Insights into Nonpilus Adhesin Functionality and the Molecular Determinants of

Nontypeable *Haemophilus influenzae* Colonization

by

Katherine Alice Rempe

Department of Molecular Genetics and Microbiology
Duke University

Date:_______________________

Approved:

___________________________
Joseph W. St. Geme, III, Supervisor

___________________________
Margarethe J. Kuehn

___________________________
Patrick C. Seed

___________________________
Raphael H. Valdivia

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor
of Philosophy in the Department of
Molecular Genetics and Microbiology in the Graduate School
of Duke University

2016
ABSTRACT

Insights into Nonpilus Adhesin Functionality and the Molecular Determinants of Nontypeable *Haemophilus influenzae* Colonization

by

Katherine Alice Rempe

Department of Molecular Genetics and Microbiology
Duke University

Date: _________________________
Approved: _________________________

___________________________
Joseph W. St. Geme, III, Supervisor

___________________________
Margarethe J. Kuehn

___________________________
Patrick C. Seed

___________________________
Raphael H. Valdivia

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Molecular Genetics and Microbiology in the Graduate School of Duke University

2016
Abstract

Bacterial colonization of the upper respiratory tract is the first step in the pathogenesis of nontypeable *Haemophilus influenzae* (NTHi) disease. Examination of the determinants of NTHi colonization process has been hampered by the lack of an appropriate animal model. To address this, we have developed a model of NTHi colonization in adult rhesus macaques that involves intranasal inoculation of $1 \times 10^5$ CFU and results in persistent colonization of the upper respiratory tract for at least three weeks with no signs of disease, mimicking asymptomatic colonization of humans. Using this model, we assessed the contributions to colonization of the HMW1 and HMW2 adhesive proteins. In competition experiments, the parent strain expressing both HMW1 and HMW2 was able to efficiently out-compete an isogenic mutant strain expressing neither HMW1 nor HMW2. In experiments involving inoculation of single isogenic derivatives of NTHi strain 12, the strains expressing HMW1 or HMW2 or both were able to colonize efficiently, while the strain expressing neither HMW1 nor HMW2 colonized inefficiently. Furthermore, colonization resulted in antibody production against HMW1 and HMW2 in one-third of the animals, demonstrating that colonization can be an immunizing event. In conclusion, we have established that NTHi is capable of colonizing the upper respiratory tract of rhesus macaques, in some cases associated with stimulation of an immune response. The HMW1 and HMW2 adhesive proteins play a major role in the process of colonization.
After establishing that the HMW1 and HMW2 proteins are colonization factors we further investigated the determinants of HMW1 function. HMW1 is encoded in the same genetic locus as two other proteins, HMW1B and HMW1C, with which HMW1 must interact in order to be functional. Interaction with HMW1C in the cytoplasm results in the glycosylation of HMW1. By employing homologues of HMW1C that glycosylate HMW1 in slightly different patterns we show that the pattern of modification is critical to HMW1 function. Structural analysis showed a change in protein structure when the pattern of HMW1 modification differed. We also identified two specific sites which must be glycosylated for HMW1 to function properly. These point mutations did not have a significant effect on protein structure, suggesting that glycosylation at those specific sites is instead necessary for interaction of HMW1 with its receptor. HMW1B is an outer membrane pore through which HMW1 is transported to reach the bacterial cell surface. We observed that HMW1 isolated from the cytoplasm has a different structure than HMW1 isolated from the bacterial cell surface. By forcing HMW1 to be secreted in a non-HMW1B dependent manner, we show that secretion alone is not sufficient for HMW1 to obtain a functional structure. This leads us to hypothesize that there is something specific in the interaction between HMW1 and HMW1B that aids in proper HMW1 folding.

The NTHi HMW1C glycosyltransferase mediates unconventional N-linked glycosylation of HMW1. In this system, HMW1 is modified in the cytoplasm by sequential transfer of hexose residues. To determine if this mechanism of N-linked
glycosylation is employed by species other than NTHi, we examined *Kingella kingae* and *Aggregatibacter aphrophilus* homologues of HMW1C. We found both homologues to be functional glycosyltransferases and identified their substrates as the *K. kingae* Knh and the *A. aphrophilus* EmaA trimeric autotransporter proteins. LC-MS/MS analysis revealed multiple sites of N-linked glycosylation on Knh and EmaA. Without glycosylation, Knh and EmaA failed to facilitate wild type levels of bacterial autoaggregation or adherence to human epithelial cells, establishing that glycosylation is essential for proper protein function.

Taken together we have shown that the HMW1 and HMW2 proteins are involved in colonization *in vivo*, that proper glycosylation and secretion of these proteins is needed for proper protein function, and that the unconventional N-linked glycosylation mechanism used by NTHi to modify HMW1 and HMW2 is also used by at least three other gram negative species to modify adhesive proteins. This work greatly enhances our understanding of NTHi adherence and colonization and demonstrates that findings regarding NTHi adherence are transferable to other species.
Contents

Abstract .............................................................................................................................................. iv

List of Tables ........................................................................................................................................ xii

List of Figures ....................................................................................................................................... xiii

Acknowledgements .............................................................................................................................. xv

1. Introduction ......................................................................................................................................... 1

1.1 Organisms Used in these Studies .................................................................................................... 1

1.1.1 Nontypeable Haemophilus Influenzae ...................................................................................... 1

1.1.2 Kingella kingae .......................................................................................................................... 4

1.1.3 Aggregatibacter aphrophilus ................................................................................................. 7

1.2 Gram-Negative Bacterial Adherence ............................................................................................. 8

1.2.1 Pilus Adhesins .......................................................................................................................... 8

1.2.1.1 Chaperone-Usher Pili ........................................................................................................... 8

1.2.1.2 Type IV pili .......................................................................................................................... 10

1.2.1.3 Curli .................................................................................................................................. 13

1.2.2 Non-Pilus Adhesins ................................................................................................................. 14

1.2.2.1 Type V Secretion ............................................................................................................... 15

1.2.2.2 Type Va (Classical) Autotransporters .............................................................................. 15

1.2.2.3 Type Vb Two Partner Secretion Systems ......................................................................... 16

1.2.2.4 Type Vc (Trimeric) Autotransporters .............................................................................. 17

1.3 Nontypeable Haemophilus influenzae Adherence ..................................................................... 18
1.3.1 Other NTHi Adhesins ................................................................. 22
1.4 Kingella kingae Adherence .......................................................... 22
1.5 Aggregatibacter aphrophilus Adherence ........................................ 24
1.6 Bacterial Protein Glycosylation ..................................................... 24
  1.6.1 O-Linked Glycosylation .......................................................... 25
  1.6.2 N-linked Glycosylation ......................................................... 25
  1.6.3 HMWC-mediated N-linked Glycosylation .................................. 26
1.7 Thesis Overview .............................................................................. 28

2. The HMW1 and HMW2 adhesins enhance the ability of nontypeable Haemophilus influenzae to colonize the upper respiratory tract of rhesus macaques ................. 30
  2.1 Introduction .................................................................................. 30
  2.2 Methods ....................................................................................... 32
    2.2.1 Bacterial Strains and Culture ................................................. 32
    2.2.2 Adherence Assays ................................................................. 33
    2.2.3 Animal Studies ....................................................................... 34
    2.2.4 Bacterial Recovery ................................................................. 35
    2.2.5 ELISA ................................................................................... 35
    2.2.6 SDS-PAGE ........................................................................... 36
    2.2.7 Statistics ............................................................................... 37
  2.3 Results .......................................................................................... 37
    2.3.1 NTHi adheres to rhesus macaque respiratory epithelial cells in vitro .......... 37
    2.3.2 NTHi is able to colonize the upper respiratory tract of rhesus macaques ...... 38
2.3.3 The HMW1 and HMW2 proteins facilitate colonization of rhesus macaques. 41
2.3.4 Colonization is an immunizing event .......................................................... 47
2.4 Discussion ........................................................................................................... 48

3. Determinants of HMW1 Function ........................................................................... 54
3.1 Introduction ........................................................................................................... 54
3.2 Methods .............................................................................................................. 57
  3.2.1 Bacterial Strains and Growth Conditions ..................................................... 57
  3.2.2 Purification of untagged HMW1 .................................................................. 58
  3.2.3 Purification of YebF fused HMW1 ................................................................. 59
  3.2.4 Western Blots ............................................................................................... 60
  3.2.5 Native Gel Electrophoresis ........................................................................... 61
  3.2.6 Circular Dichroism ....................................................................................... 61
3.3 Results .................................................................................................................. 61
  3.3.1 The Pattern of Glycosylation on HMW1 is Critical for Protein Function ....... 61
  3.3.2 HMW1 purified from the bacterial cell surface is functional while HMW1 purified from the cytoplasm is non-functional. ....................................................... 66
  3.3.3 Secretion alone is not sufficient to produce functional HMW1 .................. 69
3.4 Discussion ......................................................................................................... 71

4. Unconventional N-linked glycosylation promotes trimeric autotransporter function in Kingella kingae and Aggregatibacter aphrophilus. ........................................... 74
4.1 Introduction ......................................................................................................... 74
4.2 Methods .............................................................................................................. 77
  4.2.1 Strains and culture conditions ..................................................................... 77
4.2.2 Strain Construction ................................................................. 80
4.2.3 Homologue identification: ....................................................... 81
4.2.4 Adherence assays ................................................................. 81
4.2.5 Autoaggregation assays ......................................................... 82
4.2.6 Statistical analysis................................................................. 82
4.2.7 Isolation of outer membrane proteins ...................................... 83
4.2.8 Formic acid treatment of outer membrane proteins .................. 83
4.2.9 SDS PAGE ........................................................................... 83
4.2.10 In-gel digestion .................................................................... 83
4.2.11 Mass spectrometry analysis .................................................. 84
4.2.12 Database searching ............................................................... 85
4.2.13 Immunogold labeling and Transmission electron microscopy .... 85
4.3 Results ..................................................................................... 86
4.3.1 Kingella kingae and Aggregatibacter aphrophilus have HMW1C homologues ................................................................. 86
4.3.2 HMW1C<sub>Kk</sub> is required for <i>K. kingae</i> adherence to human epithelial cells and autoaggregation ................................................................. 89
4.3.3 HMW1C<sub>Aa</sub> glycosylates the trimeric autotransporter Knh ............ 91
4.3.4 Co-expression of Knh and HMW1C<sub>Kk</sub> in <i>E. coli</i> is sufficient to facilitate bacterial adherence to epithelial cells ................................................................. 93
4.3.5 Mutating hmw1C<sub>Kk</sub> decreases the amount of Knh on the bacterial cell surface ................................................................................. 94
4.3.6 HMW1C<sub>Aa</sub> is necessary for <i>A. aphrophilus</i> autoaggregation and adherence to human epithelial cells ................................................................. 95
4.3.7 HMW1C<sub>a</sub> glycosylates the trimeric autotransporter EmaA ........................................... 97
4.3.8 EmaA facilitates autoaggregation and bacterial adherence to host cells .......... 99
4.4 Discussion .................................................................................................................................. 100
4.5 GenBank Accession Numbers ................................................................................................. 104
5. Future Directions .......................................................................................................................... 105
  5.1 Colonization of rhesus macaques with additional strains of NTHi ....................... 105
  5.2 Protection by colonization-induced antibody .............................................................................. 107
  5.3 Structural analysis of HMW1 .................................................................................................. 108
  5.4 Investigation of additional key modified residues of HMW1 ................................. 109
  5.5 Specificity of HMW1C for HMW1 .......................................................................................... 110
  5.6 Small molecule inhibitor of HMW1C ....................................................................................... 112
References .......................................................................................................................................... 114
Biography .......................................................................................................................................... 137
Peer Reviewed Publications ............................................................................................................... 137
Book Chapters ..................................................................................................................................... 138
List of Tables

Table 1: Bacterial Strains used in Chapter 2 ..................................................................... 32
Table 2: Animal Weights over the Course of the Experiment ............................................. 40
Table 3: Bacterial Strains Used in Chapter 3 ...................................................................... 57
Table 4: Bacterial Strains used in Chapter 4 ...................................................................... 78
Table 5: Glycopeptides identified with di-hexose modification in Knh ................................. 92
Table 6: Glycopeptides identified with di-hexose modification in Knh ................................. 93
Table 7: Glycopeptides identified in EmaA ........................................................................ 99
List of Figures

Figure 1: Chaperone User Assembled Pilus Biogenesis .......................................................... 10
Figure 2: Type IV Pilus Biogenesis ...................................................................................... 13
Figure 3: Curli Biogenesis ................................................................................................... 14
Figure 4: Type Va Classical Autotransporter ..................................................................... 16
Figure 5: Type Vb Two Partner Secretion System .............................................................. 17
Figure 6: Type Vc Trimeric Autotransporter ...................................................................... 18
Figure 7: Model of K. kingae adherence ............................................................................. 23
Figure 8: Comparison of Conventional and HMW1C-Mediated N-linked Glycosylation .......................................................................................................................... 28
Figure 9: NTHi adherence to rhesus macaque bronchial epithelial cells is dependent on the HMW1 and HMW2 proteins. ................................................................................. 38
Figure 10: An inoculum of $10^5$ CFU of NTHi strain 12 results in consistent colonization of rhesus macaques. .......................................................................................................................... 39
Figure 11: HMW1 and HMW2 in NTHi colonization of rhesus macaques ....................... 43
Figure 12: HMW1 and HMW2 independently facilitate colonization of rhesus macaques ........................................................................................................................................ 46
Figure 13: Colonization with NTHi can be an immunizing event ..................................... 48
Figure 14: Pattern of glycosylation affects HMW1 function ............................................ 63
Figure 15: There are at least two specific sites that must be modified for HMW1 to function ........................................................................................................................................ 64
Figure 16: Glycosylation pattern impacts protein structure ............................................. 66
Figure 17: HMW1 obtained from the bacterial surface is structurally different than HMW1 obtained from the cytoplasm. ........................................................................................................ 68
Figure 18: Non-HMW1B depended secretion of HMW1 results in non-functional protein. ............................................................................................................................................................................................70

Figure 19: HMW1C\textsubscript{Kk} and HMW1C\textsubscript{Aa} are highly homologous to HMW1C .........................88

Figure 20: Expression of HMW1C\textsubscript{Kk} is required for adherence and autoaggregation ......90

Figure 21: HMW1C\textsubscript{Kk} is required for Knh expression on the surface of \textit{K. kingae} ............95

Figure 22: HMW1C\textsubscript{Aa} expression is required for \textit{A. aphrophilus} adherence and autoaggregation ..........................................................................................................................................................................................97

Figure 23: EmaA is an adhesin and involved in autoaggregation .................................................100

Figure 24: Plasmid Expression of HMW1C complements a HMW1C Genomic Deletion ..........................................................................................................................................................................................111
Acknowledgements

I wish to express my sincere gratitude to everyone who has helped me to succeed in this endeavor most especially my lab-mates Eric Porsch, Jessica McCann, and Susan Grass. I’d also like to acknowledge all the other scientists I worked with over the past five years at Duke and at CHOP who have been willing to share their knowledge with me, especially Dr. Steve Seeholzer, Dr. Jolaine Wilson, and Lynn Spruce. I am grateful to Dr. Philip Johnson and Dr. James Wilson who allowed me to use their non-human primates for the experiments described in Chapter 2. I’d like to thank my family and friends for encouraging me along the way. Finally, I’d like to thank my mentor, Dr. Joseph W. St. Geme III, for challenging me to be a great scientist, teaching me the skills I needed, and allowing me to investigate my own ideas.
1. Introduction

1.1 Organisms Used in these Studies

1.1.1 Nontypeable Haemophilus Influenzae

*Haemophilus influenzae* was first isolated by Richard Pfeiffer during the 1889 influenza pandemic and was believed to be the microorganism responsible for influenza (Pfeiffer 1892). In 1933 it was determined that influenza was of viral etiology rather than bacterial; however, *H. influenzae* retained its name in homage to the circumstance of its discovery. In 1917, the species was given its current name, *Haemophilus influenzae*, with the genus *Haemophilus* recognizing its affinity for blood.

*H. influenzae* is a member of Pasturellaceae and the type-species for the genus *Haemophilus* and is a non-spore forming, non-motile gram-negative coccobacillus. It is able to grow both aerobically and anaerobically and requires hemin (factor X) and nicotinamide adenine dinucleotide (NAD) (factor V) under aerobic conditions. *H. influenzae* is a human restricted organism with no known reservoir other than the human respiratory tract. Serotyping divides *H. influenzae* into 6 capsule types (a-f) and nonencapsulated or so-called nontypeable *H. influenzae* (NTHi). Among these serotypes, type b (Hib) has historically been the primary cause of morbidity and mortality and remains so in many developing nations. In wealthier nations, the introduction of the Hib conjugate vaccines have drastically decreased the disease burden attributed to this bacterium. In the United States alone, the Hib vaccines are responsible for decreasing
the number of cases of Hib invasive disease by 99% (Diseases 2015). With this decrease in Hib disease, NTHi has become the primary cause of \textit{H. influenzae} disease.

As part of the normal human respiratory flora, NTHi is spread via airborne droplets or direct contact with respiratory secretions. The organism primarily causes localized respiratory tract infections but can also cause invasive disease, including meningitis, endocarditis, pericarditis, pyogenic arthritis, osteomyelitis, epiglottitis, cellulitis, urinary tract infections, intra-abdominal infections, and occult bacteremia (St Geme 1993). NTHi is the leading cause of purulent conjunctivitis and the second leading cause of otitis media in children (Van Eldere et al. 2014, Murphy 2000). In adults, it is a common cause of community associated pneumonia and exacerbations of underlying lung disease such as chronic obstructive pulmonary disease (COPD) and bronchiectasis (Van Eldere et al. 2014, Murphy 2000). In both adults and children, NTHi is a common cause of sinusitis (Van Eldere et al. 2014, Murphy 2000). NTHi has an economic impact of over $1 billion annually in the United States alone (Hardy, Tudor, and St Geme 2003).

The pathogenesis of NTHi disease begins with colonization of the respiratory tract and then contiguous spread from the site of colonization. This spread is often precipitated by an event that damages the respiratory epithelium, such as a viral infection, allergic disease or cigarette smoking (Murphy et al. 2009, Hardy, Tudor, and St Geme 2003). NTHi encodes a variety of pathogenic factors, including multiple adhesive
proteins, IgA proteases, lipooligosaccharide, protein kinase Cε, and a
glycerophosphodiesterase called protein D (Bailey et al. 2012, Johnson and Inzana 1986,
Kanthakumar et al. 1996, Janson et al. 1999, Fan et al. 2001). Pathology is caused by
secreted LOS, which is highly immunogenic and ciliotoxic, with the toxicity located in
the lipid A moiety (Johnson and Inzana 1986, Reddy et al. 1996, Gu et al. 1995). NTHi is
able to evade the immune system using a variety of mechanisms, including antigenic
variation of LOS and surface proteins, penetration of tight junctions, entry into host

NTHi infection has historically been treated with ampicillin. However, in some
populations, up to 40% of strains are now resistant (Yanagihara et al. 2010, Zhu et al.
2015). This change in antimicrobial resistance has caused treatment protocols to be
changed to the use of second- and third-generation cephalosporins, particularly
cefotaxime and ceftriaxone. Development of an NTHi vaccine has been hampered by
the lack of capsule and the fact that NTHi isolates are not clonal (Musser et al. 1986, Kyd
and Cripps 1999). Several proteins are under investigation as vaccine antigen candidates
including P1, P2, P4, P5, P6, D15, TbpA/B, LOS, HMW1, HMW2, Hia, and pili (Poolman
et al. 2000).
1.1.2 Kingella kingae

Elizabeth O. King at the CDC first isolated *Kingella kingae* in 1960, recovering the organism from blood and joint samples. The species was initially named *Moraxella kingii*, due to its similarities with that genus (Henriksen and Bovre 1968). In 1976, it was transferred to the new genus *Kingella* and renamed *Kingella kingae* (Henriksen 1976).

A member of the Neisseriaceae, *K. kingae* is a gram-negative, β-hemolytic, facultative anaerobe that grows as bacilli in pairs or short chains (Von Graevenitz A 2003, Yagupsky, Porsch, and St Geme 2011). *K. kingae* grows well on sheep blood agar or chocolate agar and the presence of 5% CO₂. *K. kingae* appears to be a member of the normal respiratory flora in children up to 24 months, and colonization is hypothesized to be an immunizing event, as most children gain reactive antibodies without ever presenting with a *K. kingae* illness. To date *K. kingae* has only been isolated from human sources and thus appears to be a human-restricted organism.

*K. kingae* is spread via person-to-person contact, as highlighted in a recent study by Yagupsky in 2016 (Yagupsky et al. 2016). This study of two Israeli populations showed colonization clusters of genetically identical strains within groups with close contact and little to no spread to individuals with whom colonized persons had little contact. As *K. kingae* predominantly colonizes young children, the majority of spread happens within childcare settings. It is been estimated that at any given time, 6% of all
children and 27% of day care attendees under 6 years old are colonized with *K. kingae* (Yagupsky, Dagan, et al. 1995).

Less than 1% of those colonized develop invasive disease (Ceroni et al. 2012). Among patients with *K. kingae* disease, the most common manifestations are occult bacteremia or osteoarticular infections, including septic arthritis, osteomyelitis, tenosynovitis, and spondylodiscitis (Dubnov-Raz et al. 2010). In an increasing number of populations, *K. kingae* is the primary cause of joint and bone infections in children 6-36 months old (Yagupsky, Bar-Ziv, et al. 1995, Ilharreborde et al. 2009, Chometon et al. 2007). Rarely, *K. kingae* will cause meningitis, ocular infections, peritonitis, or pericarditis (Van Erps et al. 1992, Carden et al. 1991, Bofinger, Fekete, and Samuel 2007, Matta et al. 2007). *K. kingae* is also a member of the HACEK (*Haemophilus, Aggregatibacter, Cardiobacterium, Eikenella, and Kingella*) group that are unusual causes of endocarditis. Disease is generally seen in otherwise healthy children between 6 and 36 months old. Patients older than 48 months with *K. kingae* infections generally have some underlying disease, often with compromised immunity. The period of greatest susceptibility for children (6-36 months) correlates with the time between the loss of maternal antibody and the ability of a child to mount a mature immune response, thereby suggesting that *K. kingae* is readily susceptible to a mature adaptive immune system (Slonim, Steiner, and Yagupsky 2003).
Invasive disease is thought to begin with the colonization of the upper respiratory tract, as demonstrated by the fact that genetically identical organisms can be recovered from the upper respiratory tract and the site of infection (Amit et al. 2014, Yagupsky et al. 2009). Infection begins when the organism breaches the epithelial barrier and enters the bloodstream. The breach of epithelium is thought to be facilitated by a potent RTX toxin that is encoded in every K. kingae strain isolated to date, potentially influenced by viral infection-mediated damage to the epithelium (Kehl-Fie and St Geme 2007). The organism is thought to be able to then traffic through the bloodstream to distal sites, such as the joints and bones. Little is known regarding this trafficking or the mechanisms by which K. kingae causes osteoarticular infections. The organism is able to evade the immune system by means of a polysaccharide capsule and the down regulation of antigenic surface proteins (Kehl-Fie et al. 2010).

1.1.3 *Aggregatibacter aphrophilus*

*Aggregatibacter aphrophilus* was first isolated in 1940 by Khariat as the causative agent of a fatal case of endocarditis. The organism was originally assigned to the genus *Haemophilus* based on phenotypic characteristics and an observed dependence on factor X (hemin) for growth (Khairat 1940). The genus designation was somewhat contentious due to the fact that *A. aphrophilus* is capable of synthesizing NAD. The species name *aphrophilus* was given because of the affinity of the organism for CO$_2$. Following numerous reports that the bacterium was not actually dependent on factor X, but rather factor X enhanced growth, the organism was transferred to the new genus *Aggregatibacter* in 2006 (Norskov-Lauritsen and Kilian 2006).

*A. aphrophilus* is a gram-negative non-spore forming, non-motile coccobacillus and a member of the Pasturellacea. It grows best in a 5-10% CO$_2$ enriched atmosphere and poorly in anaerobic conditions. To date, it has only been isolated from humans, though there are case reports of infections following canine bites. The organism is a member of the normal human buccal flora and can be isolated from the mouth, tonsils, and respiratory secretions (Bieger, Brewer, and Washington 1978).

*A. aphrophilus* is an opportunistic pathogen and can cause endocarditis, brain abscess, sinusitis, meningitis, pneumonia, and bacteremia (Bieger, Brewer, and Washington 1978). Endocarditis and brain abscess are the most common diseases attributed to the organism (Norskov-Lauritsen 2014). *A. aphrophilus* is generally
susceptible to antibiotics active against other gram-negative organisms, including penicillin, ampicillin, ciprofloxin, and cephalosporins (Norskov-Lauritsen 2014). The literature lacks description of the mechanisms *A. aphrophilus* uses to cause disease.

### 1.2 Gram-Negative Bacterial Adherence

To establish colonization of a host surface, bacteria must be able to adhere their host. Gram-negative bacteria utilize a number of factors that facilitate this adherent interaction. These factors are broadly categorized as pilus and non-pilus adhesins.

#### 1.2.1 Pilus Adhesins

Adhesive pili, or fimbriae, are long polymeric fibers that have been described as hair-like. The fibers are composed of a single major subunit plus minor subunits with specialized functions. They can be several micrometers long and are generally no more than 10 nm in width. Pili are divided into three subcategories based on their biogenesis: chaperone-usher assembled, type IV pili, and curli.

##### 1.2.1.1 Chaperone-Usher Pili

Chaperone-usher assembled pili were the first to be described after being observed by Anderson et al. in 1949 (Anderson 1949). In this class of pili, subunits are escorted to the outer membrane by a cognate chaperone protein and then through the outer membrane by an usher protein. The fiber is formed as the subunits are translocated to the extracellular space. There are two common types of pili in this class, type 1 pili and P-pili. Type 1 pili have been most thoroughly studied in *E. coli* as the
product of the *fim* gene cluster. This gene cluster encodes *fimA-fimH*. FimA is the major pilus subunit, FimC is the chaperone, and FimD is the usher protein. The adhesive activity of type 1 pili is found in the minor subunit FimH, which is incorporated into the distal end of the pilus (Jones 1995). FimH adopts an Ig-like fold that is completed by its interaction with FimC (Choudhury et al. 1999, Zhou et al. 2001, Baorto et al. 1997). Type 1 pili are approximately 7 nm wide and 1.2 μm long (Hahn et al. 2002). They are found in most *E. coli* isolates and have an affinity for monomannose and trimannose containing glycoproteins. P-pili, or pyelonephritis-associated pili, are encoded by a pathogenicity island in uropathogenic *E. coli* (UPEC). This system is encoded by the *pap* gene cluster with PapA as the major subunit, PapD as the chaperone, and Papc C as the usher. Adhesive activity is found in the PapG minor subunit (Thanassi et al. 1998). P-pili are approximately 7 nm wide, and several μm long, and recognize glycolipids in the urinary tract (Kuehn et al. 1992, Bullitt and Makowski 1995, Hultgren, Normark, and Abraham 1991). Variation in this receptor is a factor that influences host susceptibility to UPEC urinary tract infections, while variation in PapG allows the organism to infect a wider range of individuals (Lindstedt et al. 1991, Lindstedt et al. 1989, Lund et al. 1988, Stromberg et al. 1990, Johanson, Lindstedt, and Svanborg 1992).
Figure 1: Chaperone User Assembled Pilus Biogenesis

Pili subunits FimA and FimH are secreted through the Sec section system, guided through the periplasm by the FimC chaperone, and secreted through the outer membrane by the usher protein FimD.

1.2.1.2 Type IV pili

Type IV pili are found in a wide variety of gram-negative species, with examples including *E. coli*, *Salmonella enterica* serovar *typhi*, *Pseudomonas aeruginosa*, *Legionella pneumophila*, *Neisseria gonorrhoeae*, *Neisseria meningitides*, and *Vibrio cholerae* (Craig, Pique, and Tainer 2004). The subunits of type IV pili are incorporated into the inner membrane as pre-pilins, processed by a pre-pilin peptidase and assembled into fibers in the periplasm before being exported by specialized machinery. Pilin subunits share certain characteristics, including a short hydrophobic leader sequence that is cleaved during pilus biogenesis, a methylated N-terminal residue, and two cysteine residues near the C-terminus (Strom, Nunn, and Lory 1993). Structurally, type IV pilin subunits are predicted to have an N-terminal alpha-helix and a C-terminal globular domain (Parge et
Pilins are divided into two subgroups: type IVa pilins, which have a shorter leader sequence (around 10 amino acids) and are approximately 150-160 amino acids in length, and type IVb pilins, which have a longer leader sequence (up to 30 amino acids) and can be up to 200 amino acids in length (Ramer et al. 2002, Kirn, Bose, and Taylor 2003). Type IV pili are 6-8 nm wide and can be several µm long. In some species, such as enteropathogenic *E. coli* and *Vibrio cholerae*, type IV pilus fibers aggregate to form bundles. Type IV pili are unique in their ability to be retracted by the organism. This retraction is facilitated by an inner membrane retraction ATPase capable of generating forces up to 1 nN, the strongest molecular linear motor described to date (Biais et al. 2008). Retraction is responsible for a form of motility called twitching motility, in which the organism pulls itself across a surface by extending, anchoring, and then retracting the pili. Type IV pili-mediated retraction is also involved in DNA uptake by a number of species.

Type IV pilus biogenesis and secretion involves a large number of proteins. In the well-studied type IV pilus systems of the *Neisseria* species, the leader sequence of the major pilin subunit PilE is cleaved and the resulting N-terminus is methylated by PilD (Nunn, Bergman, and Lory 1990, Lauer, Albertson, and Koomey 1993, Strom, Nunn, and Lory 1993). The pilus assembly complex is made up of the cytoplasmic protein PilM, the periplasmic proteins PilN, PilO, and PilP, the inner membrane spanning protein PilG, and the oligomeric outer membrane secretin PilQ (Drake, Sandstedt, and Koomey 1997,
Figure 2: Type IV Pilus Biogenesis

Pilus fibers are made up of the major pilin subunit PilE and a variety of minor subunits including PilK, PilH, PilI, PilJ, and PilC. These subunits are cleaved by the pre-pilin peptidase PilD in the inner membrane, and secreted through the oligomeric PilQ out membrane pore by the pilus biogenesis proteins PilF, PilG, PilN, PilO, and PilP. Energy for secretion is provided by the PilM ATPase. Energy for retraction is provided by the PilT ATPase. In the above diagram the pilus subunits are indicated by “PilX”

1.2.1.3 Curli

Curli are adhesive amyloids expressed by E. coli with no known specific ligand. These proteins are very sticky and bind a number of extracellular matrix proteins, including fibronectin and laminin (Olsen, Jonsson, and Normark 1989, Olsen et al. 1993). Curli biogenesis involves the proteins CsgA, B, E, F, and G. These proteins are secreted into the periplasm in a Sec-dependent manner. Once in the periplasm, CsgG inserts into the outer membrane and forms a pore. CsgG secretes CsgA and CsgB. Once
secreted, CsgB nucleates the amyloid fiber composed of CsgA. CsgE and CsgF act as chaperones to the CsgA and CsgB proteins in the periplasm (Hammar et al. 1995, Herwald et al. 1998).

Figure 3: Curli Biogenesis

Curli subunits are secreted through the Sec system and chaperoned through the periplasm by CsgF and CsgE. Subunits are secreted through the CsgG outer membrane pore and CsgB nucleates the folding of the CsgA subunits.

1.2.2 Non-Pilus Adhesins

Non-pilus adhesins are monomeric or oligomeric proteins that facilitate adherence. Generally, non-pilus adhesins are much shorter than pili, requiring the bacterium to be in closer contact to a surface in order to facilitate adherence. The largest class of non-pilus adhesins are those exported by the type V secretion system, though there are others, including other types of outer membrane proteins, fibronectin binding proteins, and LPS.
1.2.2.1 Type V Secretion

Members of type V secretion systems are encoded in genetic loci that include all the necessary proteins to transport the passenger domain to the bacterial cell surface. In this family there are three classes: type Va or classical autotransporters, type Vb or two partner secretion systems, and type Vc or trimeric autotransporters.

1.2.2.2 Type Va (Classical) Autotransporters

Type Va (Classical) autotransporters have an N-terminal signal sequence, internal passenger domain, and C-terminal outer membrane anchoring domain. The N-terminal signal sequence directs the protein to the Sec secretion system and is cleaved during export by signal peptidase I. Once in the outer membrane, the C-terminal domain forms a β-barrel pore through which the internal passenger domain transits. The passenger domain may remain associated with the pore or be cleaved, depending on the protein’s functionality. Autotransporters have been implicated as adhesins, proteases, peptidases, lipases, actin-motility promoting proteins, immunomodulators, and toxins (Coutte et al. 2001, Henderson and Nataro 2001, Henderson, Navarro-Garcia, and Nataro 1998, Yen et al. 2002).
1.2.2.3 Type Vb Two Partner Secretion Systems

The Type Vb secretion system members are encoded in two genes. The first gene *(tpsA)* encodes a passenger domain with an extracellular function. The second gene *(tpsB)* encodes an outer membrane pore through which the TpsA protein transits. The TpsA protein is secreted into the periplasm in a Sec-dependent manner and the N-terminal signal sequence is cleaved (Kim et al. 2007, Tommassen 2007). The resulting N-terminus encodes the TPS domain that is recognized by the periplasmic polypeptide-associated-translocator (POTRA) domain of the TpsB protein (Jacob-Dubuisson et al. 2009). Once on the surface the TpsA protein can remain anchored to the TpsB protein or be released into the extracellular space. The final location of the TpsA protein depends on its biological function.
In two partner secretion systems, the TpsB protein forms an outer membrane pore through with the TpsA protein is secreted.

1.2.2.4 Type Vc (Trimeric) Autotransporters

Similar to type Va autotransporters, type Vc autotransporters have an N-terminal signal sequence, internal passenger domain, and C-terminal outer membrane domain (Cotter, Surana, and St Geme 2005). However, the C-terminal outer membrane domain is approximately one third of the size of the domain in type Va autotransporters and therefore type Vc autotransporters must trimerize to form a functional pore. The proteins are secreted into the periplasm in a Sec-dependent manner and the signal sequence is cleaved by signal peptidase 1 (Klauser, Pohlner, and Meyer 1993, Maurer, Jose, and Meyer 1999, Suzuki, Lett, and Sasakawa 1995, Shannon and Fernandez 1999). In the outer membrane, the proteins trimerize to form a pore and the associated passenger domains are translocated to the cell surface where they trimerize by means of
a predicted coiled-coil domain (Hoiczyk et al. 2000). To date, all characterized type Vc autotransporters mediate adherence.

Figure 6: Type Vc Trimeric Autotransporter

The C-termini of trimeric autotransporters oligomerize in the outer membrane to form and pore through with their attached N-terminal passenger domains are secreted.

1.3 Nontypeable Haemophilus influenzae Adherence

NTHi is dependent on colonization of the human host for persistence, as there are no other known reservoirs besides the human upper respiratory tract. NTHi is also dependent on colonization for pathogenesis as organisms spread contiguously from the site of colonization to sites of disease. Adherence is a key determinant of colonization, and NTHi expresses multiple adhesive proteins, including chaperone-usher assembled pili, type IV pili multiple autotransporter proteins, adhesive outer membrane proteins, and LOS (St Geme 2002). These adhesins adhere preferentially to mucus, non-ciliated cells, and damaged epithelium (St Geme 2002).
NTHi chaperone-usher assembled pili are encoded by ~15% of NTHi isolates in the hif gene cluster (Geluk et al. 1998, Krasan et al. 1999). The fibers are approximately 5 nm in diameter and 450 nm in length (Stull et al. 1984, St Geme et al. 1996). The hifA gene encodes the major subunit, and hifE encodes a minor subunit with adhesive activity. The product of hifB is the chaperone protein in the system, and hifC encodes the usher protein (St Geme et al. 1996). These peritrichous fibers mediate adherence to human oropharyngeal and nasopharyngeal epithelial cells, respiratory mucus, and fibronectin (van Alphen et al. 1991, Loeb, Connor, and Penney 1988, Farley et al. 1990, Read et al. 1991, Kubiet et al. 2000, Virkola et al. 2000). The receptor on epithelial cells has been recognized as a sialylated lactosylceramide derivate, though a specific protein has not been implicated (van Alphen et al. 1991). NTHi also encodes type IV pili that appear as bundles in electron micrographs. The pilin subunits do not fit nicely into either the IVa or IVb subtypes, as the leader peptide is longer than that of IVa subunits, but shorter than that of type IVb subunits. The major subunit also has an overall shorter length than either type IVa or type IVb major subunits, with an average length of 137 residues. NTHi type IV pili were not appreciated until 2005, as they are only expressed when the organism is grown in defined media at a basic pH and are not see when NTHi is grown in rich medium (Bakaletz et al. 2005).

While a relatively small percentage of NTHi isolates (~15%) encode hif pili, virtually all isolates encode multiple non-pilus adhesins (Barenkamp 1992, Rodriguez et
al. 2003, St Geme et al. 1998, St Geme and Grass 1998). There are four type V secreted proteins that NTHi can encode: Hap, Hia, HMW1, and HMW2.

Hap is a ubiquitous IgA protease-like type Va autotransporter that mediates low level adherence, invasion, bacterial aggregation, and microcolony formation (St Geme, de la Morena, and Falkow 1994, Hendrixson et al. 1997, Hendrixson and St Geme 1998). The adhesive activity of this protein is located in the passenger domain, yet in vitro the passenger domain is autoproteolyzed and released from the bacterial cell surface. This observation initially raised questions regarding the function of Hap as an adhesin (Hendrixson et al. 1997, Fink et al. 2001). However, further study revealed that Hap autoproteolysis is inhibited by human secretory leukocyte protease inhibitor, a component of mucus, allowing the passenger domain to remain attached to the bacteria and facilitate adherence (Hendrixson and St Geme 1998). This moderation of autoproteolytic activity may allow the bacteria to maintain a balance between adherence and distribution.

Hia is a trimeric autotransporter (type Vc) expressed by roughly a quarter of NTHi isolates. It is homologous to the Hsf adhesin of encapsulated *H. influenzae* (Barenkamp and St Geme 1996, Geme 1996, St Geme and Cutter 2000). Unlike Hap, Hia is not cleaved from the bacterial cell surface (St Geme and Cutter 2000). Hia mediates adherence to respiratory epithelial cells in vitro, though the specific receptor has yet to be identified (Laarmann et al. 2002, Geme 1996).
HMW1 and HMW2 are homologous type Vb two partner secretion systems expressed by isolates that do not express Hia (approximately 75% of NTHi isolates). They facilitate high levels of adherence to different types of respiratory epithelial cells. These proteins are encoded in separate locations on the genome, and each locus also encodes a cognate outer membrane pore (HMW1B/HMW2B) through with the adhesive protein is translocated and a glycosyltransferase (HMW1C/HMW2C) that modifies the adhesive protein at multiple asparagines with mono- or di-hexose residues. The HMW1B and HMW2B proteins are 99% identical, and the HMW1C and HMW2C proteins are 97% identical, and the HMW1B/HMW2B and HMW1C/HMW2C proteins are interchangeable (Barenkamp and St Geme 1994, Grass and St Geme 2000). After being glycosylated in the cytoplasm, HMW1 and HMW2 are secreted into the periplasm in a Sec-dependent manner (Grass and St Geme 2000). In the periplasm, the proteins are recognized by the periplasmic polypeptide-associated-transport (POTRA) domain of the HMWB pores and translocated to the bacterial surface (Grass, Rempe, and St Geme 2015). HMW1 and HMW2 remain anchored in the HMWB pore by means of an intramolecular di-sulfide bond (Buscher et al. 2006). HMW1 and HMW2 are 71% identical, yet facilitate adherence to different cell types in vitro (Barenkamp and Leininger 1992). The HMW1 receptor has been identified as an N-linked oligosaccharide with sialic acid in a α2-3 configuration (St Geme 1994). The HMW2 receptor has yet to be identified.
1.3.1 Other NTHi Adhesins

In addition to Hif pili, type IV pili, and members of the type V secretion system, NTHi encodes a number of other adhesins. These include the outer membrane proteins P2 and P5, OapA, and LOS. P2 is the most abundant NTHi outer membrane protein and is capable of interacting with mucin (Hansen et al. 1988, Reddy et al. 1996). P5 is another major outer membrane protein and is similar to OmpA in *E. coli*. P5 interacts with CEACAM1 and is also capable of forming a pilus-like structure (Munson, Grass, and West 1993, Sirakova et al. 1994, Virji 2000, Hill et al. 2001). OapA is a surface associated lipoprotein that Weiser et al showed to be necessary for colonization of the nasopharynx in an infant rat model and plays a minor role in *in vitro* adherence to epithelial cells (Weiser et al. 1995, Prasadarao et al. 1999). Finally, the lipid A moiety of NTHi LOS is capable of adhering to human bronchial epithelial cells (Swords et al. 2000). Swords also showed that this binding is dependent on modification of lipid A with phosphoylcholine, a modification facilitated by protein D (Swords et al. 2000).

1.4 *Kingella kingae* Adherence

Adherence is the first step in *K. kingae* colonization and the pathogenesis of *K. kingae* disease. Invasive isolates are often genetically identical to colonizing isolates in the same patient, showing that invasion begins with colonization (Yagupsky et al. 2009). *K. kingae* is capable of adhering to human epithelial cells and synovial cells *in vitro* (Kehl-Fie, Miller, and St Geme 2008). Adherence is facilitated type IV pili and a trimeric
autotransporter (type Vb) called Knh and is inhibited by a polysaccharide capsule. Type IV pili mediate the initial adherence to host cells, likely through the pilus-associated PilC1 and PilC2 proteins (Kehl-Fie, Miller, and St Geme 2008, Porsch et al. 2013). Retraction of the pili is postulated to displace the polysaccharide capsule, allowing for Knh to mediate tight adherence to the host cell (Porsch, Kehl-Fie, and St Geme 2012). The specific receptors for *K. kingae* type IV pili and Knh have yet to be identified.

**Figure 7: Model of *K. kingae* adherence**

In the current model of *K. kingae* adherence, type IV pili mediated the initial interaction with the host cell. Retraction of the pili then displaces the polysaccharide capsule allowing for tight adherence mediated by the Knh trimeric autotransporter. (Figure from Porsch, 2012)
*K. kingae* type IV pili are regulated by PilR/PilS, a two-component system. Colonizing isolates show the highest levels of piliation, arguing that pili are important in adhering to the respiratory epithelium. Isolates from bone and joints show low to no piliation, with blood isolates being piliated at an intermediate level (Kehl-Fie et al. 2009, Kehl-Fie et al. 2010). This gradation in pilus density suggests that piliation is selected against in distal sites of infection.

1.5 *Aggregatibacter aphrophilus* Adherence

As a resident bacteria of the human upper respiratory tract, adherence to the respiratory epithelium by *A. aphrophilus* is likely necessary for colonization and pathogenesis but has not been studied to date. The first published work examining *A. aphrophilus* adherence is presented here in Chapter 4.

1.6 Bacterial Protein Glycosylation

Glycosylation, or the addition of sugar molecules to protein, is a means by which organisms can increase the functional diversity of their proteomes. Neuberger et al discovered protein glycosylation in Eukaryotes the 1930s, and for the next 40 years it was considered restricted to that domain of life (Neuberger 1938). In the 1970s, prokarytoic glycosylation was recognized on the S-layers of *Halobacterium salinarum* and *Clostridium* species (Mescher and Strominger 1976, Glaubert and Sleytr 1975, Sleytr and Thorne 1976). Modifying sugars can be single hexose residues or complex polysaccharides, and proteins can be decorated at one or multiple sites.
1.6.1 O-Linked Glycosylation

O-linked glycosylation is one of the two main forms of glycosylation and occurs on the hydroxyl oxygen of a serine or threonine residue. General protein O-glycosylation systems involve the sequential transfer of activated sugars to a protein in the cytoplasm. Research into this system is relatively new, but general O-glycosylation systems have been recognized in *Neisseria* and *Bacteroides*, though they are not yet well-characterized (Ku et al. 2009, Vik et al. 2009, Coyne et al. 2005, Krinos et al. 2001, Kalka-Moll et al. 2001, Fletcher et al. 2009). The bacterial O-glycosylation systems for the modification of flagella and pili have been more extensively studied. These systems are found in *Campylobacter, Clostridium, Pseudomonas*, and *Neisseria* (Thibault et al. 2001, Guerry et al. 2006, Ewing, Andreishcheva, and Guerry 2009). Flagellar/pilin glycosylation has been shown to be necessary for fiber biogenesis, fiber stability, and fiber-mediated autoaggregation (Guerry et al. 2006, Ewing, Andreishcheva, and Guerry 2009).

1.6.2 N-linked Glycosylation

The other main form of glycosylation, N-linked glycosylation, is better characterized than O-linked in bacteria. N-glycosylation takes place on the amide nitrogen of asparagine residues in the consensus sequence Asparagine-Xaa-Serine/Threonine, where Xaa is any amino acid besides proline. Bacterial N-glycosylation was first discovered in the 1990s in *Campylobacter jejuni* (Szymanski et al.
In this system, UDP N-acetylglucosamine is converted into \( \text{GalNAc}\cdot[\text{Glc}]\text{GalNAc}\cdot\text{diNAcBac-PP-Und} \) by the sequential action of a C6 dehydratase (PgIF), an aminotransferase (PgIE), an acetyltransferase (PgID), and four glycosyltransferases (PgIC, PgIA, PgIJ, PgIH, and PgII) (Olivier and Imperiali 2008, Rangarajan et al. 2008, Linton et al. 2005, Glover et al. 2006, Weerapana et al. 2005, Glover et al. 2005, Troutman and Imperiali 2009, Schoenhofen et al. 2006). Und-PP is a lipid carrier in the inner membrane. The sugar chain is built in the cytoplasm and is then flipped into the periplasm by PgIK (Alaimo et al. 2006, Kelly et al. 2006). Once in the periplasm, the sugar chain is a substrate for the oligosaccharyltransferase PgIB, which transfers it to the N-glycosylation consensus sequence on over 65 different substrate proteins in \textit{C. jejuni} (Wacker et al. 2002). There are PgIB homologues in at least 49 species in addition to \textit{C. jejuni}, suggesting that this method of protein modification is wide-spread.


1.6.3 HMWC-mediated N-linked Glycosylation

\textit{NTHi} employs an unusual protein glycosylation system that seemingly combines aspects of O-linked and N-linked glycosylation. The cytoplasmic proteins HMW1C and
HMW2C are glycosyltransferases that modify the TpsA proteins HMW1 and HMW2 (Grass et al. 2003, Gross et al. 2008). HMW1C/HMW2C modify HMW1/HMW2 by sequentially adding UDP-glucose or UDP-galactose to multiple sites in a sequential fashion so that each site is modified with either a mono- or di-hexose, which is similar to the system of O-glycosylation described earlier. However, HMW1/HMW2 are modified on the amide nitrogen of an asparagine residue in the conventional N-glycosylation consensus sequence (Grass et al. 2010). This mechanism of action makes HMW1C/HMW2C unique in a number of ways, including the ability to perform N-linked glycosylation in the cytoplasm, transfer sugars sequentially to the N-linked consensus sequence, and catalyze both peptide-sugar and sugar-sugar bonds.
Figure 8: Comparison of Conventional and HMW1C-Mediated N-linked Glycosylation

In conventional bacterial N-linked glycosylation, eight glycosyltransferases build a heptasaccharide change anchored in the inner membrane. This chain is then flipped across the membrane into the periplasm and transferred to an acceptor protein by the PglB oligosaccharyltransferase. HMW1C-mediated glycosylation occurs in the cytoplasm and involves the transfer of UDP-activated hexose onto an acceptor protein by the HMW1C glycosyltransferase.

1.7 Thesis Overview

Chapter 2 describes an animal model of NTHi colonization that is far superior to the current small-mammal models used. This model is then used to show the importance of the HMW1 and HMW2 proteins in colonization. Chapter 3 investigates further the roles of HMW1B and HMW1C in the expression of functional HMW1. Chapter 4 broadens the understanding of HMW1C-mediated glycosylation by showing
that this mechanism of modification is employed by species other than NTHi. Finally, Chapter 5 summarizes the implications of this work and suggests future directions.
2. The HMW1 and HMW2 adhesins enhance the ability of nontypeable *Haemophilus influenzae* to colonize the upper respiratory tract of rhesus macaques

2.1 Introduction

Nontypeable *Haemophilus influenzae* (NTHi) is a human-specific organism that colonizes the upper respiratory tract of nearly half of all children before the age of two years (Faden et al. 1996, Puig et al. 2014, Spinola et al. 1986, Trottier, Stenberg, and Svanborg-Eden 1989). In the vast majority of individuals, colonization remains asymptomatic and has no adverse effects on the host. However, under the proper circumstances NTHi will spread contiguously within the respiratory tract to produce localized respiratory tract disease. NTHi is a common cause of otitis media, sinusitis, and conjunctivitis in young children and a frequent etiology of community-acquired pneumonia and exacerbations of chronic obstructive pulmonary disease (COPD) in adults (Giebink 1989, Murphy and Sethi 1992). Colonization of the upper respiratory tract is a prerequisite for the pathogenesis of NTHi disease, underscoring the importance of understanding the bacterial and host determinants of colonization.

Animal models are widely used to study host-bacterial interactions and are valuable for examining the *in vivo* roles of putative bacterial virulence factors. Animal models that recapitulate human colonization allow investigators to assess the applicability of *in vitro* findings and facilitate efforts to prevent disease. The existing
small animal models for NTHi colonization have significant limitations, in part reflecting the differences between small animal and human anatomy and physiology and potentially also reflecting the human specificity of NTHi. Chinchillas are often used as a model for otitis media and have advantages for studying NTHi disease but are suboptimal as a model for colonization (Yang et al. 1998).

*In vitro* studies using cultured human epithelial cells have demonstrated that the NTHi HMW1 and HMW2 proteins are the major adhesins expressed by 75-80% of clinical isolates (St Geme, Falkow, and Barenkamp 1993, Barenkamp and St Geme 1996, Krasan et al. 1999, Laarmann et al. 2002, St Geme 1997a, St Geme et al. 1998). HMW1 and HMW2 are highly homologous glycoproteins that are presented on the bacterial surface by the two-partner secretion system. In strain 12, HMW1 and HMW2 are 71% identical and 80% similar and have maximal sequence divergence in the binding domain (Barenkamp and Leininger 1992). In experiments using a panel of human epithelial cell types, HMW1 and HMW2 have different cellular binding specificities, suggesting that these two proteins interact with different host cell receptors (Dawid, Grass, and St Geme 2001, St Geme, Falkow, and Barenkamp 1993).

In previous work, Weber et al. described the use of rhesus macaques as a model for *H. influenzae* type b colonization (Weber et al. 1991). Given this information, we set out to examine the ability of NTHi to colonize rhesus macaques. Here we show that NTHi is capable of colonizing the upper respiratory tract of rhesus macaques without
causing disease, similar to colonization in humans. In addition, we establish that both HMW1 and HMW2 play an important role in facilitating colonization of macaques.

### 2.2 Methods

#### 2.2.1 Bacterial Strains and Culture

Spontaneous streptomycin-resistant derivatives of NTHi strain 12 were generated by spreading a dense bacterial suspension on agar plates containing 500μg/ml streptomycin and recovering survivors. Interruption of hmw1 and/or hmw2 was achieved by insertion of a kanamycin resistance cassette.

Bacteria were grown on chocolate agar or on BHI agar supplemented with 0.1% v/v lysed horse blood as a source of hemin and 3.5μg/ml NAD (BHIs agar), with 500μg/ml streptomycin and/or 50μg/ml kanamycin as appropriate. Agar plates were incubated at 37°C with 5% CO₂, and growth was checked at 24 and 48 hours.

**Table 1: Bacterial Strains used in Chapter 2**

<table>
<thead>
<tr>
<th>Strain</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>Nontypeable Haemophilus influenzae, expresses HMW1 and HMW2</td>
<td>Clinical Isolate, middle ear fluid, acute otitis media</td>
</tr>
<tr>
<td>12hmw2</td>
<td>12hmw2::kan, expresses HMW1</td>
<td>St. Geme 1993</td>
</tr>
<tr>
<td>12hmw1/hmw2</td>
<td>12A hmw1::kan, hmw2::kan, expresses neither HMW1 nor HMW2</td>
<td>St. Geme 1993</td>
</tr>
<tr>
<td>12hmw1</td>
<td>12A hmw1::kan, expresses</td>
<td>St. Geme 1993</td>
</tr>
<tr>
<td></td>
<td>Strain Description</td>
<td>Notes</td>
</tr>
<tr>
<td>---</td>
<td>-------------------</td>
<td>-------</td>
</tr>
<tr>
<td>5</td>
<td>Nontypeable Haemophilus influenzae, expresses HMW1 and HMW2</td>
<td>Clinical Isolate, Acute Otitis Media</td>
</tr>
<tr>
<td>5hmw2</td>
<td>5hmw2::kan, expresses HMW1</td>
<td>Buscher 2004</td>
</tr>
<tr>
<td>5hmw1/hmw2</td>
<td>5A hmw1::kan, hmw2::kan, expresses neither HMW1 nor HMW2</td>
<td>Buscher 2004</td>
</tr>
<tr>
<td>5hmw1</td>
<td>15A hmw1::kan, expresses HMW2</td>
<td>Buscher 2004</td>
</tr>
<tr>
<td>15</td>
<td>Nontypeable Haemophilus influenzae, expresses HMW1 and HMW2</td>
<td>Clinical Isolate, Acute Otitis Media</td>
</tr>
<tr>
<td>15hmw2</td>
<td>15A hmw2::kan, expresses HMW1</td>
<td>Buscher 2004</td>
</tr>
<tr>
<td>15hmw1/hmw2</td>
<td>15A hmw1::kan, hmw2::kan, expresses neither HMW1 nor HMW2</td>
<td>Buscher 2004</td>
</tr>
<tr>
<td>15hmw1</td>
<td>15A hmw1::kan, expresses HMW2</td>
<td>Buscher 2004</td>
</tr>
</tbody>
</table>

### 2.2.2 Adherence Assays

Quantitative adherence assays were performed by seeding 24-well tissue culture plates with 4MBr-5 rhesus macaque bronchial epithelial cells (ATCC CCL 208), and allowing monolayers to form overnight. The following day, the spent media was removed from the wells and replaced with 300 µl of fresh tissue culture media.
Subsequently, approximately $1 \times 10^7$ CFU of bacteria were inoculated onto monolayers. Following incubation for 30 minutes at 37°C, monolayers were rinsed four times with PBS to remove non-adherent bacteria. Adherent organisms were released from the monolayer using trypsin-EDTA and were quantitated by plating dilutions on chocolate agar plates. CFU counts of adherent bacteria were compared to CFU counts of the original inoculum to determine the percentage of bacteria that were adherent.

### 2.2.3 Animal Studies

All animal work was approved by the Institutional Animal Care and use Committee at the Children’s Hospital of Philadelphia under Protocol IAC 14-001133. Animals were sedated prior to all procedures. Sedation was achieved with ketamine alone at 10mg/kg or ketamine (3-5 mg/kg) with dexmedetomidine (0.02-0.05 mg/kg) and was reversed with atipamezole (0.2-0.5 mg/kg) as soon as the procedure was over.

Prior to initiating the colonization studies, samples of the pharyngeal and nasopharyngeal flora were obtained by swabbing the nasopharynx and pharynx of sedated animals with sterile cotton swabs. For nasopharyngeal swabbing, a BD BBL CultureSwab Plus Amies gel with charcoal - flexible aluminum wire swab (product # 220125) was passed into the nares approximately to the distance of the ear. The swab was left in place for a few seconds and then slowly withdrawn with a rotating motion. For pharyngeal swabbing, a BD BBL CultureSwab Plus Amies gel with charcoal - single
swab (product # 220121) was passed through the mouth to the area near the crypts of the tonsils, rolled across the mucosa, and then withdrawn. The swabs were then streaked on *Haemophilus* agar plates (BD catalog #221954) within one hour of collection. The plates were incubated at 37°C with 5% CO₂ and checked at 24 and 48 hours for bacterial growth. Any resulting colony types were Gram stained, and the gram-negative colonies where characterized using the RapidID NH system. Using this approach, all animals were found to be culture negative for *H. influenzae* at baseline.

After confirming that animals were not naturally colonized with *H. influenzae*, they were inoculated with the relevant NTHi strain resuspended in PBS. Using 1mL syringes, 50µl was inoculated into each nare, for a total inoculation volume of 100µl. Following inoculation, the animals remained recumbent with their heads tilted back for about 60 seconds and were then placed back into their cages.

### 2.2.4 Bacterial Recovery

The heads of the swabs obtained from the animals were cut off and placed in 1mL of sterile PBS in capped 1.5 mL tubes. The tubes containing the swab heads were vortexed vigorously for 15 seconds to resuspend the bacteria, and resuspensions were serially diluted and plated on BHIs agar plates. The limit of detection was 10 CFU/mL.

### 2.2.5 ELISA

96-well high binding ELISA plates (Greiner BioOne #655061) were coated with 100 µl of 200nM purified HMW1 in bicarbonate buffer pH 9.6. Plates with coating
protein were incubated at 4°C for 16 hours. Following incubation plates were washed three times with PBS+0.05% Tween-20 (PBS-T) and blocked with 200 µl of 3% BSA in PBS-T. Wells were blocked for one hour at room temperature. Block was then removed and wells rinsed twice with PBS-T. 100 µl of serial dilutions of animal serum in PBS-T were added to the wells and the plates incubated for two hours at room temperature. Following incubation, plates were washed three times with PBS-T and 100 µl anti-rhesus macaque or anti-guniea pig antiserum was added at a dilution of 1:2,000 in 3% BSA PBS-T. The secondary antibody was incubated in the wells for one hour at room temperature. Plates were then washed three times with PBS-T and developed with TMB peroxidase ELISA substrate (Rockland #TMBE-100) for five minutes. Absorbance at 370nm was read on a PerkinElmer EnSpire plate reader.

**2.2.6 SDS-PAGE**

Bacteria were resuspended to an OD$_{600}$ of 1.0 and sonicated with 30-second pulses until the solution cleared. The resulting whole cell lysates were resolved on a 10% SDS-PAGE gel at 175V for 1 hour.

For Western analysis, proteins were transferred from SDS-PAGE gels to nitrocellulose membranes, and membranes were blocked in 5% non-fat milk/TBS for 30 minutes. Membranes were then incubated overnight at 4°C in 5% non-fat milk/TBS with rhesus macaque serum at a dilution of 1:500. Subsequently, membranes were washed with TBS-Tween and incubated for 1 hour in 5% non-fat milk/TBS and a 1:5000 dilution
of anti-rhesus macaque IgG IgA IgM, then developed using Thermo Fischer Super
Signal substrate.

**2.2.7 Statistics**

Significance for adherence assays was determined using the unpaired t-test with
Welch’s correction.

**2.3 Results**

**2.3.1 NTHi adheres to rhesus macaque respiratory epithelial cells in vitro**

As a first step to determine if rhesus macaques might be an appropriate model to
study NTHi colonization, we examined three epidemiologically unrelated prototypic
strains of NTHi in in vitro adherence assays with 4MBr05 rhesus macaque bronchial
epithelial cells (ATCC CCL 208). As shown in Figure 9, all of these strains were capable
of adherence.

To address whether the HMW1 and HMW2 proteins influence adherence to
4MBr05 cells, we compared the parent strains and isogenic mutants lacking HMW1,
HMW2, or both HMW1 and HMW2. As shown in Figure 9, adherence to the 4MBr5
cells was dependent on HMW1 and/or HMW2 expression in all three strains. With
strain 12, adherence was dependent on both HMW1 and HMW2, with HMW1
contributing more than HMW2. With strain 5, adherence was dependent only on
HMW1. With strain 15, adherence was dependent on both HMW1 and HMW2 and was
only low level when only HMW1 or HMW2 was present by itself.
Figure 9: NTHi adherence to rhesus macaque bronchial epithelial cells is dependent on the HMW1 and HMW2 proteins.

The mean adherence to rhesus macaque bronchial epithelial cells by three strains of NTHi expressing both HMW1 and HMW2, HMW1 only, HMW2 only, or neither HMW1 nor HMW2 is shown. The graph shows data from biological replicates performed in triplicate. Error bars represent standard error.

2.3.2 NTHi is able to colonize the upper respiratory tract of rhesus macaques

To determine if NTHi is capable of colonizing the upper respiratory tract of rhesus macaques, we inoculated animals with $10^2$, $10^3$, $10^4$, or $10^5$ CFU of streptomycin-resistant strain 12 intranasally, using 2 animals for each inoculum. Over a two-week period, the nasopharynx and pharynx of the inoculated animals were swabbed, and recovered NTHi were quantified. As shown in Figure 10, an inoculum of $10^5$ CFU resulted in stable colonization of both the pharynx and nasopharynx throughout the two weeks in both animals. An inoculum of $10^4$ CFU resulted in stable colonization.
throughout the two weeks in one animal and transient colonization on one day in the other animal. Inoculation with $10^3$ CFU or $10^2$ CFU failed to achieve colonization.

Figure 10: An inoculum of $10^5$ CFU of NTHi strain 12 results in consistent colonization of rhesus macaques.

The frequency at which wild type NTHi strain 12 colonizes the nasopharynx and nasopharynx of rhesus macaques when inoculated intranasally at the indicated inocula is shown. An inoculum of $10^5$ CFU resulted in colonization of both animals used by the end of two weeks.

Over the course of the experiment, the weight of the animals was recorded to assess potential distress. As shown in Table 2, there were no significant changes in weight. In addition, no other signs of distress such as redness of the mucosa, nasal discharge, or change in behavior were observed, leading us to conclude that the colonizing bacteria did not cause disease in the animals.
Table 2: Animal Weights over the Course of the Experiment

<table>
<thead>
<tr>
<th>Day Post Inoculation</th>
<th>Animal AA Weight (Kg)</th>
<th>Animal AB Weight (Kg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>9.8</td>
<td>9.2</td>
</tr>
<tr>
<td>1</td>
<td>9.8</td>
<td>9.26</td>
</tr>
<tr>
<td>2</td>
<td>9.8</td>
<td>9.18</td>
</tr>
<tr>
<td>3</td>
<td>9.8</td>
<td>9.22</td>
</tr>
<tr>
<td>4</td>
<td>9.78</td>
<td>9.1</td>
</tr>
<tr>
<td>5</td>
<td>9.8</td>
<td>9.2</td>
</tr>
<tr>
<td>8</td>
<td>9.78</td>
<td>9.2</td>
</tr>
<tr>
<td>11</td>
<td>9.84</td>
<td>9.02</td>
</tr>
<tr>
<td>14</td>
<td>9.92</td>
<td>9.04</td>
</tr>
</tbody>
</table>
2.3.3 The HMW1 and HMW2 proteins facilitate colonization of rhesus macaques

To assess the contribution of the HMW1 and HMW2 proteins to colonization of rhesus macaques, we performed a competition experiment in which six animals were inoculated with $10^5$ CFU consisting of 50% streptomycin-resistant NTHi strain 12 parent and 50% streptomycin- and kanamycin-resistant NTHi strain 12 lacking both HMW1 and HMW2 (strain 12 \textit{hmw1hmw2}). Colonization was assessed by quantitating pharyngeal and nasopharyngeal samples on days 1, 2, 3, 4, 5, 8, 11, 14, 17, and 21 post-inoculation. By plating on agar plates containing streptomycin alone versus streptomycin and kanamycin, we were able to differentiate between colonization by the strain 12 parent (streptomycin resistant) and the strain 12 mutant (12\textit{hmw1hmw2}) (streptomycin and kanamycin resistant).

On the first day post inoculation, NTHi was recovered from the nasopharynx of all six animals. Of the total recovered colonies, 86% were the strain 12 parent. Samples from animals A, C, and D contained only the strain 12 parent, and samples from animals B, E, and F contained a mix of the strain 12 parent and strain 12 \textit{hmw1hmw2}. On day 2, 95% of the total recovered colonies were the strain 12 parent, and only the sample from animal B contained strain 12 \textit{hmw1hmw2}. On days 3 and 4, over 98% of the total recovered colonies were the strain 12 parent, and again only animal B was colonized with strain 12 \textit{hmw1hmw2}. From day 5 through the end of the experiment on day 21, all recovered colonies from all animals were the strain 12 parent. Examination of
pharyngeal cultures revealed colonization of only one animal on day 1, three animals on
day 2, five animals on days 3 and 4, and finally all animals on day 5. By day 21 only two
animals remained colonized. Only the strain 12 parent was recovered from the pharynx
(Figure 11).
Figure 11: HMW1 and HMW2 in NTHi colonization of rhesus macaques

Panel A: Five days after inoculation, 100% of colonies recovered from the nasopharynx of the six animals were the parent strain, thereby demonstrating that expression of the HMW1 and HMW2 proteins facilitates colonization. Panel B: The percentage of colonies recovered from each animal on each day that were the parent strain is shown. From day 2 on, five of the six animals had only the parent strain. On day 5, the sixth animal had only the parent strain. Panel C: All bacteria recovered from the pharynx of the six animals were the parent strain, even though the pharynx of all six animals were not colonized until day 5 (panel D).

To investigate the individual contributions of the HMW1 and HMW2 proteins, animals were inoculated with $10^5$ CFU of the strain 12 parent, strain 12hmw1 (lacking HMW1 and expressing HMW2 only), strain 12hmw2 (lacking HMW2 and expressing
HMW1 only), or strain 12hmw1hmw2 (lacking both HMW1 and HMW2). Three animals
were inoculated with each strain, and colonization was assessed over a three-week
period. Two of the three animals inoculated with the strain 12 parent were colonized in
the nasopharynx and remained colonized over the course of the experiment, with
densities ranging between $2 \times 10^4$ and $5.2 \times 10^4$ CFU per swab. Among the animals
inoculated with strain 12 expressing HMW2 alone, two of three were colonized in the
nasopharynx on days 2 and 4 and all three were colonized on all other days, with
densities ranging from $3 \times 10^1$ to $1.7 \times 10^5$ and stabilizing at $\sim 5 \times 10^3$ at the end of the
experiment. Among the animals inoculated with strain 12 expressing HMW1 alone, all
three animals were colonized on all days post-inoculation, with densities ranging from
$2 \times 10^1$ to $2.8 \times 10^4$. The strain 12 derivative lacking HMW1 and HMW2 colonized the
nasopharynx of one animal, at a low density ($1 \times 10^1$-$1 \times 10^2$) during the first week and at a
higher density ($\sim 1 \times 10^5$) during the remainder of the experiment (Figure 11).

Colonization of the pharynx was more erratic. The number of animals colonized by the
strain 12 parent varied, with one animal colonized on days 3, 5, 14, and 17 and two
animals colonized on days 8, 11, and 21. The density of colonization among colonized
animals ranged between $1 \times 10^1$ and $1 \times 10^4$ CFU and was generally $<1 \times 10^3$ CFU. Strain 12
expressing HMW2 alone was recovered from the pharynx of one animal on days 2, 4, 14,
and 17, two animals on days 3, 5, 11, and 21, and three animals on day 8. The density of
colonization among colonized animals was consistently $\sim 1 \times 10^3$ CFU. Strain 12
expressing HMW1 alone colonized the pharynx most consistently, with all three animals colonized from day 5 onward, generally at a density of ~1x10^3. Strain 12 lacking both HMW1 and HMW2 colonized the pharynx of only one animal and was recovered only on days 5, 8, 14, and 17, at a density always \(\leq 1\times10^3\) (Figure 12).
Figure 12: HMW1 and HMW2 independently facilitate colonization of rhesus macaques

Panel A shows the frequency of colonization of the nasopharynx by parent strain 12 and strain 12 derivatives expressing only HMW1 (hmw2), only HMW2 (hmw1), or neither HMW1 nor HMW2 (hmw1/hmw2). Panel B shows the density of colonization of the nasopharynx by parent strain 12 and strain 12 derivatives expressing only HMW1 (hmw2), only HMW2 (hmw1), or neither HMW1 nor HMW2 (hmw1/hmw2). Panel C shows colonization frequency of the pharynx by parent strain 12 and strain 12 derivatives expressing only HMW1 (hmw2), only HMW2 (hmw1), or neither HMW1 nor HMW2 (hmw1/hmw2). Panel D shows the density of colonization of the pharynx by parent strain 12 and strain 12 derivatives expressing only HMW1 (hmw2), only HMW2 (hmw1), or neither HMW1 nor HMW2 (hmw1/hmw2). For Panels B and D, mean density is designated by a horizontal bar and error bars represent standard error.
2.3.4 Colonization is an immunizing event

To determine if colonization stimulates an immune response against the HMW1 and HMW2 proteins, we collected sera from animals before inoculation and after completion of the colonization experiment. The reactivity of paired samples against strain 12 whole cell lysates was compared by Western blot for rhesus macaque IgA, IgG, and IgM. Of the 18 animals used in our studies, post-colonization sera from six animals showed increased reactivity against strain 12 compared to the paired pre-colonization sera. For all six animals, the increased reactivity appeared as a band at the same molecular weight as HMW1 and HMW2 (Figure 13). When the sera were examined against a whole cell lysate of strain 12 *hmw1hmw2*, the additional reactive bands disappeared, suggesting that colonization resulted in generation of antibody against HMW1 and HMW2 (Figure 13).
Figure 13: Colonization with NTHi can be an immunizing event

Panel A: Sera from six of the 18 animals showed reactivity to an antigen at the same molecular weight as HMW1 or HMW2. The blot on the top left shows the correct size of HMW1. For each blot the # sign designates the 130kDa marker and the * sign designates a band at the size of HMW1. Blots labeled “Pre” were incubated with serum collected prior to colonization. Blots labeled “Post” were incubated with serum collected post colonization. Sera from NHP (non-human primate) 27, 32, 71, 75, 104, and 180 were reactive with a HMW1 sized band. NHP 227 is included as an example of serum that did not increase in reactive after colonization. Panel B: Sera from five animals reacted specifically with purified HMW1 in an ELISA. All sera were paired pre- and post-colonization. In panel B, non-reactive pairs are not shown, and the data from the pre-colonization serum samples is clustered behind the negative control.

2.4 Discussion

Characterization of the determinants of NTHi colonization of the upper respiratory tract has suffered from the lack of an ideal animal model. In this work, we
have shown that NTHi strain 12 is capable of colonizing both the nasopharynx and pharynx of rhesus macaques. Using this model, we have established that the HMW1 and HMW2 adhesins are important colonization factors \textit{in vivo}.

In recent work, Winokur et al. examined NTHi colonization of humans, inoculating the nasopharynx of 15 healthy adults with $1 \times 10^2\text{-}1 \times 10^5$ CFU of streptomycin-resistant strain 2019 and assessing colonization by plating nasal washes on media containing streptomycin. Using this protocol, they found that an inoculum of $\sim 1.5 \times 10^5$ resulted in colonization approximately 90\% of the time. None of the subjects developed fever or other severe symptoms of NTHi infection, but approximately half of the participants reported mild to moderate symptoms such as sore throat, rhinorrhea, nasal congestion, or mild cough, suggesting an inflammatory response to the inoculum (Winokur et al. 2013). In our study, we found that an inoculum of $1 \times 10^5$ CFU of streptomycin-resistant NTHi strain 12 resulted in colonization in 10 of 11 animals, mimicking the observations by Winokur et al. with human volunteers (Winokur et al. 2013). None of the animals in our study developed signs of infection, but it is difficult to exclude the possibility of sore throat or other pain. Although Winokur et al did not quantify the density of colonization in their volunteers, they ranked qualitative densities and observed substantial inter- and intra-subject qualitative variability in colonization density, similar to the variability in our colonized animals.
After establishing an inoculum of NTHI strain 12 that consistently results in colonization, we performed a competition experiment using parent strain 12 and a strain 12 derivative lacking both HMW1 and HMW2, aiming to determine if the HMW1 and HMW2 adhesins influence the process of colonization. While HMW1 and HMW2 have been studied extensively as adherence factors in vitro, prior to this study they had not been examined in an in vivo model. The data presented in this study argue strongly that HMW1 and HMW2 are important colonization factors, influencing the efficiency and the density of both early colonization and persistent colonization. In the competition experiment, within five days of inoculation, all NTHi recovered from both the nasopharynx and pharynx were the parent strain. Over the first four days of the experiment, the hmw1hmw2 mutant accounted for a small percentage of the organisms recovered from the nasopharynx of three animals. Interestingly, the double mutant was not recovered from the pharynx at any point during the experiment, suggesting either that HMW1 or HMW2 are required for colonization of the pharynx or that a HMW1 and HMW2 are required for movement from one site of colonization to another in the rhesus macaque model.

When we inoculated animals with the panel of hmw1/hmw2 mutants to determine the specific roles of HMW1 and HMW2, the strains expressing HMW1 alone or HMW2 alone were able to colonize all three inoculated animals, suggesting some level of redundancy between HMW1 and HMW2 regarding colonization. However, in vitro
experiments indicate that HMW1 and HMW2 interact with distinct receptors, suggesting that their functions are unlikely to be completely redundant in vivo (Dawid, Grass, and St Geme 2001). Additionally, the strain lacking both HMW1 and HMW2 was able to colonize one of three animals, indicating that the HMW1 and HMW2 proteins are not essential for colonization but appear to increase efficiency and density of colonization.

NTHi expresses a variety of adhesive proteins beyond HMW1 and HMW2, including type IV pili, P5, OapA, Protein E, and the autotransporter protein Hap, one or more of which might account for colonization in the absence of HMW1 and HMW2 (St Geme, de la Morena, and Falkow 1994, Brinton et al. 1989, Stull et al. 1984, St Geme et al. 1996, Loeb, Connor, and Penney 1988, Read et al. 1991, Rodriguez et al. 2003, Barenkamp and Leininger 1992, Hendrixson and St Geme 1998, Hansen et al. 1988, Reddy et al. 1996, Munson and Sasaki 1993, Sirakova et al. 1994, Hill et al. 2001, Prasadarao et al. 1999) In this experiment only two of the three animals inoculated with the parent strain became colonized, contrasting with the pilot experiment assessing inoculum and the competition experiment, in which 1x10^5 CFU of NTHi gave consistent colonization, raising the possibility of experimental error during the inoculation.

Hemagglutinating pili were indicated in *H. influenzae* type b (Hib) colonization of non-human primates by Weber et al. in 1991 (Weber et al. 1991). The NTHi strain used in this work does not express hemagglutinating pili and so we did not need to account for their adhesive activity in our study. Furthermore, HMW1 and HMW2 are expressed only by NTHi strains and therefore were not expressed by the Hib strain used by Weber (16
(Weber et al. 1991). Taken together the two studies clearly show that the determinants of colonization are different in Hib and NTHi.

In all of the animal experiments, we observed a delay of two to three days between colonization of the nasopharynx and colonization of the pharynx, possibly reflecting the need for the bacteria to travel from the site of the inoculation to the pharynx. Alternatively, while colonization density of both sites was comparable after day 3, density seemed to be more variable in the pharynx over the course of the experiment. The time lag and the variability could also indicate that the pharynx is a less preferred site of colonization, perhaps because the density of the HMW1 and HMW2 receptors is reduced, the microbiome is more competitive, or innate immune effectors are different in the pharynx versus the nasopharynx.

In six of the 18 animals used for this study, comparison of serum collected pre-inoculation and serum collected after infection revealed acquisition of reactivity to NTHi. In all six of these animals, the greater reactivity appeared as bands the size of HMW1 and HMW2 and disappeared when serum was examined against a whole cell lysate of the strain 12 derivative lacking HMW1 and HMW2. The HMW1 and HMW2 proteins are known to be highly immunogenic, and antibody to these proteins is often found in convalescent samples from patients with NTHi disease (Barenkamp and Bodor 1990). While HMW1 and HMW2 are well known to be immunogenic during disease, this study demonstrates for the first time that HMW1 and HMW2 stimulate an immune
response during colonization. We were not able to ascertain whether or not these antibodies are protective, though presence of antibody did not correlate with the ability of animals to clear colonization.

In summary, we have shown that NTHi is capable of colonizing the rhesus macaque upper respiratory tract without causing disease with characteristic very similar to colonization of humans, thereby making rhesus macaques an attractive model of colonization. Using this model, we have established that the HMW1 and HMW2 increase the efficiency and density of NTHi colonization of rhesus macaques and that colonization can result in antibody production against NTHi antigens.

This work has been submitted for publication as reference below:

3. Determinants of HMW1 Function

3.1 Introduction

NTHi is a human restricted organism that is a member of the normal human respiratory flora and an opportunistic pathogen. It is one of the two main causative agents of otitis media in children and a common cause of pneumonia and exacerbation of chronic lung conditions, such as chronic obstructive pulmonary disease (COPD), in adults. NTHi can also cause invasive diseases, including meningitis, endocarditis, pericarditis, pyogenic arthritis, osteomyelitis, epiglottitis, cellulitis, urinary tract infections, intra-abdominal infections, and occult bacteremia (St Geme 1993). The organism must adhere to host cells to establish colonization, persist, and cause disease. For this reason, NTHi encodes a number of adhesive proteins, including the type Vb secreted proteins HMW1 and HMW2, which mediate high level adherence to human respiratory epithelial cells. HMW1 and HMW2 are highly homologous and are encoded in approximately 75% of NTHi isolates. In strain 12, HMW1 has a molecular mass of 125kDa and HMW2 has a molecular mass of 120kDa in their mature forms (St Geme, Falkow, and Barenkamp 1993, Barenkamp and St Geme 1996, Krasan et al. 1999, Laarmann et al. 2002, St Geme 1997b, St Geme et al. 1998). Structural prediction software predicts that the proteins have β-helical structures with N-terminal globular domains.
HMW1 is encoded by the same genetic locus that encodes two accessory proteins: HMW1B and HMW1C. Similarly, HMW2 is encoded in a genetic locus with HMW2B and HMW2C. HMW1B and HMW2B are 99% identical and are functionally interchangeable. HMW1C and HMW2C are 97% identical and also functionally interchangeable. HMW1B forms an outer membrane pore through which HMW1 is transported and to which HMW1 is anchored. HMW1C is a cytoplasmic N-linking glycosyltransferase that modifies HMW1 with mono- and di-hexoses. Elimination of HMW1B results in the inability of HMW1 to reach the bacterial surface and thus to mediate adherence. HMW1 is still transported through HMW1B when HMW1C is eliminated but fails to anchor in the pore and is released into the extracellular space where it cannot mediate bacterial adherence. Elimination of HMW1B or HMW1C also results in lower levels of HMW1, which seems to be efficiently degraded if it is not properly processed (St Geme, Falkow, and Barenkamp 1993, Barenkamp and Leininger 1992, Barenkamp and St Geme 1994, Grass et al. 2003, Grass and St Geme 2000, St Geme and Grass 1998).

HMW1 is modified by HMW1C in the cytoplasm at 31 sites with mono- and di-hexose residues. These residues can be either glucose or galactose, though for the di-hexose modification the first sugar must be glucose. The modifications all take place on asparagines in the conventional N-linked glycosylation consensus sequence N-X-S/T where X is any amino acid except proline. At one site the consensus motif is reversed.
(Gross et al. 2008). The determinants of specificity between HMW1 and HMW1C have yet to be determined, but it is known that HMW1C does not modify any proteins in addition to HMW1 in NTHi (unpublished data).

HMW1 encodes N-terminal signal followed by a TPS domain, which is responsible for its recognition by HMW1B (Grass and St Geme 2000). After being secreted into the periplasm through the Sec system, the TPS domain of HMW1 interacts with the polypeptide associated translocation (POTRA) domains of HMW1B. These domains protrude into the periplasmic space and interact specifically with the N-terminus of HMW1, as described by Grass et al in 2015. The protein is then translocated through HMW1B in an N to C direction (Grass and St Geme 2000). As a result of secretion the TPS domain is cleaved by an as of yet unidentified protease. Two cysteine residues at the C terminal of HMW1 form a disulfide bond, which is essential for anchoring of HMW1 in the HMW1B pore (Grass and St Geme 2000). A homologous system is found in enteropathogenic E. coli. This system is made up of EtpA, EtpB, and EtpC which are homologous to HMW1, HMW1B, and HMW1C, respectively (Roy et al. 2008).

Interaction with HMW1B and HMW1C are clearly important for proper processing and presentation of HMW1. This work describes the mechanisms underlying these interactions and shows that proper modification and secretion are necessary for HMW1 to function.
3.2 Methods

3.2.1 Bacterial Strains and Growth Conditions

Bacterial strains used in this study are listed in Table 3.

Nontypeable *H. influenzae* strain 12 was used for all experiments with NTHi. Strain 12 was grown on chocolate agar or in BHI supplemented with 0.1% v/v lysed horse blood as a source of hemin and 3.5µg/ml NAD. Plates were incubated at 37°C with 5% CO₂ and liquid cultures were grown at 37°C with agitation.

*E. coli* BL21 was used to express the heterologous proteins. In all cases, HMW1 and HMW1B were expressed in the pT7-7 plasmid and the glycosyltransferases were expressed in the pACYC184 plasmid under their native promoters. *E. coli* was grown on LB plates or in LB broth supplemented with 100µg/ml ampicillin and 25µg/ml chloramphenicol as appropriate. Plates were incubated at 37°C with 5% CO₂ and liquid cultures were grown at 37°C with agitation.

Table 3: Bacterial Strains Used in Chapter 3

<table>
<thead>
<tr>
<th>Strain</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>12B</td>
<td>Nontypeable <em>Haemophilus influenzae</em> Strain 12</td>
<td>St. Geme 1993</td>
</tr>
<tr>
<td></td>
<td><em>hmw2::kan</em></td>
<td></td>
</tr>
<tr>
<td>BL21 hmw1ABC</td>
<td><em>E. coli</em></td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>pT7-7 hmw1A/hmw1B</td>
<td></td>
</tr>
<tr>
<td></td>
<td>pACYC hmw1C</td>
<td></td>
</tr>
<tr>
<td>BL21 HMW1AB EtpC</td>
<td><em>E. coli</em></td>
<td>This study</td>
</tr>
</tbody>
</table>
\begin{tabular}{lll}
\textbf{BL21 Hmw1AB Hmw1C} & \textit{E. coli} & \textbf{Rempe et. al 2015} \\
pT7-7 hmw1A/hmw1B & pACYC hmw1C & \\
pACYC hmw1Ck & \\
\end{tabular}

\begin{tabular}{lll}
\textbf{BL21 Hmw1ABC 484} & \textit{E. coli} & \textbf{This study} \\
pT7-7 hmw1A N484Q / hmw1B & pACYC hmw1C & \\
\end{tabular}

\begin{tabular}{lll}
\textbf{BL21 Hmw1ABC 806} & \textit{E. coli} & \textbf{This study} \\
pT7-7 hmw1A N806Q / hmw1B & pACYC hmw1C & \\
\end{tabular}

\begin{tabular}{lll}
\textbf{BL21 Hmw1AC} & \textit{E. coli} & \textbf{This study} \\
pT7-7 hmw1A & pACYC hmw1C & \\
\end{tabular}

\begin{tabular}{lll}
\textbf{BL21 yebF-hmw1AC} & \textit{E. coli} & \textbf{This study} \\
pACES40 yebF-hmw1A fusion & pACYC hmw1C & \\
\end{tabular}

\subsection*{3.2.2 Purification of untagged Hmw1}

For all of the purification procedures, an overnight culture was back diluted 1:50 into 9 L of the media and grown to an OD\textsubscript{600} of 1.2. Cultures were then centrifuged for 10 minutes at 18,598xg, and the supernatant was discarded. The pellets were frozen at -20°C for a minimum of 1 hour.
To release HMW1 from the cell surface, the frozen pellets were thawed and resuspended 1/100th of the starting volume in resuspension buffer (0.5M NaCl, 10mM EDTA, 10mM Tris pH 7.5, 50 µM 1,10 phenanthroline). The resuspended bacteria were incubated on ice for one hour and then centrifuged for 10 minutes at 18,598xg. The supernatant containing the released HMW1 was saved, and the pellet was discarded.

To isolate HMW1 from the cytoplasm, the pellets were resuspended in 1/50th of the starting volume in resuspension buffer (0.5M NaCl, 10mM EDTA, 10mM Tris pH 7.5, 50 µM 1, 10 phenanthroline) and were sonicated until the suspension reached maximum clarity. The sonicate was then centrifuged for 10 minutes at 18,598 xg. The supernatant containing the released HMW1 was saved, and the pellet was discarded.

The saved supernatants were dialyzed into 20mM MES, 10mM NaCl pH 5.5 and were loaded on a RESOURCE-S ionic exchange AKTA column. The bound protein was eluted with 20mM MES, 1M NaCl pH 5.5. Fractions containing HMW1 were combined and concentrated before loading onto a SuperDex 16/60 size exclusion AKTA column equilibrated with 20mM MES, 150mM NaCl, pH 5.5 + 5% glycerol. The fractions containing HMW1 were resolved on SDS-PAGE gels and stained with Coomassie blue to ensure HMW1 purity.

3.2.3 Purification of YebF fused HMW1

DNA sequence encoding the HMW1 protein residues 442-1536 was cloned into the pAES40 vector in frame with the yebF gene. The resulting plasmid expressed a fusion
protein containing YebF fused to mature HMW1. \textit{E. coli} DH5\textalpha \ containing this plasmid along with the pACYC \textit{hmw1C} was grown overnight and then back diluted 1:10 into 500 mL of Power Broth (AthenaES product number 0106) and grown to an OD$_{600}$ of 0.6. The culture was then induced with 50 \textmu M IPTG and grown at 30°C for 20 hours. The culture was then centrifuged for 10 minutes at 18,598xg, and the supernatant was saved and ammonium sulfate added to 70\%. The ammonium sulfate containing supernatant was incubated at 4°C for two hours and was then centrifuged for 30 minutes at 18,598xg. The resulting precipitate was resuspended in 20mM MES, 10mM NaCl pH 5.5 and dialyzed against the same buffer for 18 hours with two buffer changes. The solution was then loaded on a RESOURCE-S ionic exchange AKTA column. The bound protein was eluted with 20mM MES, 1M NaCl pH 5.5. Fractions containing HMW1 were combined and concentrated. These fractions were incubated for 16 hours at 4°C with Enterokinase (1 unit enzyme/50ug protein) to cleave YebF. After incubation, the fractions were loaded onto a SuperDex 16/60 size exclusion AKTA column equilibrated with 20mM MES, 150mM NaCl, pH 5.5 + 5\% glycerol. The fractions containing HMW1 were resolved on SDS-PAGE gels and stained with Coomassie blue to ensure HMW1 purity.

\textbf{3.2.4 Western Blots}

For Western analysis, proteins were transferred from SDS-PAGE gels to nitrocellulose membranes, and membranes were blocked in 5\% non-fat milk/TBS for 30 minutes. Membranes were then incubated for one hour with a guniea pig antiserum
raised against full length HMW1 at a 1:1000 dilution (GP104). Subsequently, membranes were washed with TBS-Tween and incubated for 1 hour in 5% non-fat milk/TBS and a 1:10000 dilution of anti-guinea pig IgG, then developed using Thermo Fischer Super Signal substrate.

3.2.5 Native Gel Electrophoresis

Purified protein was run on a 4-15% gradient gel in Native-PAGE buffer (25mM Tris 192mM Glycine) at 200V for 45 minutes. Gels were then silver stained.

3.2.6 Circular Dichroism

Circular dichroism analysis was performed on 25 µM of purified protein in 10mM MES, 5mM NaCl pH 5.5. The spectrum from 190-250 nM was read and results are the average of three runs. Estimation of protein secondary structure was done with the K2D2 program (http://cbdm-01.zdv.uni-mainz.de/~andrade/k2d2/).

3.3 Results

3.3.1 The Pattern of Glycosylation on HMW1 is Critical for Protein Function

We assayed the ability of differentially modified HMW1 to mediate adherence to cultured human epithelial cells and determined that bacteria expressing HMW1 modified by the HMW1C homologue EtpC were significantly less adherent than bacteria expressing HMW1 modified by HMW1C. Even more dramatic, bacteria expressing HMW1 modified by the HMW1C homologue HMW1C_{Kk} were non-adherent. (Figure 14B) To further explore HMW1 function we examined the ability of the purified proteins
to block HMW1-mediated bacterial adherence to human epithelial cells. Purified HMW1 modified by HMW1C efficiently blocked adherence, while purified HMW1 modified by EtpC resulted in only modest inhibition of HMW1-mediated bacterial adherence (Figure 14C).

To investigate the relationship between the pattern of glycosylation of HMW1 and adhesive activity of the protein, we heterologously expressed HMW1, HMW1B, and either HMW1C, EtpC, or HMW1C_{Kk}, in *E. coli* BL21. We then purified HMW1 and used LC-MS/MS to identify the pattern of glycosylation (Figure 14A). In control experiments, we compared glycosylation of HMW1 purified from NTHi and *E. coli* BL21. As shown in Figure 14A, there was no difference in the glycosylation pattern.

LC-MS/MS analysis of EtpC-modified and HMW1C_{Kk}-modified HMW1 showed that EtpC modified fewer sites than did HMW1C, with some of the differentially modified sites located in the binding domain. HMW1C_{Kk} modified even fewer sites. These results suggest a relationship between the pattern of glycosylation and the ability of the HMW1 protein to facilitate adherence (Figure 14A).
Figure 14: Pattern of glycosylation affects HMW1 function.

Panel A: Identified sites of glycosylation are represented by the vertical numbers in the top row. Each number corresponds to the location of an asparagine residue in the polypeptide chain. The modifying enzyme and source of the protein are in the left-hand column. HMW1 had the same pattern of modification by HMW1C in NTHi as it had when heterologously expressed along with HMW1C in *E. coli*. EtpC and HMW1C\textsubscript{Kk} modified HMW1 in different patterns, though at no additional sites. N/C means no coverage, the peptide with that asparagine was not identified in the mass spectrometry run. Panel B: The ability of *E. coli* expressing HMW1 to adhere to human respiratory epithelial cells varies depending on the glycosylation pattern of HMW1. Panel C: The ability of purified HMW1 to block adherence of HMW1 expressing *E. coli* differs depending on the pattern of glycosylation on the purified protein. HMW1 modified by HMW1\textsubscript{Kk} was not included in this assay, as it is not present on the bacterial cell surface. Panel D: Coomassie stained SDS-PAGE gel showing purified proteins. Lane 1 is HMW1 modified by HMW1C purified from the surface of *E. coli*. Lane 2 is HMW1 modified by EtpC purified from the surface of *E. coli*. 
Following our observance of a correlation between the adhesive ability of HMW1 and the pattern of glycosylation, we explored if there are any individual glycosylation sites that are necessary for protein function. We mutated the modified asparagine residue to glutamine at eight sites, starting from the N-terminus of the mature protein. From these eight sites, we identified two sites for which the NQ mutation ablated HMW1-mediated adherence. The defect was abolished with when the the glutamine was changed back to asparagine. (Figure 15)

**Figure 15:** There are at least two specific sites that must be modified for HMW1 to function.

Panel A: Mutation of HMW1 residues 484 or 806 from asparagine to glutamine results in a decrease in adherence. This phenotype is reversed when the glutamine is changed back to asparagine. Panel B: A Western blot for HMW1 shows that the protein is still expressed in the N806Q mutant but not expressed in the N484Q mutant. Panel C: Purified HMW1 N806Q is still able to block adherence of HMW1 expressing *E. coli* to human respiratory epithelial cells, though not to the same extent as the wild type protein. Panel D: Coomassie blue stained SDS-PAGE gel showing purified HMW1 (lane 1) and HMW1 N806Q (lane 3).
We compared the conformation of purified HMW1 and HMW1 N806Q modified by HMW1C and HMW1 and HMW1 modified by EtpC from the surface of Bl21 E. coli using native-PAGE and circular dichroism. HMW1 modified by HMW1C_KK is not surface located and thus was not included in the comparison. The mutation at asparagine 484 abolished surface localization of HMW1 (Figure 15B) and therefore precluded any further studies on this version of the protein. Differences in protein structure were not apparent by native-PAGE; however, circular dichroism revealed similar structures in the HMW1C-modified HMW1 and HMW1C-modified HMW1 N806Q proteins while the structure of EtpC-modified HMW1 differed significantly (Figure 16). Estimation of secondary structure using the CD spectra revealed HMW1C-modified HMW1 and HMW1C-modified HMW1 N806Q to be 1.86% α-helix and 49.03% β-sheet, while EtpC-modified HMW1 is 52.46% α-helix and 9.17 β-sheet.
Figure 16: Glycosylation pattern impacts protein structure.

Panel A: A Native-PAGE gel. Lane 1 has purified HMW1 modified by HMW1C. Lane 2 shows purified HMW1 modified by EtpC. Lane 3 shows purified HMW1 in which the asparagine at residue 806 has been mutated to glutamine and therefore is no longer modified. The major band for all three protein runs at the same size, indicating similar structure. Panel B: Circular dichroism traces show similarity between the structures of HMW1C-modified HMW1 and HMW1C-modified HMW1 N806Q while the structure of EtpC-modified HMW1 is significantly different.

3.3.2 HMW1 purified from the bacterial cell surface is functional while HMW1 purified from the cytoplasm is non-functional.

The pore diameter of HMW1B is 2.7nm, which is too small to translocate a folded protein (Duret et al. 2008). Therefore, we have assumed that HMW1 remains unfolded until it reaches the cell surface, potentially needing to interact with HMW1B to fold properly. Based on the well-documented correlation between structure and function, we hypothesized that HMW1 needs to be fully folded in order to facilitate adherence. To test this hypothesis, we compared the ability of HMW1 purified from the cytoplasm and HMW1 purified from the cell surface to block the binding of whole bacteria expressing
HMW1 to host cells. Protein purified from the surface of *E. coli* BL21 significantly inhibited the adherence of NTHi to human epithelial cells. In contrast, protein purified from the cytoplasm had no inhibitory effect, suggesting that localization of the protein and potentially secretion and folding are important to function. We compared the structures of the extracellularly and cytoplasmically purified proteins by native-PAGE and circular dichroism. Both assays demonstrated significant differences between the two proteins. On the native-PAGE gel the extracellular protein migrated more slowly than the cytoplasmic protein, indicating that increased structure retarded its migration. This observation was corroborated by circular dichroism, which showed a clear difference between the structure of HMW1 isolated from the bacterial cell surface and HMW1 isolated from the cytoplasm (Figure 17). Estimation of protein secondary structure from the CD spectra showed the extracellular protein to be 1.86% α-helix and 49.03% β-sheet, while the cytoplasmic protein was 38.39% α-helix and 2.25% β-sheet.
Figure 17: HMW1 obtained from the bacterial surface is structurally different than HMW1 obtained from the cytoplasm.

Panel A: Purified HMW1 from the surface of *E. coli* (lane 1) and HMW1 from the cytoplasm of *E. coli* (lane 2). Panel B: HMW1 purified from the surface of *E. coli* is able to block the adherence of HMW1 expressing *E. coli* to human respiratory epithelial cells. HMW1 purified from the cytoplasm is not able to block adherence. Panel C: A Native-PAGE gel. Lane 1 has purified HMW1 from the surface of *E. coli*. Lane 2 shows purified HMW1 from the cytoplasm of *E. coli*. The major band in lanes 1 and 2 migrated differently indicating different amounts of protein structure. Panel D: Circular dichroism traces show significant differences between the structure of HMW1 isolated from the surface of *E. coli* and HMW1 isolated from the cytoplasm.
3.3.3 Secretion alone is not sufficient to produce functional HMW1

To determine if secretion itself is sufficient to promote appropriate HMW1 structure, we fused the mature HMW1 protein (residues 442-1536) to YebF, an extracellular protein secreted by *E. coli*. Using the ACES™ YebF Protein Export Kit from AthenaES, we expressed the YebF-HMW1 fusion protein and purified it from the culture supernatant. YebF was then removed using enterokinase to leave mature HMW1 that had been secreted in a non-HMW1B dependent manner. This protein was unable to block adherence of HMW1 expressing *E. coli* to human respiratory epithelial cells. We also used native-PAGE and circular dichroism to evaluate any structural differences between HMW1B-secreted HMW1 and HMW1 secreted in a non-HMW1B dependent manner. These assays both showed significant differences in structure between the two proteins (Figure 18), suggesting that secretion alone is not sufficient to promote proper HMW1 folding. Estimation of protein secondary structure using the CD spectra showed HMW1B-secreted HMW1 to be 1.86% $\alpha$-helix and 49.03% $\beta$-sheet, while YebF-secreted HMW1 was 37.46% $\alpha$-helix and 3.02% $\beta$-sheet.
Figure 18: Non-HMW1B depended secretion of HMW1 results in non-functional protein.

Panel A: Purified HMW1B-secreted HMW1 (Lane 1) and YebF-secreted HMW1 (lane 2). Panel B: HMW1B-secreted HMW1 is able to block the adherence of HMW1 expressing E. coli to human respiratory epithelial cells. YebF-secreted HMW1 is not able to block adherence. Panel C: A Native-PAGE gel. Lane 1 has purified HMW1 from the surface of E. coli. Lane 2 shows purified HMW1 from the cytoplasm of E. coli. The major band in lanes 1 and 2 migrated differently indicating different amounts of protein structure. Panel D: Circular dichroism traces show significant differences between the structure of HMW1B-secreted HMW1 and YebF-secreted HMW1.
3.4 Discussion

Efficiency in protein production is necessary as it decreases wasted energy. Therefore, it is not surprising that HMW1 must interact specifically with HMW1B and HMW1C in order to be functional, as any unnecessary or inefficient interactions would most likely have been lost during evolution. It is also widely recognized that protein structure and function are intrinsically linked, making it logical that the specific interactions with HMW1B and HMW1C produce functional proteins by influencing HMW1 folding.

N-linked glycosylation, especially in eukaryotic cells, is responsible for proper protein function, and there are a number of congenital disorders due to absent or aberrant N-linked glycosylation of proteins. These include phosphomannoulose 2 deficiency, which results in hypoglycosylation of a number of proteins and causes nervous system problems, dysmorphology, hepatomegaly, skeletal abnormalities, and hypogonadism (Lefeber et al. 2009). Glycosylation abnormalities can also be caused by phosphomannose-isomerase deficiency in which patients cannot make GDP-mannose, a common starting glycan, and therefore hypoglycosylate many proteins. This disease results in hepatic-intestinal disorders (Romano et al. 2009). Aberrant glycosylation is also a cause of transmissible spongiform encephalopathies. Differential glycosylation of a single site can allow prions to form (Atkinson 2004).
Glycosylation deficiencies have not been well studied in bacteria, but glycoforms of proteins with different functions have been described. The *E. coli* TibA autotransporter is O-glycosylated and is responsible for bacterial autoaggregation and adherence to host cells. Both glycosylated and non-glycosylated forms of the protein are found naturally. Interestingly, glycosylation seems to be a way for the organism to regulate TibA function, as non-glycosylated TibA is locked into a conformation that is not suitable for autoaggregation or adherence (Cote, Charbonneau, and Mourez 2013).

In considering the above examples, the observation that differentially glycosylated HMW1 fails to function properly is not surprising. This is the first published work to look at differing patterns of N-linked glycosylation of a bacterial protein and their effect on protein function. The structural difference between HMW1C and EtpC modified HMW1 indicate that the pattern of glycosylation is involved in protein folding. The lack of HMW1 on the bacterial cell surface when modified by HMW1C\textsubscript{K8} suggests that the pattern of glycosylation also effects protein stability and/or secretion. The lack of structural differences between HMW1 and HMW1 N806Q suggests that modification at that specific residue is important in receptor recognition rather than structure. As more bacterial proteins are found to be glycosylated, the potential for these results to be applied to a number a range of proteins is intriguing as it could help elucidate the roles of glycans at specific locations.
Multiple lines of research support the hypothesis that TpsA proteins, like HMW1, remain unfolded until they reach the surface of the cell. The size of the TpsB pore is too small to allow a folded TpsA protein to transit through (Duret et al. 2008). The TPS domain responsible for interacting with the TpsB POTRA domains is recognized best when it is in an unfolded conformation, suggesting that the protein is unfolded in the periplasm prior to transport through TpsB (Hodak et al. 2006, Grass, Rempe, and St Geme 2015). Finally, when TpsB is absent, TpsA proteins are rapidly degraded, consistent with an unfolded protein response (Hodak et al. 2006). This is the first work, however, to investigate whether secretion in general or secretion through a TpsB pore is necessary for proper TpsA folding. The difference in structure between extracellular and cytoplasmic HMW1 argues against the ability of the protein to properly fold spontaneously, suggesting that secretion is indeed an important part of the folding process. Furthermore, secretion by itself is not sufficient for proper HMW1 structure, as HMW1 secreted in a non-HMW1B dependent manner did not obtain the same structure as HMW1 secreted through HMW1B. We have not yet been able to ascertain whether HMW1 must be secreted through a TpsB pore in general or HMW1B specifically in order to function, but experiments are on-going to address this.
4. Unconventional N-linked glycosylation promotes trimeric autotransporter function in *Kingella kingae* and *Aggregatibacter aphrophilus*.

4.1 Introduction

Glycosylation provides a means by which a cell can further specify a particular function for a protein. Once thought to be unique to eukaryotic organisms, glycosylation of proteins is now recognized to be common in prokaryotes (Szymanski and Wren 2005). Modification of bacterial proteins promotes proper folding, adhesive activity, solubility, antigenic variation, and protection against proteases (Grass et al. 2003, Marceau and Nassif 1999, Doig et al. 1996, Herrmann et al. 1996, Cote, Charbonneau, and Mourez 2013). There are two major forms of protein glycosylation: O-linked, in which the modification is attached to the side chain oxygen of a serine or threonine residue, and N-linked, in which the modification is attached to the amide nitrogen of an asparagine residue. Rarely, proteins can be glycosylated on other residues beyond serine, threonine, or asparagine (Hofsteenge et al. 1994).

O-linked glycosylation in bacteria generally takes place in the cytoplasm and involves a series of glycosyltransferases that sequentially build polysaccharide chains on the acceptor protein. These glycosyltransferases are not membrane bound and are commonly able to transfer only a single specific activated saccharide molecule (Iwashkiw et al. 2013). In a limited number of cases, O-glycosylation has been shown to
occur in the periplasm. Specifically, the O-linked modification of pili in *Pseudomonas aeruginosa* and *Neisseria meningitidis* involves an inner membrane bound oligosaccharyltransferase (Faridmoayer et al. 2007). To date there is no recognized consensus sequence for modification by O-linked glycosylation, though existing evidence suggests that there may a structural element in the acceptor peptide that directs glycosylation (Charbonneau et al. 2012).

The best studied mechanism of bacterial protein N-linked glycosylation is the *pgl* system used by the gram-negative pathogen *Campylobacter jejuni*. In this system, undecaprenyl linked sugars are used to assemble branching heptasaccharide chains attached to a lipid anchor on the cytoplasmic face of the bacterial inner membrane. Once assembled, the chains are flipped across the membrane into the periplasm and are transferred to proteins by an inner membrane bound oligosaccharyltransferase. Glycans are attached to the substrate protein at the consensus site Asn-Xaa-Ser/Thr, where Xaa can be any amino acid except proline (Szymanski et al. 1999). The *pgl* system mimics the N-linked glycosylation system found in eukaryotic cells in which branching sugar chains are built on the cytoplasmic face of the endoplasmic reticulum and flipped into the lumen and then transferred to substrate proteins by a membrane bound oligosaccharyltransferase at the same Asn-Xaa-Ser/Thr consensus site (Szymanski et al. 1999).
Nontypeable *Haemophilus influenzae* (NTHi) uses a glycosylation system with elements of both the N-linked and O-linked processes. In this system, the N-linking glycosyltransferase HMW1C transfers glucose and galactose to the acceptor protein HMW1 using the activated UDP forms of the sugars (Grass et al. 2010). This transfer occurs in the cytoplasm, and sugar chains are built on HMW1, similar to O-linked glycosylation. However, the sugars are added at the Asn-Xaa-Ser/Thr N-linked glycosylation consensus site (Gross et al. 2008). HMW1 is modified at at least 31 sites with either mono- or di-hexoses, and HMW1C is responsible for this modification (Grass et al. 2003, Grass et al. 2010, Gross et al. 2008).

HMW1 is an adhesive protein that mediates interaction between bacterial cells and human epithelial cells and is encoded by a two-partner secretion system gene cluster that also encodes HMW1B (an outer membrane pore-forming protein that facilitates the translocation of HMW1 to the bacterial cell surface) and HMW1C. (St Geme, Falkow, and Barenkamp 1993, St Geme and Grass 1998). Without glycosylation, HMW1 is released from the bacterial cell surface and is unable to facilitate adherence (Grass et al. 2003). A two-partner secretion system genetic locus with similar organization is found in enterotoxigenic *Escherichia coli* (ETEC). This locus contains the *etpB*, *etpA*, and *etpC* genes, with the HMW1C homologue EtpC responsible for modifying the adhesive protein EtpA in a manner similar to HMW1C modification of HMW1 (Fleckenstein et al. 2006).
Due to the unique nature of HMW1C-mediated glycosylation, we wondered if this system is used by other gram-negative pathogenic species beyond NTHi and ETEC. In this study, we identified a number of HMW1C homologues in other species using bioinformatic approaches. Interestingly, these HMW1C homologues were not always encoded by a genetic locus that also encodes a potential acceptor protein as in NTHi and ETEC. To look specifically at HMW1C homologues that are not encoded near a gene for a potential acceptor protein, we examined *Kingella kingae* and *Aggregatibacter aphrophilus*. Our results demonstrate that the HMW1C homologues in these species modify adhesive proteins with mono- and di-hexoses and establish that modification is needed for normal function of the acceptor protein.

### 4.2 Methods

#### 4.2.1 Strains and culture conditions

The bacterial strains are listed in Table 4. *K. kingae* strains were cultured at 37°C with 5% CO₂ on chocolate agar supplemented with 50 µg/mL kanamycin, 1 µg/mL erythromycin, or 0.75 µg/mL chloramphenicol, as needed. *A. aphrophilus* strains were cultured at 37°C with agitation (225 rpm) in tryptic soy broth (TSB) supplemented with NAD (3.5 µg/mL) and lysed horse blood (1:1000 dilution), using 50 µg/mL kanamycin for selection as necessary.
Table 4: Bacterial Strains used in Chapter 4

<table>
<thead>
<tr>
<th>Strain</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>DH5α</td>
<td><em>E. coli</em> F- ϕ80lacZΔM15 Δ(lacZYA-argF)U169 deoR recA1 endA1 hsdR17(ρ− mK+) phoA supE441 thi-1 gyrA96 relA</td>
<td></td>
</tr>
<tr>
<td>269-492 (KK03)</td>
<td>Clinical isolate, spontaneous spreading, corroding colony variant of <em>K. kingae</em> isolate 269-492</td>
<td>Kehl-Fie 2007</td>
</tr>
<tr>
<td>KK03 hmw1Ckk::kan</td>
<td>Interruption of <em>hmw1C</em>&lt;sub&gt;kk&lt;/sub&gt; with kanamycin resistance cassette</td>
<td>This study</td>
</tr>
<tr>
<td>KK03 hmw1Ckk::kan (hmw1Ckk)</td>
<td>Reversion of interrupted <em>hmw1C</em>&lt;sub&gt;kk&lt;/sub&gt; to wild type gene</td>
<td>This study</td>
</tr>
<tr>
<td>KK03 crtA::erm</td>
<td>Interruption of <em>crtA</em> with erythromycin resistance cassette</td>
<td>This study</td>
</tr>
<tr>
<td>KK03 pilA1::cm</td>
<td>Interruption of <em>pilA1</em> with chloramphenicol resistance cassette</td>
<td>This study</td>
</tr>
<tr>
<td>KK03 crtA::erm pilA1::cm</td>
<td>Interruption of <em>crtA</em> with erythromycin resistance cassette and <em>pilA1</em> with chloramphenicol resistance cassette</td>
<td>This study</td>
</tr>
<tr>
<td>KK03 crtA::erm pilA1::cm hmw1Ckk::kan</td>
<td>Interruption of <em>crtA</em> with erythromycin resistance cassette, <em>pilA1</em> with chloramphenicol</td>
<td>This study</td>
</tr>
<tr>
<td><strong>KK03</strong> ctrA::erm pilA1::cm knh::kan</td>
<td>Interruption of ctrA with erythromycin resistance cassette, pilA1 with chloramphenicol resistance cassette, and knh with kanamycin resistance cassette</td>
<td>Porsch, <em>et al.</em>, 2012</td>
</tr>
<tr>
<td><strong>KK03</strong> ctrA::erm pilA1::cm hmw1Cka::kan (hmw1Cka)</td>
<td>Interruption of ctrA with erythromycin resistance cassette and pilA1 with chloramphenicol resistance cassette, and reversion of interrupted hmw1Cka to wild type gene</td>
<td>This study</td>
</tr>
<tr>
<td><strong>CCUG 11575</strong></td>
<td>Clinical isolate of <em>Aggregatibacter aphrophilus</em></td>
<td>Culture Collection of the University of Gothenburg</td>
</tr>
<tr>
<td><strong>CCUG 11575</strong> hmw1Cka::kan</td>
<td>Interruption of hmw1Cka with kanamycin resistance cassette</td>
<td>This study</td>
</tr>
<tr>
<td><strong>CCUG 11575</strong> hmw1Cka::kan (hmw1Cka)</td>
<td>Reversion of interrupted hmw1Cka to wild type gene</td>
<td>This study</td>
</tr>
<tr>
<td><strong>CCUG 11575</strong> emaA::kan</td>
<td>Interruption of emaA with kanamycin resistance cassette</td>
<td>This study</td>
</tr>
<tr>
<td><strong>CCUG 11575</strong> emaA::kan (emaA)</td>
<td>Reversion of the interrupted emaA to wild type gene</td>
<td>This study</td>
</tr>
<tr>
<td><em>E. coli</em> BL21 pACYC HMW1Cka + pBAD</td>
<td><em>E. coli</em> expressing HMW1Cka</td>
<td>This study</td>
</tr>
</tbody>
</table>
4.2.2 Strain Construction

Disruption of hmw1CKk (the *K. kingae* hmw1C homologue) was achieved via random transposon mutagenesis and identified by nucleotide sequencing (Kehl-Fie and St Geme 2007). All other gene disruptions in *K. kingae* and *A. aphrophilus* were generated by insertion of an antibiotic resistance cassette into the gene of interest by homologous recombination. To do this plasmid-based gene disruption constructs were made in *E. coli*, linearized, and introduced into the appropriate species via natural transformation. Transformants were recovered after plating on media containing the appropriate antibiotic. Correct localization of the interruption was confirmed by PCR and nucleotide sequencing.

Reversions of gene disruptions were generated by PCR amplifying a wild type copy of the gene and using the wild type gene to replace the disrupted copy via natural transformation. Transformants were screened for loss of resistance to the appropriate antibiotic, and sequence integrity was confirmed by nucleotide sequencing.
4.2.3 Homologue identification:

Homologues of HMW1C were identified by using the NCBI BLAST program (http://blast.ncbi.nlm.nih.gov/Blast.cgi) with the nontypeable Haemophilus influenzae HMW1C protein (GenBank ID: ADO96126.1) as the query sequence. Results were organized by percent similarity and were assessed for relevance of the host organism to human health. Identity and similarity were determined using the EMBOSS Needle program (http://www.ebi.ac.uk/Tools/psa/emboss_needle/).

4.2.4 Adherence assays

Quantitative adherence assays were performed using Chang human epithelial cells (ATCC CCL 20.2) (Kehl-Fie and St Geme 2007). Cells were seeded into 24-well tissue culture plates and allowed to form a monolayer overnight. The following day the spent media was removed and replaced with 300 µl of fresh tissue culture media. Subsequently, approximately 6.5 x10⁶ CFU of bacteria were inoculated onto the monolayer. Bacteria were allowed to adhere to the cells for 30 minutes. Monolayers were then rinsed four times with PBS to remove non-adherent bacteria. Adherent bacteria were released from the monolayer using trypsin-EDTA and were quantitated by plating dilutions on chocolate agar plates. CFU counts of adherent bacteria were compared to CFU counts of the original inoculum to determine the percentage of bacteria that were adherent.
Qualitative adherence assays were performed as quantitative adherence assays with the following modifications: human epithelial cells were seeded onto glass cover slips in a 24-well tissue culture plate; cover slips were removed and stained with Giemsa stain; stained coverslips were observed at 400x with a light microscope.

For adherence assays with \textit{K. kingae}, human epithelial cells were fixed with 2\% glutaraldehyde in 0.1 M sodium phosphate buffer prior to the assay to prevent disruption of the monolayer by the \textit{K. kingae} RTX toxin.

### 4.2.5 Autoaggregation assays

For autoaggregation assays with \textit{K. kingae}, bacteria were grown overnight on chocolate agar at 37\(^\circ\)C with 5\% CO\(_2\). The next day, the bacteria were resuspended in tubes containing brain heart infusion broth to an OD\(_{600}\)=1.0. Tubes were allowed to stand at room temperature, and OD\(_{600}\) was measured after 30, 60, 90, 120, and 180 minutes.

For autoaggregation assays with \textit{A. aphrophilus}, bacteria were grown overnight with agitation in TSB supplemented with NAD and heme. The following day, cultures were removed from agitation, were allowed to stand at room temperature, and the OD\(_{600}\) was measured hourly for three hours.

### 4.2.6 Statistical analysis

Statistical analyses were done using the GraphPad Prism 6 software. Values were compared using the unpaired t-test. P values of less than 0.05 were considered to be statistically significant.
4.2.7 Isolation of outer membrane proteins

Bacteria were resuspended in 10 mM HEPES pH 7.5 and sonicated until the suspension cleared. Suspensions were then centrifuged for 2 minutes at 21,000 x g, 4°C to remove any intact cells. The supernatant was centrifuged again in a Beckman Coulter Optima MAX-TL Ultracentrifuge using a TLA-5 rotor for 1 hour at 100,000 x g, 4°C to pellet membranes. The membrane pellet was treated with 1% sarkosyl to solubilize the inner membrane fraction, and the solution was centrifuged again for 1 hour at 100,000 x g, 4°C to pellet the outer membrane fraction.

4.2.8 Formic acid treatment of outer membrane proteins

Outer membrane fractions were incubated with 75% formic acid overnight, at 25°C in the dark. Samples were then lyophilized three times to remove the formic acid and the dried protein resuspended in 1.5M Tris pH 8.8.

4.2.9 SDS PAGE

Samples were boiled with Laemmli SDS-PAGE loading buffer for five minutes and were then loaded and resolved on 7.5% SDS-PAGE gels.

4.2.10 In-gel digestion

Coomassie stained samples were excised from gels, cut into 1 mm cubes destained with 50% methanol/1.25% acetic acid, reduced with 5 mM Dithiothreitol (Thermo), and alkylated with 20 mM iodoacetamide (Sigma) (Shevchenko et al. 1996). Gel pieces were then washed with 20 mM ammonium bicarbonate (Sigma) and
dehydrated with acetonitrile (Fisher). Trypsin (Promega) (5ng/mL in 20 mM ammonium bicarbonate) was added to the gel pieces, and proteolysis was allowed to proceed overnight at 37°C. Peptides were extracted with 0.3% trifluoroacetic acid (J.T.Baker), and then 50% acetonitrile. Extracts were combined, and the volume was reduced by vacuum centrifugation.

4.2.11 Mass spectrometry analysis

Tryptic digests were analyzed by LC-MS/MS on a hybrid LTQ Orbitrap Elite mass spectrometer (Thermo Fisher Scientific San Jose, CA) coupled with a nanoLC Ultra (Eksigent). Peptides were separated by reverse phase (RP)-HPLC on a nanocapillary column, 75 μm ID x 15 cm Reprosil-pur 3 um, 120A (Dr. Maisch, Germany) in a nanoflex chip system (Eksigent). Mobile phase A consisted of 1% methanol (Fisher)/0.1% formic acid (Thermo) and mobile phase B consisted of 1% methanol/0.1% formic acid/80% acetonitrile. Peptides were eluted into the mass spectrometer at 300 nL/min, with each RP-LC run comprising a 90 minute gradient from 10 to 25% B in 65 min, 25-40% B in 25 min. The mass spectrometer was set to repetitively scan m/z from 300 to 1800 (R = 240,000 for LTQ-Orbitrap Elite) followed by data-dependent MS/MS scans on the twenty most abundant ions, with a minimum signal of 1500, dynamic exclusion with a repeat count of 1, repeat duration of 30s, exclusion size of 500 and duration of 60s, isolation width of 2.0, normalized collision energy of 33, and waveform injection and dynamic exclusion enabled. FTMS full scan AGC target value was 1e6, and MSn AGC
was 1e4. FTMS full scan maximum fill time was 500 ms, and ion trap MSn fill time was 50 ms; microscans were set at one. FT preview mode, charge state screening, and monoisotopic precursor selection were all enabled with rejection of unassigned and 1+ charge states.

4.2.12 Database searching

The tandem mass spectra were extracted using ProteoWizard (v3.0.5047). Charge state deconvolution and deisotoping were not performed. All MS/MS samples were analyzed using SEQUEST (Thermo Fisher Scientific; version 1.0). SEQUEST was set up to search the UniProt *Kingella* sp. complete proteome database (20140306, 10238 entries) or *Aggregatibacter aphrophilus* complete proteome database (20140701, 2407 entries) appended with common contaminants, assuming a full tryptic digestion with the possibility of two missed cleavages. SEQUEST was searched with a fragment ion mass tolerance of 1 Da and a parent ion tolerance of 15 ppm. S-carbamamidomethyl of cysteine was specified in SEQUEST as a fixed modification. Oxidation of methionine and hexose on asparagine were specified in SEQUEST as variable modifications.

4.2.13 Immunogold labeling and Transmission electron microscopy

Bacteria were grown overnight on chocolate agar 37°C and resuspended in PBS to on OD$_{600}$ of 0.8. Samples were then fixed using 1% paraformaldehyde at room temperature for 30 minutes. Following fixing samples were washed twice with TBS. Samples were blocked for 30 minutes using 2% naive guinea pig serum and 0.1% BSA in
PBS. Following blocking, samples were incubated with guinea pig serum raised against Knh (1:250) for 1 hour. Samples were then washed with PBS and incubated with goat pAb raised against guinea pig IgG and conjugated to 10nm gold particles for 1 hour. Following washing ton PBS, bacteria were resuspended in 0.2M ammonium acetate. Samples were negatively stained with uranyl acetate, and transmission electron microscopy was performed using a FEI-Technai 12 microscope.

4.3 Results

4.3.1 Kingella kingae and Aggregatibacter aphrophilus have HMW1C homologues

A NCBI standard protein BLAST search using the HMW1C polypeptide sequence as the query sequence and default parameters revealed 21 bacterial species encoding a HMW1C homologue with identity over 50%. From these results, we chose to investigate the homologues in *K. kingae* and *A. aphrophilus*, in particular because of their high homology to HMW1C (Fig. 19), the fact that they are encoded in the genome away from any gene encoding an obvious acceptor protein, and the relevance of *K. kingae* and *A. aphrophilus* to human health. The *K. kingae* homologue is encoded by a gene located between an upstream gene encoding a putative protein-disulfide reductase and a downstream gene encoding a putative 2-C-methyl-D-erythritol 4-phosphate cytidylytransferase, gene products that are predicted cytoplasmic proteins and are not suspected to be targets of glycosylation. Similarly, the *A. aphrophilus* homologue is encoded between genes encoding predicted cytoplasmic proteins, namely an upstream
gene encoding a putative ribosomal protein S12 methyltransferase and a downstream gene encoding a putative low-molecular weight protein-tyrosine-phosphatase downstream. The fact that the flanking genes do not encode proteins that are targets of glycosylation or are involved in glycosylation suggests that any phenotypes observed when the hmw1C homologue genes are interrupted are not due to polar effects. The K. kingae and A. aphrophilus HMW1C homologues will henceforth be referred to as HMW1C_{kk} encoded by hmw1C_{kk} and HMW1C_{aa} encoded by hmw1C_{aa}. 
Figure 19: HMW1C\textsubscript{Kk} and HMW1C\textsubscript{Aa} are highly homologous to HMW1C

The amino acid sequences of the \textit{K. kingae} strain 269-492 and \textit{A. aphrophilus} strain NJ8700 HMW1C homologues are aligned to the sequence from nontypeable \textit{H. influenzae} strain R2846. The plus signs (+) below the sequences denote conserved residues. The three conserved residues involved in binding UDP are highlighted in gray.

Previous work identified residues in HMW1C that are needed to bind UDP-hexose, including lysine-467, asparagine-547, and aspartic acid-551. Aligning HMW1C to HMW1C\textsubscript{Kk} and HMW1C\textsubscript{Aa} revealed that these residues are conserved, suggesting that the homologues may be capable of binding and transferring UDP-hexose (Fig. 19).
4.3.2 HMW1C<sub>KK</sub> is required for <i>K. kingae</i> adherence to human epithelial cells and autoaggregation

<i>K. kingae</i> derivatives with an interruption of <i>hmw1C<sub>KK</sub></i> or reversion of the interrupted <i>hmw1C<sub>KK</sub></i> gene were compared with the wild type strain for their ability to adhere to cultured human epithelial cells. The interruption of <i>hmw1C<sub>KK</sub></i> resulted in a marked reduction in adherence (Fig. 20). Wild type levels of adherence were restored when the interrupted <i>hmw1C<sub>KK</sub></i> was reverted to the wild type sequence (Fig. 20A).

A number of bacterial species are capable of autoaggregation (adherence of bacteria to themselves). To assess whether <i>K. kingae</i> autoaggregates, we performed tube settling assays, measuring OD<sub>600</sub> over time. We found that <i>K. kingae</i> is capable of autoaggregation when the polysaccharide capsule (ctrA) and type IV pili (pilA1) are lacking, suggesting that these two surface factors mask one or more other surface factors responsible for autoaggregation. Interruption of <i>hmw1C<sub>KK</sub></i> in a capsule-deficient, type IV pili-deficient mutant abolished autoaggregation, and reversion of the interrupted <i>hmw1C<sub>KK</sub></i> gene reversed this phenotype. Interestingly, elimination of the trimeric autotransporter Knh by interruption of the knh gene abolished the ability of <i>K. kingae</i> to autoaggregate, suggesting that Knh may be a substrate of HMW1C<sub>KK</sub>. These results are illustrated in figure 20B, where it can be seen that the wild-type, pilA1, ctrA, pilA1 ctrA knh, and pilA1 ctrA <i>hmw1C<sub>KK</sub></i> curves overlap and the pilA1 ctrA and pilA1 ctrA <i>hmw1C<sub>KK</sub></i>(<i>hmw1C<sub>KK</sub></i>) curves are indistinguishable.
Figure 20: Expression of HMW1C_kk is required for adherence and autoaggregation

Panel A shows mean adherence of *K. kingae* with the wild type *hmw1C_kk* gene, an interrupted *hmw1C_kk* gene, and the interrupted *hmw1C_kk* gene reverted to wild type. Panel B shows mean autoaggregation of *K. kingae*. The *pilA1* mutant does not express in type IV pili, the *ctrA* mutant lacks polysaccharide capsule, the *knh* mutant lacks the trimeric autotransporter protein Knh, and the *hmw1C_kk* mutant lacks HMW1C_kk.

Panel C shows a Western blot of the formic acid treated outer membrane fraction of wild type *K. kingae* (lane 1), *K. kingae* with an interrupted *hmw1C_kk* gene (lane 2), and the interrupted *hmw1C_kk* gene reverted to wild type (lane 3) using an antiserum raised against the Knh protein. Panel D shows a qualitative adherence assay using *E. coli* strain BL21 expressing HMW1C_kk (frame 1), Knh (frame 2), or HMW1C_kk and Knh (frame 3). Each graph shows data from biological replicates performed in triplicate. Error bars represent standard error. An asterisk denotes a p value of < 0.05 obtained from a paired t-test.
4.3.3 **HMW1C\textsubscript{Kk} glycosylates the trimeric autotransporter Knh**

In previous work on *K. kingae*, we established that type IV pili and the trimeric autotransporter Knh are responsible for mediating bacterial adherence to human epithelial cells (Kehl-Fie and St Geme 2007, Porsch, Kehl-Fie, and St Geme 2012). In an effort to determine the mechanism by which interruption of *hmw1C\textsubscript{Kk}* interferes with *K. kingae* adherence, we examined the PilA1 major subunit of type IV pili and the Knh protein in the wild type, mutant, and reverted strains by Western analysis. There was no difference in mobility of the PilA1 protein (data not shown). However, Western analysis using Knh antiserum demonstrated that Knh migrated at a lower molecular mass and was less abundant in the *hmw1C\textsubscript{Kk}* mutant compared to the wild type strain (Fig. 20C). Both species of Knh were excised from a Coomassie blue-stained gel of outer membrane fractions from *K. kingae* and *K. kingae* lacking HMW1C\textsubscript{Kk} and were analyzed by LC-MS/MS on a hybrid LTQ Orbitrap Elite mass spectrometer.

We obtained 57% coverage of the Knh protein from the wild type strain. With this coverage, we identified 32 sites of glycosylation (Table 5). All 32 of these sites were modified with at least a single hexose, and four of the 32 sites were modified with either a single hexose or a di-hexose (Table 6). Potential sites of glycosylation were sorted by comparing the expected and observed m/z ratios. Sites for which the expected and observed m/z ratios were within 1ppm of each other were considered true sites of modification. All hexose modifications were on asparagine residues, though only 81% of
these sites were in the recognized Asn-Xaa-Ser/Thr consensus sequence for N-linked glycosylation. In contrast, there was no evidence of glycosylation of the Knh protein from the \textit{hmw1}C\textsubscript{k} mutant (for which we obtained 50\% coverage).

Table 5: Glycopeptides identified with di-hexose modification in Knh

<table>
<thead>
<tr>
<th>Peptide sequence(^a)</th>
<th># mono hexose</th>
<th>Observed m/z</th>
<th>(z)</th>
<th>Calculate d m/z</th>
<th>Delta Mass (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>STDGP(\text{NT} \text{TTTDADINK})</td>
<td>1</td>
<td>685.6501</td>
<td>3</td>
<td>685.6502</td>
<td>-0.16</td>
</tr>
<tr>
<td>STDGP(\text{NT} \text{TTTDADINK})</td>
<td>2</td>
<td>739.6662</td>
<td>3</td>
<td>739.6678</td>
<td>-2.17</td>
</tr>
<tr>
<td>DGINAG(\text{NATISNVK})</td>
<td>2</td>
<td>849.4078</td>
<td>2</td>
<td>849.4074</td>
<td>0.48</td>
</tr>
<tr>
<td>DTLP(\text{N} \text{GDP} \text{NSVGNPVTK})</td>
<td>1</td>
<td>663.3198</td>
<td>3</td>
<td>663.32</td>
<td>-0.28</td>
</tr>
<tr>
<td>DGSGP(\text{VNAS} \text{NTSQISAIK})</td>
<td>2</td>
<td>1128.5387</td>
<td>2</td>
<td>1128.5399</td>
<td>-1.07</td>
</tr>
<tr>
<td>GV(\text{D}\text{NVTLLTNGTTT} \text{TVNK})</td>
<td>1</td>
<td>1056.0389</td>
<td>2</td>
<td>1056.0393</td>
<td>-0.38</td>
</tr>
<tr>
<td>VNVTT(\text{TPTNTPDANN} \text{NTINEAGK})</td>
<td>2</td>
<td>899.0973</td>
<td>3</td>
<td>899.0979</td>
<td>-0.71</td>
</tr>
<tr>
<td>DN(\text{VTFKNV} \text{NVTM} \text{TVGSKPDNAVNFK})</td>
<td>2</td>
<td>1051.4902</td>
<td>3</td>
<td>1051.4907</td>
<td>-0.46</td>
</tr>
<tr>
<td>KV(\text{NVT} \text{TTPNTPD} \text{ANN} \text{NTINEAGK})</td>
<td>2</td>
<td>941.795</td>
<td>3</td>
<td>941.7962</td>
<td>-1.34</td>
</tr>
<tr>
<td>VV(\text{APTN} \text{T} \text{VAGNT} \text{FLTAK})</td>
<td>2</td>
<td>705.692</td>
<td>3</td>
<td>705.695</td>
<td>-4.18</td>
</tr>
<tr>
<td>(\text{NIN} \text{ITQTG} \text{STITVATK})</td>
<td>2</td>
<td>993.5063</td>
<td>2</td>
<td>993.5099</td>
<td>-3.62</td>
</tr>
<tr>
<td>SN(\text{VSN} \text{Y} \text{VA} \text{V} \text{D} \text{N} \text{K} \text{TT} \text{QIV} \text{DEKGNPLV})</td>
<td>2</td>
<td>1052.8657</td>
<td>3</td>
<td>1052.8647</td>
<td>1.01</td>
</tr>
<tr>
<td>TN(\text{AP} \text{TALN} \text{ITSADGK})</td>
<td>2</td>
<td>949.9558</td>
<td>2</td>
<td>949.9575</td>
<td>-1.82</td>
</tr>
<tr>
<td>GT(\text{LENT} \text{SV} \text{A} \text{NAD} \text{GK} \text{N} \text{G} \text{AATGI} \text{GTEVK})</td>
<td>1</td>
<td>947.1305</td>
<td>3</td>
<td>947.1316</td>
<td>-1.2</td>
</tr>
<tr>
<td>TT(\text{TLDV} \text{STEPMT} \text{ANN} \text{NTPAGK})</td>
<td>2</td>
<td>829.7133</td>
<td>3</td>
<td>829.7145</td>
<td>-1.47</td>
</tr>
<tr>
<td>AQ(\text{GEN} \text{ASVVN} \text{PGTV} \text{DMK})</td>
<td>2</td>
<td>1049.4771</td>
<td>2</td>
<td>1049.4783</td>
<td>-1.15</td>
</tr>
<tr>
<td>IN(\text{VSTNTT} \text{GGAN} \text{Y} \text{D} \text{SV} \text{N} \text{TGK})</td>
<td>2</td>
<td>865.0835</td>
<td>3</td>
<td>865.0836</td>
<td>-0.14</td>
</tr>
<tr>
<td>G(\text{DLTV} \text{AGNTTVK})</td>
<td>1</td>
<td>669.3434</td>
<td>3</td>
<td>669.3434</td>
<td>-0.47</td>
</tr>
<tr>
<td>N(\text{FTVN} \text{NPNST} \text{V} \text{NM} \text{GGNK})</td>
<td>1</td>
<td>936.4244</td>
<td>2</td>
<td>936.4256</td>
<td>-1.34</td>
</tr>
</tbody>
</table>

\(\text{aUnderlined residues represent hexosylated sites}\)
### Table 6: Glycopeptides identified with di-hexose modification in Knh

<table>
<thead>
<tr>
<th>Peptide sequence</th>
<th># di hexose</th>
<th>Observed m/z</th>
<th>z</th>
<th>Calculated m/z</th>
<th>Delta mass (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AQGENASVNPAGTVDMK</td>
<td>1</td>
<td>705.3198</td>
<td>3</td>
<td>705.3196</td>
<td>0.36</td>
</tr>
<tr>
<td>GDLTVAGNTTVK</td>
<td>1</td>
<td>750.3694</td>
<td>2</td>
<td>750.3698</td>
<td>-0.53</td>
</tr>
<tr>
<td>FIAGDNLNTQTGSNFTYSLNK</td>
<td>1</td>
<td>914.7706</td>
<td>3</td>
<td>914.7713</td>
<td>-0.69</td>
</tr>
<tr>
<td>SLTINSTGAIIDVK</td>
<td>1</td>
<td>872.4447</td>
<td>2</td>
<td>872.4409</td>
<td>4.32</td>
</tr>
<tr>
<td>SNVSYNVAVDNK</td>
<td>2</td>
<td>979.4284</td>
<td>2</td>
<td>979.4284</td>
<td>0</td>
</tr>
<tr>
<td>NFTVNPNSTVNMGGNK</td>
<td>1</td>
<td>1017.452</td>
<td>2</td>
<td>1017.452</td>
<td>-0.04</td>
</tr>
</tbody>
</table>

*aUnderlined residues represent hexosylated sites*

For samples from the wild type and *hmw1C*Kk mutant strains, control database searching for O-linked glycosylation on serine or threonine residues showed nothing above the false discovery rate.

### 4.3.4 Co-expression of Knh and HMW1C*Kk* in *E. coli* is sufficient to facilitate bacterial adherence to epithelial cells

To further demonstrate the dependence of Knh function on glycosylation, we expressed both Knh and HMW1C*Kk* in non-adherent BL21 *E. coli*. When Knh or HMW1C*Kk* was expressed alone, the bacteria remained non-adherent (Fig 2D). However, when Knh and HMW1C*Kk* were co-expressed, the bacteria were capable adhering to human epithelial cells, establishing that Knh is capable of facilitating adherence only when glycosylated.
4.3.5 Mutating \textit{hmw1C}_{Kk} decreases the amount of Knh on the bacterial cell surface

In an effort to assess whether HMW1C_{Kk} influences surface structures on \textit{K. kingae}, we immunogold labeled Knh on the wild type and \textit{hmw1C}_{Kk} mutant strains and examined them by transmission electron microscopy. In order to avoid interference by type IV pili, we used strains lacking type IV pili. As shown in Figure 3, we saw labeling on the surface of the type IV pili mutant (Fig. 21A) that disappeared when the \textit{knh} gene was interrupted (Fig. 21B). Labeling was less abundant in the \textit{hmw1C}_{Kk} mutant strain (Fig. 21C) and became more abundant when \textit{hmw1C}_{Kk} was reverted to wild type (Fig. 21D). Average number of gold particles on the type IV pili mutant strain was 19.1; when the \textit{knh} gene was interrupted there was an average of 1.1 particles per cell; when \textit{hmw1C}_{Kk} was interrupted there was an average of 6.7 particles per cell; and when \textit{hmw1C}_{Kk} was reverted to wild-type, we saw an average of 11.1 particles per cell.

Bacteria were also examined without labeling, as Knh fibers can be appreciated on the surface.
Figure 21: HMW1C<sub>Kk</sub> is required for Knh expression on the surface of <i>K. kingae</i>

Images show immunogold labeled Knh on the surface of <i>K. kingae ΔpilA1 ΔctrA</i> (Panel A), <i>ΔpilA1 ΔctrA Δknh</i> (Panel B), <i>ΔpilA1 ΔctrA Δhmw1C<sub>Kk</sub></i> (Panel C), and <i>ΔpilA1 ΔctrA Δhmw1C<sub>Kk</sub> (hmw1C<sub>Kk</sub>)</i> (Panel D).

4.3.6 HMW1C<sub>Aa</sub> is necessary for <i>A. aphrophilus</i> autoaggregation and adherence to human epithelial cells

To expand our knowledge of species that use HMW1C-mediated glycosylation, the <i>hmw1C<sub>Aa</sub></i> gene in <i>A. aphrophilus</i> was interrupted and adherence by this strain was compared with adherence by the wild type strain. The interruption of <i>hmw1C<sub>Aa</sub></i> resulted in a statistically significant reduction in adherence (Fig. 22A). Wild type levels of
adherence were restored when the interrupted *hmw1C* was reverted to the wild type sequence (Fig. 22A).

Experiments with *A. aphrophilus* have demonstrated that this species autoaggregates rapidly in liquid culture in the absence of agitation. To determine if interruption of *hmw1C* affects the ability of *A. aphrophilus* to autoaggregate, the OD$_{600}$ was measured over the course of eight hours for isogenic strains containing the wild type, interrupted, and reversion of the interrupted *hmw1C* gene. As shown in Figure 22, the *hmw1C* mutant initially autoaggregated at a slower rate, but was able to obtain the same degree of autoaggregation as the wild type strain over the course of the experiment. Reversion of the interrupted *hmw1C* to wild type restored the rate of autoaggregation to wild type levels (Fig. 22B).
Figure 22: HMW1C\textsubscript{Aa} expression is required for \textit{A. aphrophilus} adherence and autoaggregation

Panel A shows mean adherence of \textit{A. aphrophilus} with the wild type \textit{hmw1C\textsubscript{Aa}} gene, an interrupted \textit{hmw1C\textsubscript{Aa}} gene, and the interrupted \textit{hmw1C\textsubscript{Aa}} gene reverted to wild type. Panel B shows mean autoaggregation of \textit{A. aphrophilus} with the wild type \textit{hmw1C\textsubscript{Aa}} gene, an interrupted \textit{hmw1C\textsubscript{Aa}} gene, and the interrupted \textit{hmw1C\textsubscript{Aa}} gene reverted to wild type. Panel C shows a Commassie blue-stained 7.5\% SDS-PAGE gel of the outer membrane fractions of \textit{A. aphrophilus} with the wild type \textit{hmw1C\textsubscript{Aa}} gene (lane 1), an interrupted \textit{hmw1C\textsubscript{Aa}} gene (lane 2), and the interrupted \textit{hmw1C\textsubscript{Aa}} gene reverted to wild type (lane 3). Each graph shows data from biological replicates performed in triplicate. Error bars represent standard error. An asterisk denotes a p value of < 0.05 obtained from a paired t-test.

4.3.7 HMW1C\textsubscript{Aa} glycosylates the trimeric autotransporter EmaA

Adherence factors have not been described previously for \textit{A. aphrophilus}. To determine the mechanism by which interruption of \textit{hmw1C\textsubscript{Aa}} interferes with adherence
in this species, we examined outer membrane fractions from isogenic strains of *A. aphrophilus* with wild type, interrupted, or reverted *hmw1C<sub>Aa</sub>* on a Coomassie blue-stained SDS-PAGE gel (Fig. 22C). We noticed a band at approximately 180 kDa that had increased mobility in the strain with an interrupted *hmw1C<sub>Aa</sub>* and a band larger than 300 kDa that disappeared in the strain with an interrupted *hmw1C<sub>Aa</sub>* (Fig. 22C). These bands were excised and analyzed by LC-MS/MS on a hybrid LTQ Orbitrap Elite mass spectrometer.

In the samples from *A. aphrophilus*, both the 180 kDa band and the >300 kDa band were identified as the trimeric autotransporter protein EmaA, with the 180 kDa band corresponding to the monomer form and the >300 kDa band likely representing a multimer. Coverage of 82% of the EmaA sequence in the sample from wild type *A. aphrophilus* revealed six sites of glycosylation with a single hexose. (Table 7) All six of the identified sites were asparagine residues and were in the Asn-Xaa-Ser/Thr consensus site for N-linked glycosylation. We obtained 23% coverage of the EmaA protein from the strain with an interrupted *hmw1C<sub>Aa</sub>* gene and detected no modifications. Peptides obtained from the mutant corresponded to the peptides that were modified in the wild type sample.
Table 7: Glycopeptides identified in EmaA

<table>
<thead>
<tr>
<th>Peptide Sequence</th>
<th>Asparagine Position</th>
<th>Calculated m/z</th>
</tr>
</thead>
<tbody>
<tr>
<td>INLNTLDLGSSGSIK</td>
<td>687</td>
<td>904.96</td>
</tr>
<tr>
<td>VSGTSPITVNK</td>
<td>819</td>
<td>632.83</td>
</tr>
<tr>
<td>TTNNGVDDYAVSFNGTEAAK</td>
<td>834</td>
<td>1118.49</td>
</tr>
<tr>
<td>AAVAGTPVNGANGTDGKDGVATQNV</td>
<td>907</td>
<td>1062.53</td>
</tr>
<tr>
<td>VDALNK</td>
<td>1140</td>
<td>727.33</td>
</tr>
<tr>
<td>VSLGGDNGNTTEK</td>
<td>1592</td>
<td>697.34</td>
</tr>
</tbody>
</table>

*aUnderlined residues represent hexosylated sites

4.3.8 EmaA facilitates autoaggregation and bacterial adherence to host cells

The identification of EmaA as the substrate of HMW1C α and the phenotypes demonstrated by an *hmw1C α* mutant suggested that EmaA is involved in *A. aphrophilus* adherence to human epithelial cells and autoaggregation. To determine if EmaA does indeed mediate these functions, we insertionally inactivated the *emaA* gene with a kanamycin resistance cassette. The resulting mutant was unable to adhere to human epithelial cells or autoaggregate (Fig. 23). The dependence of bacterial adherence and autoaggregation on EmaA was confirmed by reverting the *emaA* gene to wild type by removing the kanamycin resistance cassette (Fig. 23).
Figure 23: EmaA is an adhesin and involved in autoaggregation

Panel A shows mean adherence to human epithelial cells by A. aphrophilus with the wild type emaA gene, an interrupted emaA gene, and the interrupted emaA gene reverted to wild type. Panel B shows mean autoaggregation of A. aphrophilus with the wild type emaA gene, an interrupted emaA gene, and the interrupted emaA gene reverted to wild type. Each graph shows data from biological replicates performed in triplicate. Error bars represent standard error. An asterisk denotes a p value of < 0.05 obtained from a paired t-test.

4.4 Discussion

In this study we describe two homologues of the H. influenzae HMW1C protein that are functional glycosyltransferases. Unlike HMW1C, HMW1Ck in K. kingae and HMW1CAa in A. aphrophilus are not encoded in the same genetic locus as a two-partner
secretion system. Rather, these enzymes modify type V secretion proteins encoded by genes at distant locations in the genome. The acceptor protein for HMW1C\textsubscript{Kk} was identified as the trimeric autotransporter protein Knh, and the acceptor protein for HMW1C\textsubscript{Aa} was found to be EmaA, also a trimeric autotransporter.

Autotransporters are a family of adhesive proteins characterized by a C-terminal domain that forms an outer-membrane beta-barrel pore and facilitates surface localization of the N-terminal passenger domain of the protein. A number of autotransporters are modified by O-glycosylation, including Ag43, AIDA-I, and TibA of \textit{E. coli}. Glycosylation of these proteins affects stability (AIDA-I, TibA), oligomerization (TibA), and adhesive activity (AIDA-I, TibA) (Sherlock et al. 2006, Benz and Schmidt 2001, Cote, Charbonneau, and Mourez 2013). In the case of trimeric autotransporters, the protein must trimerize in order to form a fully functional pore. To date all known trimeric autotransporters have adhesive properties, and only one has been identified as being glycosylated, namely the \textit{Aggregatibacter actinomycetemcomitans} EmaA protein (Tang and Mintz 2010).

Knh is a trimeric autotransporter adhesin expressed by the pediatric pathogen \textit{Kingella kingae} and has been shown in previous work to be essential for bacterial adherence to human epithelial cells (Porsch, Kehl-Fie, and St Geme 2012). Our data in this study show that the ability of Knh to mediate adherence and autoaggregation is dependent on expression of \textit{hmw1C}\textsubscript{Kk}. 

101
Interestingly, only 81% of the identified modified residues in Knh were in the N-linked glycosylation consensus sequence. This lack of sequence specificity is reminiscent of O-linking glycosyltransferases, for which no glycosylation consensus sequence has been identified, leading to the hypothesis that O-linked glycosylation sites are specified by structural cues (Charbonneau et al. 2012). It is possible that there is a structural component to the selection of modification sites on Knh. Recent work showed that the HMW1C homologue from *Actinobacillus pleuropneumoniae* is able to modify asparagine residues outside of the NXS/T consensus sequence, including NXA and NXV, albeit with low efficiency (Naegeli et al. 2014). This variation is much more limited than the sites identified in Knh, which included NXG, NXV, N XF, NXF, and NXN.

Glycosylation can protect against protease activity, stabilize protein structure, and facilitate interaction between a protein and a binding partner. Any of these mechanisms could explain the phenotypes that we observed when Knh is not glycosylated. The Western blot of outer membrane fractions revealed that Knh was present in the outer membrane of the strain with the interrupted *hmw1C*Kk gene, albeit at reduced levels. To further investigate the relationship between the amount of Knh on the bacterial surface and glycosylation, we used transmission electron microscopy. We observed surface fibers on wild type *K. kingae* and significantly fewer fibers on a derivative with an interrupted *hmw1C*Kk gene. When the *knh* gene was deleted, we observed no surface fibers. This result is consistent with observation that the *hmw1C*Kk
mutant is deficient in adherence and autoaggregation, similar to a knh mutant (Porsch 2012). The observed decrease of Knh on the surface of the hmw1Cκ mutant could be a result of either a decrease in efficiency in trafficking to the outer membrane or a change in protein stability.

The A. aphrophilus EmaA protein is also a trimeric autotransporter adhesin. The contribution of EmaA to bacterial adherence and/or autoaggregation had not been characterized prior to this report. By interrupting the emaA gene, we showed that both adherence to human epithelial cells and autoaggregation are dependent on expression of EmaA. In the presence of HMW1Cκ, EmaA is modified with a monosaccharide at at least six sites. All six of these sites were at the NXS/T N-linked glycosylation consensus motif, similar to observations with HMW1C and HMW1 and different from HMW1Cκ and Knh. The strain lacking HMW1Cκ displayed reduced adherence to cultured epithelial cells and delayed autoaggregation. Unlike the strain lacking EmaA, the strain lacking HMW1Cκ was still capable of low level of adherence to human epithelial cells, suggesting that unmodified EmaA is still functional, though to a lesser extent than glycosylated EmaA.

The EmaA protein in A. aphrophilus is 69% identical to the EmaA protein in Aggregatibacter actinomycetemcomitans, a protein that has been studied extensively and facilitates adherence to collagen (Mintz 2004). Interestingly, A. actinomycetemcomitans EmaA is O-glycosylated in a process involving enzymes used in the biosynthesis of the
O-polysaccharide of LPS (Naegeli et al. 2014). The proposed modification is a nonasaccharide rather than the monosaccharide that we identified on *A. aphrophilus* EmaA. It is interesting to note that *A. actinomycetemcomitans* does not encode an HMW1C homologue.

In summary, this work describes the first examples of HMW1C-like enzyme that glycosylates a protein that is not encoded in the same gene cluster as the enzyme itself and is not a member of a two-partner secretion system. Furthermore, Knh and EmaA are the first examples of trimeric autotransporters that are N-glycosylated. To date all known targets of HMW1C-facilitated modification are members of type V secretion systems and function as adhesins, with adhesive activity dependent on HMW1C-mediated glycosylation. As adherence to host tissue is a critical first step in infection, the ability to interrupt HMW1C-mediated glycosylation may have therapeutic potential.

### 4.5 GenBank Accession Numbers

Accession number for *hmw1C*~*Aa*~: KR131724

This work has been published as reference below:

5. Future Directions

5.1 Colonization of rhesus macaques with additional strains of NTHi

The investigation of NTHi colonization of the upper respiratory tract has been hampered by the lack of an appropriate animal model. NTHi appears to be a human-restricted organism as it has not been found as a microbiome component of any other animal hosts. Furthermore, the differences between the human respiratory tract and immune system and those of small mammals have complicated the development of an animal model. We have established a model of NTHi colonization using rhesus macaques that appears to mimic human colonization. Using this model we have shown the importance of the adhesive molecules HMW1 and HMW2 in the colonization process.

Further characterization is needed to make this a robust model. To do this we can determine the ability or a diverse set of NTHi strains to colonize the rhesus macaque upper respiratory tract. The experiments to date have been conducted with only NTHi strain 12, an isolate from the middle ear fluid of a child with acute otitis media. Experiments attempting to colonize the animals with a variety of strains will allow us to strengthen our conclusion that rhesus macaques are a good model of NTHi colonization by showing that the results described in Chapter 2 are not strain 12 specific. We will attempt to colonize animals with four other common lab strains: strain 5, strain 11, strain 15, and N187.
Strain 12 expresses the HMW1 and HMW2 adhesins, like approximately 75% of NTHi isolates. The other 25% express another adhesin called Hia. In order to examine the relative colonization efficiency of strains expressing these different adhesins, we can express HMW1/HMW2 or Hia in an isogenic background and use those two strains in our animal model. By comparing the ability of these two strains to colonize the rhesus macaque upper respiratory tract, we can determine if the expression of one or the other adhesin results in an increased ability to colonize. We can also mark these two strains with different antibiotic resistances and use them in a competition experiment. The majority of NTHi isolates express HMW1/HMW2 suggesting that there may be an advantage to the organism in expressing these adhesins rather than Hia. Therefore, it is possible that we will observe that the isogenic Hia-expressing strain is able to colonize fewer animals and is out competed in our competition experiment. Alternatively, we may not observe any differences in the ability of the strains to colonize the animals, thus suggesting that both HMW1/HMW2 and Hia are equally able to facilitate colonization.

As described in Chapter 2, we used the rhesus macaque model to confirm that the HMW1 and HMW2 adhesins are important for colonization. These two proteins have been studied extensively in vitro and are notated as putative colonization factors. There are many other potential NTHi colonization factors that have been studied in vitro but are yet to be confirmed in vivo, a current problem in the NTHi field that can be addressed using our animal model. Factors that would be especially interesting to test
are protein D, protein E, protein F, and LuxS or RbsB. Protein D is surface factor and glycerophosphodiesterase involved in the scavaging and uptake of choline from host cells (Munson and Sasaki 1993). This choline is then added to LOS as phosphorylcholine, which allows LOS to facilitate adherence to bronchial cells (Swords et al. 2000, Swords et al. 2001). Protein E and Protein F are involved in immune evasion by binding to host factors to disguise the surface of the bacterial cell. Finally, LuxS and RbsB are involved in quorum sensing and biofilm development: LuxS is responsible for the production of the quorum-signaling molecule DPD, and RbsB is the cognate DPD sensor (Harrison et al. 2005, Armbruster et al. 2011). I hypothesize that elimination of any of these factors would have a detrimental effect on colonization. To test this we can delete the genes for these factors assess the mutants’ ability to colonize rhesus macaques. If a factor does effect colonization, then the strain missing that factor should colonize fewer animals than the wild type strain and/or be cleared from the upper respiratory tract sooner than the wild type strain.

5.2 Protection by colonization-induced antibody

Some, but not all, of the animals in our studies developed antibody reactive to the HMW1 and HMW2 proteins. We can determine if these antibodies are protective against secondary colonization by strain 12 by attempting to re-colonize the animals that developed anti-HMW1 antibody. The ability of strain 12 to re-colonize animals that have anti-HMW1 antibody would be compared to the ability of the strain to re-colonize
animals that did not develop antibody, as well as animals that have not been previously colonized. These results will be informative for vaccine development, as HMW1 and HMW2 are potential vaccine antigens. If these antibodies prove to be protective against strain 12, we can then attempt to re-colonization with a heterologous strain.

Furthermore, the secondary antiserum used to detect HMW1 reactivity in the rhesus macaque serum samples is reactive against IgG, IgA, and IgE. Therefore, it would be informative to determine which of these three classes the anti-HMW1/HMW2 antibodies belong to. To do this, we can obtain and use antiserum specific to each antibody class. Serum was collected from blood, so the antibodies are most likely IgG. However, it is possible that they could be from the other two classes or a combination of classes.

**5.3 Structural analysis of HMW1**

The work in Chapter 3 indicates that the structure of HMW1 is critical for function and that this structure is obtained through interaction with HMW1B and HMW1C. This relationship between proper folding/structure and function is one of the principles of protein biology.

Solving the crystal structure of HMW1 with and without glycosylation and before and after secretion through HMW1B would complement the results presented in this work. Furthermore, it will allow us to appreciate the specific conformational changes in a way that the techniques used in Chapter 3 could not. CD and extrinsic
fluorescence measurements show a change in overall conformation, while a crystal structure will show exactly where the conformational changes take place.

Our current methods of purifying HMW1 from the surface and cytoplasm of NTHi results in large amounts of clean protein. Work to solve the crystal structure of glycosylated and unglycosylated HMW1 using these purifications is ongoing by our collaborator Guoyu Meng.

5.4 Investigation of additional key modified residues of HMW1

HMW1 is modified by glycosylation at 36 sites (Gross et al. 2008). To date, modification of residues 484 and 806 have been determined to be necessary for HMW1 function. Most likely, there are other critical sites beyond residues 484 and 806, as described in Chapter 3. Using the differential modifications of HMW1 by HMW1Ck as a guide, additional sites of glycosylation can be modified to determine glycosylation of which, if any, are also necessary. Twenty of the 36 sites modified by HMW1C are not modified by HMW1Ck, and HMW1Ck-modified HMW1 is non-functional. This suggests that some of the differentially modified residues may be key to protein function. Of the 20 differentially modified residues, two have been mutated, 636 and 806, and mutation of residue 806 produced non-functional protein. Each of the other 18 differentially modified sites can be mutated individually, and the ability of the resulting protein to facilitate adherence to epithelial cells evaluated. In this way, additional critical modifications can be identified.
5.5 Specificity of HMW1C for HMW1

HMW1 is the only protein in NTHi modified by HMW1C; however, the basis for this specificity has yet to be uncovered. *In vitro*, HMW1C is capable of modifying a variety of peptides, some of which were derived from HMW1, and others derived from other proteins (Grass et al. 2010, Naegeli et al. 2014). Additionally, there is evidence that heterologously expressing HMW1C in *E. coli* results in some off-target modifications (Gawthorne et al. 2014). Therefore it would be interesting to determine what leads to the specificity of HMW1C for HMW1 in NTHi.

There are at least three potential explanations of the specificity of HMW1C for HMW1: (1) the HMW1 signal sequence that is somehow important for the specificity of HMW1C, (2) the spacial proximity of the *hmw1* and *hmw1C* genes that somehow leads to the spacial proximity of their respective proteins, or (3) the stoichiometric relationship between HMW1 and HMW1C so that HMW1C is unavailable for the glycosylation of other proteins.

The use of a specific signal sequence is a common way in which two proteins can achieve a specific interaction. HWM1 contains at least two signal sequences; one that directs secretion through the Sec secretion system and another that is specifically recognized by HMW1B. To test whether there is a third sequence that directs interaction between HMW1 and HMW1C, we can create fusion proteins containing the N- or C-terminal of HMW1 and test whether the fusion protein is glycosylated.
HMW1C is clearly not dependent on a single sequence as it can modify dis-similar peptides \textit{in vitro}. Furthermore, alignment of the targets of glycosylation by HMW1C, HMW1C\textsubscript{Kk}, and HMW1C\textsubscript{Aa} show no similarities that would suggest a conserved signal sequence.

HMW1 and HMW1C are encoded in the same genetic locus in NTHi. If there is a mechanism that retains the translated HMW1C enzyme near the HMW1 locus, then the specificity of HMW1C for HMW1 could be based on the proximity of the proteins. Expressing HMW1C off of a plasmid in NTHi and seeing if it then modifies any off-target proteins can test this hypothesis. Preliminary data suggests that HMW1C expressed in such a way is capable of modified HMW1, though the efficiency of HMW1 modification or the existence of any off-target modifications are unknown.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure24.png}
\caption{Plasmid Expression of HMW1C complements a HMW1C Genomic Deletion}
This western blot using antiserum raised against HMW1 shows HMW1 in a wild-type strain (lane 1), an absence of HMW1 in a strain with a genomic deletion of HMW1C (lane 2) and the recovery of HMW1 in strains with a genomic deletion of HMW1C complemented with HMW1C expressed off of a plasmid (lanes 3-7).

Finally, it is possible that there is a finely-tuned stoichiometric relationship between HMW1 and HMW1C that prevents HMW1C from modifying other proteins. If HMW1C has a higher affinity for HMW1 than for any other protein and there is enough HMW1 present to saturate the HMW1C population, then off-target proteins would not
be modified. Examining the glycosylation of proteins in a HMW1 mutant can test this hypothesis. If stoichiometry is the key to HMW1-specific modification, I would expect to find many off-target modifications in a HMW1 mutant strain.

5.6 Small molecule inhibitor of HMW1C

Adherence is the first step in disease and the ability to disrupt this step could be an ideal way to prevent and control disease. The studies outlined in Chapter 4 showing that three pathogenic species use HMW1C-like N-linking glycosylation to modify confirmed and potential colonization factors suggest that HMW1C may be a viable target for therapeutic intervention. In NTHi, *K. kingae*, and *A. aphrophilus*, HMW1C-like proteins modify adhesive proteins. Without glycosylation these proteins are unable to facilitate adherence. A small-molecule inhibitor of HMW1C function could be used to prevent disease caused by these three organisms and potentially others with a functional HMW1C homologue. Such an inhibitor can be found via a high-throughput screen using a commercially available small-molecule library. HMW1C is capable of modifying an acceptor protein *in vitro* when UDP-glucose or UDP-galactose is present. For a high throughput screen HMW1C, an acceptor protein, UDP-glucose, and the reagents from the Promega UDP-Glo Glycosyltransferase Assay kit will be combined in 96 well plates. Potential inhibitors from the small molecule library will also be added to the wells. The UDP-Glo Glycosyltransferase kit reacts with free UDP to create a fluorescent product. The plates will be read to reveal any wells showing a decrease in fluorescence, thereby
indicating a decrease in HMW1C activity. The small molecule responsible for the decrease will be further studied to determine any therapeutic potential.
References


amplification and sequencing of the 16S rRNA gene."


Pfeiffer, R. 1892. "Aus dem Institut für Infektionskrankheiten. II. Vorläufige Mittheilungen über die Erreger der Influenza." Deutsche medicinische Wochenschrift no. 18 (28).


Biography

Katherine Rempe was born in Port Townsend, WA on January 5th of 1990 and grew up in Puyallup, WA. Katherine attended the Honors College at Washington State University in Pullman, WA as a Washington State Scholar and WSU Regents Scholar. In December of 2010, she graduated summa cum laude with a Bachelor’s of Science in microbiology and a minor in Spanish.

Katherine completed her graduate work at Duke University and the Children’s Hospital of Philadelphia under Dr. Joseph W. St. Geme III. Her graduate studies were partially funded by the National Science Foundation’s Graduate Research Fellowship. Katherine also received the Chancellor’s Award from Duke University and the Richard and Mary Finkelstein Student Travel Grant to present her work at the 2015 American Society for Microbiology (ASM) General Meeting. She was an active member of the student chapter of the Eastern Pennsylvania Brach of the ASM and served as secretary for 18 months. Katherine graduated with her PhD in Molecular Genetics and Microbiology in May of 2016.

Peer Reviewed Publications

Rempe KA, Wilson JM, St. G.eme JW 3rd. The HMW1 and HMW2 adhesins enhance the ability of nontypeable Haemophilus influenzae to colonize the upper respiratory tract of rhesus macaques. Infection and Immunity. Submitted


Rempe KA, Hinz AK, Vadyvaloo V. Hfq regulates biofilm gut blockage that facilitates flea-borne transmission of Yersinia pestis. Journal of Bacteriology 194(8):2036-40, April 2012

**Book Chapters**
