GFP (pCafA<sup>SP</sup>). **(C)** Measurement of fluorescence intensity from the supernatant (SN) or whole cell (WC) fraction from normalized cultures in B grown to midlog phase measured in triplicate and repeated in three independent experiments.

Both of the established Tat-dependent reporter assays revealed that the CafA signal peptide is likely not utilizing the Tat pathway for secretion. Although, the RR2 mutant was defective for secretion, it was unstably and poorly expressed. To test this, I mutated the arginine residues to glutamine. Glutamine maintains a similar hydrophobicity and size as arginine, but is not positively charged and is also not sufficient to mediate Tat transport (Fig. 6-3A) (150,151). However, the mutation of RR2 to QQ was sufficient to restore translocation of CafA (Fig. 6-3B).

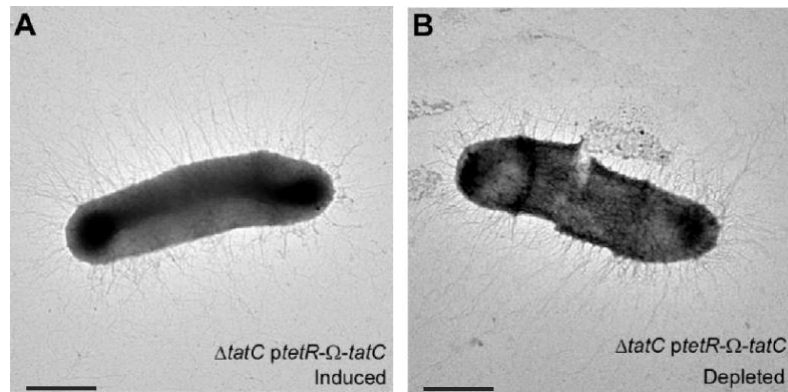
I next sought to determine the role of the CafA signal peptide since it was not used for Tat secretion. I used fluorescence microscopy to visualize the GFP alone or fused to CafA, FimA, and DagA signal peptides. The CafA signal peptide GFP fusion presented an interesting phenotype. Unlike the *gfp* control, the *cafA<sup>SP</sup>-gfp* had punctate localization at the poles and septum (Fig. 6-3C). There are two major possibilities directing this localization, 1) the CafA signal peptide mediates specific localization of the fluorescent protein, or 2) the CafA signal peptide drives the fusion protein into insoluble aggregates.



**Figure 6-3: Requirement of bulky hydrophobic residues for protein stability and examination of GFP expressing strains by fluorescence microscopy. (A)** Schematic of full-length CafA indicating the signal peptide (SP) and cell wall sorting signal (CWS). The pCafA<sub>ΔCWS</sub>-H6 variant lacks the CWS and contains a hexa-histidine tag at the C-terminus (blue rectangle). The SP is wild-type (WT), mutated to Ala as indicated in RR2 or Gln as indicated in QQ2. **(B)** Western blot analysis of the supernatant fraction from *A. oris* ΔcafA strains harboring an empty vector (-), expressing wild-type CafA signal peptide, RR2, or QQ2 signal peptide variants reacted with an α-His monoclonal antibody conjugated to HRP. **(C)** Fluorescence micrographs of strains expressing GFP signal peptide reporter fusion proteins from 6-2C. The scale bar represents 0.5 μM.

To further investigate whether Tat was playing a role in pilus formation, I decided to delete components of the Tat machine. During the course of attempting to delete *tatA* and *tatC* genes individually and together, I found that it was not possible to recover a *tat* mutant strains. Deletions in *A. oris* are made by first cloning the upstream and downstream regions of the gene of interest into a knockout vector carrying an antibiotic resistance cassette (93). The knockout vector replicates in *E. coli*, but is non-replicative in *A. oris*. Therefore, it must integrate into the *A. oris* to confer the resistance phenotype. The knockout vector also carries a gene, which is used for negative selection, *galK*. The *galK* gene encodes galactose kinase and can phosphorylate galactose or 2-deoxygalactose (2-DG). Phosphorylation of these sugar molecules retains the sugar within the cell, and in the case of 2-DG, this is lethal. Depending on the orientation of the first integration event (crossover), the outcome of the second crossover, which serves to eliminate the plasmid from the chromosome, should have a 50% chance of recovering the wild-type and 50% chance of recovering the mutant strain, if no deleterious effects are associated with the deletion. However, we were unable to recover *tatA*, *tatC* or *tatAC* deletions.

Next we constructed a conditional mutant to examine the effects of the loss of the twin arginine translocon, where the gene of interest is under the control of a tet-inducible promoter and a theophylline-responsive riboswitch. Since there is a small colony phenotype compared to wild-type in the absence of inducer for the *tatC* conditional mutant. By electron microscopy, the *tatC* conditional mutant strain appeared to have an increase in cell shearing compared to the wild-type (Fig. 6-4A and 6-4B), but piliation did not appear to be affected.



**Figure 6-4: Analysis of the *tatC* conditional mutant by electron microscopy.** The *tatC* conditional mutant was grown to midlog phase in the presence (A) or absence (B) of inducers. The cells were then immobilized on nickel grids and stained with 1% uranyl acetate prior to viewing with an electron microscope. Scale bar represents 0.5  $\mu\text{m}$ .

## Discussion

This study determined that the Tat machinery is likely not responsible for transport of CafA. However, my findings suggest that the CafA signal peptide contributes in novel ways to CafA maturation. The CafA signal peptide has 56 amino acids and is therefore much longer than an average signal peptide. A study, which examined signal peptide variety, found that Gram-positive signal peptides are an average of 30 amino acids in length with only 4.5% of signal peptides having greater than 40 amino acids. In addition, the majority having a positive net charge (52). When compared to the established actinobacterial Tat signal peptide from *S. violaceoruber* DagA, the CafA signal peptide has twice as many positively charged residues per sequence length.

The arginine residues R11 and R12 are likely required for stabilizing the CafA protein, because I was unable to detect CafA in the cytoplasmic fraction when I mutated these to alanine (A), but when I mutated these residues to glutamine (Q), which should abolish translocation if it was Tat mediated, I was able to functionally restore translocation (Fig. 6-3B). When the *cafA<sup>SP</sup>-gfp* reporter was visualized with fluorescence microscopy, it displayed punctate localization at the cell poles and cell septum compared to signal peptide lacking control reporter (Fig. 6-3C). Therefore, rather than serving as a Tat signal sequence, this signal peptide may serve to localize CafA at distinct sites in the membrane, which are also known to be where nascent peptidoglycan is deposited in actinobacteria (65). Additionally, CafA has 12 cysteines, which potentially form six disulfide bonds. It has been demonstrated in actinobacteria that disulfide bond containing proteins are folded on the exoplasmic face of the membrane by a protein called MdbA, a process called oxidative folding. Oxidative folding requires proteins to be in an unfolded state during secretion, which is not compatible with the Tat machine (63).

Following several attempts at deletion of the twin-arginine translocon (Tat), it seems likely that Tat secretion machine is essential for cell viability. Although, we were

able to construct an inducible mutant, the depletion is not robust enough to produce a viability phenotype, even with a subset of the cells displaying a “shearing” phenotype. It will be interesting in the future to determine the cause of the Tat essentiality, as it is typically non-essential.

**CHAPTER VII**  
**Perspectives and Future Implications**



## **A Model for Bacterial Type I Signal Peptidases**

*A. oris* expresses two distinct bacterial type I signal peptidases (SPases), LepB1 and LepB2. Using FimA as a model substrate I have shown that LepB2 demonstrates specificity for pilin proteins. Deletion of *lepB2* dramatically reduces pilus assembly and affects pilus-related phenotypes, such as biofilm formation. Conversely, LepB1 and LepB2 can both act on a separate substrate, GspA (80). Having only two enzymes and a semi-defined substrate repertoire makes the *A. oris* system ideal to study type I SPase enzymes and to address some of the remaining questions in the field. Firstly, how do SPases select specific substrates? This is especially intriguing when the SPase enzymes are not temporally separated and I have shown that *lepB1* and *lepB2* are expressed independently during logarithmic growth. Are the specificity determinants present in the signal peptides, the SPase itself, or do both contribute? Finally, *lepB1* and *lepB2* are the only traditional SPase enzymes present in the *A. oris* genome, but these genes are not essential for cell viability and loss of both produces ragged termini in the secreted substrates. So, does an uncharacterized secretion stress pathway exist in *A. oris*?

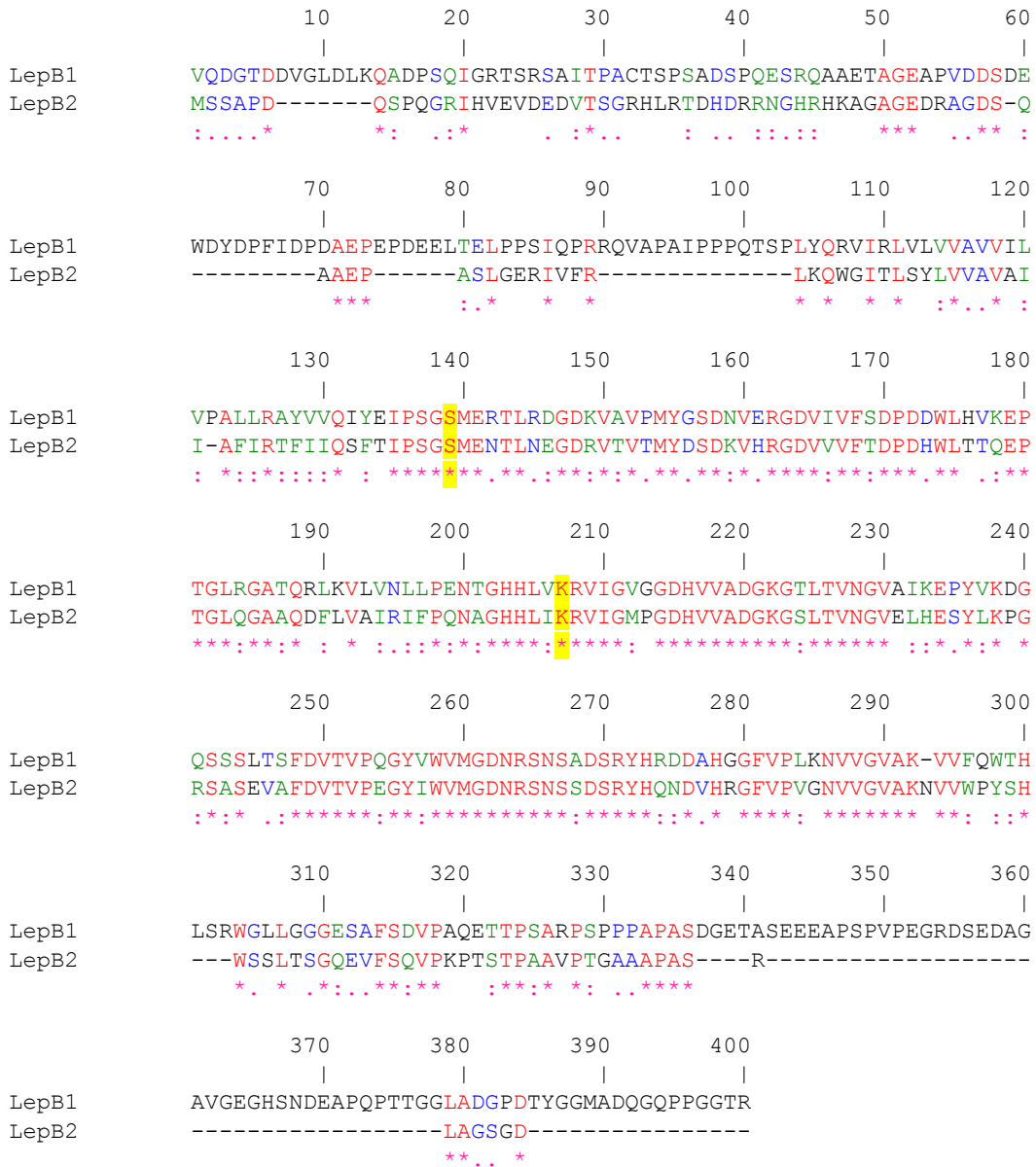
### ***Specificity of bacterial type I signal peptidases***

The signal peptides of the pilin proteins do not differ significantly from the signal peptide of GspA. In fact, in bacteria most signal peptides do not exhibit sequence conservation aside from small neutral amino acids at positions P-1, P-3, and a helix-breaking residue in the P-6 position that are required for cleavage of all type I SPase substrates (53). Additionally, LepB1 and LepB2 exhibit high similarity at the amino acid level with the exception that LepB1 has an extended C-terminus, but these regions do not correspond to specificity clefts present in the well-studied *E. coli* type I SPase (Fig. 7-1) (52). Gram-positive bacteria commonly encode numerous SPase I enzymes (60). For example, *B. subtilis* encodes five SPases, and *Streptomyces lividans* encodes four

of these proteins (109,152). Still, it remains unknown why multiple SPases exist in Gram-positive bacteria. It has been shown that in *S. lividans* and *A. oris* SPases have non-identical substrate preference suggesting that these SPases have specialized to process different substrates (80,153). Moreover, what determines SPase substrate specificity remains an open question. The *A. oris* dual SPase arrangement is useful for addressing this question in particular, because there are definitive phenotypes associated with the deletion of pilin specific LepB2 that do not affect cell viability.

A first step toward determining factors involved in signal peptidase specificity would be to assess the secreted protein profile of a *lepB2* deletion strain compared to wild-type. Similar studies have been successfully performed in *S. lividans*, which also has clear substrate specificity and no cell viability phenotype, but specificity profiling focused on only a single substrate (108,153). A total secretome including proteins from the supernatant, cell wall, and membrane could be analyzed in *A. oris* by performing two-dimensional protein electrophoresis on the *lepB2* deletion and comparing it to the wild-type secretome. LepB2 specific processing could be confirmed similarly to the methods that I described for FimA (80). I also found previously that surface exposure decreases and the profile of pilus polymers changes in the absence of *lepB2*, and these substrates can serve as a control. Identification of additional LepB2-specific substrates could lead to a greater understanding about SPase I substrate selection. After analyzing the secretome of LepB2 specific substrates, similarities between these proteins can be analyzed to look for specificity determinants.

FimA serves as a fully characterized substrate of LepB2 and may be used to determine whether features in the signal peptide or signal peptidase affect substrate preference. A screen for biofilm formation could be performed with a large, random pool of FimA signal peptide mutants. Transfer and sequencing of the biofilm defective mutants would reveal sequences important for type I SPase processing. Also, analysis



**Figure 7-1: Sequence alignment of *A. oris* LepB1 and LepB2.** Sequence alignment of *A. oris* LepB1 and LepB2 was performed using CLUSTALW (112). The conserved Ser-Lys catalytic dyad is highlighted in yellow. Asterisks (\*) indicate fully conserved residues and the colon (: ) indicates conserved residue properties.

of the SPase itself could be done by creating a chimeric LepB2 with the LepB1 C-terminal extension or eliminating the extension from LepB1. If these constructs modulate FimA processing, further analysis can determine how the C-terminal extension affects substrate selection.

FimA serves as a fully characterized substrate of LepB2, and may be used to determine whether features in the signal peptide or signal peptidase affect substrate preference. A screen for biofilm formation could be performed with a large, random pool of FimA signal peptide mutants. Transfer and sequencing of the biofilm defective mutants would reveal sequences important for type I SPase processing. Also, analysis of the SPase itself could be done by creating a chimeric LepB2 with the LepB1 C-terminal extension, or eliminating the extension from LepB1. If these constructs modulate FimA processing, further analysis can determine how the C-terminal extension affects substrate selection.

Finally, analysis of the substrate pool and identifying factors contributing to specificity may provide insight into the glycosylation pathway of GspA. Initially, *lepB2* was isolated because it could suppress *srtA* essentiality (50). Surprisingly, it was found that this was not due to a direct effect on the GspA substrate, but rather it modified the glycosylation profile and accumulated low molecular mass GspA products in the absence of *lepB2* (80). Therefore, one of the substrates of LepB2 may be a glycosylation enzyme, and absence of specific cleavage could decrease enzyme efficiency. This enzyme may be part of the changed pool upon isolation of LepB2 specific substrates.

### ***Secretion stress pathway***

Deletion of both *lepB1* and *lepB2* from *A. oris* has a minor effect on growth kinetics and produces a ragged N-terminus in the shared substrate GspA (Fig. 3-1C and Table 3-1) (80). Endopeptidases, such as signal peptidases, cleave precisely generating

a specific fragment. Ragged termini are typically indicative of an exopeptidase activity, as they trim the polypeptide processively, but may also result from a new cleavage site from an intramembrane protease (154). Thus our results suggest that there may be a robust, yet undefined pathway for *A. oris* to cope with secretion stress resulting from a block in signal peptide cleavage. A secretion stress system is present in *S. aureus*, where the *ayrRABC* operon has been identified as encoding an alternative path that can be utilized in the absence of a traditional type I SPase. Indeed, expression of *ayrABC* results in imperfect cleavage sites like those identified for GspA. This operon was identified under treatment with the antibiotic arylomycin that acts by inhibiting type I SPase enzymes (155). Arylomycin works synergistically with  $\beta$ -lactam antibiotics on methicillin-resistant *S. aureus* by specifically targeting and inhibiting SPase I SpsB activity (156).

If arylomycin is effective against *A. oris* LepB1 and LepB2 and does not produce a growth defect, it could be utilized to screen for synthetic lethal mutants against a transposon library. By patch plating onto plates with and without arylomycin, those mutants that cannot grow in the presence of arylomycin would be selected for sequencing. The identified genes could be analyzed, especially for their effect on secreted proteins and whether their gene products exhibit proteolytic activity.

### **A New Paradigm for an LCP Enzyme**

LcpA is a unique LCP enzyme because of its ability to modify the protein substrate GspA with glycopolymers. This is not a conserved feature for all LCP enzymes, even those present in *A. oris*, and in fact it has never been demonstrated before in any organism (Fig. 4-1B) (81). My work has uncovered that LcpA acts as a phosphotransferase to transfer a lipid-linked glycopolymer to GspA, and this represents the terminal step of GspA glycosylation. The composition of the glycan transferred to GspA remains completely unknown. Beyond that, what genes are involved in synthesizing these

glycopolymers? It is important to address these questions in *A. oris*, because although the activity of LcpA is functionally similar to other LCP enzymes, how recognition of the acceptor substrate occurs prior to glycan transfer is not well understood. Characterizing LcpA could uncover information applicable to all LCP enzyme acceptor selection.

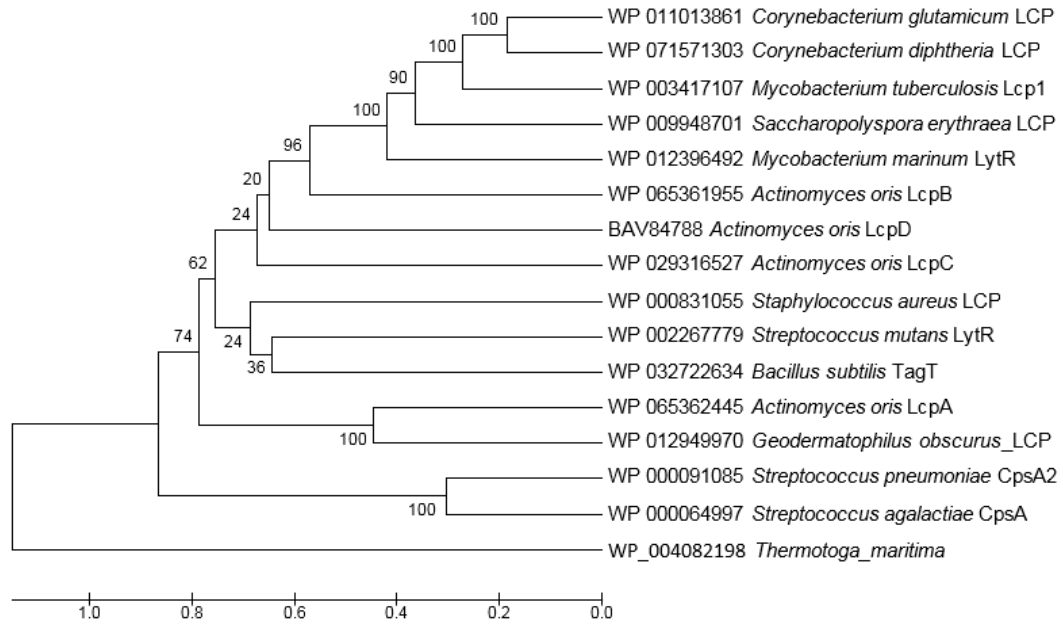
### ***LcpA selects novel acceptor substrate***

LcpA seems to be distantly phylogenetically related to currently characterized enzymes and even other enzymes encoded by *A. oris* (Fig. 7-2). The closest match to LcpA by BLAST search appears to be an LCP enzyme encoded by the actinobacterium *Geodermatophilus obscurus*. It may be interesting to determine whether the *G. obscurus* LCP enzyme can be expressed heterologously in *A. oris* and whether an LcpA-GspA like relationship exists in this species. LcpA may serve as the founding member of a new type of LCP, but until additional members are identified it will be difficult to elucidate the evolutionary changes that drove LcpA to select a new acceptor.

Regardless of whether LcpA is the sole member of its group, it is still unknown how LcpA specifically selects the GspA protein acceptor. The catalytic core of the LcpA enzyme exhibits high structural similarity with other enzymes, especially *B. subtilis* TagT (Fig. 4-2). Many of the biochemical activities first described for wall teichoic LCP enzymes are also conserved in LcpA and require similar molecular features. Therefore, it is likely that the substrate selection domain exists outside of this conserved region. Remodeling these domains or targeting mutagenesis to the sequences encoding structurally dissimilar regions may reveal how GspA is selected by LcpA. Additionally, the antibodies and pyrophosphatase assay presented in this thesis can be adapted to ensure LcpA stability and activity remain functional in the mutant constructs. GspA glycosylation can be simply assessed by immunoblotting for GspA glycosylation.

### ***GspA glycan profile***

Another unknown feature regarding GspA is the glycopolymer identity, which unfortunately makes it difficult to recreate an accurate substrate for *in vitro* reactions or to understand the function of the glycosylated GspA. Lectin-based profiling is a technique that would be useful to begin characterizing the composition of the glycan present on the GspA low and high molecular mass moieties. Lectins are proteins that recognize sugars with high specificity. This methodology has been useful and successful in other applications, and high throughput arrays have been adapted from the low throughput technology (157). Many lectins are available conjugated to either a fluorescent or peroxidase molecule to facilitate their application. To narrow down which glycans are present in GspA LMM and HMM, a technique called lectin-blotting could be accomplished using these labeled lectins (158). Purified GspA LMM or HMM separated by SDS PAGE would be reacted with a series of lectins that each recognize a unique sugar moiety. This information would then be incorporated to further analyze the glycan using enzymatic digestion and mass spectrometry techniques (159).



**Figure 7-2: Phylogenetic analysis of LCP proteins.** The tree was rooted with the HD-domain containing protein from *Thermotoga maritima* (WP\_004082198) and constructed by the mega6 program. Numbers at nodes represent percentage levels of bootstrap support based on the unweighted pair group method with arithmetic mean of 1000 resampled datasets.



## ***GspA glycan synthetic pathway***

My study highlights several similarities between *A. oris* LcpA and WTA synthesis. However, no genes upstream of LcpA involved in the synthesis of GspA glycopolymers have been identified. We may exploit knowledge and tools derived from the well-studied *S. aureus* wall teichoic acid synthesis pathway, which begins in the cytoplasm and is mediated by *tar* genes to determine whether a parallel pathway exists in *A. oris*. Chemical inhibitors and their targets have been identified that affect synthesis of the teichoic acid in *S. aureus* (160,161). Early steps in the pathway, initiated by *tarO* and *tarA*, are non-essential for cell viability. *tarO* transfers a GlcNAc molecule onto Und, and *tarA* transfers a ManNAc sugar onto the GlcNAc-Und precursor. The GlcNAc-ManNAc priming sugars on undecaprenol (Und) is a product that can be recycled if the process is stalled. The first step, transfer of GlcNAc by *tarO*, can be inhibited with sub-lethal concentrations of tunicamycin (160). The remaining steps responsible for construction of the teichoic acid using the *tarOA* Und-GlcNAc-ManNAc platform, involve genes *tarBFLGH*. These downstream genes are essential for cell viability likely because the Und can no longer be recycled at this stage, which would stall other cellular processes. The *tarBFL* genes are responsible for adding glycerol-3-phosphate and ribitol-5-phosphate to synthesize the variable length lipid-linked teichoic acid. The completed teichoic acid is subsequently flipped across the membrane by *tarGH* gene products. A second WTA inhibitor compound called targosil targets *tarG* inhibiting teichoic acid flipping and subsequently eliminating its incorporation into the cell wall, where it accumulates at the cytoplasmic side of the membrane (161). Potentially, the WTA inhibitors tunicamycin and targosil can be used in *A. oris* to assess whether these inhibitors affect GspA glycosylation mirroring WTA synthesis. If the molecules can inhibit GspA glycosylation, they can be used to identify genes in the synthesis pathway, which will likely have similar characteristics to the *S. aureus* genes.

Many of the features that I have uncovered regarding LcpA structure and activity are similar to those described for LCP proteins with a role in WTA synthesis. Therefore, future work should focus on how the differences displayed by LcpA drive its novel substrate selection. However, given that WTA synthesis is well-understood and there are more tools available to study it, using the properties from this pathway as a starting point will undoubtedly contribute to determining the overall GspA glycosylation pathway.

### **Interkingdom Adhesin with a Pivotal Role in Oral Biofilm Formation**

CafA is a pilus-associated adhesin that binds a receptor polysaccharide (RPS) present on some oral streptococci, and I found that it can also mediate adherence to human cells under certain conditions (25). *A. oris* is primary colonizer of the oral biofilm and CafA plays a major role the niche selection. What structural characteristics confer CafA the ability to bind RPS and why does binding require extension on the pilus protein? The tertiary structure of CafA is unknown, in part because CafA cannot be stably purified from a traditional *E. coli* over-expression system. I utilized an alternative approach of comparison of CafA amino acid sequences from a strain that displays a negative coaggregation phenotype. With this approach, I identified a several residues important for coaggregation. Are these residues present in a conserved domain and do they similarly affect human cell binding? Finally, does the signal peptide of CafA contribute to localization and pilus assembly in a way which has not yet been explored?

### ***Structural identity of CafA***

A CafA structure would be useful to better understand the structural components required for RPS binding as well as assigning the identified residues necessary for CafA adherence. There have been technical challenges purifying CafA from an *E. coli* overexpression system likely because of the large size and number of cysteines that

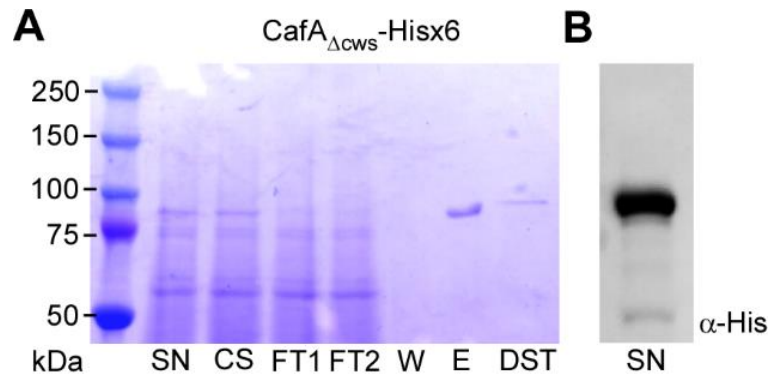
potentially form disulfide bonds. To simplify purification of full length CafA, I developed a native expression system; although it is plagued by low yield, it produces pure, stable, full-length protein (Fig. 7-3). A structure of CafA would also be informative about the nature of the disulfide bond scheme and how these contribute to maintaining the structural integrity of this protein.

### ***CafA strain N11A12 glycan specificity***

It is unknown whether the changes present in CafA from *A. oris* strain N11A12 result in reduced binding capacity for all glycans, or whether the mutations confer a new glycan specificity. To test this, a glycan spot array using purified proteins from *A. oris* MG1 and *A. oris* N11A12 could be employed to determine whether a new specificity of N11A12 exists. This would be analogous to *E. coli*, where the chaperone-usher tip pilins FimH and FmlH have a similar structure but exhibit an entirely different glycan binding specificity. FmlH adheres only under inflammatory conditions, when the sialic acid production is downregulated and the receptor is uncovered (135). There could be a similar mechanism for CafA, where the N11A12 has a new target.

### ***CafA signal peptide localization and pilus assembly***

From the experiments outlined here, the CafA signal peptide does not seem to have the capacity to translocate proteins through the Tat pathway. However, using fluorescence microscopy, it is clear that when GFP is fused to the CafA signal peptide it displays a punctate localization. In contrast, the Tat-targeted GFP and the GFP alone have a diffuse localization. Therefore, the CafA signal peptide may be localizing GFP to areas of nascent peptidoglycan synthesis, where under native conditions CafA would initiate synthesis of a CafA-containing type 2 pilus. However, these puncta could also be a result of protein aggregation in the cell. To distinguish between these possibilities, the *cafA<sup>SP</sup>-gfp* strain could be fractionated into the soluble and insoluble fractions,



**Figure 7-3. Purification of CafA from *A. oris* supernatant.** (A) Coomassie blue stained total supernatant (SN), cleared supernatant (CS), flow-through (FT), wash (W), eluate (E) and pre-concentration, post-desalt (DST) fractions from *A. oris* expressing pCafA<sub>ΔCWS</sub>-Hisx6 separated by SDS-PAGE. (B) Immunoblot of the supernatant (SN) fraction with an α-His monoclonal antibody.

immunoblotted with an  $\alpha$ -GFP antibody, and compared to the signal from the *gfp* only expressing strain. If *cafA<sup>SP</sup>-gfp* is soluble, it would suggest genuine localization by the CafA signal peptide. The *cafA<sup>SP</sup>-gfp* could then be immunoprecipitated by  $\alpha$ -GFP to determine whether there are additional factors, like chaperones, involved in localizing this protein to the midcell and poles. Then, confirmed with immunofluorescence of the native CafA protein.

### **Essentiality of the Twin-Arginine Translocon**

From the work presented here, the twin-arginine translocon (Tat) appears to be essential in *A. oris*. Essentiality of the twin arginine transporter has been demonstrated previously in the actinobacterium *Mycobacterium tuberculosis*, but was never linked to a particular substrate (162). By bioinformatic analysis of secreted proteins, there are 43 potential Tat substrates in *A. oris* based on the presence of a signal peptide and a twin arginine motif within the first 50 amino acids (Table 7-1). Which of these predicted substrates actually requires the Tat machine for secretion? Appending the signal peptides to the agarase reporter with those from proteins on the list may provide a better view about what constitutes a Tat signal peptide in *A. oris*, especially because the CafA signal peptide is a predicted to be a Tat signal peptide, yet failed to display Tat-dependency.

Of these potential Tat substrates, which contribute to the essentiality of the Tat pathway in *A. oris*? The NADH dehydrogenase (*ana\_0411*) was identified on this list and has been demonstrated to be a bona fide Tat substrate and is a potential cause for loss of aerobic growth in *Corynebacterium glutamicum*, but is not essential for viability (163). Additionally, several subunits of the NADH dehydrogenase complex could be deleted from *A. oris* in a previous study (125). Confirming which of the signal peptides can indeed

**Table 7-1: Predicted Tat-dependent proteins in the *A. oris* genome.** The *A. oris* MG1 proteome was searched for consecutive “RR” within first 50 amino acids, then filtered through TatP to confirm Tat consensus and signal peptide cleavage site.

<b>Annotation</b>	<b>Los Alamos ID</b>
OmpA/MotB	ana_0003
OmpA/MotB	ana_0004
OmpA/MotB	ana_0005
OmpA/MotB	ana_0008
Lipoprotein, putative	ana_0009
OmpA family	ana_0013
LPXTG-motif protein (FimB)	ana_0023
LPXTG-motif protein	ana_0196
Zn/Mn transport substrate binding protein	ana_0218
Oxidoreductase	ana_0238
Tannase	ana_0255
Glycosyl/glycerophosphate transferase teichoic acid biosynthesis	ana_0303
Polysaccharide deacetylase	ana_0318
Peptidase, M23/M37 family	ana_0332
Hypothetical protein	ana_0387
NADH dehydrogenase	ana_0411
ABC-type multidrug transport, ATPase	ana_0458
Peptidylprolyl isomerase, FKBP type	ana_0589
Virulence factor MVIN family	ana_1001
Hypothetical protein	ana_1354
ABC-type Co/Fe siderophore transport	ana_1460
Von Willebrand factor, type A domain containing	ana_1486
Mannose-6-phosphate isomerase	ana_1537
Hypothetical protein	ana_1540
ABC-type proline/glycine betaine transport system	ana_1633
ABC-type multiple sugar transport system	ana_1639
Peptide/Ni transport system	ana_1692
Hydrolase, alpha/beta fold family	ana_1727
Multicopper oxidase, type 3	ana_1755
Lipoprotein	ana_1762
Hly-III family	ana_1803
Tat pathway signal sequence domain	ana_1831
Hypothetical protein	ana_1864
Polysaccharide deacetylase	ana_1914
Glycoside hydrolase	ana_1967
Membrane-fusion protein	ana_2031
Cobalamin/Fe siderophore transport system	ana_2051
CnaB domain containing (CafA)	ana_2235
Hypothetical protein	ana_2283
Usher-like protein precursor (FimQ)	ana_2509
Type 1 fimbrial major subunit (FimP)	ana_2510
Exo-alpha-sialidase	ana_2709
ABC-type cobalamin/Fe <sup>3+</sup> siderophores transport system	ana_2758

mediate Tat transport from the predictions using the agarase reporter would significantly narrow down which of the proteins on the list may be responsible for Tat essentiality. Together these considerations make it worth exploring the role of these gene products in bacterial physiology, and the strict requirement for the machine may explain why actinobacteria utilize the apparatus for export of critical products.

### **Final Remarks**

Cell surface proteins and glycoconjugates have a significant impact on bacterial physiology. My studies have addressed several aspects of the molecular assembly and display of these proteins. Cell surface proteins require secretion prior to fulfilling their function. I have found that a specific type I SPase is necessary for pilus polymerization, and also applied tools to investigate Tat secretion in *A. oris*. I have found that in the exoplasmic space, LcpA utilizes a conserved mechanism to mediate glycosylation of a novel acceptor substrate GspA. Finally, I demonstrated that *Actinomyces* clinical isolates exhibit differential CafA display and exploited these differences to identify key amino acid changes that contribute to the binding of *A. oris* MG1 CafA.

Studies of the fascinating microbes in the oral communities continue to reveal their importance in human health and disease. These works illuminate the complex components necessary for *A. oris* to fulfill its function as primary colonizer of the oral biofilm. These studies have also uncovered new tools and information, which is applicable to principles of biofilm formation, interspecies interactions, glycoconjugate formation, and bacterial pathogenesis.

## BIBLIOGRAPHY

1. Huang, R., Li, M., and Gregory, R. L. (2011) Bacterial interactions in dental biofilm. *Virulence* **2**, 435-444
2. Rickard, A. H., Gilbert, P., High, N. J., Kolenbrander, P. E., and Handley, P. S. (2003) Bacterial coaggregation: an integral process in the development of multi-species biofilms. *Trends Microbiol* **11**, 94-100
3. Howell, A., Jr., Murphy, W. C., 3rd, Paul, F., and Stephan, R. M. (1959) Oral strains of *Actinomyces*. *J Bacteriol* **78**, 82-95
4. Kolenbrander, P. E. (1993) Coaggregation of human oral bacteria: potential role in the accretion of dental plaque. *J Appl Bacteriol* **74 Suppl**, 79S-86S
5. Kolenbrander, P. E., Ganeshkumar, N., Cassels, F. J., and Hughes, C. V. (1993) Coaggregation: specific adherence among human oral plaque bacteria. *FASEB J* **7**, 406-413
6. Dewhirst, F. E., Chen, T., Izard, J., Paster, B. J., Tanner, A. C., Yu, W. H., Lakshmanan, A., and Wade, W. G. (2010) The human oral microbiome. *J Bacteriol* **192**, 5002-5017
7. Kolenbrander, P. E., Palmer, R. J., Jr., Periasamy, S., and Jakubovics, N. S. (2010) Oral multispecies biofilm development and the key role of cell-cell distance. *Nat Rev Microbiol* **8**, 471-480
8. Periasamy, S., and Kolenbrander, P. E. (2009) Mutualistic biofilm communities develop with *Porphyromonas gingivalis* and initial, early, and late colonizers of enamel. *J Bacteriol* **191**, 6804-6811
9. Mark Welch, J. L., Rossetti, B. J., Rieken, C. W., Dewhirst, F. E., and Borisy, G. G. (2016) Biogeography of a human oral microbiome at the micron scale. *Proc Natl Acad Sci U S A* **113**, E791-800
10. Marsh, P. D. (2005) Dental plaque: biological significance of a biofilm and community life-style. *J Clin Periodontol* **32 Suppl 6**, 7-15



11. Hojo, K., Nagaoka, S., Ohshima, T., and Maeda, N. (2009) Bacterial interactions in dental biofilm development. *J Dent Res* **88**, 982-990
12. Hajishengallis, G. (2015) Periodontitis: from microbial immune subversion to systemic inflammation. *Nat Rev Immunol* **15**, 30-44
13. Petersen, P. E., and Ogawa, H. (2005) Strengthening the prevention of periodontal disease: the WHO approach. *J Periodontol* **76**, 2187-2193
14. Offenbacher, S., Jared, H. L., O'Reilly, P. G., Wells, S. R., Salvi, G. E., Lawrence, H. P., Socransky, S. S., and Beck, J. D. (1998) Potential pathogenic mechanisms of periodontitis associated pregnancy complications. *Ann Periodontol* **3**, 233-250
15. Beck, J. D., and Offenbacher, S. (2005) Systemic effects of periodontitis: epidemiology of periodontal disease and cardiovascular disease. *J Periodontol* **76**, 2089-2100
16. Dige, I., Raarup, M. K., Nyengaard, J. R., Kilian, M., and Nyvad, B. (2009) *Actinomyces naeslundii* in initial dental biofilm formation. *Microbiology* **155**, 2116-2126
17. Brennan, M. J., Cisar, J. O., Vatter, A. E., and Sandberg, A. L. (1984) Lectin-dependent attachment of *Actinomyces naeslundii* to receptors on epithelial cells. *Infect Immun* **46**, 459-464
18. Wu, C., Mishra, A., Yang, J., Cisar, J. O., Das, A., and Ton-That, H. (2011) Dual function of a tip fimbrillin of *Actinomyces* in fimbrial assembly and receptor binding. *J Bacteriol* **193**, 3197-3206
19. Nyvad, B., and Kilian, M. (1987) Microbiology of the early colonization of human enamel and root surfaces in vivo. *Scand J Dent Res* **95**, 369-380
20. Nyvad, B., and Fejerskov, O. (1987) Scanning electron microscopy of early microbial colonization of human enamel and root surfaces in vivo. *Scand J Dent Res* **95**, 287-296

21. Ton-That, H., Das, A., and Mishra, A. (2011) *Actinomyces oris fimbriae: An adhesive principle in bacterial biofilms and tissue tropism*, ASM Press
22. Ritz, H. L. (1967) Microbial population shifts in developing human dental plaque. *Arch Oral Biol* **12**, 1561-1568
23. Socransky, S. S., Manganiello, A. D., Propas, D., Oram, V., and van Houte, J. (1977) Bacteriological studies of developing supragingival dental plaque. *J Periodontal Res* **12**, 90-106
24. Zijngel, V., van Leeuwen, M. B., Degener, J. E., Abbas, F., Thurnheer, T., Gmur, R., and Harmsen, H. J. (2010) Oral biofilm architecture on natural teeth. *PLoS One* **5**, e9321
25. Reardon-Robinson, M. E., Wu, C., Mishra, A., Chang, C., Bier, N., Das, A., and Ton-That, H. (2014) Pilus hijacking by a bacterial coaggregation factor critical for oral biofilm development. *Proc Natl Acad Sci U S A* **111**, 3835-3840
26. Cisar, J. O., McIntire, F. C., and Vatter, A. E. (1978) Fimbriae of *Actinomyces viscosus* t14v: their relationship to the virulence-associated antigen and to coaggregation with *Streptococcus sanguis* 34. *Adv Exp Med Biol* **107**, 695-701
27. Suzuki, N., Yoneda, M., and Hirofujii, T. (2013) Mixed red-complex bacterial infection in periodontitis. *Int J Dent* **2013**, 587279
28. Girard, A. E., and Jacius, B. H. (1974) Ultrastructure of *Actinomyces viscosus* and *Actinomyces naeslundii*. *Arch Oral Biol* **19**, 71-79
29. Yeung, M. K., Donkersloot, J. A., Cisar, J. O., and Ragsdale, P. A. (1998) Identification of a gene involved in assembly of *Actinomyces naeslundii* T14V type 2 fimbriae. *Infect Immun* **66**, 1482-1491
30. Cisar, J. O., Barsumian, E. L., Curl, S. H., Vatter, A. E., Sandberg, A. L., and Siraganian, R. P. (1981) Detection and localization of a lectin on *Actinomyces viscosus* T14V by monoclonal antibodies. *J Immunol* **127**, 1318-1322

31. Clark, W. B., Wheeler, T. T., and Cisar, J. O. (1984) Specific inhibition of adsorption of *Actinomyces viscosus* T14V to saliva-treated hydroxyapatite by antibody against type 1 fimbriae. *Infect Immun* **43**, 497-501
32. Cisar, J. O., Sandberg, A. L., and Mergenhagen, S. E. (1984) The function and distribution of different fimbriae on strains of *Actinomyces viscosus* and *Actinomyces naeslundii*. *J Dent Res* **63**, 393-396
33. Cisar, J. O., Curl, S. H., Kolenbrander, P. E., and Vatter, A. E. (1983) Specific absence of type 2 fimbriae on a coaggregation-defective mutant of *Actinomyces viscosus* T14V. *Infect Immun* **40**, 759-765
34. Revis, G. J., Vatter, A. E., Crowle, A. J., and Cisar, J. O. (1982) Antibodies against the Ag2 fimbriae of *Actinomyces viscosus* T14V inhibit lactose-sensitive bacterial adherence. *Infect Immun* **36**, 1217-1222
35. Ton-That, H., and Schneewind, O. (2003) Assembly of pili on the surface of *Corynebacterium diphtheriae*. *Mol Microbiol* **50**, 1429-1438
36. Ton-That, H., Marraffini, L. A., and Schneewind, O. (2004) Sortases and pilin elements involved in pilus assembly of *Corynebacterium diphtheriae*. *Mol Microbiol* **53**, 251-261
37. Silhavy, T. J., Kahne, D., and Walker, S. (2010) The bacterial cell envelope. *Cold Spring Harb Perspect Biol* **2**, a000414
38. Tytgat, H. L., and Lebeer, S. (2014) The sweet tooth of bacteria: common themes in bacterial glycoconjugates. *Microbiol Mol Biol Rev* **78**, 372-417
39. Weidenmaier, C., and Peschel, A. (2008) Teichoic acids and related cell-wall glycopolymers in Gram-positive physiology and host interactions. *Nat Rev Microbiol* **6**, 276-287
40. Schneewind, O., and Missiakas, D. (2014) Sec-secretion and sortase-mediated anchoring of proteins in Gram-positive bacteria. *Biochim Biophys Acta* **1843**, 1687-1697

41. Vollmer, W., Blanot, D., and de Pedro, M. A. (2008) Peptidoglycan structure and architecture. *FEMS Microbiol Rev* **32**, 149-167
42. Neuhaus, F. C., and Baddiley, J. (2003) A continuum of anionic charge: structures and functions of D-alanyl-teichoic acids in gram-positive bacteria. *Microbiol Mol Biol Rev* **67**, 686-723
43. Marquis, R. E., Mayzel, K., and Carstensen, E. L. (1976) Cation exchange in cell walls of gram-positive bacteria. *Can J Microbiol* **22**, 975-982
44. Kawai, Y., Marles-Wright, J., Cleverley, R. M., Emmins, R., Ishikawa, S., Kuwano, M., Heinz, N., Bui, N. K., Hoyland, C. N., Ogasawara, N., Lewis, R. J., Vollmer, W., Daniel, R. A., and Errington, J. (2011) A widespread family of bacterial cell wall assembly proteins. *EMBO J* **30**, 4931-4941
45. Sadovskaya, I., Vinogradov, E., Courtin, P., Armalyte, J., Meyrand, M., Giaouris, E., Palussiere, S., Furlan, S., Pechoux, C., Ainsworth, S., Mahony, J., van Sinderen, D., Kulakauskas, S., Guerardel, Y., and Chapot-Chartier, M. P. (2017) Another Brick in the Wall: a Rhamnan Polysaccharide Trapped inside Peptidoglycan of *Lactococcus lactis*. *MBio* **8**
46. Eberhardt, A., Hoyland, C. N., Vollmer, D., Bisle, S., Cleverley, R. M., Johnsborg, O., Havarstein, L. S., Lewis, R. J., and Vollmer, W. (2012) Attachment of capsular polysaccharide to the cell wall in *Streptococcus pneumoniae*. *Microb Drug Resist* **18**, 240-255
47. Harrison, J., Lloyd, G., Joe, M., Lowary, T. L., Reynolds, E., Walters-Morgan, H., Bhatt, A., Lovering, A., Besra, G. S., and Alderwick, L. J. (2016) Lcp1 Is a Phosphotransferase Responsible for Ligating Arabinogalactan to Peptidoglycan in *Mycobacterium tuberculosis*. *MBio* **7**
48. Chan, Y. G., Kim, H. K., Schneewind, O., and Missiakas, D. (2014) The capsular polysaccharide of *Staphylococcus aureus* is attached to peptidoglycan by the LytR-CpsA-Psr (LCP) family of enzymes. *J Biol Chem* **289**, 15680-15690

49. Koster, S., Upadhyay, S., Chandra, P., Papavinasasundaram, K., Yang, G., Hassan, A., Grigsby, S. J., Mittal, E., Park, H. S., Jones, V., Hsu, F. F., Jackson, M., Sasseti, C. M., and Philips, J. A. (2017) Mycobacterium tuberculosis is protected from NADPH oxidase and LC3-associated phagocytosis by the LCP protein CpsA. *Proc Natl Acad Sci U S A* **114**, E8711-E8720
50. Wu, C., Huang, I. H., Chang, C., Reardon-Robinson, M. E., Das, A., and Ton-That, H. (2014) Lethality of sortase depletion in *Actinomyces oris* caused by excessive membrane accumulation of a surface glycoprotein. *Mol Microbiol* **94**, 1227-1241
51. Siegel, S. D., Reardon, M. E., and Ton-That, H. (2017) Anchoring of LPXTG-Like Proteins to the Gram-Positive Cell Wall Envelope. *Curr Top Microbiol Immunol* **404**, 159-175
52. Choo, K. H., and Ranganathan, S. (2008) Flanking signal and mature peptide residues influence signal peptide cleavage. *BMC Bioinformatics* **9 Suppl 12**, S15
53. Auclair, S. M., Bhanu, M. K., and Kendall, D. A. (2012) Signal peptidase I: cleaving the way to mature proteins. *Protein Sci* **21**, 13-25
54. Natale, P., Bruser, T., and Driessen, A. J. (2008) Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane--distinct translocases and mechanisms. *Biochim Biophys Acta* **1778**, 1735-1756
55. Freudl, R. (2013) Leaving home ain't easy: protein export systems in Gram-positive bacteria. *Res Microbiol* **164**, 664-674
56. Palmer, T., and Berks, B. C. (2012) The twin-arginine translocation (Tat) protein export pathway. *Nat Rev Microbiol* **10**, 483-496
57. Goosens, V. J., Monteferrante, C. G., and van Dijl, J. M. (2014) The Tat system of Gram-positive bacteria. *Biochim Biophys Acta* **1843**, 1698-1706
58. Cline, K. (2015) Mechanistic Aspects of Folded Protein Transport by the Twin Arginine Translocase (Tat). *J Biol Chem* **290**, 16530-16538

59. Paetzel, M. (2014) Structure and mechanism of Escherichia coli type I signal peptidase. *Biochim Biophys Acta* **1843**, 1497-1508
60. van Roosmalen, M. L., Geukens, N., Jongbloed, J. D., Tjalsma, H., Dubois, J. Y., Bron, S., van Dijl, J. M., and Anne, J. (2004) Type I signal peptidases of Gram-positive bacteria. *Biochim Biophys Acta* **1694**, 279-297
61. Kol, S., Nouwen, N., and Driessen, A. J. (2008) Mechanisms of YidC-mediated insertion and assembly of multimeric membrane protein complexes. *J Biol Chem* **283**, 31269-31273
62. Reardon-Robinson, M. E., Osipiuk, J., Chang, C., Wu, C., Jooya, N., Joachimiak, A., Das, A., and Ton-That, H. (2015) A Disulfide Bond-forming Machine Is Linked to the Sortase-mediated Pilus Assembly Pathway in the Gram-positive Bacterium *Actinomyces oris*. *J Biol Chem* **290**, 21393-21405
63. Reardon-Robinson, M. E., and Ton-That, H. (2015) Disulfide-Bond-Forming Pathways in Gram-Positive Bacteria. *J Bacteriol* **198**, 746-754
64. Luong, T. T., Reardon-Robinson, M. E., Siegel, S. D., and Ton-That, H. (2017) Reoxidation of the Thiol-Disulfide Oxidoreductase MdbA by a Bacterial Vitamin K Epoxide Reductase in the Biofilm-Forming Actinobacterium *Actinomyces oris*. *J Bacteriol* **199**
65. Reardon-Robinson, M. E., Osipiuk, J., Jooya, N., Chang, C., Joachimiak, A., Das, A., and Ton-That, H. (2015) A thiol-disulfide oxidoreductase of the Gram-positive pathogen *Corynebacterium diphtheriae* is essential for viability, pilus assembly, toxin production and virulence. *Mol Microbiol*
66. Comfort, D., and Clubb, R. T. (2004) A comparative genome analysis identifies distinct sorting pathways in gram-positive bacteria. *Infect Immun* **72**, 2710-2722
67. Suree, N., Liew, C. K., Villareal, V. A., Thieu, W., Fadeev, E. A., Clemens, J. J., Jung, M. E., and Clubb, R. T. (2009) The structure of the *Staphylococcus aureus*

- sortase-substrate complex reveals how the universally conserved LPXTG sorting signal is recognized. *J Biol Chem* **284**, 24465-24477
68. Mandlik, A., Swierczynski, A., Das, A., and Ton-That, H. (2008) Pili in Gram-positive bacteria: assembly, involvement in colonization and biofilm development. *Trends Microbiol* **16**, 33-40
  69. Dramsi, S., Caliot, E., Bonne, I., Guadagnini, S., Prevost, M. C., Kojadinovic, M., Lalioui, L., Poyart, C., and Trieu-Cuot, P. (2006) Assembly and role of pili in group B streptococci. *Mol Microbiol* **60**, 1401-1413
  70. Budzik, J. M., Marraffini, L. A., and Schneewind, O. (2007) Assembly of pili on the surface of *Bacillus cereus* vegetative cells. *Mol Microbiol* **66**, 495-510
  71. Sillanpaa, J., Chang, C., Singh, K. V., Montealegre, M. C., Nallapareddy, S. R., Harvey, B. R., Ton-That, H., and Murray, B. E. (2013) Contribution of individual Ebp Pilus subunits of *Enterococcus faecalis* OG1RF to pilus biogenesis, biofilm formation and urinary tract infection. *PLoS One* **8**, e68813
  72. Nallapareddy, S. R., Singh, K. V., Sillanpaa, J., Garsin, D. A., Hook, M., Erlandsen, S. L., and Murray, B. E. (2006) Endocarditis and biofilm-associated pili of *Enterococcus faecalis*. *J Clin Invest* **116**, 2799-2807
  73. Telford, J. L., Barocchi, M. A., Margarit, I., Rappuoli, R., and Grandi, G. (2006) Pili in gram-positive pathogens. *Nat Rev Microbiol* **4**, 509-519
  74. Barocchi, M. A., Ries, J., Zogaj, X., Hemsley, C., Albiger, B., Kanth, A., Dahlberg, S., Fernebro, J., Moschioni, M., Massignani, V., Hultenby, K., Taddei, A. R., Beiter, K., Wartha, F., von Euler, A., Covacci, A., Holden, D. W., Normark, S., Rappuoli, R., and Henriques-Normark, B. (2006) A pneumococcal pilus influences virulence and host inflammatory responses. *Proc Natl Acad Sci U S A* **103**, 2857-2862
  75. Kang, H. J., Paterson, N. G., Kim, C. U., Middleditch, M., Chang, C., Ton-That, H., and Baker, E. N. (2014) A slow-forming isopeptide bond in the structure of the

- major pilin SpaD from *Corynebacterium diphtheriae* has implications for pilus assembly. *Acta Crystallogr D Biol Crystallogr* **70**, 1190-1201
76. Kang, H. J., Coulibaly, F., Clow, F., Proft, T., and Baker, E. N. (2007) Stabilizing isopeptide bonds revealed in gram-positive bacterial pilus structure. *Science* **318**, 1625-1628
77. Deivanayagam, C. C., Rich, R. L., Carson, M., Owens, R. T., Danthuluri, S., Bice, T., Hook, M., and Narayana, S. V. (2000) Novel fold and assembly of the repetitive B region of the *Staphylococcus aureus* collagen-binding surface protein. *Structure* **8**, 67-78
78. Echelman, D. J., Alegre-Cebollada, J., Badilla, C. L., Chang, C., Ton-That, H., and Fernandez, J. M. (2016) CnaA domains in bacterial pili are efficient dissipaters of large mechanical shocks. *Proc Natl Acad Sci U S A* **113**, 2490-2495
79. Mishra, A., Wu, C., Yang, J., Cisar, J. O., Das, A., and Ton-That, H. (2010) The *Actinomyces oris* type 2 fimbrial shaft FimA mediates co-aggregation with oral streptococci, adherence to red blood cells and biofilm development. *Mol Microbiol* **77**, 841-854
80. Siegel, S. D., Wu, C., and Ton-That, H. (2016) A Type I Signal Peptidase Is Required for Pilus Assembly in the Gram-Positive, Biofilm-Forming Bacterium *Actinomyces oris*. *J Bacteriol* **198**, 2064-2073
81. Amer, B. R., and Clubb, R. T. (2014) A sweet new role for LCP enzymes in protein glycosylation. *Mol Microbiol* **94**, 1197-1200
82. Yeung, M. K., and Kozelsky, C. S. (1994) Transformation of *Actinomyces* spp. by a gram-negative broad-host-range plasmid. *J Bacteriol* **176**, 4173-4176
83. Hanahan, D. (1985) Techniques for transformation of *E. coli*. in *DNA Cloning: A Practical Approach* (Glover, D. M. a. H., B.D. ed.). pp 109



84. Mishra, A., Das, A., Cisar, J. O., and Ton-That, H. (2007) Sortase-catalyzed assembly of distinct heteromeric fimbriae in *Actinomyces naeslundii*. *J Bacteriol* **189**, 3156-3165
85. Yoshida, Y., Ganguly, S., Bush, C. A., and Cisar, J. O. (2006) Molecular basis of L-rhamnose branch formation in streptococcal coaggregation receptor polysaccharides. *J Bacteriol* **188**, 4125-4130
86. Bessette, P. H., Aslund, F., Beckwith, J., and Georgiou, G. (1999) Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm. *Proc Natl Acad Sci U S A* **96**, 13703-13708
87. Taton, A., Unglaub, F., Wright, N. E., Zeng, W. Y., Paz-Yepes, J., Brahmsha, B., Palenik, B., Peterson, T. C., Haerizadeh, F., Golden, S. S., and Golden, J. W. (2014) Broad-host-range vector system for synthetic biology and biotechnology in cyanobacteria. *Nucleic Acids Res* **42**, e136
88. Wu, C., and Ton-That, H. (2010) Allelic exchange in *Actinomyces oris* with mCherry fluorescence counterselection. *Appl Environ Microbiol* **76**, 5987-5989
89. Donnelly, M. I., Zhou, M., Millard, C. S., Clancy, S., Stols, L., Eschenfeldt, W. H., Collart, F. R., and Joachimiak, A. (2006) An expression vector tailored for large-scale, high-throughput purification of recombinant proteins. *Protein Expr Purif* **47**, 446-454
90. Corrigan, R. M., and Foster, T. J. (2009) An improved tetracycline-inducible expression vector for *Staphylococcus aureus*. *Plasmid* **61**, 126-129
91. Rohrschneider, L. R., Custodio, J. M., Anderson, T. A., Miller, C. P., and Gu, H. (2005) The intron 5/6 promoter region of the *ship1* gene regulates expression in stem/progenitor cells of the mouse embryo. *Dev Biol* **283**, 503-521
92. Schmittgen, T. D., and Livak, K. J. (2008) Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc* **3**, 1101-1108

93. Wu, C., Mishra, A., Reardon, M. E., Huang, I. H., Counts, S. C., Das, A., and Ton-That, H. (2012) Structural determinants of Actinomyces sortase SrtC2 required for membrane localization and assembly of type 2 fimbriae for interbacterial coaggregation and oral biofilm formation. *J Bacteriol* **194**, 2531-2539
94. Kaplan, A., Kaplan, C. W., He, X., McHardy, I., Shi, W., and Lux, R. (2014) Characterization of aid1, a novel gene involved in Fusobacterium nucleatum interspecies interactions. *Microb Ecol* **68**, 379-387
95. Broadway, M. M., Rogers, E. A., Chang, C., Huang, I. H., Dwivedi, P., Yildirim, S., Schmitt, M. P., Das, A., and Ton-That, H. (2013) Pilus Gene Pool Variation and the Virulence of Corynebacterium diphtheriae Clinical Isolates during Infection of a Nematode. *J Bacteriol* **195**, 3774-3783
96. Montealegre, M. C., La Rosa, S. L., Roh, J. H., Harvey, B. R., and Murray, B. E. (2015) The Enterococcus faecalis EbpA Pilus Protein: Attenuation of Expression, Biofilm Formation, and Adherence to Fibrinogen Start with the Rare Initiation Codon ATT. *MBio* **6**, e00467-00415
97. Baumgart, M., Schubert, K., Bramkamp, M., and Frunzke, J. (2016) Impact of LytR-CpsA-Psr Proteins on Cell Wall Biosynthesis in Corynebacterium glutamicum. *J Bacteriol* **198**, 3045-3059
98. Ekwunife, F. S., Singh, J., Taylor, K. G., and Doyle, R. J. (1991) Isolation and purification of cell wall polysaccharide of Bacillus anthracis (delta Sterne). *FEMS Microbiol Lett* **66**, 257-262
99. Sieling, P. A., Thomas, M. J., and van de Rijn, I. (1992) Characterization of the Streptococcus adjacens group antigen structure. *J Bacteriol* **174**, 349-354
100. Larson, T. R., and Yother, J. (2017) Streptococcus pneumoniae capsular polysaccharide is linked to peptidoglycan via a direct glycosidic bond to beta-D-N-acetylglucosamine. *Proc Natl Acad Sci U S A* **114**, 5695-5700

101. Chae, H. J., Byun, J. O., Chae, S. W., Kim, H. M., Choi, H. I., Pae, H. O., Chung, H. T., and Kim, H. R. (2005) p38 MAPK and NF-kappaB on IL-6 release in human gingival fibroblasts. *Immunopharmacol Immunotoxicol* **27**, 631-646
102. Widdick, D. A., Dilks, K., Chandra, G., Bottrill, A., Naldrett, M., Pohlschroder, M., and Palmer, T. (2006) The twin-arginine translocation pathway is a major route of protein export in *Streptomyces coelicolor*. *Proc Natl Acad Sci U S A* **103**, 17927-17932
103. Widdick, D. A., Eijlander, R. T., van Dijk, J. M., Kuipers, O. P., and Palmer, T. (2008) A facile reporter system for the experimental identification of twin-arginine translocation (Tat) signal peptides from all kingdoms of life. *J Mol Biol* **375**, 595-603
104. Mazmanian, S. K., Liu, G., Jensen, E. R., Lenoy, E., and Schneewind, O. (2000) *Staphylococcus aureus* sortase mutants defective in the display of surface proteins and in the pathogenesis of animal infections. *Proc Natl Acad Sci U S A* **97**, 5510-5515
105. Garandeau, C., Reglier-Poupet, H., Dubail, I., Beretti, J. L., Berche, P., and Charbit, A. (2002) The sortase SrtA of *Listeria monocytogenes* is involved in processing of internalin and in virulence. *Infect Immun* **70**, 1382-1390
106. Paterson, G. K., and Mitchell, T. J. (2006) The role of *Streptococcus pneumoniae* sortase A in colonisation and pathogenesis. *Microbes Infect* **8**, 145-153
107. Gaspar, A. H., Marraffini, L. A., Glass, E. M., Debord, K. L., Ton-That, H., and Schneewind, O. (2005) *Bacillus anthracis* sortase A (SrtA) anchors LPXTG motif-containing surface proteins to the cell wall envelope. *J Bacteriol* **187**, 4646-4655
108. Geukens, N., Parro, V., Rivas, L. A., Mellado, R. P., and Anne, J. (2001) Functional analysis of the *Streptomyces lividans* type I signal peptidases. *Arch Microbiol* **176**, 377-380

109. Parro, V., Schacht, S., Anne, J., and Mellado, R. P. (1999) Four genes encoding different type I signal peptidases are organized in a cluster in *Streptomyces lividans* TK21. *Microbiology* **145 ( Pt 9)**, 2255-2263
110. Caspi, R., Altman, T., Billington, R., Dreher, K., Foerster, H., Fulcher, C. A., Holland, T. A., Keseler, I. M., Kothari, A., Kubo, A., Krummenacker, M., Latendresse, M., Mueller, L. A., Ong, Q., Paley, S., Subhraveti, P., Weaver, D. S., Weerasinghe, D., Zhang, P., and Karp, P. D. (2014) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic Acids Res* **42**, D459-471
111. Jakubovics, N. S., Robinson, J. C., Samarian, D. S., Kolderman, E., Yassin, S. A., Bettampadi, D., Bashton, M., and Rickard, A. H. (2015) Critical roles of arginine in growth and biofilm development by *Streptococcus gordonii*. *Mol Microbiol* **97**, 281-300
112. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673-4680
113. Mishra, A., Devarajan, B., Reardon, M. E., Dwivedi, P., Krishnan, V., Cisar, J. O., Das, A., Narayana, S. V., and Ton-That, H. (2011) Two autonomous structural modules in the fimbrial shaft adhesin FimA mediate *Actinomyces* interactions with streptococci and host cells during oral biofilm development. *Mol Microbiol* **81**, 1205-1220
114. Guttilla, I. K., Gaspar, A. H., Swierczynski, A., Swaminathan, A., Dwivedi, P., Das, A., and Ton-That, H. (2009) Acyl enzyme intermediates in sortase-catalyzed pilus morphogenesis in gram-positive bacteria. *J Bacteriol* **191**, 5603-5612
115. Zahner, D., and Scott, J. R. (2008) SipA is required for pilus formation in *Streptococcus pyogenes* serotype M3. *J Bacteriol* **190**, 527-535

116. Young, P. G., Kang, H. J., and Baker, E. N. (2013) An arm-swapped dimer of the *Streptococcus pyogenes* pilin specific assembly factor SipA. *J Struct Biol* **183**, 99-104
117. Young, P. G., Proft, T., Harris, P. W., Brimble, M. A., and Baker, E. N. (2014) Structure and activity of *Streptococcus pyogenes* SipA: a signal peptidase-like protein essential for pilus polymerisation. *PLoS One* **9**, e99135
118. Nothaft, H., and Szymanski, C. M. (2010) Protein glycosylation in bacteria: sweeter than ever. *Nat Rev Microbiol* **8**, 765-778
119. Symersky, J., Patti, J. M., Carson, M., House-Pompeo, K., Teale, M., Moore, D., Jin, L., Schneider, A., DeLucas, L. J., Hook, M., and Narayana, S. V. (1997) Structure of the collagen-binding domain from a *Staphylococcus aureus* adhesin. *Nat Struct Biol* **4**, 833-838
120. Siegel, S. D., Liu, J., and Ton-That, H. (2016) Biogenesis of the Gram-positive bacterial cell envelope. *Curr Opin Microbiol* **34**, 31-37
121. Percy, M. G., and Grundling, A. (2014) Lipoteichoic acid synthesis and function in gram-positive bacteria. *Annu Rev Microbiol* **68**, 81-100
122. Brown, S., Santa Maria, J. P., Jr., and Walker, S. (2013) Wall teichoic acids of gram-positive bacteria. *Annu Rev Microbiol* **67**, 313-336
123. Fitzgerald, S. N., and Foster, T. J. (2000) Molecular analysis of the tagF gene, encoding CDP-Glycerol:Poly(glycerophosphate) glycerophosphotransferase of *Staphylococcus epidermidis* ATCC 14990. *J Bacteriol* **182**, 1046-1052
124. Luong, T. T., Reardon-Robinson, M. E., Siegel, S. D., and Ton-That, H. (2017) Reoxidation of the Thiol-Disulfide Oxidoreductase MdbA by a Bacterial Vitamin K Epoxide Reductase in the Biofilm-Forming Actinobacterium *Actinomyces oris*. *J Bacteriol* **199**, e00817-00816
125. Sanchez, B. C., Chang, C., Wu, C., Tran, B., and Ton-That, H. (2017) Electron Transport Chain Is Biochemically Linked to Pilus Assembly Required for

Polymicrobial Interactions and Biofilm Formation in the Gram-Positive  
Actinobacterium *Actinomyces oris*. *MBio* **8**

126. Becker, S., Frankel, M. B., Schneewind, O., and Missiakas, D. (2014) Release of protein A from the cell wall of *Staphylococcus aureus*. *Proc Natl Acad Sci U S A* **111**, 1574-1579
127. Lanz, N. D., Pandelia, M. E., Kakar, E. S., Lee, K. H., Krebs, C., and Booker, S. J. (2014) Evidence for a catalytically and kinetically competent enzyme-substrate cross-linked intermediate in catalysis by lipoyl synthase. *Biochemistry* **53**, 4557-4572
128. Liszewski Zilla, M., Chan, Y. G., Lunderberg, J. M., Schneewind, O., and Missiakas, D. (2015) LytR-CpsA-Psr enzymes as determinants of *Bacillus anthracis* secondary cell wall polysaccharide assembly. *J Bacteriol* **197**, 343-353
129. Yokoyama, K., Miyashita, T., Araki, Y., and Ito, E. (1986) Structure and functions of linkage unit intermediates in the biosynthesis of ribitol teichoic acids in *Staphylococcus aureus* H and *Bacillus subtilis* W23. *Eur J Biochem* **161**, 479-489
130. Reardon-Robinson, M. E., and Ton-That, H. (2016) Disulfide-Bond-Forming Pathways in Gram-Positive Bacteria. *J Bacteriol* **198**, 746-754
131. Kline, K. A., Falker, S., Dahlberg, S., Normark, S., and Henriques-Normark, B. (2009) Bacterial adhesins in host-microbe interactions. *Cell Host Microbe* **5**, 580-592
132. Li, J., Helmerhorst, E. J., Leone, C. W., Troxler, R. F., Yaskell, T., Haffajee, A. D., Socransky, S. S., and Oppenheim, F. G. (2004) Identification of early microbial colonizers in human dental biofilm. *J Appl Microbiol* **97**, 1311-1318
133. Kolenbrander, P. E., Palmer, R. J., Jr., Rickard, A. H., Jakubovics, N. S., Chalmers, N. I., and Diaz, P. I. (2006) Bacterial interactions and successions during plaque development. *Periodontol 2000* **42**, 47-79

134. Kolenbrander, P. E., and Andersen, R. N. (1986) Multigeneric aggregations among oral bacteria: a network of independent cell-to-cell interactions. *J Bacteriol* **168**, 851-859
135. Conover, M. S., Ruer, S., Taganna, J., Kalas, V., De Greve, H., Pinkner, J. S., Dodson, K. W., Remaut, H., and Hultgren, S. J. (2016) Inflammation-Induced Adhesin-Receptor Interaction Provides a Fitness Advantage to Uropathogenic *E. coli* during Chronic Infection. *Cell Host Microbe* **20**, 482-492
136. Spaulding, C. N., Klein, R. D., Schreiber, H. L. t., Janetka, J. W., and Hultgren, S. J. (2018) Precision antimicrobial therapeutics: the path of least resistance? *NPJ Biofilms Microbiomes* **4**, 4
137. Putnins, E. E., and Bowden, G. H. (1993) Antigenic relationships among oral *Actinomyces* isolates, *Actinomyces naeslundii* genospecies 1 and 2, *Actinomyces howellii*, *Actinomyces denticolens*, and *Actinomyces slackii*. *J Dent Res* **72**, 1374-1385
138. Henssge, U., Do, T., Radford, D. R., Gilbert, S. C., Clark, D., and Beighton, D. (2009) Emended description of *Actinomyces naeslundii* and descriptions of *Actinomyces oris* sp. nov. and *Actinomyces johnsonii* sp. nov., previously identified as *Actinomyces naeslundii* genospecies 1, 2 and WVA 963. *Int J Syst Evol Microbiol* **59**, 509-516
139. Mergenhagen, S. E., Sandberg, A. L., Chassy, B. M., Brennan, M. J., Yeung, M. K., Donkersloot, J. A., and Cisar, J. O. (1987) Molecular basis of bacterial adhesion in the oral cavity. *Rev Infect Dis* **9 Suppl 5**, S467-474
140. Dame-Teixeira, N., Parolo, C. C., Maltz, M., Tugnait, A., Devine, D., and Do, T. (2016) *Actinomyces* spp. gene expression in root caries lesions. *J Oral Microbiol* **8**, 32383

141. Cisar, J. O., Kolenbrander, P. E., and McIntire, F. C. (1979) Specificity of coaggregation reactions between human oral streptococci and strains of *Actinomyces viscosus* or *Actinomyces naeslundii*. *Infect Immun* **24**, 742-752
142. Kolenbrander, P. E., Inouye, Y., and Holdeman, L. V. (1983) New *Actinomyces* and *Streptococcus* coaggregation groups among human oral isolates from the same site. *Infect Immun* **41**, 501-506
143. Kolenbrander, P. E., and Williams, B. L. (1983) Prevalence of viridans streptococci exhibiting lactose-inhibitable coaggregation with oral actinomycetes. *Infect Immun* **41**, 449-452
144. Sakellaris, H., Munson, G. P., and Scott, J. R. (1999) A conserved residue in the tip proteins of CS1 and CFA/I pili of enterotoxigenic *Escherichia coli* that is essential for adherence. *Proc Natl Acad Sci U S A* **96**, 12828-12832
145. Douillard, F. P., Rasinkangas, P., von Ossowski, I., Reunanen, J., Palva, A., and de Vos, W. M. (2014) Functional identification of conserved residues involved in *Lactobacillus rhamnosus* strain GG sortase specificity and pilus biogenesis. *J Biol Chem* **289**, 15764-15775
146. Okura, M., Osaki, M., Fittipaldi, N., Gottschalk, M., Sekizaki, T., and Takamatsu, D. (2011) The minor pilin subunit Sgp2 is necessary for assembly of the pilus encoded by the srtG cluster of *Streptococcus suis*. *J Bacteriol* **193**, 822-831
147. Berks, B. C. (2015) The twin-arginine protein translocation pathway. *Annu Rev Biochem* **84**, 843-864
148. Feilmeier, B. J., Iseminger, G., Schroeder, D., Webber, H., and Phillips, G. J. (2000) Green fluorescent protein functions as a reporter for protein localization in *Escherichia coli*. *J Bacteriol* **182**, 4068-4076
149. Thomas, J. D., Daniel, R. A., Errington, J., and Robinson, C. (2001) Export of active green fluorescent protein to the periplasm by the twin-arginine translocase (Tat) pathway in *Escherichia coli*. *Mol Microbiol* **39**, 47-53



150. Buchanan, G., Sargent, F., Berks, B. C., and Palmer, T. (2001) A genetic screen for suppressors of Escherichia coli Tat signal peptide mutations establishes a critical role for the second arginine within the twin-arginine motif. *Arch Microbiol* **177**, 107-112
151. Stanley, N. R., Palmer, T., and Berks, B. C. (2000) The twin arginine consensus motif of Tat signal peptides is involved in Sec-independent protein targeting in Escherichia coli. *J Biol Chem* **275**, 11591-11596
152. Pragai, Z., Tjalsma, H., Bolhuis, A., van Dijl, J. M., Venema, G., and Bron, S. (1997) The signal peptidase II (Isp) gene of Bacillus subtilis. *Microbiology* **143 ( Pt 4)**, 1327-1333
153. Palacin, A., Parro, V., Geukens, N., Anne, J., and Mellado, R. P. (2002) SipY Is the Streptomyces lividans type I signal peptidase exerting a major effect on protein secretion. *J Bacteriol* **184**, 4875-4880
154. Marino, G., Eckhard, U., and Overall, C. M. (2015) Protein Termini and Their Modifications Revealed by Positional Proteomics. *ACS Chem Biol* **10**, 1754-1764
155. Smith, P. A., and Romesberg, F. E. (2012) Mechanism of action of the arylomycin antibiotics and effects of signal peptidase I inhibition. *Antimicrob Agents Chemother* **56**, 5054-5060
156. Therien, A. G., Huber, J. L., Wilson, K. E., Beaulieu, P., Caron, A., Claveau, D., Deschamps, K., Donald, R. G., Galgoci, A. M., Gallant, M., Gu, X., Kevin, N. J., Lafleur, J., Leavitt, P. S., Lebeau-Jacob, C., Lee, S. S., Lin, M. M., Michels, A. A., Ogawa, A. M., Painter, R. E., Parish, C. A., Park, Y. W., Benton-Perdomo, L., Petcu, M., Phillips, J. W., Powles, M. A., Skorey, K. I., Tam, J., Tan, C. M., Young, K., Wong, S., Waddell, S. T., and Miesel, L. (2012) Broadening the spectrum of beta-lactam antibiotics through inhibition of signal peptidase type I. *Antimicrob Agents Chemother* **56**, 4662-4670

157. Hirabayashi, J., Kuno, A., and Tateno, H. (2011) Lectin-based structural glycomics: a practical approach to complex glycans. *Electrophoresis* **32**, 1118-1128
158. Dan, X., Liu, W., and Ng, T. B. (2016) Development and Applications of Lectins as Biological Tools in Biomedical Research. *Med Res Rev* **36**, 221-247
159. Morelle, W., and Michalski, J. C. (2007) Analysis of protein glycosylation by mass spectrometry. *Nat Protoc* **2**, 1585-1602
160. Hancock, I. C., Wiseman, G., and Baddiley, J. (1976) Biosynthesis of the unit that links teichoic acid to the bacterial wall: inhibition by tunicamycin. *FEBS Lett* **69**, 75-80
161. Swoboda, J. G., Meredith, T. C., Campbell, J., Brown, S., Suzuki, T., Bollenbach, T., Malhowski, A. J., Kishony, R., Gilmore, M. S., and Walker, S. (2009) Discovery of a small molecule that blocks wall teichoic acid biosynthesis in *Staphylococcus aureus*. *ACS Chem Biol* **4**, 875-883
162. Saint-Joanis, B., Demangel, C., Jackson, M., Brodin, P., Marsollier, L., Boshoff, H., and Cole, S. T. (2006) Inactivation of Rv2525c, a substrate of the twin arginine translocation (Tat) system of *Mycobacterium tuberculosis*, increases beta-lactam susceptibility and virulence. *J Bacteriol* **188**, 6669-6679
163. Oertel, D., Schmitz, S., and Freudl, R. (2015) A TatABC-type Tat translocase is required for unimpaired aerobic growth of *Corynebacterium glutamicum* ATCC13032. *PLoS One* **10**, e0123413

## VITA

Sara Danielle Siegel was born in Panama City, Florida, the daughter of Frank Jesse Siegel and Sandra Ellen Romano. After completing her work at Rutherford High School, Panama City, Florida in 2007, she entered Gulf Coast College in Panama City, Florida. She received the degree of Associates of Arts with a major in biology. She then attended the University of Florida and graduated with a degree in microbiology and cell science in Gainesville, Florida in December, 2011. For the next year, she worked as a laboratory technician in the Department of Pediatrics Glycogen Storage Disease program at the University of Florida. In August 2013 she entered The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences.

Permanent address:

2825 Farlow Lane  
Manvel, Texas 77578