FUNCTIONAL CHARACTERIZATION OF TWO *ANAPLASMA PHAGOCYTOPHILUM* GENES REQUIRED FOR INFECTION OF MAMMALIAN HOST AND TICK VECTORS, RESPECTIVELY

A DISSERTATION

SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL

OF THE UNIVERSITY OF MINNESOTA

BY

Adela Sarahí Oliva Chávez

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Dr. Ulrike G. Munderloh

January, 2014

Acknowledgements

I want to thank to my adviser, Dr. Uli Munderloh, for all these years that she has helped me to become a better researcher. Her guidance and encouragement is what has helped me to finish. I am very thankful with Dr. Kurtti as well. His kindness and advice was also essential to get through all the hard days.

I want to thank my committee members Dr. Fallon and Dr. Redig for their help and revisions of this document.

I also want to recognize Nicole, Rod, Curt, and Mike who helped me endlessly with my research, teaching me techniques and how to do things. I know I am a thick head and sometimes it may have been hard. I want to thank Kendra, Bryten, Chan, Jon, Hyo, and Geoff for help and a precious friendship in the lab. I will miss you guys once I leave.

Also, my friends in the department Theresa, Renatita, Amy, and Thelma for the many times you guys held my hands, or gave me a kind word or just some company.

This research was partially funded by the Lydia and Alexander Anderson Fellowship from Graduate School and NIH grant R01-AI042792 to Dr. Munderloh.

Dissertation Abstract

Anaplasma phagocytophilum is a tick-borne pathogen and the causative agent of Human Granulocytic Anaplasmosis (HGA). *A. phagocytophilum* is not transovarially transmitted from the mother to the progeny of infected ticks and therefore needs to survive in both a mammalian and the arthropod vector in order to complete its life cycle. To adapt to different environments, *A. phagocytophilum* relies on differential gene expression as well as the post-translational modification of proteins. However, little is known about what *A. phagocytophilum* genes and enzymes are required for the infection of human or tick cells. I used random mutagenesis to generate knock-out strains of *Ap* for functional genomics studies. One of the mutated strains presented an insertion within the coding region of an *O*-methyltranferase (OMT) family 3, which affects the ability of *A. phagocytophilum* to infect tick cells. Studies in the function of this enzyme suggest that it is involved in the methylation of an outer membrane protein (Major Surface Protein 4), which appears to be involved in bacterial binding and entry. The second mutant presents an insertion within the coding region of a hypothetical protein in the locus *APH_0906* and is unable to infect HL-60 cells and is impaired in its ability to grow in endothelial cells. Localization analysis of the protein showed that the protein is secreted into the cytoplasm and then translocated into the nucleus of host cells. Bioinformatic analyses demonstrated differences in Nuclear Localization Signals (NLSs) as well as binding residues within the protein homolog of human and non-human strains. Herein, I present the results from the functional analyses of both gene products.

Table of Contents

Page

List of Tables

Chapter 2. Identification of two genes involved in the infection process of the tick-borne pathogen *Anaplasma phagocytophilum*

Chapter 3. Methylation of an *Anaplasma phagocytophilum* outer membrane protein is required for adhesion to tick cells

Chapter 4. The *Anaplasma phagocytophilum* hypotethical protein APH_0906 is secreted into the cytoplasm of mammalian host cells

List of Figures

Chapter 2. Identification of two genes involved in the infection process of the tick-borne pathogen *Anaplasma phagocytophilum*

Chapter 1

Important aspects on the biology of *Anaplasma phagocytophilum* and findings on the study of genes important for the infection of human and tick hosts

Introduction

Human Granulocytic Anaplasmosis (HGA)

HGA, caused by *Anaplasma phagocytophilum,* is the second most frequently diagnosed tick-borne disease in United States after Lyme disease [1]. Like the Lyme disease agent, *Borrelia burgdoferi,* the etiologic agent of HGA, is maintained in nature in a transmission cycle involving small rodents such as mice, chipmunks and squirrels, and *Ixodes* spp. ticks [2,3]. *A. phagocytophilum*, the etiologic agent of HGA, infects peripheral blood phagocytes, specifically neutrophil granulocytes, and their progenitors in the bone marrow. *A. phagocytophilum* was previously known as *Rickettsia phagocytophila, Cytoecetes phagocytophila,* the Human Granulocytic Ehrlichiosis (HGE) agent, *Ehrlichia equi or Ehrlichia phagocytophila* before it was finally assigned to the genus *Anaplasma* [4,5]. HGA shares many symptoms, signs, and laboratory findings with diseases caused by other members of the family *Anaplasmataceae*, including high fever, headache, myalgia, malaise, thrombocytopenia, leukopenia, and elevated serum transaminases due to mild liver injury [6]. Additionally, respiratory and gastrointestinal manifestations are frequently reported. Following diagnosis, half of all patients require hospitalization and 7% of these require intensive care. The mortality rate is 0.5% - 1% , affecting mainly immunocompromised, elderly, or other people suffering from a pre-existing condition [6-8]. Complications in HGA patients can include shock, clotting disorder, respiratory distress, myocarditis, renal failure, hemorrhage, and opportunistic infections [6,7].

Severity of the disease is independent of parasite load, and many of the signs and complications are due to the host's inflammatory response [4,9], which causes lesions observed during HGA, and has been related to the over-expression of pro-inflammatory cytokines such as interferon gamma (IFN-γ) [9,10],[11]. The inflammatory response, accompanied by activation of NKT cells, is triggered by the polar lipid components of the bacterial outer membrane, and partially explains the histopathologic findings observed during infection with *A. phagocytophilum* [12]. Dysregulation of cytokine and chemokine production further affects differentiation of blood cells in the bone marrow during haematopoesis, and leads to the thrombocytopenia and leukocytopenia characteristic of the disease [13].

HGA is increasingly diagnosed in the USA, and has recently been diagnosed in Canada for the first time [14]. There has been a steady increase in the number and incidence of HGA cases [15], from 348 annual cases in 2000 to 1,761 reported in 2010, representing a rise from 1.4 cases per million to 6.1 cases per million [8]. The CDC [16] estimates that 88 % of HGA cases reported in the USA come from only 6 states, one of them Minnesota where a peak number of 788 cases was diagnosed in 2011 [17]. Clinical human cases have also been reported in Asia and Europe [6]. The first confirmed European case was diagnosed in Slovenia in 1997 [18], but the disease has remained relatively rare in that region, with just over 70 cases of HGA reported in Europe to date [reviewed in [19]. In some regions in Central Eastern Europe the sero-prevalence can be as high as 15.4 % [20], demonstrating frequent exposure of people living in endemic areas. In China and other Asian countries such as South Korea, seroprevalence of up to

20 % has been reported, and several HGA cases have been identified in individuals living within high risk areas [21-23].

The most common drugs for treatment of HGA are the tetracyclines, especially doxycycline that are considered the drugs of choice for all patients including pregnant women and children [1]. The CDC [1] recommends that treatment should be started before laboratory confirmation is obtained as complications are more likely to occur when treatment is delayed.

Anaplasmosis in domestic and wild animals

The principal natural mammalian reservoir of *A. phagocytophilum* in the USA are small rodents, including white-footed mice (*Peromyscus leucopus*), chipmunks, squirrels and woodrats [3,44, 45]. However, *A. phagocytophilum* has been detected in other rodents such as southern red backed voles, short tailed shrews, meadow jumping mice, and meadow voles. Analysis of 16S rDNA and other genes (e.g., *ankA*) suggested that mouse and human isolates were similar, but those from woodrats were distinct. Moreover, *A. phagocytophilum* variants found in woodrats did not infect horses, whereas chipmunk and human origin *A. phagocytophilum* did [24].

Tick-borne fever (TBF) in sheep and pasture fever in cattle caused by *A. phagocytophilum* has been known in Europe for at least 200 years [25,26]. In the UK, it was estimated that at least 300,000 lambs develop TBF per year [26], over half of all lambs in Southern Norway tested seropositive [27], and it ranked third as a cause of abortion in sheep in northern Spain [28], indicating it is wide-spread in Europe. In addition, molecular and immunological evidence of *A. phagocytophilum* infection in

wild-ruminants, especially deer, has been widely reported in several countries in Europe, North America and Asia [26,29,30]. Ruminant variants of *A. phagocytophilum* associated with TBF in Europe have not been found associated with human disease, which is similar to the Ap-Variant 1 isolated from ticks in Rhode Island and Minnesota that has been found to infect only ruminants [31-35]. Indeed, a human isolate of *A. phagocytophilum* (NY-18) was unable to persistently infect white-tailed deer when injected intravenously [33], whereas deer injected with the Ap-Variant 1, an *A. phagocytophilum* strain associated with deer, remained positive for 65 days by IFA and 25 days by PCR [33]. *A. phagocytophilum* strains similar to Ap-Variant 1 appear to be more abundant in ticks when compared to human strains. This has led to the hypothesis that the high presence of ruminant variants may explain the low rates of human infections in Europe and in some endemic regions in the USA [30].

A. phagocytophilum also infects other domestic animals such as dogs, cats, and horses [25,26], and wild animals, such as wild boar, red and gray foxes, opossums, and raccoons [36] [25]. The highest prevalence of *A. phagocytophilum* in dogs in the USA has been documented in the upper Midwest (6.7 %) and the Northeast (5.5 %), and can reach 50 % in some regions of Wisconsin, Minnesota, and Massachusetts [37]. Infection of dogs with *A. phagocytophilum* has also been detected in other countries, including Brazil, Slovakia, France, Germany, and South Korea [38-42]. Most infected dogs develop fever, lethargy, anorexia, and sometimes diarrhea and coughing [reviewed in [43]. 30 % of the domestic cats from endemic areas in the northeastern USA had serum antibodies against *A. phagocytophilum* [44], but prevalence of clinical disease in cats is actually low (4.3 %) [45]. Even though experimentally cats developed fever and 8 % of neutrophils

contained morulae, anorexia and lethargy were not present [46]. Phylogenetic analysis of the variants infecting horses, dogs, and cats, showed that they were closely related to human pathogenic variants, although analysis of the abundance of these variants in tick populations demonstrated that they were rare in the regions where the cases were reported [47].

Anaplasma phagocytophilum

A. phagocytophilum is a small, gram negative bacterium in the family *Anaplasmataceae* within the order *Rickettsiales* around 0.4 X 2.0 μm in size that is enveloped by two membranes (reviewed in [48]). Members of the family *Anaplasmataceae* have lost the genes involved in biosynthesis of lipopolysaccharide (LPS) and peptidoglycan, thus they escape recognition by Toll-like receptors (TLR) and do not present pathogen-associated molecular patterns (PAMPs) that stimulate the innate immune response of mammals and invertebrates [49]. *A. phagocytophilum* grows in membrane-bound inclusions called "morulae", unlike members of the family *Rickettsiaceae* that replicate directly in the cytosol of cells [49]. The most abundant protein of *Anaplasma phagocytophilum* is outer membrane protein MSP2 (P44), which is involved in a mechanism of antigenic variation that facilitates persistent infection [50]. MSP2 (P44) has been proposed to act as a porin [51] that is formed by dimers and oligomers of several interacting P44s [52], and may act as an adhesin to human granulocytes [53]. This protein is highly upregulated during growth in mammalian cells as compared to growth in tick cells (>10 fold up-regulation) [54], suggesting that the function of this protein is specific to mammalian cells. Other proteins that have been

detected as part of the outer membrane of *A. phagocytophilum* included APH_0441, OMP85, APH_0404, APH_0405, several P44 paralogs, and OMP1-a [55]. Two of these proteins, APH_0404 and APH_0405, were shown to be involved in infection as pre-treatment with monoclonal antibodies against these proteins prevented internalization of the bacteria [55].

In addition to proteins, *A. phagocytophilum* incorporates cholesterol obtained from the host cell into its outer membrane, since it lacks genes involved in the synthesis and/or modification of cholesterol [56]. Host derived cholesterol was required for the integrity of the outer membrane of both *A. phagocytophilum* and the related tick-borne pathogen *Ehrlichia chaffeensis.* Treatment of host cell-free bacteria with Methyl-β-cyclodextrin (MβCD) to extract cholesterol disrupted the bacterial membrane integrity and prevented invasion of host cells [56]. The replacement of cholesterol with NBD-cholesterol, a non-functional fluorescent analogue, also affected the ability of the bacteria to infect and survive within host cells [56].

A. phagocytophilum has a relatively small genome with only 1.47 Mb and 1,369 open reading frames (ORFs), 45 % of which encode hypothetical proteins (82 ORFs represent conserved hypothetical proteins and 458 ORFs encode hypothetical proteins unique to this organism). There is no evidence that *A. phagocytophilum* harbors plasmids or transposons and only a few phage components are found in its genome. Several of the hypothetical proteins have been suggested to be part of the outer membrane or to be components of the type 4 secretion system (T4SS) of the bacteria (reviewed in [48]). Like other members of the *Rickettsiales*, *A. phagocytophilum* has undergone reductive evolution, which explains the need for a host cell to provide essential functions [57]. *A.*

phagocytophilum has evolved to survive in two completely different environments, the mammalian host and the tick vector. This has been accomplished, in part, through differential use of its genome and expression of certain genes only during residence in a specific environment [54]. The differences in development and gene expression in each host cell type are described below.

A. phagocytophilum in neutrophils

A. phagocytophilum infects mammalian neutrophil granulocytes and their progenitors by using the α-(1,3) fucosylated P-selectin glycoprotein ligand-1 (PSGL-1) for entry [58]. *A. phagocytophilum* binds to the N-terminal region of PSGL-1 and to sialyl Lewis x (sLe^x) through interactions with α2,3- sialic acid and α1,3- fucose. However, pretreatment of HL-60 cells with antibodies against PSGL-1 (KPL1), the N-terminal amino acid sequence of PSGL-1 (mAb PL1), and sLe^x mAbs, or treatment with sialidase did not completely prevent binding of *A. phagocytophilum*, indicating that an alternative adhesin exists that interacts with a different ligand [58,59]. This was confirmed using HL-60 cells deficient in sLe^x (HL-60 $sLe^{x\text{-}low}$) to enrich for bacteria that do not require this ligand [59]. These sialic acid-independent bacteria showed a reduced ability to bind to wild-type HL-60 cells, however, when HL-60 cells were pretreated with the antibodies mentioned above, sialic acid-independent *A. phagocytophilum* exhibited greater binding and internalization than wild-type populations [59]. This sialic acid-independent *A. phagocytophilum* population supports the notion that an alternate adhesin protein exists but is underrepresented in the bacterial outer membrane. Two proteins have been identified as possible adhesins interacting with PSGL-1, OmpA [60] and Asp14

(Aph_0248) [61]. Biotinylation of the outer membrane of *A. phagocytophilum* indicated that these proteins were abundant in dense core (DC) bacteria along with P44 [61]. Expression of both proteins was up-regulated in bacteria binding to neutrophils, and preincubation of the bacteria with serum against either protein interfered with infection [60,61]. Likewise, incubation of HL-60 cells with recombinant versions of OmpA and Asp14 reduced the number of bacteria infecting HL-60 cells by 50 % [60,61]. Although three possible invasins have been identified, it is very likely that as yet unidentified proteins are involved in adhesion of *A. phagocytophilum* to neutrophils and endothelial cells [reviewed in [62].

It has been suggested that *A. phagocytophilum* does not enter neutrophils by phagocytosis, but instead actively invades cells via caveolae mediated endocytosis [63]. There is evidence that PSGL-1 interacts with a cytoplasmic tyrosine kinase (Syk) and ROCK1, a kinase that co-localizes with caveolae, to support entry of *A. phagocytophilum* into the cell [64] as cells treated with *rock1* siRNA did not efficiently internalize *A. phagocytophilum* [64]. Alternatively, activation of Syk may promote cytoskeleton rearrangement that induces phagocytic uptake of the bacteria [65]. Also, *A. phagocytophilum* requires lipid rafts or proteins concentrated in lipid microdomains, such as Caveolin-1, Glycosylphosphatidylinositol anchored proteins (GAPs) and flotillin 1 for internalization (reviewed in [48]). A recent study proposed that *A. phagocytophilum* was able to sequester Rab proteins, such as Rab10, Rab11A, Rab14, and other Rab proteins to acquire amino acids and possibly cholesterol, and to coat the vacuolar membrane of the inclusion in which the bacterium develops, thus preventing maturation of the vacuole and

lysosomal fusion [66]. Rab proteins are involved in regulating endocytic recycling and transport of vesicles in the slow clathrin-independent pathway [66].

Like other members of *Anaplasmataceae*, such as *Anaplasma marginale* and *Ehrlichia chaffeensis*, *A. phagocytophilum* undergoes a biphasic developmental cycle, alternating between dense-cored (DC) forms 0.6-0.8 μm in width, which possess a dense nucleoid and a ruffled outer membrane, and larger reticulated forms (RC) [65]. Only the DC form of the bacteria bound to and was internalized by HL-60 cells into cell-derived vacuoles. By 12 hr DC forms were less frequent and most of the pathogens occupying vacuoles had changed into RC. RC bacteria replicated, presented dispersed nucleoids and smoother outer membranes, and measured between 0.69±0.36 μm and 1.04±0.44 μm in size. By 24 hr, differentiation into DC cells occurred, and bacteria were released from host cells, completing the developmental cycle of *A. phagocytophilum* in neutrophils. However, the mechanism whereby bacteria are released from the host cell remains unknown.

The vacuole where *A. phagocytophilum* develops presents hallmarks of an autophagosome, such as a double-lipid bilayer and colocalization with LC3-II and Beclin 1 [67]. Moreover, treatment of HL-60 cells with 3-MA, a pharmacological inhibitor of autophagosome formation, inhibited of *A. phagocytophilum* replication [67]. Three hypothetical proteins from *A. phagocytophilum* have been shown to be associated with the formation of the vacuole APH_0233, APH_1387, and APH_0032 [reviewed in [48]. Both APH_0032 and APH_1387 are acidic proteins that contain several repeats. APH_0032 possesses eight repeats in its C-terminal domain that consist of nearly identical sequences of 33 to 35 amino acids [68], whereas APH_1387 has only 3 tandem

repeats of 93 – 130 amino acids consisting mostly of amphipathic alpha helices with similarity to chlamydial Inc proteins [69]. Both proteins are secreted into the vacuole, however, APH_1387 is expressed by RC bacteria, whereas APH_0032 is expressed by DC organisms [68,69]. Both proteins appear to colocalize with the vacuole membrane in mammalian and tick cells, but their function is not known [68,69]. APH_0233, also known as AptA, interacts with vimentin, a filament protein associated with phosphorylated Erk1/2 that is localized around the bacterial inclusions [70]. AptA is preferentially expressed during infection of mammalian cells [54,70], is incorporated into both bacterial and vacuolar membranes, and has been proposed to aid in the acquisition of lipids from the host cell [54,70]. All these proteins probably interact with host cell proteins and modify the vacuole to promote replication of the bacteria.

One of the host proteins that seems to be associated with the bacterial vacuole appears to be ubiquitinated. Ubiquitin typically modifies lysine residues on target eukaryotic proteins after translation [71]. The ubiquitination of vacuole proteins starts 4 hr p.i. during vacuole formation and increases for up to 18 hr decreasing after 24 hr p.i. probably after the morulae has become "mature." Similar to Rab GTPases, colocalization of ubiquitinated proteins with the bacterial inclusion depends on *A. phagocytophilum* protein synthesis as ubiquinization is reduced with tetracycline treatment. Mono-ubiquitination of the inclusion membrane is also observed in RF/6A cells, but is reduced in ISE6 cells [71]. Ubiquitination may avoid lysosomal targeting of the morula by the host cell, preventinf destruction of the bacteria.

A. phagocytophilum has evolved ways to escape destruction by the phagocytic machinery of neutrophils [72]. Infected neutrophils are unable to adhere to endothelium,

transmigrate and generate phagocyte oxidase, and are deficient in anti-microbial activity [73]. Infection with *A. phagocytophilum* stimulates NADPH oxidase production in neutrophils but this does not result in an increase of reactive oxygen species (ROS). Apparently, the bacteria are able to scavenge superoxide (O_2^-) by some heat-labile surface protein [63], a mechanism that is important during early entry of *A. phagocytophilum* into the neutrophil when the bacteria stimulate cellular activation and O_2^- release [63]. After internalization, *A. phagocytophilum* is relocated into a vacuole that excludes NADPH oxidase and does not acquire $gp91^{phox}$ and $p22^{phox}$ [63]. Infection with *A. phagocytophilum* resulted in the reduction of PU.1 and Interferon Regulatory Factor 1 (IRF1) concentrations in nuclear extracts of HL-60 cells, in turn causing enhanced binding of the CCAAT displacement protein (CDP) that acts as a repressor in the promoter region, resulting in inhibition of $gp91^{phox}$ expression [74]. Although all members of the *Rickettsiales* contain *sodB*, encoding an iron superoxide dismutase that protects against intracellular ROS produced during metabolic activities, and which is probably injected into the host cytosol by the type IV secretion system [75], it is not involved in defense against extracellular ROS.

A. phagocytophilum infection of neutrophils furthermore reduces their ability to adhere to endothelial cells, which results from the degranulation induced by *A. phagocytophilum* [76]. bacterial binding to E-selectin during infection reduces PSGL-1 and L-selectin expression in neutrophils and decreases the calcium flux that normally occurs during PMN activation during rolling [77,78]. Reduced E-selectin-mediated adherence to endothelium suppresses p38 MAPK phosphorylation, which is necessary for arrest and directional transmigration [77]. It has been suggested that the resulting

increased availability of infected neutrophils circulating in the blood stream enhances acquisition by ticks during feeding [78]. *A. phagocytophilum* are able bind to neutrophils even after they have engaged in rolling under shear in a flow chamber, which may reflect how bacteria and uninfected neutrophils initially make contact in a mammalian host [77]. Inflammation of microvascular endothelium at the tick bite site attracts neutrophils that would come into contact with bacteria that were released into the bite wound by feeding ticks and had infected microvascular endothelial cells in the area. This is plausible considering that neutrophils do not return to the lumen of blood vessels once they have completed extravasation, and provides a mechanism for infection of neutrophils that does not require their transmigration [79].

There is some disagreement about the transcriptional response of of neutrophils during infection with *A. phagocytophilum*, however upregulation of cytokines/chemokines, such as IL-8 (involved in inflammation response) and IL-1β (involved in cell adhesion to endothelial cells), as well as transcriptional regulator molecules has been reported by several researchers [72,78,80,81]. A central theme observed is global inhibition of pro-apoptotic pathways through manipulation of mitochondrial responses including the p38MAPK and ERK signaling pathways, as well as inhibition of the Fas-induced pathway of apoptosis involving TNFAIP3, CFLAR, and SOD2. At the same time, there is induction of anti-apoptotic genes collectively allowing prolonged circulation of infected neutrophils in peripheral blood to make them available for acquisition by feeding ticks [72,80,81]. Thus *A. phagocytophilum* appears to manipulate both the intrinsic (mitochondrial) and extrinsic (death receptor) pathways [81].

Other cells infected by *A. phagocytophilum*

Experimental infection of mice has demonstrated dissemination of *A. phagocytophilum* to multiple tissues including skin, lungs, kidney, liver, spleen, lymph nodes, and bone marrow, suggesting that cells other than granulocytes become infected, e.g., endothelial cells, megakaryocytes and mast cells [82]. In support of this hypothesis, *A. phagocytophilum* causes persistent infections in ruminants with periodically recurrent bacteremias [83]. Also endothelial cell lines, such as RF/6A (American Type Culture Collection, Manasssas, VA, USA; ATCCCRL-1780) from the retina choroid of a normal fetal rhesus monkey (*Macaca mulatta*) and the human microvascular endothelial cell line HMEC-1, support continuous replication of *A. phagocytophilum* [84]. The appearance of *A. phagocytophilum* in endothelial cells differs from that observed in HL-60 cells or neutrophils, since colonies of bacteria within endothelial cells are smaller and more numerous, and individual bacteria are always distinguishable [84]. Transfer of *A. phagocytophilum* between infected HMEC-1 cells and human neutrophils occurred rapidly and was highly effective [85]. The closely related bovine pathogen *A. marginale* infected endothelial cells in vitro, and in the kidney of an experimentally infected calf *A. marginale* colocalized with von Willebrand factor VIII, an endothelial marker [86]. Even though *in vivo* infection of blood vessel endothelial cells with *A. phagocytophilum* has not been demonstrated, it is very likely that this bacterium does in fact infect microvascular endothelium, especially around the tick bite site.

A. phagocytophilum in ticks

A. phagocytophilum is transmitted exclusively by members of the *Ixodes ricinus* complex. In the Northeast and Midwest and the Westcoast states of the USA, *Ixodes scapularis* (the black-legged tick) and *Ixodes pacificus* (the western black-legged tick), respectively, are the vectors, and in Europe and Asia, *A. phagocytophilum* is transmitted by *Ixodes ricinus* and other ticks in the *Ixodes persulcatus* complex, including *I. persulcatus Ixodes ovatus* [87]. However, it has also been found associated with other tick species, such as *Haemaphysalis punctata, Riphicephalus sanguineus,* and *Dermancentor reticulatus* [reviewed in [25], [88]]. It has been estimated that in endemic areas of the USA, 10 – 50 % of *I. scapularis* ticks are infected with *A. phagocytophilum*, and the proportion of infected *I. ricinus* in Europe is similar, at 2 – 45 % [reviewed in [8]]. *Ixodes* ticks cannot transmit *A. phagocytophilum* transovarially from one generation to the next, and immature ticks must reacquire the bacteria with an infected blood meal to maintain natural transmission cycles [reviewed in [8]]. A comparison of the genomes of the *Anaplasmataceae* with those of the *Rickettsiaceae* that are transmitted transovarially, revealed absence of a class II aldolase/adducing domain protein in the *Anaplasmataceae*, which was hypothesized to enable transovarial passage in the *Rickettsiaceae* [75]. Transovarial transmission of *A. phagocytophilum* was reported in *Dermacentor albipictus* (winter tick), as 42 % of larval progeny of infected ticks collected from deer were also infected, as demonstrated by PCR and sequencing. However, transovarial transmission to F2 larvae did not occur [89]. The *A. phagocytophilum* strains infecting these ticks were WI-1 and WI-2, which are believed to be non-pathogenic to humans, similar to Ap-Variant 1 [89]. This is the only evidence supporting transovarial transmission of *A.*

phagocytophilum, although the low efficiency of transovarial maintenance confirms the need for the pathogen to infect mammalian hosts in order to maintain its life cycle in nature.

Acqusition of *A. phagocytophilum* began within the first 24 hr of tick feeding, but the percentage of infected ticks continued to increase substantially until time of de-attachment. Subsequently, the number of infected ticks in a cohort declined, while the number of bacteria within individual ticks increased [90]. Under controlled experimental conditions, the time required for transmission of *A. phagocytophilum* to naïve mice was 40 to 48 hrs, suggesting that bacteria must replicate to reach optimal numbers for transmission to occur [90]. The development of *A. phagocytophilum* and its tissue tropism in the vector has not been described, and only one study demonstrated *A. phagocytophilum* in gut muscle, salivary glands and the synganglion of adult ticks [33]. Although the presence of *A. phagocytophilum* in hemocytes was interpreted to indicate their involvement in the transport of the bacteria to different organs, it is also possible that this resulted from phagocytosis as a defense against the bacteria [91]. The only detailed description of the development of *A. phagocyophilum* in tick cells so far is available from in vitro studies using a tick cell line, ISE6, derived from the vector. The development of *A.phagocytophilum* in tick cell culture was similar to that observed in human cell culture, and proceeded from invading dense forms (DC) to RC forms that replicated inside large morulae until differentiation into DC and release from cells by 4 days when the cycle began anew [92]. Major differences in the development in tick cells versus human cells included that individual inclusions were much larger, there were only a few (usually one to three) in each host cell, and completion of the developmental cycle

took nearly twice as long [65]. Some morulae within tick cells contained so many *A. phagocytophilum* that individual bacteria were hard to discern [92].

A. phagocytophilum recognizes and binds to specifically glycosylated ligands on neutrophil granulocytes during infection of mammalian cells [58,93]. It is interesting that *A. phagocytophilum* utilizes platelet selectin glycoprotein-1 (PSGL-1) that must be both α-2,3-sialylated and α1,3-fucosylated to invade human neutrophils. The as yet unknown mouse-receptor is distinct from PSGL-1, but must also be α1,3-fucosylated while α-2,3-sialylation is not required [93]. Similarly, α1,3-fucosylation of an unidentified ligand was required for acquisition of *A. phagocytophilum* by *I. scapularis* ticks, and for colonization of *I. ricinus* cells, but not for tick-transmission to mice [94]. Possibly, utilization of PSGL-1 is a recently acquired trait in *A. phagocytophilum* and may play a role in pathogenesis of human anaplasmosis by altering the pro-inflammatory response of neutrophils [95]. By comparison, utilization of α1,3-fucosylated moieties that participate in infection of diverse organisms by *A. phagocytophilum* appears to be a more anciently conserved mechanism.

A. phagocytophilum was shown to interact with or take advantage of *I. scapularis* proteins to infect and probably mobilize within the tick. Two of these proteins, P11 (DQ066011) and Salp16, were upregulated in infected *vs* uninfected ticks. They specifically affected transfer of *A. phagocytophilum* to salivary glands but not colonization of the midgut, and appeared to be secreted by salivary gland cells or possibly hemocytes (P11) [48] While these studies demonstrated the specific requirement of the proteins for colonization of tick salivary glands by *A. phagocytophilum* using RNAi, their function in ticks and the ligands or receptors

involved in their interaction with *A. phagocytophilum* remain undefined. Not unexpectedly, proteins that were required for efficient tick engorgement, oviposition, or survival, such as subolesin, also impacted *A. phagocytophilum* infection of tick salivary glands [96]. It is likely that these interactions were not specific to the pathogen.

An analysis of data obtained from infected *I. scapularis* processed during transmission feeding revealed a small set of *A. phagocytophilum*-transcripts using liquid chromatography tandem mass spectrometry (LC MS/MS) and RNA-Seq. The authors hypothesized that the reason for this limited number of active genes detected was due to the fact that at that time, only DC bacteria were present [97]. However, whether DC *A. phagocytophilum* are truly inactive awaits conclusive analysis, and it is also possible that there were not sufficient numbers of bacteria in the samples. The fact that several structural ribosomal proteins were identified supports this. Further, the biphasic development of the intracellular bacterial pathogen, *Chlamydia trachomatis*, includes a dense form referred to as "elementary body" that had long been assumed to be metabolically dormant, but was recently shown to be as active as *C. trachomatis* RC cells [98]. Nonetheless, several *A. phagocytophilum* outer membrane proteins thought to be involved in mammalian cell invasion, such as OMP-1a and MSP2/P44, as well as stress response proteins were detected in the feeding ticks, and likely represented a subset of those mediating the transition from the tick vector to the mammalian host [97].

It is difficult to identify specifically interacting ligands and receptors, or other proteins and factors involved in infection and development of *A. phagocytophilum* using whole ticks. Our laboratory applied transcriptomics based on whole genome tiling microarrays to identify 41 *A. phagocytophilum* genes that were only expressed during

infection of tick cells but not mammalian cells in vitro [54]. Three of these were known such as *msp4 (Aph_1240)* encoding major surface protein 4, but over 90 % were identified as genes encoding hypothetical proteins, demonstrating how little is known about the biology and the proteins involved in tick-*A. phagocytophilum* interactions [54]. For this study, RNA was harvested from cultures in which ≥ 90 % of cells were infected, representing transcripts from late phases of infection (during intracellular development and exit from cells) [54] and thus genes involved in early phases of infection, such as binding and cell entry, were likely missed. Thus, there is a big gap in knowledge of the factors that are important during tick infection, especially those important during binding and cell entry, which are part of the focus of the studies presented in this dissertation.

A. phagocytophilum effectors

Several pathogens have evolved the ability to hijack the host cell machinery in order to ensure their survival. For example, the pathogen *Legionella pneumophila* possesses 270 effectors that have been verified experimentally and many of which have been shown to modulate host immune response, regulation of cell death, and gene expression [99]. Likewise, *A. phagocytophilum* controls apoptosis, reduces neutrophil phagocytosis and adherence to vascular endothelium, and affects other host cell functions by interfering with signaling pathways and by manipulating gene expression with secreted factors, i.e., effectors [100]. To secrete these effectors bacteria have evolved different secretion systems [reviewed in [48]], outer membrane vesicles [101], and translocases, such as TAT systems (Twin Arginine Translocation) [102]. One such effector is AnkA, a protein that is secreted into the host cell cytoplasm, probably by the

type 4 secretion system (T4SS, encoded by the VirB/D4 locus) of *A. phagocytophilum.* AnkA subsequently translocates to the nucleus to control host innate immune responses by binding diverse partners such as nuclear proteins, ATC-rich sequences, regulatory regions of genes including CYBB (encoding gp91(phox)), and to internucleosomal linkers which could stabilize DNA to interfere with progression of apoptosis [100]. Optimal AnkA delivery depended on engagement of PSGL-1, resulting in phosphorylation of this protein within minutes of attachment to a cell [103]. *A. phagocytophilum* AnkA decreased acetylation of histone 3 (H3) by stimulating the expression of HDAC1 and HDAC2, leading to the hypothesis that AnkA may induce epigenetic changes in the host by altering chromatin structure [104]. This hypotheis is supported by the similarity of AnkA with SATB-1, a protein known to interact with HDAC1, and the fact that AnkA binds to promoters within the H3 deacetylated region [104].

It is thought that *A. phagocytophilum* encodes many other effectors responsible for manipulation of host cell's processes [105]. However, only one additional effector protein has been described in *A. phagocytophilum*, i.e., APH_0859, designated Ats-1 (*Anaplasma* translocated substrate 1). The protein accumulated within the *A. phagocytophylum* inclusion starting one day after infection, and was secreted into the cytoplasm of HL-60 cells 32 hr pi [105], whereupon Ats-1 was transported to the mitochondria. This process was independent of any other bacterial factors as the protein also co-localized with mitochondria when expressed in transfected RF/6A cells. Instead, translocation depended on the presence of an N-terminal sequence to direct it to mitochondria, as its deletion resulted in diffuse distribution within the cytoplasm [105].

However, not all of Ats-1 is directed to mitochondria, but was shown to sequester Beclin1, and subsequently recruited endoplasmic reticulum (ER) proteins to induce formation of autophagosomes that transported nutrients to the *A. phagocytophilum* vacuole in HL-60 cells and RF/6A cells [106].

Both AnkA and Ats-1 are secreted through the T4SS, and although *A. phagocytophilum* also encodes T1SS, Sec, and TAT systems, no substrates that utilize these have been identified (reviewed in [48]). Even though conserved motifs of secreted proteins have not been demonstrated in either AnkA or Ats-1, both have been shown to share characteristics with secreted proteins in other bacterial species such as a net positive charge and characteristic hydropathy profiles [107]. These proteins represent two mechanisms that *A. phagocytophilum* exploits to assure its own survival and extend the life span of its host cell, thus increasing the chances of acquisition by feeding ticks to complete its life cycle. It is very likely that *A. phagocytophilum* possesses many other effector molecules that have not been identified to date. Identification of these effectors and their function during host infection is crucial for understanding the mechanisms of pathogenesis of human anaplasmosis, and for the discovery of alternative treatments.

A. phagocytophilum functional genomics

To date, research on *A. phagocytophilum* biology has relied on proteomics analysis involving procedures such as immune-precipitation, biotinylation of outer membranes, and Western blots. In addition, microarrays and RNA-Seq have shed light on gene expression of the bacteria during replication in various host cells. However, these techniques may overlook proteins that are not as abundant as outer membrane proteins or

genes that are only expressed at specific times during development. For example, expression arrays will likely miss activity of important non-coding regions in the genome, like promoters, enhancers, and silencers that may be important for the pathogenesis of human anaplasmosis. Because of the intracellular nature of *A. phagocytophilum,* functional genomics studies based on mutagenesis are difficult. However, recent progress in transposon mutagenesis of *A. phagocytophilum* using the *himar*1 transposase [108], has opened the possibility of applying random mutagenesis to study gene function in *A. phagocytophilum*. Transposon mutagenesis has led to identification of a gene encoding a dihydrolipoamide dehydrogenase 1 (Ldpa1), APH_0065, that facilitates infection of neutrophils, and stimulates secretion of IFN-γ and macrophage activation during mouse infection. This mutation led to a stronger inflammatory response and greater ROS production in mouse macrophages, but did not affect pathogen load in tissues. This indicated that Ldpa1 was responsible for reducing ROS and pro-inflammatory cytokine production in macrophages, which was shown to play a role in pathogenesis of human anaplasmosis [10]. This was the first report of a mutation of an *A. phagocytophilum* gene that allowed identification of its potential role during infection of and survival in the mammalian host. Since then, the *himar*1 transposase system has been used to generate additional mutants, and here I present my results from experiments to determine the exact function of two genes during infection of human and tick cells.

Methyltransferases and infection

The importance of protein modifications for virulence of bacteria and other pathogens has been a subject of increased interest, especially since it has been discovered that the properties and functions of proteins change depending on the modifications they have undergone [reviewed in [109]. Such modifications include phosphorylation, acetylation, pupylation (modification by a prokaryotic ubiquitin-like protein), carboxylation, glycosylation, deamination, and methylation [reviewed in [109]. During methylation, a methyltransferase adds a methyl moiety obtained most commonly from S-adenosine-methionine (SAM) to a receiving atom [110]. Methyltransferases modify proteins, DNA, RNA, co-factors, xenobiotics, lipids, and other types of molecules in plants and mammalian, however, the importance of this modification in bacterial and viral infection has only been discovered recently [111-114]. Protein methylation occurs in two different forms, *N*-methylation of lysines, arginines, and glutamine and *O*-methylation of glutamic acid [as reviewed in [109], and has been shown to be important in several bacterial processes, including chemotaxis and signal transduction [111], binding to and invasion of host cells [115], and toxin production [116].

Two examples of *O*-methyltransferases are CheR and CheB methyltransferases, a *O*-methyltransferase involved in the modification and activation of chemoreceptors in *E. coli, Salmonella thyphimurium,* and *Pseudomonas putida* to facilitate chemotaxis, as well as biofilm and pili formation [111]. Chemoreceptors and chemotaxis are important for bacterial adaptation to different environments, and bacteria use methylated sensors in their outer membrane to detect external stimuli [111]. Lateral transfer of genes encoding methyltransferases may enable development of pathogenicity, or, bacterial evolution towards pathogenicity may shape methyltransferases in similar ways. Pathogenic strains

of bacteria harbor *O*-methyltransferases that are more similar to those in unrelated pathogenic bacterial species than to those found in non-pathogenic relatives, thus linking the presence of these enzymes with virulence and pathogenesis [117]. *O*-methylation of bacterial adhesins is an important modification, and methylation of an outer protein in *Leptospira interrogans* was necessary for adhesion to and colonization of epithelial cells in hamster kidneys [118]. Similarly, methylation of OmpB is necessary for binding to endothelial cells by *Rickettsia prowazekii* and *Rickettsia typhi,* even though this appears to involve a different kind of methyltransferase [119,120]. However, little is known about the role of methylation in the virulence and pathogenesis of *A. phagocytophilum* or any member of the *Anaplasmataceae* family.

Conclusion

Human Granulocytic Anaplasmosis (HGA) has rapidly become the second most important tick-borne disease, and probably the third most important vector-borne disease, in the USA, and is gaining importance in Europe and Asia. This disease represents an important risk especially for people with immunological deficiencies. Furthermore, this disease not only represents an important burden in human health, but also affects the sheep and cattle industries of several European countries. There is no vaccine, and treatment relies on tetracycline antibiotics. Therefore, it is important to find other possible ways of treatment and prevention of the disease, which can only be obtained through the understanding of the pathogen's biology and life cycle and the molecules that are involved in the infection process. Important advances have been obtained in this regard, but they are minimal when compared to other pathogens, largely because of the

obligately intracellular life of *A. phagocytophilum*. This is especially true in our understanding of the vector-pathogen-host interfaces of the disease. Adaptation of the *himar*1 transposase tool for random mutagenesis of *A. phagocytophilum* may help overcome this lack of understanding. This dissertation describes studies to elucidate the functions of two genes, which were mutated using the *himar*1 transposon tool, in adhesion, invasion and replication of *A. phagocytophilum* in host cells. The two mutant strains are deficient in their abilities to infect either human or tick cells. One of the mutants has insertion into an *O*-methyltransferase gene (*omt*) that prevents the pathogen from binding to and invading tick cells. The analysis presented herein suggests that the OMT methylates an outer membrane protein of *A. phagocytophilum* that mediates adhesion to tick cells. The second mutant has an insertion into a gene encoding a hypothetical protein, APH_0906, which may be an effector that is secreted into the host cytoplasm and translocated into the nucleus. Mutational analysis of gene function in *A. phagocytophilum* is now available for mechanistic studies to determine how this bacterium is able to cycle between tick vectors and mammalian hosts.

Chapter 2

Identification of two genes involved in the infection process of the tick-borne pathogen *Anaplasma phagocytophilum*

Introduction

Anaplasma phagocytophilum is an obligately intracellular bacterium classified in the order Rickettsiales and is the causative agent of Human Granulocytic Anaplasmosis (HGA) [121]. HGA is characterized by high fevers, rigors, generalized myalgias, and severe headache in most cases. Generally, HGA is a potentially life-threatening disease, and 36% of patients diagnosed with HGA require hospitalization, 7% need urgent care, and around 1% die [122]. The incidence of HGA has been increasing steadily, from 1.4 cases per million (348 cases) in 2000 when it first became a reportable disease to the CDC reaching 6.1 cases per million (1761 cases) in 2010 [123]. Similar trends are evident in other countries in Europe and Asia [reviewed in [124]]. In addition, *A. phagocytophilum* strains that infect humans also affect a wide range of domestic and wild mammals from dogs, cats, and horses to different rodents [125]. Several factors including increases in *Ixodes* ticks densities, changes in human recreational practices, global warming, and abundance of wild hosts for the ticks result in higher risk of disease for humans as well as domestic animals [126]. *Anaplasma phagocytophilum* infects neutrophils, granulocytes and their progenitors, resulting in cytotoxic lymphocyte activation and overproduction of cytokines such as IFN-γ and IL-10, which negatively affects the immune system of its host [122,127]. The bacteria reside within granulocyte endosomes that fail to mature to phagosomes by preventing the fusion of lysosomes into the vacuole where it develops , thus avoiding destruction by the host [127].

A. phagocytophilum is transmitted by ticks of the *Ixodes persulcutus* complex, with *I. scapularis* and *I. pacificus* being the most important vectors in the USA [127].

Transovarial transmission does not occur in these ticks, and has only been reported in tick species and *A. phagocytophilum* strains that are not implicated in the transmission cycle affecting humans and domestic animals [89]. The natural transmission cycle involves acquisition of the pathogen from small wild rodents by *I. scapularis* larvae, and transstadial transmission to nymphs and adults that may infect a mammalian host during a subsequent bloodmeal. Therefore, the ability of *A. phagocytophilum* to cycle between ticks and mammalian hosts is imperative for bacterial survival [89]. Survival in dissimilar hosts requires rapid adaptation of the bacterium to environments that present important physiological and other biological differences. This rapid adaptation to different environments most likely involves proteins and other molecules that are differentially expressed or produced during the infection of the tick vector and/or the mammalian host [54].

A. phagocytophilum possesses a small genome of only 1.47 million bases pairs and 1,369 ORFs [128], which is why the organism needs to make efficient use of its genome in order to infect and survive in the mammalian host and its arthropod vector. This includes the expression of specific genes only during infection of a specific host cell type [54] or only at specific phases of infection [51,129]. Using a complete tiling array of the complete *A. phagocytophilum* genome, Nelson *et al.* [54] determined that 117 ORFs were differentially transcribed during infection of HL-60 cells, a promyelocytic leukemia cell line typically used to culture *A. phagocytophilum*, when compared to the expression in ISE6 cells, a tick cell line derived from *I. scapularis* embryos. Of these 117 ORFS, 76 were upregulated during infection of HL-60 cells and 41 were upregulated in ISE6 cells [54]. Some of these upregulated genes have shown evidence of involvement in important

processes during infection of human and tick cells, such as adhesion to host cells and vacuole membrane formation [51,129,130]. However, a high percentage of the upregulated genes encode for hypothetical proteins (54% in HL-60 cells and 93% in ISE6 cells) [54] and their role in the life cycle of *A. phagocytophilum* is still unknown. Proteome and transcriptome analysis of the proteins and genes expressed during the transmission of *A. phagocytophilum* from the salivary glands to mice also detected up to 21 hypothetical proteins (15% of the detected proteins) that have not been characterized and several others that have been characterized during growth in HL-60 cells but their function in the tick vector remains unknown [97]. The high number of upregulated genes that encode proteins with unknown function highlights how little is known about the proteins involved in vector-pathogen interactions and about the pathogenesis of this organism.

Functional genomics in *A. phagocytophilum* has focused on the study of gene expression and proteomics [97]. The identification of the proteins present during a specific phase of infection or during infection of a specific host is important in understanding the biology of the organism. However, it does not reveal the exact function and that these proteins fulfill. The intracellular nature of *A. phagocytophilum*, has made it difficult to study the function of genes involved in intracellular invasion and replication using genetic techniques such as site-directed mutagenesis via homologous recombination. Nevertheless, Felsheim et al. [54] developed a method for random mutagenesis of *A. phagocytophilum* using the Himar1 transposase system to transform *A. phagocytophilum* to express Green Fluorescent Protein (GFPuv) and a red fluorophore,

mCherry [131]. One of these tranformants has resulted in a gene knock-out that affects pathogenesis of *A. phagocytophilum* [10].

Herein, I report the phenotypical changes that have resulted from two gene knock-outs in two different mutants. One of these mutants presents an insertion in *aph_0906* that encodes a hypothetical protein. This gene is highly upregulated during infection of HL-60 cells[54], and in fact the mutation abrogates survival of *A. phagocytophilum* in these cells. A second transformant has an insertion into *aph_0584*, a member of the o-methyltransferase family 3. *aph_0584* does not appear to be expressed during late phases of infection and transcripts are barely detectable in any of the host cells tested [54]. However, the mutant is unable to grow in tick (ISE6) cells, which indicates that the gene product facilitates some aspect of infection in tick cells. In this manuscript I 1) identify and describe the phenotype resulting from mutations in the APH_0906 and the APH_0584 genes with respect to infection of and growth in vitro and in vivo, 2) determine the expression pattern of these genes to identify the exact phase of infection in which they are important during infection of human and tick cells, 3) examine the presence of putative functional motifs in both proteins (APH_0906 and APH_0584) and the similarity of the protein products to other proteins with known function, and 4) establish the effects of a methyltransferase inhibitor on the entry and invasion of tick cells by *A. phagocytophilum*.

Materials and Methods

Identification and growth of the mutant strains in ISE6 or HL-60 cells

Uninfected ISE6 cells were maintained in L-15C300 medium with an osmotic pressure of approximately 315 mOsm/liter [92], supplemented with 5 % fetal bovine serum (FBS), 5 % tryptose phosphate broth (TPB), and 0.1% lipoprotein concentrate (LPC). Fresh medium was supplied once a week and cultures were maintained at 34°C. The *A. phagocytophilum* isolate HGE2 was originally cultured from a human patient in Minnesota [92]. Several transformants were generated as described [54] to express Green Fluorescent Protein (GFPuv) or red fluorescent protein (mCherry). These transformants were generated in both HL-60 and ISE6 cells. The ability of the transformants to growth in the reciprocal cell line in which they were produced (HL-60 to ISE6 or ISE6 to HL-60) was tested by inoculating purified cell-free bacterias as well as whole infected cells into 25 cm^2 flask containing the recipient cell line. The growth and development of the transformants in the reciprocal cell line was evaluated by fluorescent microscopy using an inverted Nikon Diaphot microscope (New York) and giemsa staining. Two transformants were identified as having a defect in their ability to infect and survive within a different cell species from those in which it was originally produced. One of the transformants (designated as the APH0906-mutant), had a transposon insertion into *aph_0906.* The APH0906-mutant was continuously grown in ISE6 cells as described in [54] by passing one tenth of infected culture onto an uninfected cell layer maintained in L-15C300 supplemented with 10 % FBS, 5 % TPB, and 0.1 % LC, and the pH was adjusted to 7.5 - 7.7 with sterile 1N NaOH (referred to herein as *A. phagocytophilum*

medium). Medium was changed twice a week. Medium was changed twice a week. Spectinomycin and streptomycin (100 μg/ml each) were added to the cultures for selection of mutants carrying the inserted *aadA* resistance gene, and cultures were monitored several times a week by fluorescence microscopy to detect *A. phagocytophilum* expressing green fluorescent protein GFPuv [79].

The second mutant was derived from the *A. phagocytophilum* isolate HZ, which was originally cultured from a New York patient [132], and had an insertion into the *aph_0584* gene (designated as the OMT-mutant). The OMT-mutant was cultivated in HL-60 cells maintained in RPMI 1640 medium (Gibco, New York) supplemented with 10% FBS and 2mM glutamine at 37°C with 5% CO_2, in humidified air [92]. Spectinomycin (100 μg/ml) was added to the medium for selection of mutants. The bacteria were subcultured every 5 days by inoculating 3.2 x 10^3 infected HL-60 cells into a new flask containing around 3.6 x 10^5 uninfected cells and fresh medium. Uninfected HL-60 cells were maintained in the same medium and under conditions described for infected cells with the exception that antibiotics were not added. Cultures were monitored weekly for presence of *A. phagocytophilum* expressing GFPuv by epifluorescence microscopy using an inverted Nikon Diaphot (New York) with Sapphire GFP filters as described previously [21].

Determination of insertion sites

The number of insertion sites in each of the mutant populations was determined by Southern blots. DNA was purified from a 25 cm^2 flask of HL-60 cells infected with the OMT-mutant and from a 25 cm^2 flask of ISE6 cells infected with the APH0906-

mutant, using the Puregene Core Kit A (Qiagen, Maryland). An extra phenol/chloroform extraction was performed to remove any remaining protein. The aqueous phase containing the nucleic acids was separated from the organic phase containing the proteins using tubes preloaded with Phase Lock Gel Heavy (5 Prime, Maryland). DNA concentration was measured with a BioPhotometer (Eppendorf, New York). Additionally, DNA was extracted from HZ and HGE2 wildtype bacteria. DNA (100 ηg) from the APH0906-mutant and wild type HGE2 were digested with *Bgl*II, *Hind*III and *EcoR*I. The same amounts of DNA from the OMT-mutant and wild-type HZ were digested with *Bgl*II and *EcoR*V. The samples were then electrophoresed in 1% agarose gels. DNA was transferred and probed as described in [79], using digoxigenin-labeled probes specific for GFPuv prepared with the PCR DIG Probe Synthesis kit (Roche, Indiana). The positive control consisted of 10 pg and 1 pg of the plasmid construct pHIMAR1-UV-SS originally used for the first *A. phagocytophilum* transformation [54].

To determine the insertion sites into each mutant population, 5 µg of DNA from the mutants was digested with *Bgl*II. Digested DNA from the HGE2 mutant was cleaned using a phenol/chloroform extraction as described above for the Southern blots, whereas the HZ mutant was cleaned with DNA clean & concentrator (Zymo Research, California). Cleaned DNA was ligated into the pMOD plasmid and then the construct was electroporated into ElectroMAX DH5α cells (Invitrogen, New York). Bacterial cells containing the transposon were selected on YT plates with 50 µg/ml of spectinomycin and streptomycin. DNA from transformed ElectroMAX DH5α cells was purified by phenol/chloroform extraction and then sequenced at the BioMedical Genomics Center (University of Minnesota).

Effects of the mutations on the growth of A. phagocytophilum as determined by qPCR

The APH0906-mutant and wild-type HGE2 were grown in 25 cm^2 flasks containing ISE6 cells as described above. Bacteria of each isolate were purified from two flasks of infected ISE6 cells when infections reached >90%, and then separated from host cells by passing the cell suspension through a bent 27 Ga needle and filtration of the lysate through a 2 µm pore size filter. Purified bacteria were transferred into two 15 ml tubes and centrifuged at 2000 x g for 5 minutes and then resuspended in 3 ml of RPMI medium supplemented as described above. Five 25 cm^2 flasks containing uninfected HL-60 cells were inoculated with different dilutions of purified bacteria. The dilutions consisted of 0.5 ml of re-suspended bacteria in 20 ml of RPMI medium (1:40), 200 µl of re-suspended bacteria in 20 ml of RPMI medium (1:100) and 50 µl of re-suspended bacteria in 20 ml of RPMI medium (1:400). The experiment was repeated in triplicates. Inoculated cultures were incubated at 37°C as described above for infected HL-60 cells. Samples of 1.5 µl were taken from each dilution everyday for a 5 day period and DNA was extracted from each sample as described before for DNA samples used in Southern blots. The growth of the OMT-mutant in HL-60 cells was determined using the same procedure.

To determine the growth curve of the APH0906-mutant in ISE6 cells, bacteria were purified as described for HL-60 cells but resuspended in L15C300 and inoculated into 12.5 cm^2 flasks containing confluent cell layers of ISE6 cells. The dilutions used were: 0.5 ml of re-suspended bacteria in 2.5 ml of L15C300 medium (1:6), 0.25 ml of re-suspended bacteria in 2.75 ml of L15C300 medium (1:12), and 125 µl of re-suspended bacteria in 2.875 ml of L15C300 medium (1:24). Four flasks were inoculated with each

dilution for each sampling day and the experiment was done in triplicate. The mutant bacteria grown in ISE6 cells were sampled every 3 days for 12 days by taking 1.5 ml of infected cells. DNA was extracted as described before. Likewise, the OMT-mutant was grown in HL-60 cells as described above. Bacteria were purified from four flasks containing 25 ml of an >90 % infected cell suspension. Purified bacteria were divided into two 15 ml tubes and centrifuged at 2000 x g for 5 min. Supernatant was discarded and cell free bacteria were resuspended either in 3 ml of RPMI or L15C300. Bacteria were diluted and growth analyzed as described for the APH0906-mutant.

The number of bacteria per sample was determined by qPCR using the primers msp5 fwd and msp5 rev (Table 1) that amplify a fragment of the single copy number *msp5* gene. qPCR reactions were performed in an Mx3005 pro (Agilent, California) cycler, using Brilliant II SYBR Green Low ROX QPCR Master Mix (Agilent, California) under the following conditions: an initial cycle of 10 minutes at 95°C, 40 cycles of 30 seconds at 95°C, 1 minute at 50°C, and 1 minute at 72°C, and a final cycle of 1 minute at 95°C, 30 seconds at 50°C, and 30 seconds at 95°C. A standard curve was developed using the *msp5* fragment cloned into the pCR4-TOPO vector (Invitrogen, New York).

Effect of the mutation on bacterial binding

OMT-mutant and wildtype *A. phagocytophilum* were purified from HL-60. Host cell-free bacteria purified from one fully infected 25 cm^2 were added to about $2.5x10^5$ ISE6 cells in 50 µl of *A. phagocytophilum* medium in a 0.5 mL centrifuge tube. The bacteria were incubated with host cells for 30 min at room temperature, flicking the tube every 5 minutes to enhance contact between bacteria and cells. The cells were washed

twice with unsupplemented L15C300 and centrifuged at 300 x g for 5 minutes to remove unbound bacteria. The cell pellet was re-suspended in PBS and spun onto microscope slides for 5 minutes at 60 x g, using a Cytospin 4 (Thermo Shandon, Pennsylvania). Slides were fixed by submersion in methanol for 5 minutes and dried at 50°C for 10 minutes. Bound bacteria were labeled using an indirect immuno-fluorescence assay (IFA) with dog polyclonal antibody against *A. phagocytophilum* and FITC-labeled anti-dog secondary antibodies. DAPI was used to stain the host cell DNA and aid in host cell visualization. The number of bacteria bound to 300 cells was counted for each sample. This was repeated in triplicate, and differences were evaluated using Student's t-test to assess significance. A similar experiment was done with the APH0906-mutant. HGE2 wildtype and the APH0906-mutant were grown in ISE6 cells as described, and bacterial adhesion to uninfected HL-60 cells was determined as described for the OMT-mutant and ISE6 cells.

Infection of hamsters with the APH0906-mutant and re-isolation from blood

Nine, 2-week old hamsters were injected intraperitoneally with 500 µl of ISE6 cells infected with the APH0906-mutant. After 7 days, a tail blood smear stained-with Giemsa's stain was examined microscopically to confirm infection. Additionally, two hamsters were euthanized by CO_2 overexposure, and 800 µl of blood was removed by sterile heart puncture using a 25 Ga needle and syringe containing EDTA. 100 µl of blood was inoculated into HL-60 cell culture, 300 µl into ISE6 cell culture, and DNA was extracted from the remaining 400 µl. ISE6 and HL-60 cultures inoculated with the blood were maintained under the same conditions as described for cell cultures infected with

mutant strains, and monitored daily for bacterial growth using an inverted Nikon Diaphot (New York) fitted for fluorescence microscopy and a Sapphire GFP filter as described [133].

The seven remaining hamsters were kept for 21 days, and then blood was extracted after euthanasia as described above, using the Qiagen Puregene Core A kit blood protocol and treated with Proteinase K (Roche, Indiana) overnight at 50°C. DNA was subjected to a phenol/chloroform/isoamyl alcohol extraction to remove Proteinase K, using the Phase Lock Gel 2.0 tubes (5 Prime, Maryland). Nested primer pairs EE1-EE2 and EE3-EE4 were used to PCR-amplify a 928 bp product by the first set of primers and a 867 bp product from *A. phagocytophilum* 16S rDNA by second as described [134] using FlexiTaq polymerase (Promega, Wisconsin). Amplified products were run in agarose gels and visualized using GelGreen™ (Biotium, California).

Growth of APH0906-mutant in RF/6A and HMEC

The ability of the APH0906-mutant to infect rhesus choroid-retina endothelial (RF/6A) cells and Human Microvascular Endothelial (HMEC-1, [135]) cells was tested. RF/6A cells were maintained in *A. phagocytophilum* medium (defined above) at 37°C with 5% CO_2 in humidified air. HMEC-1 cells were maintained in RPMI medium as described for HL-60 cells. RF/6A and HMEC-1 cells were subcultured by trypsinization (Gibco, New York) as follow: cell layers were exposed to 1 ml of trypsin with EDTA for 1 minute at 37°C, trypsin was aspirated, and cell layers incubated for an additional 4 minutes at 37°C. Cells were resuspended in 5 ml of their respective growth medium, and

then diluted 10-fold in fresh growth medium. Both RF/6A and HMEC cells were subcultured every two weeks.

The APH0906-mutant was grown in ISE6 cells as described above. When infection in ISE6 cell cultures reached >90%, the bacteria were purified from one 25 cm^2 flask and inoculated onto confluent RF/6A and HMEC-1 cell layers. Cultures were monitored weekly for presence of green fluorescent *A. phagocytophilum* by epifluorescence as stated above. The levels of infection were monitored weekly for a month in three replicates by diluting 100 µl of trypsinized cell suspension in 900 µl PBS. Samples were spun onto microscope slides as described above fixed in methanol and stained with Giemsa's stain (Gibco, New York). A total of 300 cells were counted and the number of infected cells was determined in each sample.

Bioinformatics-based analysis of proteins encoded by genes identified in the mutants

The amino acid sequence of the hypothetical protein encoded by *aph_0906* was analyzed using motif finder: MOTIF (http://www.genome.jp/tools/motif/), Motif Scan (http://myhits.isb-sib.ch/cgi-bin/motif_scan), and Pfam (http://pfam.sanger.ac.uk/search). PSI-BLAST was performed to exclude any members of the Anaplasmatacea family to find Position Specific Conserved Motifs that were present in proteins of unrelated organisms. To determine if specific regions of the protein presented high similarity to known motifs, regions of 300 amino acids were blasted separately. Additionally, the protein was analyzed to establish if putative DNA binding motifs were present within the complete protein sequence, using BindN (http://bioinfo.ggc.org/bindn/), GYM 2.0 (http://users.cis.fiu.edu/~giri/bioinf/GYM2/welcome.html), and DNABindR

(http://turing.cs.iastate.edu/PredDNA/). To establish if a signal peptide was also present in the protein sequence, the prediction SignalP 4.0 (http://www.cbs.dtu.dk/services/SignalP/) and PrediSi (http://www.predisi.de/) were used. Subcellular of the protein within the bacteria was assessed using the programs CELLO (http://cello.life.nctu.edu.tw/) and Psortb v3.0 (http://www.psort.org/psortb/index.html).

To determine the origin of the *A. phagocytophilum* OMT, a protein sequence available in GenBank (gi|88607321) was used for a PSI-BLAST that excluded members of the Anaplasmataceae family. The OMTs identified as the most similar to the *A. phagocytophilum* OMT were aligned using ClustalW from MacVector 12.0 (MacVector, Inc, North Carolina). A Minimun Evolution phylogenetic tree of all the OMTs was generated using MEGA 4.0 [136]. To look for conserved motifs in similar OMTs, the protein sequences from the four closest, non-Anaplasmataceae OMTs along with the *A. phagocytophilum* OMT were input into MEME (http://meme.nbcr.net). Phyre^2 (http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index) was used to determine the putative tertiary structure and also to identify OMTs in other bacteria with a similar structure.

Relative expression of APH_0906 and OMT

A. phagocytophilum HGE2 wild-type was purified from ISE6 cells and inoculated into four 25 cm^2 flasks containing either uninfected ISE6 or HL-60 cells. Bacteria inoculated into ISE6 cultures were incubated at 34°C for 4, 24, 48, or 72 hours. Total RNA from bacteria and host was extracted from whole infected cell cultures, using the

Absolutely RNA Miniprep Kit (Agilent, California) according to the manufacturer's specifications. Likewise, cell-free, wild-type *A. phagocytophilum* strain HZ was purified from HL-60 cells and inoculated into four 25 cm^2 flasks containing either uninfected ISE6 or HL-60 cultures. Bacteria inoculated into HL-60 cells were incubated at 37°C for 30, 60, 120, or 240 minutes and RNA was extracted as described above. RNA was DNAse treated with 0.5 units TURBO DNase (Ambion, New York), and incubated at 37°C for 30 minutes. The DNAse treatment was repeated once. RNA concentrations were measured using a Biophotometer as described for DNA.

Expression of *aph_0906* was measured in extracts from the HGE2 strain and expression of the OMT was determined from HZ strain extracts. Gene expression was normalized against expression of the *rpoB* and *msp5* genes that had been found to be significantly expressed in both cell types during tiling array analysis [54]. qRT-PCR reactions were carried out using Brilliant II QRT-PCR SYBR Green Low ROX Master Mix (Agilent, California) using primers for the *omt*, *aph_0906, msp5,* and *rpoB* described in Table 1. The reactions were run as follow: one cycle at 50°C for 30 min, one denaturing cycle at 95°C for 10 min, 40 cycles that consisted of 30 secs at 95°C, 1 min at 50°C, and 1 min at 72°C, and a final cycle of 1 min at 95°C and 30 secs at 50°C. Ct values were established during amplification and the dissociation curve was determined during the final denaturation cycle. Expression of the APH_0906 and the *omt* genes was analyzed using the $2^{-\Delta\Delta ct}$ method, and significant differences in expression of the genes were determined using Student's t-test with SigmaStat (Systat Software, California).

Effects of the Inhibition of SAM dependent Methyltransferases on the binding of A. phagocytophilum to ISE6 cells

To confirm the effect of the *omt* mutation on binding, adenosine dialdehyde (adenosine periodate, or AdOx), which inhibits S-Adenosine Methionine (SAM) dependent methyltransferases by increasing the concentration of S-adenosyl-L-homocysteine [114], was added to wild-type cultures to abrogate o-methyltransferase activity. Wild-type and OMT-mutant bacteria were purified from 75 cm^2 flasks containing infected HL-60 cells as previously described. Purified wild-type bacteria were incubated with AdOx at 20 ηM, 30 ηM, and 40 ηM final concentration for 1 hr at 34°C before adding them to $2x10^5$ uninfected ISE6 cells. Controls consisted of wild-type and OMT-mutant bacteria purified the same way, but incubated at 34°C for 1 hr without addition AdOx.

The bacteria and ISE6 cells were incubated in 200 µl of *A. phagocytophilum* medium at 34°C for an additional hr to allow binding. Cells were washed 3 times with *A. phagocytophilum* medium by centrifugation at 600 xg to remove unbound bacteria. After the final wash, cells were resuspended in 1 ml of *A. phagocytophilum* medium and 50 µl of the suspension was spun onto microscope slides using a Cytospin 4 centrifuge as described above. Cells were fixed in methanol for 10 min. The bacteria were incubated with dog anti *A. phagocytophilum* serum and labeled with FITC conjugated anti-dog antibodies. To aid in host cell visualization, nuclei were stained with DAPI. Bacteria were observed using a 100x oil immersion objective on a Nikon Eclipse E400 microscope. Bacteria were counted as described above for regular binding assays, and

differences in the number of bacteria per cell were evaluated using the Student-Newman-Keuls one-way ANOVA on ranks, using SigmaStat.

Results

Disruption of omt and aph_0906 genes results in phenotypic changes

Analysis of the phenotypic changes in *A. phagocytophilum* transformants was done by testing their ability to infect the reciprocal host-cell type in which they were propagated, e.g. transformants generated in ISE6 cells were tested in HL-60 cells and *vice versa*. Of several transformants, two were identified as unable to infect either ISE6 cells or HL-60 cells. The OMT-mutant, generated and maintained in HL-60 cells, was unable to infect and grow in ISE6 cells. In contrast, the APH0906-mutant was generated and maintained in ISE6 cells and was unable to grow in HL-60 cells. Both mutants express Green Fluorescent Protein (GFPuv) from a Himar1 transposon [54] and Southern blot analysis identified a single insertion site within each genome (Figure 1A and B). Digestion of *omt*-mutant DNA with *EcoR*V showed several smaller bands that are believed to be artifacts resulting from nonspecific digestion since digestion with *Bgl*II showed a single band, suggesting a clonal population (Figure 1B).

Recovery of the transposon along with flanking sequences from OMT-mutant DNA by restriction enzyme digestion and cloning indicated transposition into APH_0584 (Gene ID: 3930223; o-methyltransferase 3 family member) occurred between nucleotide positions 612707 - 612706 of the HZ genome sequence ([128]; Figure 2A). The second mutant, obtained from the HGE2 isolate, resulted from transposition into the coding

region of the hypothetical protein gene *aph_0906* (Gene ID: 3930051) at nucleotide positions 965962 – 965963 (Figure 2B) of the HZ genome (the HGE2 genome has not been sequenced). Several *E. coli* colonies from each mutant were examined and all revealed transposition into the same location, supporting a single insertion event. The single insertion event in both strains suggests that the changes in phenotype are due to the knockout of those particular genes.

Mutation of the OMT and APH_0906 genes reduces A. phagocytophilum binding, growth, and survival

To determine the mechanisms by which the mutations affected the phenotype of *A. phagocytophilum,* I analyzed several parameters. I developed growth curves of the mutant strains in HL-60 and ISE6 cells using qPCR. The growth curves from the mutant strains were compared to the growth curves from wild-type bacteria grown under the same conditions. The APH0906-mutant persisted in HL-60 cells for up to 5 days post-inoculation without a reduction in bacterial numbers (Figure 3A). This trend was observed in all three dilutions and was significantly different (P=0.004) from the growth of wild-type bacteria in HL-60 cells (Figure 3B), which showed a steady increase throughout the 5 days of the experiment. On the other hand, growth of the APH0906-mutant in ISE6 cells was not impaired (Figure 3F) and was not statistically different (P=0.734) from the growth of wild-type HGE2 bacteria (Figure 3C), demonstrating that the mutation only affected the growth of the mutant in HL-60 cells.

By comparison, the OMT-mutant was not able to grow in ISE6 cells and *msp5* gene copy numbers decreased, resulting in a negative slope in its growth curve from all of

the dilutions tested (Figure 3D). Degradation of the bacteria occurred early during infection, starting at day 1 and continuing during the 12 days of the experiment. Gene copy numbers of the OMT-mutant were statistically different from those of wild-type bacteria in ISE6 cells (P=0.008). This is in contrast to their growth in HL-60 cells, in which the OMT-mutant bacteria are able to infect and multiply irrespective of the knock-out of the gene. This was demonstrated by the positive slope in their growth curve when replicating in HL-60 cells (Figure 3E). When compared to wild-type bacteria grown in HL-60 cells (Figure 3B), the mutant did not present any significant difference in growth (P=0.504).

Because of the defects in growth for both of the mutants present at the first 24 hrs of incubation with the host cells, I tested the ability of the mutants to bind to ISE6 or HL-60 cells during incubation for 30 min at 34°C or 37°C, depending on the host cell line. The OMT-mutant presented a significant reduction in binding (t-value= -4.1011; P= 0.0001) to ISE6 cells from 0.3 bacteria per cell in the wild type to 0.12 mutant bacteria per cell (Figure 4A), which amounts to a >50 % reduction in binding. The APH0906-mutant, on the other hand, did not present such a marked reduction in the binding, with an average of 0.8 bound bacteria per HL-60 cell compared to the wild-type, in which I observed 1.2 bacteria bound/cell (Figure 4B). This difference in binding represents ~30 % reduction from the wild-type, which is statistically significant with a t-value of 3.901 and P=0.008. Thus mutations had a negative effect on binding, however the *omt* knock-out showed a stronger effect.

Development in HMEC and RF/6A cells is impaired, but the 0906-mutant is able to infect hamsters

Since the APH0906-mutant is unable to productively infect HL-60 cells, we tested its capacity to infect mammals. Inoculation of hamsters with the APH0906-mutant and amplification of an 867 bp fragment of *A. phagocytophilum* 16s rDNA from blood by nested PCR, demonstrated that 7 out of 9 hamsters (infection rate of 78 %) became infected, and 5 of them remained infected for 21 days (Figure 5). Blood samples from mice injected with uninfected ISE6 cells did not show amplification of any product (negative control). Blood smears from Hamster 2 also contained morulae within neutrophils (data not shown). Additionally, two blood samples from PCR positive hamsters (Hamster 1 and 2) were inoculated into two different cultures of ISE6 cells and maintained under the same conditions as described for infected cells. After 5 weeks of incubation, only one culture was infected. The APH0906-mutant isolated from the hamster did not show any difference in development when compared to the mutant bacteria that were not passed through a rodent (data not shown). These results demonstrated the ability of the APH0906-mutant to infect rodents.

To test if this mutant was able to infect human and non-human primate cells, HMEC-1 and RF/6A cells were inoculated with cell-free APH0906-mutant bacteria. The bacteria were able to bind to and probably enter into both cell types, as a large number of individual bacteria were observed attached to both RF/6A and HMEC cells (Figure 6A and E). However, the number of bacteria appeared to decrease with time as fewer bacteria were detected by day 6 (Figure 6B and F) and day 10 (Figure 6C and G). However, more bacteria were present in RF/6A cells (Figure 6B and C) than observed in HMEC cells

(Figure 6F and G) on both day 6 and 10, suggesting a more rapid decline in the human endothelial cells. Likewise, at day 14, more bacterial colonies were observed in RF/6A cells (Figure 6D) than in HMEC cells (Figure 6H). Nevertheless, the mutant colonies in both cell lines were very small and contained few individual bacteria when compared to the wild-type strain (data not shown).

OMT gene (APH_0584) and APH_0906 are differentially expressed in ISE6 and HL-60 cells

Because of the OMT-mutant defects on infection during the first 24 hrs and the deficiency on binding detected after 30 min incubation with ISE6 cells, I tested the expression of the *omt* gene during early stages of *A. phagocytophilum* wild-type development in ISE6 cells. The results showed that at 30 min to 1 hr pi expression of the *omt* gene is not significantly upregulated in ISE6 (average 0.4 and 0.3 *omt/rpoB* ratio of expression, respectively) (Figure 7A). However, as infection of ISE6 cells progresses, OMT expression is upregulated and by 2 hours, the ratio of expression (*omt/rpoB*) reaches an average of 10.3 (Figure 7A), which corresponds to ~30 fold increase in the expression in ISE6 cells (Figure 7B). Subsequently it decreases to a 0.7 ratio of expression by 4 hr pi (Figure 7A), corresponding to only 1.2 fold upregulation in ISE6 cells (Figure 7B). According to the statistical analysis, the differences in expression of the *omt* at 2 and 4 hrs were significant (P=0.01). Two and four hours correspond to the time in which the bacteria are binding and entering into tick cells [92], which suggests the importance of methylation during the binding and internalization of *A. phagocytophilum*. By contrast, *omt* gene expression in HL-60 cells in which I measured ratios of expression

(*omt/rpoB*) at 0.2, 0.2, 0.4, and 0.4 at 30 min, 1 hr, 2 hr, and 4 hr, respectively (Figure 7A), did not significantly change throughout the time points tested in this study. However, it is important to point out that even though the differences in ratios between HL-60 cells and ISE6 at 30 min and 1 hour was not statistically, there was still an upregulation in the gene at these time points as it is shown by the fold change of 3 and 4 (Figure 7B). The steady increase in regulation and in the fold change probably represents the progression in the binding and interaction of the bacteria with ISE6 cells.

Expression of *aph_0906* in ISE6 vs HL-60 cells was compared as well. However, because of the results from the growth curve and binding assays which showed that the bacteria had a reduction in binding but were able to remain in the cells without replicating, early time points during the infection in HL-60 and ISE6 were not tested. Unlike the *omt* gene, the *aph_0906* gene was shown to be expressed in both cell lines, but at different time points. At 4 hours pi, the expression ratio (*aph_0906/rpoB*) in ISE6 was 0.15 whereas in HL-60 cells the ratio was 0.004 (Figure 7C). Expression of *aph_0906* in ISE6 was reduced by 24 hr and remained low by 48 and 72 hr (ratios of 0.009, 0.032, and 0.016 (*aph_0906/rpoB*), respectively) (Figure 7C), whereas it was upregulated in HL-60 cells with ratios of 0.0009, 0.013, and 0.15 (*aph_0906/rpo*B) at 24, 48, and 72 hr pi (Figure 7C). These ratios represent fold changes of 0.03, 0.09, 0.38, and an increase in the fold change of 9.5 in HL-60 cells compared to ISE6 (Figure 7D), which demonstrates the upregulation of the gene at late time points.

In-silico characterization of APH_0906 and the OMT

Analysis of the putative motifs present in the hypothetical protein APH_0906 indicated the presence of several DNA binding residues in the translated protein sequence (Figure 8A). These residues were detected using BindN and DNABindR. The second half of the protein presented long stretches of amino acids with potential DNA binding properties starting around amino acid 1000 until the end of the protein at aa 1528 (Figure 8A), specifically the sequence KPPRPARRGAKSSSAHSVAGV that was determined to have DNA binding properties by all 3 programs used (Figure 8B). Additionally, the first half of the protein contained a section of around 100 aa that showed high similarity (29 - 33%) to members of the antirepressors BRO-N family, according to the BLAST results, with a high percentage of residues that represent a synonymous change as well (46 – 50% positive residues) (Figure 8C). Many of the conserved aa residues are predicted to be DNA binding residues (Figure 8B). Because of the presence of DNA binding motifs in the protein, we examined the APH_0906 protein sequence for the presence of putative signal peptides. However, neither SignalP 4.0 nor PrediSi detected signal peptides in the sequence (data not shown) and both localization prediction programs, CELLO and Psortb 3.0v, suggested that APH_0906 is an outer membrane protein (Table 2).

To try to understand the possible relationship of the *A. phagocytophilum* OMT with other members of this family of enzymes, PSI-BLAST was used to search for similar OMTs in other organisms. Among *α-proteobacteria,* only members of the family *Anaplasmataceae* and *Candidatus* Midichloria mitochondrii (order *Rickettsiales*) encoded OMTs with close relationship to *A. phagocytophilum* OMT (Figure 9A). Other OMTs that are closely related to *A. phagocytophilum* OMT occurred in the Δ-proteobacteria

(Figure 9A). Furthermore, when the motif sites were compared using MEME, the Δ-proteobacteria OMTs appeared to be more similar to *A. phagocytophilum* OMT than the *Candidatus* Midichloria mitochondrii OMT (Figure 9B). Motif 1 presented 48% identity and 66% similarity to *Haliangium ochraceum* OMT, whereas *C.* M. mitochondrii OMT had 33% identity and 53% similarity. Motif 2 exhibited 60% identity and 74% similarity with *Bdellovidrio bacteriovorus* OMT compared to 52% and 67% with *C.*M. mitochondrii OMT, and Motif 3 showed 44% identity and 72% similarity to *B. bacteriovorus* OMT compared to 33% identity and 67% similarity with *C.* M. mitochondrii OMT (Figure 9B). These results corroborated the scores obtained with PSI-BLAST (Table 3). The predicted secondary and tertiary structure of the protein presented high resemblance with other bacterial OMTs family 3 members, showing α-helices surrounding a complex of β-strands (Figure 9C) [110].

Adenosine-peroxidase (AdOx) reduces the binding of wild-type bacteria

An inhibitor of SAM-dependent Methyltransferases was used to confirm the effects of methylation on binding to and infection of ISE6 cells by *A. phagocytophilum.* Wild-type bacteria were incubated with different concentrations (20 ηM, 30 ηM, and 40 ηM) of adenosine periodate (AdOx) for 1 hour. Binding of AdOx pre-treated wildtype bacteria to ISE6 cells was evaluated after allowing bacteria to attach for 1 hour and was compared to untreated bacteria. All concentrations of AdOx affected the ability of *A. phagocytophilum* to bind to ISE6 cells significantly ($P<0.001$) (Figure 10A). In controls, an average of 0.529 bacteria bound per cell, compared to the OMT mutant with only 0.156 bacteria per cell. 20 ηM Adox decreased binding to 0.307 bacteria per cell, which

was less than the reduction in attachment observed in the OMT mutant (Figure 10A and B). However, as the concentration of Adox increased to 30 ηM and 40 ηM, the effects were stronger than in the OMT mutant with only 0.093 and 0.086 bacteria bound per cell, respectively (Figure 10A and B). These results corroborate the effects that the inhibition of methylation had in *A. phagocytophilum* binding.

Discussion

In this manuscript, I report for the first time the effects of the mutations in specific genes on the ability of *A. phagocytophilum* to infect and grow in tick or mammalian host cells. Previously, Chen *et al*. [10] described an *A. phagocytophilum* mutant with an insertion in the dihydrolipoamide dehydrogenase 1 (*lpda1*) gene at the APH_0065 locus. This mutation affected the inflammatory response during infection of mice by increasing the production of reactive oxygen species from NADPH oxidase and NF-κB *in vitro* [10]. However, mutations with a direct effect on growth and infection *in vitro* have not been reported until now.

The analysis of several mutants led to the discovery of two that had deficient growth in either ISE6 cells or HL-60 cells. The first mutant had an insert into the APH_0584 gene which encodes an o-methyltransferase (OMT) family 3 (GI: 88598384). The knock-out of this gene diminished the ability of the bacteria to infect and replicate within ISE6 cells, but had no effect on the growth and infection of HL-60 cells or mice (data not shown). This suggested that the mutation affected a process required for invasion of tick cells. Analysis of the development of the mutant in ISE6 cells demonstrated a rapid decrease in bacterial numbers as early as 1 day pi (Figure 3D), which suggested an effect during early stages of infection. Indeed, the mutation reduced

the binding of *A. phagocytophilum* to ISE6 cells by over 50% compared to the wild-type bacteria. I considered that the lack of in methylation of a residue important for adhesion to and invasion of tick cells was responsible for the reduction in binding and internalization. This would explain the rapid decrease in bacterial numbers as *A. phagocytophilum* cannot survive extracellularly for long periods of time. To confirm that the decrease in binding is resulted from the defective methylation, an indirect inhibitor of methylation (Adox) was used to decrease the number of wild-type bacteria bound to ISE6 cells 5-fold over the untreated control (Figure 10A). Even, in the lowest concentration of Adox reduced the binding of *A. phagocytophilum* to ISE6 cells by around 30%, suggesting that methyltransferases played a role in the binding process. Methylation proteins involved in the adhesion to host cells has been reported in other members of the Rickettsiales [115,137]. Methylation of *Rickettsia prowazekii* OmpB was shown to be important for adhesion to and infection of endothelial cells and for virulence [137], and involved two different lysine methyltransferases [115]. These authors discovered a mutation in a hypothetical methyltrasnferase gene in the Madrid E strain and that pathogenesis was restored in the *R. prowazekii* Evir (derived from Madrid E) after the methyltransferase was restored [138]. In *V. vulnificus,* mutation of a type IV leader peptidase/N-methyltransferase resulted in lower pathogenicity and its inability to adhere to epithelial cells *in vitro* [139]. There is evidence that methyltransferases play a role in viral infectivity as well. Incubation of HIV-1 with Adox resulted in a reduction in viral infectivity, likely due to inhibition of methylation of one or more proteins [114], similar to the effect this methylation inhibitor had on binding of *A. phagocytophilum* to ISE6 cells.

It is interesting that the search for OMTs similar to that encoded by *aph_0584* of *A. phagocytophilum* only identified an OMT in one *Rickettsiales* organism outside the family *Anaplasmataceae,* i.e., in *C.* M. mitochondrii. *C.* M. mitochondrii is an intracellular organism that develops in the mitochondria of *I. ricinus* ticks, and is a member of a new family suggested to be most closely related to the *Anaplasmataceae* in the *Rickettsiales* [140]. No members of the OMT family 3 were found in any *Rickettsiaceae,* and the only enzyme to show slight similarities was a bifunctional methyltransferase (m7G46) present in some *Rickettsia* species, albeit with really high e-values (0.28 – 1.2) (data not shown). This suggested that these OMTs family 3 members existing in *Anaplasmataceae* and the new family "*Candidatus* Midichloriaceae" are not required by other members of the *Rickettsiales,* or that their function is carried out by a different methyltransferase.

Interestingly, the PSI-BLAST search assigned a better e-value (4e-40) to an OMT from the pathogen *B. bacteriovorus* than the e-value obtained for *C.* M. mitochondrii (2e-36) OMT, suggesting that that the former enzyme is more similar to *A. phagocytophilum* OMT. Also, according to structural analyses, small molecule methyltransferases (SMMT), such as OMTs, posses N-terminal extensions, discrete domains and active sites that are critical for substrate binding and the positioning of the methyl accepting atom [110]. It is interesting to point out that motif 3, which correlates with the N- terminal of the *A. phagocytophilum* OMT, is involved in substrate binding in 900 OMTs from studied plants [110], and is more similar to the OMT from *Δ-proteobacterium B. bacteriovorus* than to that from *C.* M. mitochondrii. The higher similarities between *B. bacteriovorus* OMT and *A. phacytophilum* OMT were also apparent from the PSI-

BLAST search results (Table 3). *B. bacteriovorus* is a predatory bacterium that attacks gram-negative bacteria and, like *A. phagocytophilum,* presents a biphasic life cycle with an "attack form" that attaches to the host cell and a "dividing form" that occurs only in the periplasm of its host [141] Like *A. phagocytophilum, B. bacteriovorus* develop in a compartment formed by host and bacterial proteins, but in the periplasm of the bacteria instead of the cytoplasm of cells, and differentially express genes depending on their environment [141,142]. *B. bacteriovorus* possesses an unknown binding protein that interacts with the outer membrane of its prey, and encodes two OMTs, one that is up-regulated during growth phases and another during early phases of invasion [142]. Thus, it is possible that an ancestor of the families *Anaplasmataceae* and "*Candidatus* Midichloriaceae" obtained their OMTs from members of the *Δ-proteobacteria*.

The second *A. phagocytophilum* mutant with a transposition event into the APH_0906 locus displayed a significant reduction in binding to HL-60 cells, and was unable to productively infect HL-60 cells. According to the growth curves, replication of the bacteria appeared to be arrested early during infection as bacterial numbers did not change during 5 days pi when >90% of the HL-60 cells in a culture became infected by wild-type bacteria [65]. Similarly, though the mutant was able to infect non-human primate RF/6A cells for up to 10 days, the number of bacteria did not significantly change within the 14 days of cultivation, suggesting impaired replication. Similar results were observed in the human endothelial cell line HMEC-1 in which fluorescent bacteria declined as observed by microscopy (Figure 5). However, the mutation of APH_0906 did not affect the ability of the bacteria to infect rodent hosts as mutants were able to infect and persist in 78 % of the hamsters infected in this study. It would be interesting to

analyze the immune response of the host to observe the potential use of the mutant as a vaccine strain.

Bioinformatic analysis of the hypothetical protein encoded by APH_0906 identified putative DNA binding sites within several parts of the sequence. A BLAST search revealed similarities of the first half of the protein with BRO-like protein domains from *Lactobacillus* phage Lv-1 and BRO *Enterococcus faecium* Com15 (YP_002455816 and ZP_05678756, respectively). BRO-proteins are also found in *Bombyx mori* nucleopolyhedrovirus that infect insects in the order *Lepidoptera* as well as in several other dsDNA viruses of insects [143] and in bacteriophages and transposable elements in bacteria [144]. In infected cells, BRO-proteins are exported into the nucleus and the cytoplasm [143]. The first 100 – 150 aa in the N-terminus of the BRO-like proteins are conserved and have been shown to bind nucleic acids. Related coding sequences have been found in viruses, bacteriophages or prophages integrated into bacterial genomes, including that of *Yersinia pestis* [144]. This region and the region from amino acids 120 – 220 in the protein sequence of APH_0906 show high similarity. According to our PSI-BLAST search results, the only other proteins with similarity to APH_0906 are hypothetical proteins with unknown function in *A. central* and *A. marginale*. Because of this, hypotheses of the potential function of APH_0906 are difficult and more studies are necessary.

This manuscript represents the first report of random mutagenesis of two *A. phagocytophilum* genes to which we were able to assign a role in the infection process of tick or mammalian host cells based on their mutant phenotype. Although the exact function of both proteins remain to be elucidated, evidence presented here identifies

APH_0584 as a member of the o-methyltransferase family 3 that methylates *A. phagocytophilum* MSP4 and possibly other targets, and APH_0906 as a putative DNA-binding protein with evidence that it may be secreted, and transported into the nucleus or cytoplasm to modulate host responses. Biochemical and mechanistic studies to determine their exact function will be the focus of future studies.

Figure and Tables

Table 1. Primers used for qPRC or qRT-PCR to quantify genes or expression of genes of the pathogen *A. phagocytophilum*

Gene target	Primer identifier	Sequence	Piece size
Major Surface Protein 5 (Msp5)	*Msp5 fw* *Msp5 rev*	5'-GCTTATGTTCGGCTTTTCTTC-3' 5'-CCTTCACTATCAGATTCACCTACG- 3'	320 bp
RNA polymerase subunit B (rpoB)	*rpoB fw* *rpoB rev*	5'-CTTTATCCTGCTTTAGAACAACATC-3' 5'-GGTCCGTATGGTCTGGTTACTC-3'	286 bp
o-methyltransferase family 3 (omt)	*Omt fw* *Omt rev*	5'-CAGTTGGGTGCTCATCAAATAC-3' 5'-GCAAACAAATCCTCATACCTTG-3'	121 bp
APH_0906	*Aph_0906 fw* *Aph_0906 rev*	5'-AGCTCATATGCACAACTATG-3' 5'-AAACGGACACTCTCTATGTT-3'	247 bp

Figure 1. Southern blots showing the insertion events within the OMT and APH_0906 mutants.

A) The APH0906-mutant digested with *Bgl*II, *EcorR*I, and *Hind*III presented a single band (as shown by white arrows) when labeled with UVSS probe that binds to *GFPuv* gene sequence. Wildtype HGE2, also digested with same enzymes, did not show any band, ruling out the possibility of nonspecific binding. B) The OMT-mutant was digested using *Bgl*II and *EcoR*V. *EcoR*V digestion of the mutant presents several lighter bands, however these are probably artifacts from the digestion since *Bgl*II digestion presents a single band. Wild-type HZ DNA digested with *Bgl*II was used as negative control. Bands representing real insertion events are indicated with black arrows.

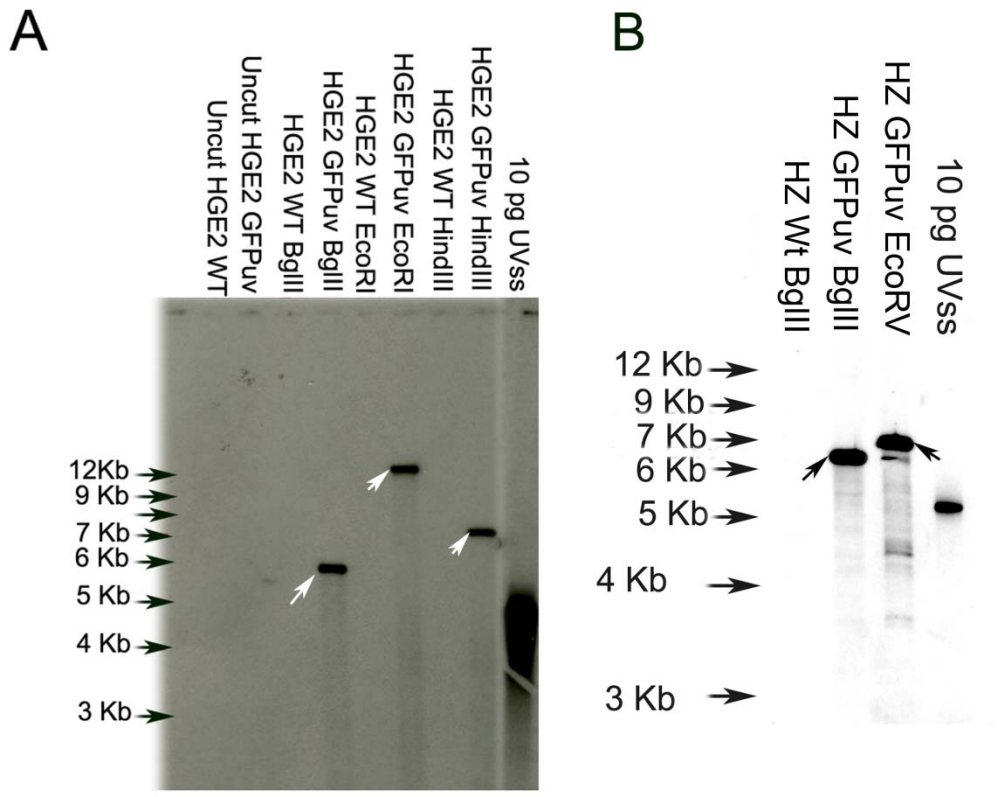


Figure 2. Insertion sites in A. phagocytophilum mutant strains that present phenotypic changes.

A) Insertion site of the construct containing the GFPuv protein gene (light green), Spectinomycin resistance gene (blue), and the Am tr promoter (red) within the himar repeats (dark red). The coding region of the O-methyltransferase family member 3 gene *aph_0584* (dark yellow) was interrupted at positions 612707 – 612706. B) Insertion site of the construct containing the GFPuv protein gene (dark green), Spectinomycin resistance gene (blue), and the Am tr promoter (red) within the himar repeats (dark red). The coding region of the hypothetical protein APH_0906 (light green) was interrupted at positions 965962 – 965963.

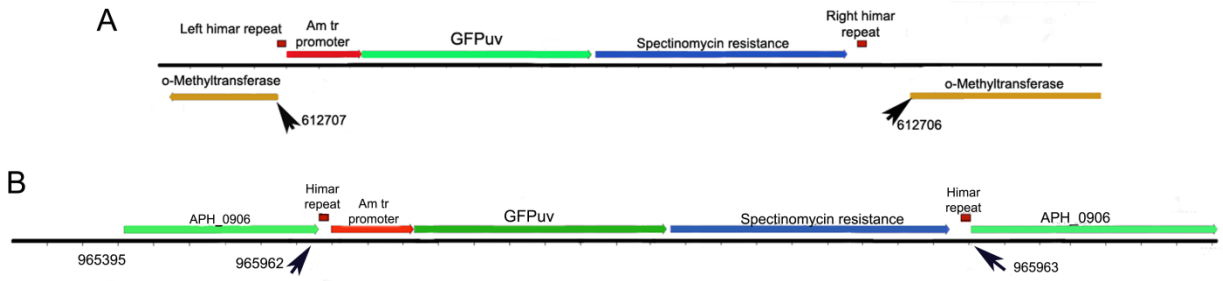


Figure 3. Effects of the mutation on the growth of A. phagocytophilum during infection of tick and human cells.

Growth curves representing the development of (A) the APH0906-mutant in HL-60 cells, (B) wild-type HZbacteria in HL-60 cells, (C) wild-type HGE2 in ISE6 cells, (D) the OMT-mutant in ISE6 cells, (E) the OMT-mutant in HL-60 cells, and (F) the APH0906-mutant in ISE6 cells. Each point in the graph represents the average number of bacteria from triplicates. Each line color corresponds to the dilution used to inoculate the uninfected cells. Red stands for the dilution 1:6 (ISE6 cells) or 1:40 (HL-60 cells), green represents the dilutions 1:12 (ISE6 cells) or 1:100 (HL-60), and purple stands for 1:24 (ISE6 cells) or 1:400 (HL-60 cells). The linear bars in each time point represent the standard error. The Y axis is shown in Log_2 to represent growth.

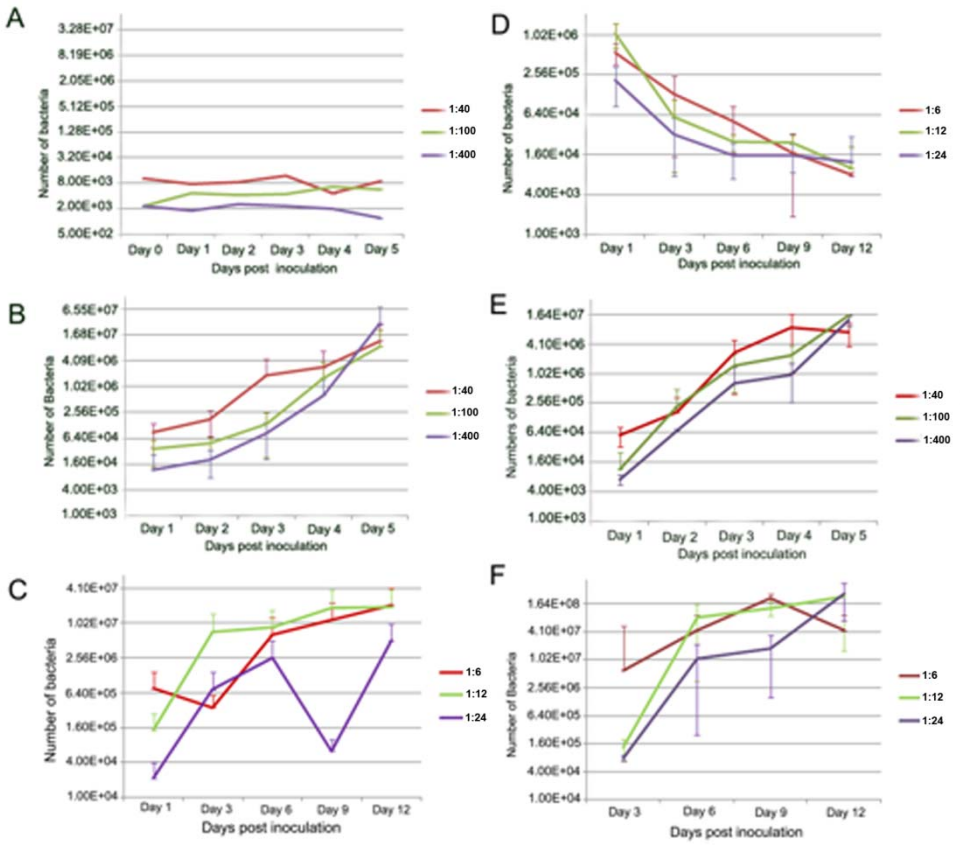


Figure 4. Reduction in the binding of A. phagocytophilum mutant strains to tick or human cells.

A) Average of the HZ bacteria bound per ISE6 cell. Red bar represents the average OMT-mutant bacteria bound to ISE6 cells, whereas the yellow bar represents the average of wildtype HZ bacteria bound. B) Average of the HGE2 bacteria bound per HL-60 cell. Red bar represents the average APH0906-mutant bacteria bound to HL-60 cells, whereas the yellow bar represents the average of wild type HGE2 bacteria bound. The lines above each bar represent the standard error.

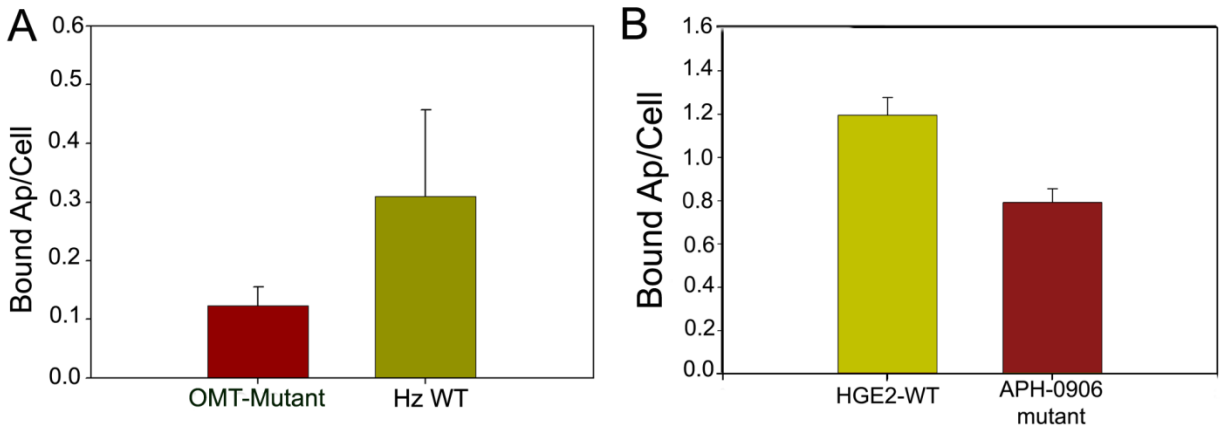


Figure 5. Experimental infection of hamsters with the APH0906-mutant

Negative control consisted of blood from hamsters injected with uninfected ISE6 cells (lanes A), positive control consisted of DNA from APH0906-mutant grown in ISE6 cells (lanes B), lanes 1 – 9 represent the PCR products from each of blood sample taken from the nine inoculated hamsters. Lanes 1 and 2 show the products from the blood samples taken after the first week of infection. The remaining lanes, 3 – 9, show the products from blood samples taken 21 days p.i. The 100 bp ladder (New England Biolabs, Massachusetts) is shown in the first lane of each gel. Molecular markers from 500 bp and 900 bp are indicated by the arrows.

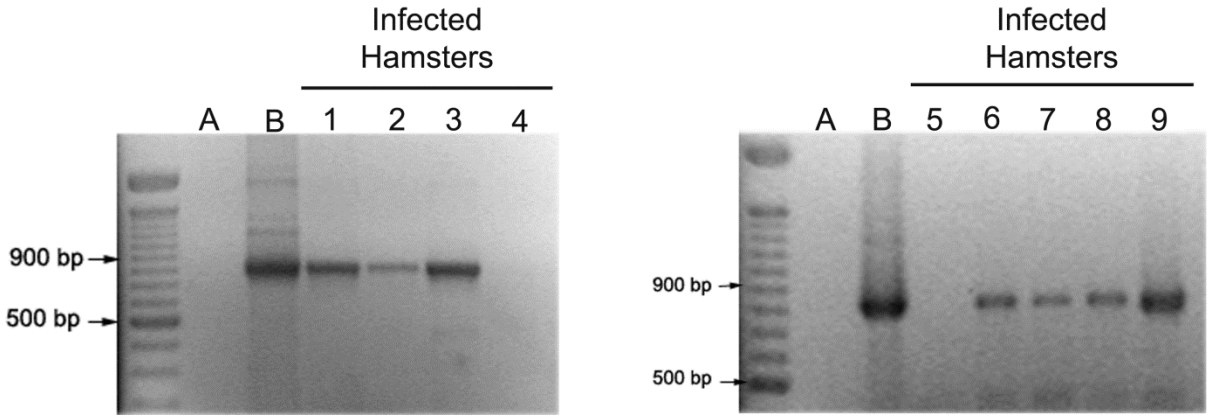


Figure 6. Development of the APH0906-mutant in RF/6A and HMEC cells

Fluorescence microscopy with UV light and Sapphire filter images of the APH0906-mutant during infection of RF/6A cells at day 3 (A) and day 10 (B). C) Cytospin showing a Giemsa staining preparation of RF/6A cells infected with the APH0906-mutant at day 14. Fluorescence microscopy with UV light and Sapphire filter images of the APH0906-mutant during infection of HMEC cells at day 3 (D) and day 10 (E). F) Cytospin showing a Giemsa staining preparation of HMEC cells infected with the 0906-mutant at day 14. APH0906-mutant in fluorescent microscopy images is shown by yellow arrows and by black arrows in Giemsa stain images.

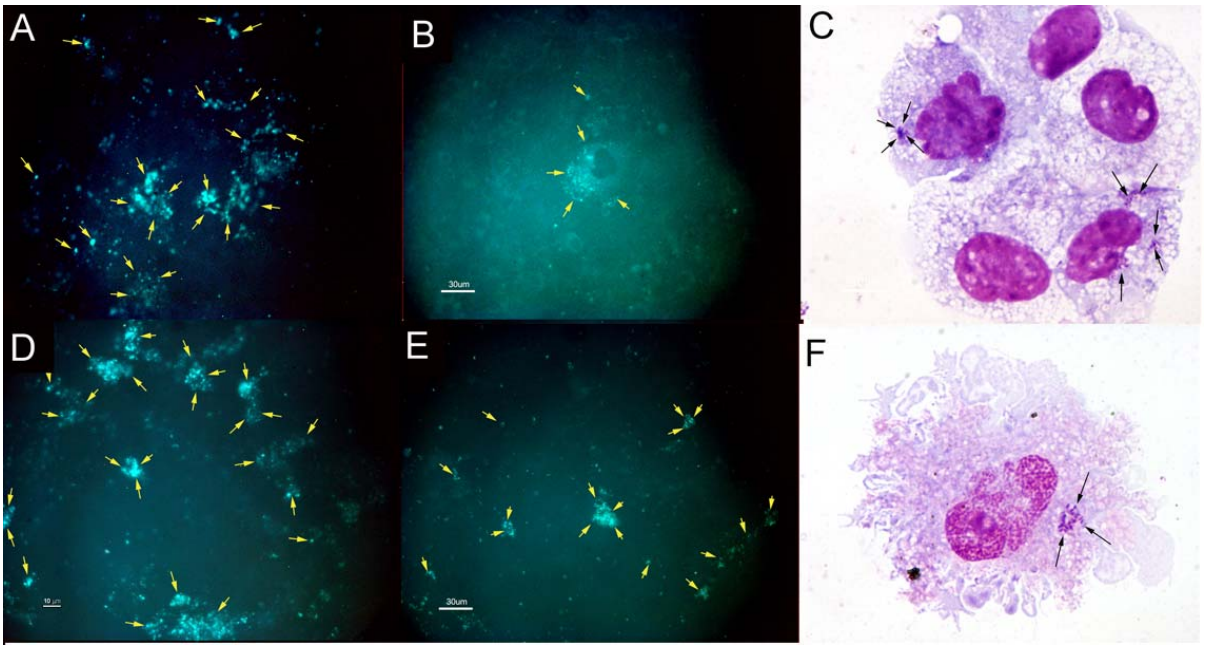


Figure 7. Differential expression of the omt gene and aph_0906 during infection of ISE6 and HL-60 cells.

A) Expression of the *omt* gene during infection of ISE6 and HL-60 cells. Red bars represent the expression of the *omt* gene during *A. phagocytophilum* HZ wildtype infection of ISE6 cells, whereas yellow bars represent the expression of the *omt* gene during infection of this same strain in HL-60 cells. B) Fold change in the expression of the *omt* gene comparing ISE6/HL-60. The bars represent the average of the fold change and the lines represent the standard errors. C) Expression of *aph_0906* gene during infection of ISE6 and HL-60 cells. Like in graph A, red bars represent the expression of *A. phagocytophilum* in ISE6, whereas yellow represents the expression in HL-60 cells; using HGE2 wildtype bacteria. The length of the solid bars represents the average expression and the lines above the bars represent the standard error of each time point. D) Fold change in the expression of the *aph_0906* gene comparing HL-60/ISE6. The bars represent the average of the fold change and the lines represent the standard errors.

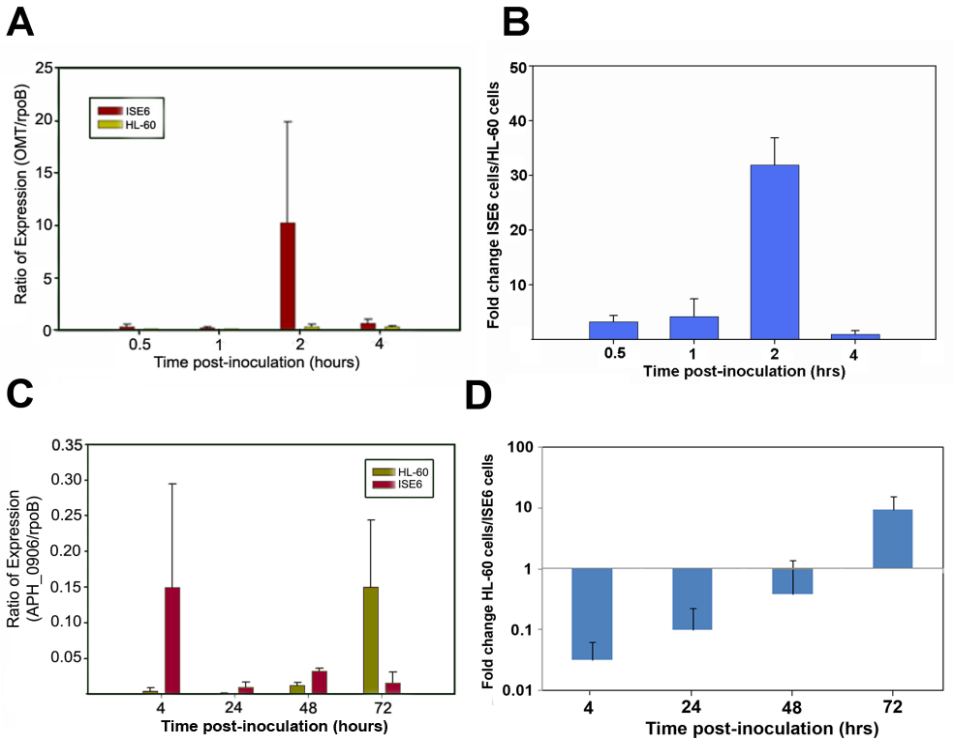


Figure 8. Putative DNA binding residues detected in the protein APH_0906.

A) Linear representation of the amino acid sequence of APH_0906 showing all the putative residues detected. The black bar symbolizes the complete amino acid sequences of APH_0906, positions in the sequence are shown by the red numbers. The pink bars show the localization of DNA binding residues detected using BindN and DNABindR. The blue line represents the position of the DNA motif detected by GYM 2.0. The red line represents the position the residues that presented high similarity with BRO motif of the antirepressor protein from the Lactobacillus phage LV-1 detected by PSI-BLAST. B) Amino acid sequence of APH_0906 showing specific residues that present putative DNA binding capabilities. The residues marked with red asterisks were detected using BindN and DNABindR, all other amino acids indicated by the pink bars were detected by one of the program as putative binding residues. C) NCBI representation of the antirepressor protein from Lactobacillus phage LV-1. The blue bar represents the BRO domain in the protein and the blue line indicates the residues that show high similary and percentage of positives with APH_0906. The exact function of BRO proteins is unknown, however it is suggested they are DNA binding proteins that influence host DNA replication and transcription.

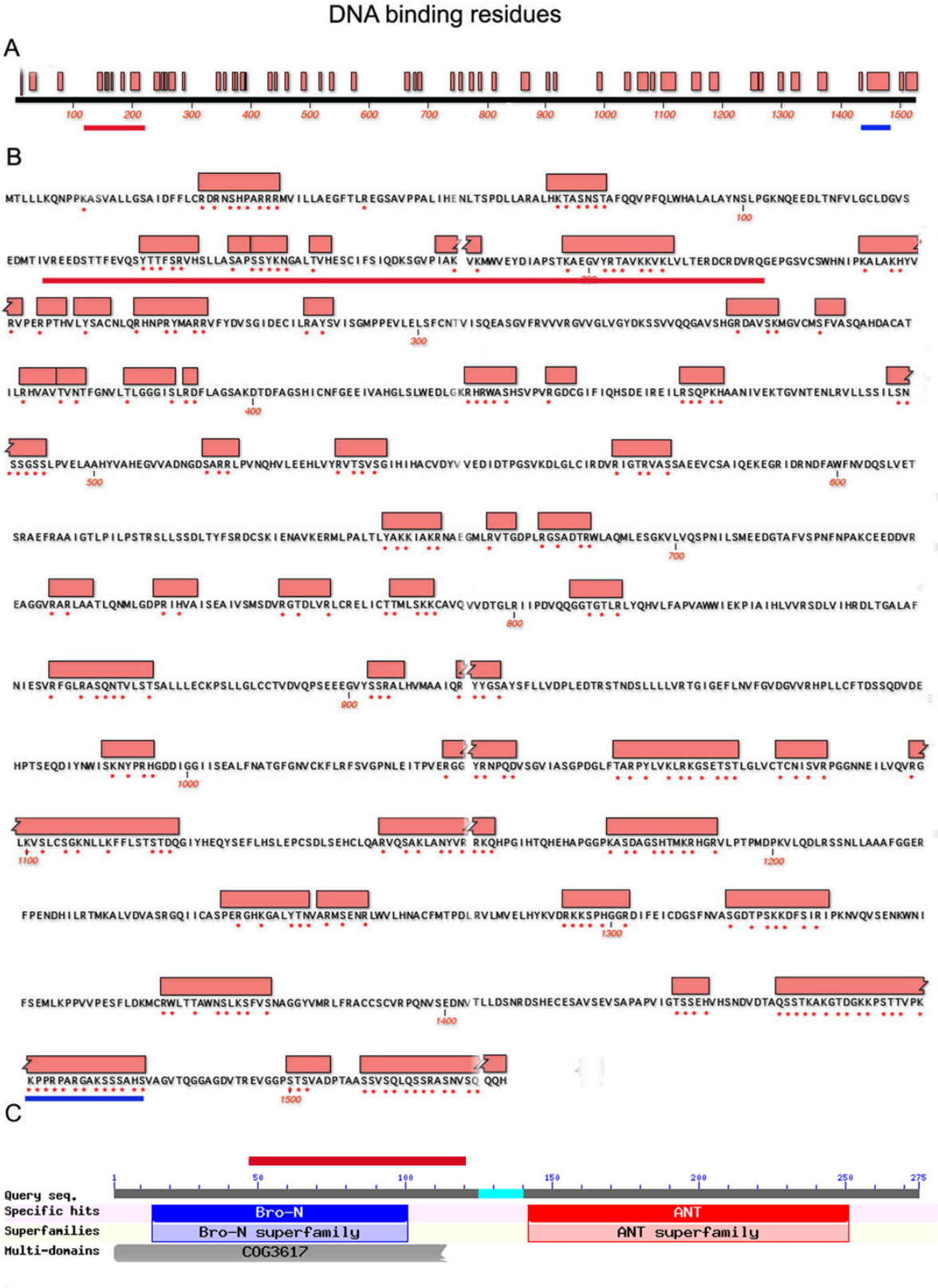

DNA binding residues
A
B
C
Query seq.
Specific hits
Superfamilies
Multi-domains
Bro-N
Bro-N superfamily
ANT
ANT superfamily
COG3617

Table 2. Results from the predicted localization of the APH_0906 protein in *A. phagocytophilum*

CELLO

SeqID: APH_0906

Analysis Report:

SVM	LOCALIZATION	RELIABILITY
Amino Acid Comp.	OuterMembrane	0.805
N-peptide Comp.	Extracellular	0.357
Partitioned seq. Comp.	OuterMembrane	0.878
Physico-chemical Comp.	InnerMembrane	0.370
Neighboring seq. Comp.	Cytoplasmic	0.389

CELLO Prediction:

	OuterMembrane	2.735 *
	Cytoplasmic	0.763
	InnerMembrane	0.632
	Extracellular	0.602
	Periplasmic	0.268

PSORTb 3.0v

```
Localization Scores:
  Cytoplasmic                0.03
  CytoplasmicMembrane        0.01
  Periplasmic                0.09
  OuterMembrane              9.49
  Extracellular              0.38
Final Prediction:
  OuterMembrane              9.49
```

Figure 9. Similarities between the A. phagocytophilum OMT to the enzymes present in other bacteria

A) Phylogenetic tree generated with the amino acid sequence of the 30 enzymes most similar to the *A. phagocytophilum* OMT. The relationship between the OMTs was inferred using the Minimum Evolution method. The bootstrap values from1000 replicates are shown next to the branches. The evolutionary distances were computed using the total number of differences. All positions containing gaps and missing data were eliminated from the dataset (Complete deletion option). Members of the Anaplasmatacaea family are shown in maroon, members of the α-proteobacteria are shown in red, Δ-proteobacteria are shown in green, Cyanobacteria are shown in light blue, Actinobacteria are shown in pink, Bacteriodetes are shown in dark purple, Chloroflexi are shown in medium blue, Acidobacteria are shown in dark green, Firmicutes are shown in light purple, and γ-proteobacteria are shown in dark blue. B) Alignment of the amino acid sequences from the 5 non-Anaplasmatacaea members closest related to the *A. phagocytophilum* OMT. Motifs found using MEME are represented by the blocks on the alignments. The conserved residues on the motif sequences from each of the OMT sequences are shown in the alignment. C) *A. phagocytophilum* OMT putative tertiary structure as predicted by Phyre2. The C and N-terminus are shown by the arrows.

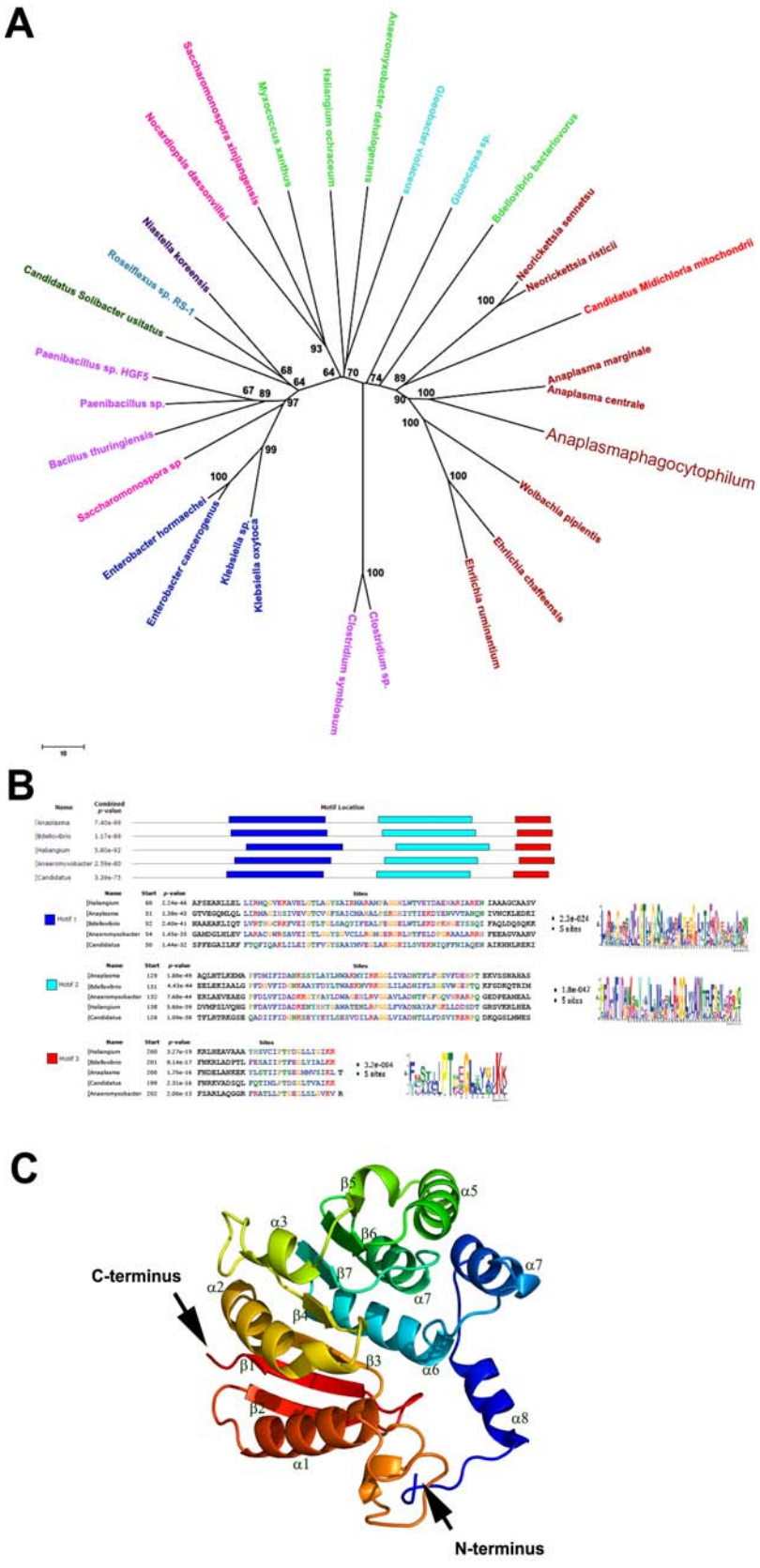
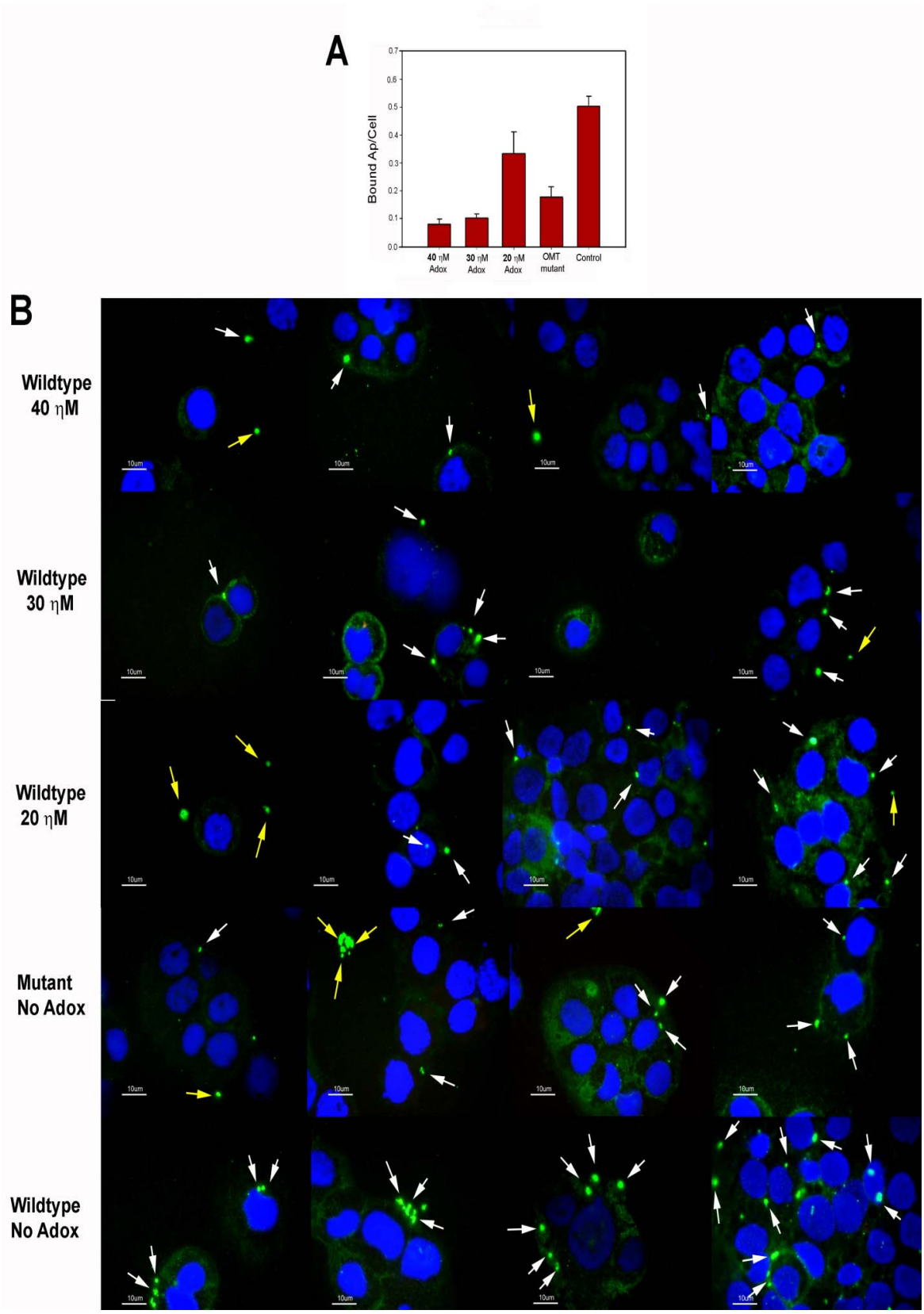

A
Saccharomonospora xinjiangensis
Nocardiopsis dassonvillei
Myxococcus xanthus
Halangium ochraceum
Anaeromyxobacter dehalogenans
Gloeobacter violaceus
Gloeocapsa sp.
Bdellovibrio bacteriovorus
Neorickettsia sennetsu
Neorickettsia risticii
Candidatus Midichloria mitochondrii
Niastella koreensis
Roseiflexus sp. RS-1
Candidatus Solibacter usitatus
Paenibacillus sp. HGF5
Paenibacillus sp.
Bacillus thuringiensis
Saccharomonospora sp
Enterobacter hormaechei
Enterobacter cancerogenus
Klebsiella sp.
Klebsiella oxytoca
Clostridium symbiosum
Clostridium sp.
Ehrlichia ruminantium
Ehrlichia chaffeensis
Wolbachia pipientis
Anaplasma marginale
Anaplasma centrale
Anaplasmaphagocytophilum
B
C
C-terminus
N-terminus
α1
α2
α3
α4
α5
α6
α7
α8
β1
β2
β3
β4
β5
β6
β7

Table 3. Protein identity of the subjects with the highest identity and similitude to *A. phagocytophilum* retrieved with PSI-BLAST

Organism	Protein	Protein ID	E-value	Score (bits)
Anaplasma marginale	hypothetical protein AMF_370	GI:222475076	4E-80	249
Anaplasma centrale	O-methyltransferase	GI:269958872	2E-79	248
Ehrlichia ruminantium	O-methyltransferase	GI:58617067	3E-68	219
Ehrlichia chaffeensis	O-methyltransferase	GI:88657893	4E-68	219
Wolbachia endosymbiont of Culex quinquefasciatus	O-methyltransferase family protein	GI:190571337	2E-63	207
Wolbachia endosymbiont of Drosophila melanogaster	O-methyltransferase family protein	GI:42520865	5E-61	200
Wolbachia endosymbiont of Onchocerca volvulus	O-methyltransferase	GI:111035805	5E-58	192
Neorickettsia sennetsu	O-methyltransferase family protein	GI:88608149	4E-42	152
Neorickettsia risticii	O-methyltransferase family protein	GI:254796950	3E-41	148
Bdellovibrio bacteriovorus	O-methyltransferase	GI:426402377	4E-40	146
Haliangium ochraceum	O-methyltransferase family protein	GI:262196431	1E-39	143
Anaeromyxobacter sp. K	O-methyltransferase family protein	GI:197124171	1E-36	138
Candidatus Midichloria mitochondrii IricVA	O-methyltransferase family protein	GI:339319550	2E-36	137
Gloeocapsa sp. PCC 7428	O-methyltransferase family 3	GI:434391334	2E-32	126
Klebsiella oxytoca	hypothetical protein HMPREF9686_01422	GI:423102816	7E-32	125
Gloeobacter violaceus PCC 7421	O-methyltransferase	GI:37522385	3E-31	124
Roseiflexus sp. RS-1	O-methyltransferase family protein	GI:148657074	3E-31	123
Enterobacter cancerogenus	O-methyltransferase	GI:261339745	2E-30	121
Enterobacter hormaechei	O-methyltransferase	GI:334124011	1E-29	119
Saccharomonospora xinjiangensis	putative O-methyltransferase	GI:383828057	3E-29	118
Nocardiopsis dassonvillei	O-methyltransferase family 3	GI:297559607	4E-29	118
Bacillus thuringiensis	O-methyltransferase	GI:75758590	6E-29	117
Clostridium sp.	O-methyltransferase family protein	GI:283796163	9E-29	117
Niastella koreensis	O-methyltransferase family protein	GI:375143786	1E-28	117
Coniophora puteana	O-methyltransferase family 3 protein	GI:392596339	3E-28	115

Figure 10. Repression of SAM-dependent methyltransferases before inoculation results in a reduction of binding to ISE6 cells that is concentration dependent

A) Bars represent the average number of bacteria bound to ISE6 cells in each treatment. The lines above the bars show the standard error from four replicates. B) Images of the bacteria bound to ISE6 cells in the different treatments. Bacteria were labeled with FITC (green fluorescence) while the DNA was labeled with DAPI (blue fluorescence) to count the number of host cells. The yellow arrows point to bacteria that are not bound to ISE6 cells, whereas the white arrows show the bacteria that are attached to cells.


A
Bound Ap/Cell
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
40 ηM Adox
30 ηM Adox
20 ηM Adox
OMT mutant
Control
B
Wildtype 40 ηM
Wildtype 30 ηM
Wildtype 20 ηM
Mutant No Adox
Wildtype No Adox
10um

Chapter 3

Methylation of an *Anaplasma phagocytophilum* outer membrane protein is required for adhesion to tick cells

Introduction

Methyltransferases are involved in important bacterial activities such as cell signaling, cell invasion, and gene expression, as well as in metabolic pathways and pathogenesis [110,117,145]. They participate in the modification of membrane components, cofactors, signaling and defense compounds [110] and have been linked with virulence of several bacteria [139,146,147], fungi [148], and viruses [114]. Methyltransferases have been proposed to be important virulence factors conserved among several bacterial taxa [139,149,150], and methylation of outer membrane proteins and other membrane components has been shown to be an essential part of invasion and pathogenesis in several pathogens, including bacteria and viruses [114,115,117,139,148]. For example, inhibition of methyltransferases reduced the ability of HIV-1 to infect and replicate within CEMT- and HEK293T cells transfected with the HIV-1 proviral plasmid pNL4, [114], possibly by inhibiting virus binding and cell entry. Transfected HEK293T cells produce pseudoviruses and are used to study envelope proteins and molecules important for the study of invasion and infection process of HIV- viruses [151]. According to Willensen *et al.* [114], AdOx affected the methylation of the envelope proteins (Env), which reduced the ability of the virus and pseudovirus to enter host cells. Likewise, it has been shown that methylation of the outer membrane protein B (OmpB) is necessary for binding to and invasion of endothelial cells by *Rickettsia prowazekii,* the etiologic agent of epidemic typhus. Bacteria with a frame shift or other mutation in a lysine methyltransferase (rRP027-028) gene are attenuated and have been used as a vaccine strain previously [146]. Likewise, the methylation of glutamic acid residues in the outer membrane protein OmpL32 of *Leptospira interrogants* has been suggested to

be involved in the virulence of the pathogen and its ability to colonize kidney cells in hamsters [118].

O-methyltransferases (OMTs) affect the virulence of pathogens from bacteria to fungi [148,152], e.g., through modification of the outer membrane components or toxins. OMTs have been proposed to be virulence factors in pathogenic strains of *Mycobacterium* sp., *Francisella* sp., *Legionella* sp., and in *Coxiella burnetii* in which they are highly conserved, but are less related to those from non-pathogenic strains [117]. Moreover, the enzymes appear to be closely related to the OMTs present in their hosts [117]. Mutation of a 3-o-methyltransferase gene in *Mycobacterium avium* prevented the methylation of glycopeptidolipids that form part of the surface of the bacteria, resulting in the loss of pathogenicity for mice and immune clearance [152]. Nodulation of *Rhizobium* sp. and specificity of the bacteria for certain host plants is regulated by Nod factors that are modified by an o-methyltransferase [153]. Likewise, the mutation of an o-methyltransferase gene, *Aph_0584*, in the rickettsial pathogen *Anaplasma phagocytophilum* has negative effects on the binding of this bacterium to the tick embryonic cell line ISE6, derived from *Ixodes scapularis*.

A. phagocytophilum is a tick-borne pathogen that causes Human Granulocytic Anaplasmosis (HGA), a disease that has been gaining importance in the USA and Europe [154], and is expanding into areas where the pathogen was not present before [154-156]. It is carried and transmitted by ticks of the *I. ricinus* group [157], although the pathogen has been found in several other ticks, including species in the genera *Amblyomma, Dermacentor, and Haemaphysalis, but which have not been shown to be vectors [as reviewed in [158]]*. Likewise, this pathogen infects several mammalian hosts including

ruminants, rodents, wolves, dogs, cats, horses, and humans [as reviewed in [158]]. It develops in neutrophil granulocytes, neutrophil progenitor cells and less frequently eosinophils, and possibly in endothelial cells [85].

The development of *A. phagocytophilum* in *Ixodes* sp. vector ticks remains unknown but has been described in tick cell culture where it is biphasic [92]. *A. phagocytophilum* development in ISE6 cells is very different from that observed in HL-60 and endothelial cells [92], and although many of the molecular events involved in the infection of mammalian cells are known [159], little is known in the tick counterpart. Studies to understand vector-pathogen interactions have focused on the tick responses and tick factors important for the infection process of the pathogen in ticks [160-162], but little is known about which *A. phagocytophilum* genes and proteins are important for development in tick cells and ticks. Some studies have looked at the specific gene expression of *A. phagocytophilum* during infection of *I. scapularis* ticks or ISE6 cells, but have focused on certain periods, such as transmission feeding or late phases of infection [54,97]. As a result, little is known about what proteins and their modifications are necessary for the early phases in successful infection of tick vector cells by *A. phagocytophilum*. Here, I report on methylation of the Major Surface Protein 4 (Msp4) and the possible role that this modification plays in the bacterium's ability to bind to and infect ISE6 cells.

Materials and Methods

Recombinant OMT (rOMT) protein was produced using the pET29a expression vector (Novagen, Germany) by amplifying the entire coding region with the primers rOMT fw and rOMT rv (Table 1). The product was amplified using *pfu* DNA polymerase (Promega, Madison, WI), under the following conditions: one denaturating cycle at 94 °C for 3 min, 10 cycles with a denaturating step at 94 °C for 1 min, 40 °C for 1 min for annealing, and extension at 72 °C for 2 min, then 20 additional cycles with a denaturating step at 94 °C for 1 min, 47 °C for 1 min for annealing, and extension at 72 °C for 2 min, and a final extension step of 5 min at 72 °C. The amplified product was then digested with the restriction enzymes *SalI* and *EcorV,* followed by the ligation into the vector at 15°C overnight. The plasmid was cloned into One Shot® TOP10 competent cells (Invitrogen, Grand Island, NY) for replication. Plasmid was purified using the High Pure plasmid isolation kit (Roche, Indianapolis, IN). Integrity of the plasmid was checked by sequencing the plasmid with the T7 promoter (5'- TAA TAC GAC TCA CTA TAG GG – 3') and the T7 terminator (5'- GCT AGT TAT TGC TCA GCG G – 3') primers at the Biomedical Genomics Center of the University of Minnesota. Sequences were reviewed using MacVector 12.5 (MacVector, Inc., Cary, NC) and when the integrity was confirmed, plasmids were transfected into BL21(D3) *E. coli* (New England Biolabs, Ipswich, MA) for expression. BL21(D3) *E. coli* were inoculated into 100 ml of Superior Broth (AthenaES, Baltimore, MD), induced with 200 µM IPTG, and incubated at 37°C overnight with constant shaking. Protein was purified using Ni-NTA Fast Start Kit

columns (Qiagen, Germantown, MD). Protein concentrations were measured using the BCA protein assay kit (Pierce, Rockford, IL).

Functional rOMT was produced using the expression vector pET29a and the procedure described previously but amplifying the full coding region of the gene with the primers rOMTns Fw and rOMTns Rv (Table 1), using the following conditions: one denaturating cycle at 94 °C for 3 min, 10 cycles with a denaturating step at 94 °C for 1 min, 45°C for 1 min for annealing, and extension at 72 °C for 2 min, then 20 additional cycles with a denaturating step at 94 °C for 1 min, 54 °C for 1 min for annealing, and extension at 72 °C for 2 min, and a final extension step of 5 min at 72 °C. The amplified product was then digested with *NdeI* and *EcoR1*, and then ligated into pET29a to produce rOMTns to eliminate the S-tag present in the plasmid. Ligated pET29- rOMTns plasmid was then cloned into One Shot® TOP10 competent cells, integrity of the plasmid was determined as previously described. The plasmid was cloned into Rosetta 2(DE3) pLysS (Novagen, Germany) cells. Transformed *E. coli* were incubated in 100 ml of superior broth and induced with 1 mM IPTG at 37 °C for 5 hours. Proteins were purified with Ni-NTA Fast Start Kit columns as previously described, and protein concentration was measured as described previously.

Serum production

Recombinant non-functional rOMT was dialyzed in a 3 ml Slide-A-Lyzer Dialysis Cassette 10K MWCO (Thermo Scientific, Rockford, IL) and tris-buffered saline (TBS), changing the TBS after two hours and dialyzing in fresh TBS overnight. Four 6 – 8 weeks old C57BL/6J mice (Jackson Laboratories, Bar Harbor, ME) were immunized by

injection of 100 µg of the recombinant non-functional dialyzed rOMT in 50 µl elution buffer combined with an equal volume of TiterMax Research adjuvant (CytRx Co., Norcross, GA). The vaccine was divided into 100 µl doses given subcutaneously (s.c.) on the back at the base of the tail. Mice were given a second and third booster injection with 100 µg of the rOMT 14 days and 24 days later, respectively. Blood was collected 10 days after the third booster by heart puncture after euthanasia using CO_2 overexposure. Mice from the same cohort, but that were not immunized with the recombinant protein, were used as controls. Serum was prepared by incubating the blood at 37 °C for 60 min. After incubation, the clot that formed was separated from the walls of the centrifuge tubes followed by an overnight incubation at 4 °C. The tubes containing the samples were centrifuged at 3000 RPM for 10 min and serum was collected from tubes and frozen at -20°C.

Localization assays

Wild-type bacteria from the isolate HZ (HZWT) were grown in 75 mL flasks containing HL-60 cells until >90% of the cells were infected. Around 5 x 10^7 infected cells were used to obtain cell free bacteria by vortexing the infected cells with 60/90 grit silicon carbide (Lortone, inc., Mukilteo, WA) for 30 seconds followed by filtration through a 2.0 µm pore size filter and centrifugation at 700 x g for 5 min to remove remaining cell debris. The percent of infected cells in the culture was calculated by counting the number of infected and uninfected cells in two different Giemsa-stained preparations from the same flask that had been centrifuged onto slides using a Cytospin 4 centrifuge (Thermo Shandon, Asheville, NC) for 5 min at 1,000 xg. The total number of

cells was determined by counting using a hematocytometer. Cell free bacteria were co-cultivated with 2.5 x 10^5 uninfected ISE6 cells in 500 µl of *A. phagocytophilum* medium (described in Chapter 2) and incubated for 2 hours at 34°C to allow binding and expression of the OMT. Unbound bacteria was removed with three washes using *A. phagocytophilum* medium and centrifugation at 800 x g to collect only those bacteria bound to tick cells. Resuspended whole cell samples (100 µl) were taken and diluted into 900 µl of fresh medium. Samples were spun onto a microscope slide at 1000 rpm for 5 min using a Cytospin 4 centrifuge.

Cell spots were fixed for 10 min in methanol, and incubated with anti-rOMT serum diluted 1:250 in phosphate-buffered saline (PBS) containing 3 % bovine serum albumin (BSA) for 2.5 hours at room temperature. The slides were washed 3 times in PBS and blocked in PBS with 3% BSA for 10 min at room temperature. OMT expressing bacteria were then labeled with anti-mouse antibodies conjugated with TRITC (1:500 dilutions) for 1 hour at room temperature. All *A. phagocytophilum* were labeled with anti-*A. phagocytophilum* dog serum diluted 1:500 followed by incubation with anti-dog IgG conjugated to FITC for contrast, using the same procedure. Tick cell nuclei were labeled using DAPI present in the VectaShield mounting medium (Vector laboratories, Burlingame, CA) and samples were viewed with a Nikon Eclipse E400 microscope (Nikon Instruments, Melville, NY). Images were recorded with a Nikon digital camera DXM1200 (Nikon Inc., Melville, NY), using Nikon ACT-1 software.

Measurement of the differential expression of proteins and determination of methylation

HZ wild-type bacteria and OMT-mutant HZ bacteria were grown in two 250 ml flask containing 50 ml of RPMI medium with HL-60 cells each. The number of infected cells was determined as previously described. Cell free bacteria, mutant and wildtype, purified from 2.5 $x10^9$ infected cells that were ruptured by repeated passage through a 27G needle and centrifuged at 600 x g to remove cell debris, were inoculated into an entire 25 cm^2 flask containing 6.4 x 10^7 cells. Needle purification was used to avoid introducing small silicone particles in the MS/MS apparatus that could result in damage. This procedure was replicated three times. Cells and bacteria were incubated for 4 hours at 34 °C without agitation and bacteria were purified from ISE6 cells as described above, using the needle and syringe to rupture cells. Bacteria were washed in un-supplemented L15Cd medium three times by centrifugation at 16,000 xg for 5 min at 4 ^oC to remove FBS, and the final bacteria pellet was extracted for mass spectrometry.

Protein Extraction and Preparation*

Protein concentrations were determined by Bradford assay using two aliquots for each sample. All samples were prepared as follows at the Center for Mass Spectrometry and Proteomics at the University of Minnesota: cell pellets were reconstituted with 120 µl of protein extraction buffer [7 M urea, 2 M thiourea, 0.4 M triethylammonium bicarbonate (TEAB) pH 8.5, 20% methanol and 4 mM tris(2-carboxyethyl)phosphine (TCEP)] while on ice. Samples were sonicated at 30% amplitude for 7 seconds with a Branson Digital Sonifier 250 (Emerson, Danbury, CT). The sample was transferred to a

PCT tube, which is a shredder used to homogenize protein samples, with a 50 µl cap for the Barocycler NEP2320 (Pressure Biosciences, Inc., South Easton, MA) and cycled between 35k psi for 30 sec. and 0 k psi for 15 sec. for 40 cycles at 37 °C. The PCT tube was uncapped and 200 mM methyl methanethiosulfonate (MMTS) was added to a final concentration of 8 mM MMTS, recapped, inverted several times and incubated 15 min at room temperature. The sample was transferred to a new 1.5 ml Eppendorf LoBind Protein microcentrifuge tube (Eppendorf, New York, NY).

In-solution proteolytic digestions were performed as follow: a 200 µg aliquot of each sample was transferred to a new 1.5 ml microfuge tube and brought to the same volume with protein extraction buffer plus 8 mM MMTS. All samples were diluted four fold with ultra-pure water and trypsin (Promega, Madison, WI) was added at a 1:35 ratio of trypsin to total protein. Samples were incubated for 16 hr at 37 °C after which they were frozen at -80 °C for 30 min and dried in a vacuum centrifuge. Each sample was then cleaned using a 4 ml Extract Clean™ C18 SPE cartridge (Grace-Davidson, Deerfield, IL), and eluates were vacuum dried and resuspended in dissolution buffer (0.5M triethylammonium bicarbonate, pH8.5) to a final 2 µg/µl concentration. For each iTRAQ® 4-plex (AB Sciex, Foster City, CA), two 50 µg replicates for each sample were labeled with iTRAQ® reagent (AB Sciex). After labeling, the samples were multiplexed together and vacuum-dried. The multiplexed sample was cleaned with a 4 mL Extract Clean™ C18 SPE cartridge (Mandel Scientific Company Inc., Guelph, Canada) and the eluate was dried in vacuo.

Peptide Liquid Chromatography Fractionation & Mass Spectrometry*

The iTRAQ® labeled samples were resuspended in Buffer A (10 mM ammonium formate pH 10 in 98:2 water:acetonitrile) and fractionated offline by high pH C18 reversed-phase (RP) chromatography [163]. A MAGIC 2002 HPLC (Michrom BioResources, Inc., Auburn, CA) was used with a C18 Gemini-NX column, 150 mm x 2 mm internal diameter, 5 μm particle, 110 Å pore size (Phenomenex, Torrence, CA). Buffer A was 10 mM ammonium formate, pH 10 in 98:2 water:acetonitrile and Buffer B was 10 mM ammonium formate, pH 10 in 10:90 water:acetonitrile. The flow rate was 150 μl/min with a gradient from 5-35% Buffer B over 60 min, followed by 35-60% over 5 min. Fractions were collected every 2 min and uv absorbances were monitored at 215 nm and 280 nm. Peptide containing fractions were divided into two equal numbered groups, labeled "early" and "late". The first "early" fraction was concatenated with the first "late" fraction by combining the two fractions, and so on. Concatenated samples were dried in vacuo, resuspended in load solvent (98:2:0.01, water:acetonitrile:formic acid) and 1-1.5 μg aliquots were run on a Velos Orbitrap mass spectrometer (Thermo Fisher Scientific, Inc., Waltham, MA) as described previously [164] with the exception that the Higher-energy Collisional Dissociation (HCD) activation energy was 20 ms..

The mass spectrometer RAW data (Proteowizard files) were converted to mzXML using MSconvert software and to MGF files using TINT raw-to-mgf converter. ProteinPilot 4.5 (AB Sciex, Foster City, CA) searches were performed against the NCBI reference sequence for the I. scapularis (taxon 6945; November 14, 2011 version) protein FASTA database with canonical and isoform sequences (20468 proteins), to which a NCBI reference sequence A. phagocytophilum str. HZ (taxon 212042; November 14,

2011; 1267 protein) and a contaminant database (thegpm.org/crap/index, 109 proteins) was appended. Search parameters were: cysteine

MMTS; iTRAQ 8plex (Peptide Labeled); trypsin; instrument Orbi MS (1–3ppm) Orbi MS/MS; biological modifications ID focus; thorough search effort; and False Discovery Rate analysis (with reversed database).

* These sections were written by the staff at the Mass Spectometry Laboratory at the University of Minnesota.

Function and Pathways of Differentially Expressed A. phagocytophilum Proteins.

The putative function of all the differentially expressed proteins identified by iTRAQ was determined using the databases in NCBI (www.ncbi.nlm.nih.gov) to identify conserved domains, Uniprot (http://www.uniprot.org/uniprot/), EMBL-EBI (http://www.ebi.ac.uk/interpro/IEntry?ac=IPR000866), and OMA (http://omabrowser.org). Pathways which involved proteins differentially expressed during infection with the OMT-mutant compared to wildtype were identified with KEGGS (http://www.genome.jp/kegg/tool/map_pathway1.html).

Identification of Differentially Methylated Proteins

To identify the peptides that differed in their methylation, a text file of data for all peptides was imported into Xcel (Microsoft, Redmond, WA), and entered into TextWrangler (Bare Bones Software, North Chelmsford, MA). Peptides that were less abundant in the mutant when compared to the wild-type were selected and the intensity of spectra was analyzed using Protein Pilot to confirm differences.

Production and Testing of Functional Recombinant OMT

A functional version of the OMT without the sTag (rOMTns) was produced as described previously. After eluting the protein was dialyzed as described for non-functional protein. rOMTns was used in methylation assays using *A. phagocytophilum* and ISE6 cells lysates as substrates to evaluate the activity of the enzyme. *A. phagocytophilum* lysates were produced from bacteria purified from 75 ml flasks of infected HL-60 cells and lysed with ProFound Lysis Buffer (Thermo Scientific, Rockford, IL) for 2 hours at 4 °C. ISE6 cell lysates were produced from confluent cell layers in 25 cm^2 flasks using the same procedure as for *A. phagocytophilum* lysate. Concentration of the protein in lysates was measured using the BCA protein assay. Lysates were dialyzed as described for recombinant proteins and then frozen at -70 °C to inactivate any possible methyltransferase activity.

The SAM-fluoro: SAM Methyltransferase Assay (Gbiosciences, St. Louis, MO) was used to measure the activity of the enzyme. A total of 50 ηg rOMTns was used along with a total of 40 ηg of protein from the lysates. Methylation was measured by the excitation of resorufin after the generation of AdoHcy (adenosine homocysteine) by SAM-dependent methyltransferases. Resorufin production was determined in a Synergy H1 Hybrid microplate reader (Biotek, Winooski, VT) during a kinetic run measuring the fluoresce at an excitation wavelength of 530nm and an emission wavelength of 585nm every two min for 2 hr at 34°C. Reactions were carried out in a 96 well EIA/RIA plate flat bottom plate (Costar, Corning, NY) covered with a MicroAmp optical adhesive film (Applied Biosystems, Grand Island, NY) to protect samples from evaporation. Negative

controls included the lysates or the rOMTns alone without the enzyme or substrate, respectively, in the reaction buffer with AdoMet. Positive controls included the addition of AdoHcy with the rOMTns and the positive control enzyme provided by the manufacturer. Statistical significance was measured with One-Way ANOVA in repeated measurements (SigmaStat).

Production of recombinant proteins for substrate testing

Possible substrates identified by iTRAQ where produced as recombinant proteins using the pET29a vector (Novagen). Primers to amplify partial or complete coding sequences where designed with restrictions sites for *NdeI* and *XhoI* enzymes, using the primer sequences listed in Table 1. Inserts were amplified using *pfu* enzyme under conditions listed in Table 2. Products were purified, replicated, and the integrity of the sequences in the plasmids was confirmed by Sanger sequencing as described. Plasmid constructs were then cloned into BL21(D3) *E. coli* (New England Biolabs, Ipswich, MA) for expression. Proteins were produced in 150 ml of superior broth induced with 200 µM IPTG, above, and used in methylation assays as described below. Proteins concentrations were measured using BCA micro protein assay kit (Pierce).

Methylation assay of potential substrates

Methylation of different proteins in the two lysates was tested with the SAM*fluoro*: SAM Methyltransferase Assay, as previously described in the section *Production and testing of functional recombinant OMT*. Assays were performed with 50 ng of the enzyme and 40 ng total proteins for substrates and lysates, as specified by the

manufacturer. Kinetic measurements were done every 2 min for 4 hrs. Statistical significance was measure with One-Way ANOVA in repeated measurements (SigmaStat).

Putative tertiary structure and localization of methylated residues

Phyre2 (http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index) was used to determine the putative tertiary structure of the proteins that were identified as being methylated by the iTRAQ analysis as well as the in vitro methylation assay. The putative localization of the modified residues was determined from the protein sequence and the structure generated from Phyre2. Phobius (http://phobius.sbc.su.se/cgi-bin/predict.pl) was used to determine where the modified residues were located within the membrane of the bacteria since Msp4 is a surface protein [165].

OMT is expressed by A. phagocytophilum interacting with tick cells

According to the results from the previous analyses, OMT expression correlated with the binding of the bacteria to ISE6 cells and the mutation of the gene encoding the enzyme affected the ability of the bacteria to bind to host cells (previous chapter). Therefore, I was interested in the expression and localization and the expression of the protein during binding. Antiserum against the recombinant version of the complete OMT was produced to localize the protein during binding to ISE6, using Immuno Fluorescent Assays (IFA). The expression and localization of wild-type HZ OMT during incubation with ISE6 cells for 2 h was analyzed using a binding assay and immunofluoresence labeling as described above. Expression of the OMT was detected using mouse anti-OMT

serum and anti-mouse IgG conjugated to TRITC (red fluorescence) whereas bacteria were labeled with dog anti-anaplasma serum and FITC-conjugated anti-dog IgG (green fluorescence). Thus, co-localization of the enzyme with the bacteria resulted in a yellow signal (Figure 1A-B). Unbound bacteria or bacteria not interacting with ISE6 cells did not express the OMT and only green fluorescence was observed (Figure 1B). The results demonstrated that only bacteria interacting with ISE6 cells expressed the OMT (Figure 1A-B). Bacteria incubated with pre-immune serum did not fluoresce red (Figure 1C) and uninfected ISE6 cells incubated with anti-OMT followed by TRITC-conjugated anti-mouse IgG did not show any red fluorescence either (negative control; Figure 1D), demonstrating that the serum specifically interacted with the OMT. This correlates with the previous results and corroborates the hypothesis that the enzyme methylates a bacterial protein that is important during binding to ISE6 cells.

Mutation of OMT results in the differential expression of several A. phagocytophilum proteins

To identify the proteins that were differentially expressed in the *omt*-mutant compared with the wild-type bacteria, I used a proteomic approach. Peptides in each sample were labeled with isotopic tags of known mass. Both mutant and wild-type bacteria were incubated with ISE6 cells and allowed to bind to the cells for 4 hrs at 34°C without agitation. Proteins were extracted from pellets of bacteria purified from host cells by passage through a small-bore needle as described earlier, and triplicate samples were analyzed by tandem mass spectrometry (MS/MS). In each replicate, multiple *A. phagocytophilum* proteins were identified that appeared to be differentially expressed in the mutant as compared to wild-type bacteria (data not shown). However, only 24 *A.*

phagocytophilum proteins (Table 3) were identified as differentially regulated in all replicates. Among these, 5 proteins were downregulated (hypothetical protein APH_0406, major surface protein 4, glutamine synthetase type 1, anti-oxidant AhpCTSA family protein, and ankyrin (GI88607707)), whereas the remaining 19 proteins were upregulated (Table 3). Glutamine synthetase type 1, hypothetical protein APH_0406, and major surface protein 4 (Msp4) presented the lowest relative expression ratios (>0.2), suggesting that they are highly expressed in HZ wild-type during binding to ISE6 cells compared to the mutant (Table 3). Several proteins known to be highly expressed [54], or known to be involved in infection of mammalian cells [55,130], were upregulated in the mutant (Table 3). These proteins included several membrane proteins (P44-18ES, an OmpA family protein, P44-1 Outer membrane protein, an OMP85 family outer membrane protein, and hypothetical protein APH_0405) as well as stress response proteins (co-chaperone GrpE, chaperonin GroEl, and chaperone DnaK) (Table 3). This suggests that unlike HZ wild-type, the *omt* mutant possibly did not respond to the interaction with ISE6 cells in a host cell specific manner and did not change the repertoire of proteins in its outer membrane as a result. An analysis of proteins expressed by HZ wild-type *vs omt* mutant bacteria when binding to or replicating in HL-60 cells could answer this question. Additionally, it is possible that selection favored a population in which the repertoire of expressed outer membrane proteins differs from that of the parent for reasons unrelated to the mutation. It is also possible that lack of OMT disrupted an environmentally responsive regulatory mechanism or sensor that prepares *A. phagocytophilum* for changes in hosts. An analysis of the pathways that were affected by the *omt* mutation showed that several proteins involved in transcription and protein

metabolism were also upregulated, suggesting that the mutant was metabolically active; whereas only proteins involved in purine and pyrimidine metabolism were downregulated (Table 4).

Identification of possible OMT substrates by iTRAQ

Anaplasma and host cell peptides marked in iTRAQ results as having a methyl modification were analyzed to identify those that were less abundant in the *omt* mutant, and in mutant-infected cells. After examination of the spectra of each peptide in *A. phagocytophilum*, only peptides with a <0.7 ratio between the wild-type and mutant were considered to display reduced methylation. Those peptides with >0.7 were not considered as significant for further analysis since the spectra from mutant and wild-type displayed similar levels (Figure 2). This analysis identified eight *A. phagocytophilum* proteins with reduced methylation of 8 corresponding peptides (Table 4). Two of the proteins, Msp4 and APH_0406, were downregulated in the mutant as well as being less methylated, highlighting their importance during tick cell infection. The affected amino acids were the glutamic acid resides (E) in the Msp4 peptide VEVEVGYK, and the asparagine residue (N) in the APH_0406 peptide NVVLGGMLK (Figure 3A and B). These were also the same peptides with the greatest reduction in abundance in the mutant compared to HZ wild-type (Table 5). However, the non-methylated versions presented less of a difference in abundance, on average yielding a ratio of 0.2199 for the Msp4 peptide in the mutant versus the HZWT, and 0.9243 for APH_0406 (Figure 4A and B). This suggested that in both proteins the non-methylated version of these peptides was more abundant, especially in APH_0406 in which the difference was minimal. This result confirmed the

low level of confidence in the identification of a non-methylated APH_0406 peptide (>1%), whereas confidence in the identification of the Msp4 non-methylated peptides was from 77 – 99% (Figure 4A and B).

Likewise, several proteins in tick host cells displayed reduced methylation when infected with the *omt* mutant vs when infected with HZ wild-type (Table 6). A total of 15 proteins and 17 peptides (two proteins presented changes in two peptides) showed reduced methylation. However, only one peptide of these *I. scapularis* proteins yielded a ratio of <0.4, in contrast to *A. phagocytophilum* proteins (Table 6). This peptide was present in a heat shock protein (GI:241263225) and showed a ratio of 0.2999 in the second glutamic acid residue of DGILTIEAPLPALEAPNR (Table 6). All other peptides showed a less significant reduction in methylation. However, this protein was only found to be downregulated in one of the iTRAQ replicates (data not shown), questioning its significance as one of the substrates. Prolyl 4-hydroxylase alpha subunit (GI:240974259) and flavonol reductase/cinnamoyl-CoA reductase (GI:241703753) were the only two *I. scapularis* proteins to be downregulated in all replicates and to present peptides with a reduction in their methylation (Table 6). However, the reduction in methylation of the peptides in this protein was not as significant, with ratios of ~0.65 for flavonol reductase/cinnamoyl-CoA reductase and ~0.70 for prolyl 4-hydroxylase alpha subunit (Table 6).

Recombinant OMT methylates A. phagocytophilum and I. scapularis lysates

To test if the proteins identified with the iTRAQ were in fact methylated by the OMT, a recombinant version of the OMT was produced in *E. coli* using the complete

coding sequence of the gene *aph_0584* and removing the S-tag from the pET29a vector. The purity of rOMT was verified by gel electrophoresis (SDS-PAGE) of the eluted protein and Coomassie blue staining (Figure 5). The molecular weight (MW) of the purified rOMT corresponded with the predicted MW of ~24KDa (Figure 5). Two higher molecular weight proteins present in the uninduced *E. coli* lysate were also purified during the elution. Methylation assays using only the eluted rOMT along with all the reagents except for the substrate (negative control) did not demonstrate any detectable increase in fluorescence of these contaminant proteins, discarding the chance that they affected the results obtained in the other samples (Figure 6). On the other hand when the rOMT was incubated along with the positive control consisting of Adenosine Homocystein as acceptor of the methyl group, there was a rapid increase of the fluorescence (Figure 6). Likewise, when the assay was done using lysates from purified *A. phagocytophilum* or uninfected ISE6 cells lysates, there was an increased detection of fluorescence overtime (Figure 6), with increments in fluorescence as efficient as those detected with the positive control (Figure 6). Both results, however, indicate that the rOMT was active and suggested that possible substrates were present in the tick and *A. phagocytophilum* lysates.

In vitro methylation assay confirms the methylation of Major Surface Protein 4 (Msp4)

Recombinant versions of some of the proteins identified with the iTRAQ as differentially methylated between the mutant and the wild-type bacteria were produced in *E. coli.* The proteins were purified and used in an *in vitro* methylation assay as described above. The purity of the proteins was confirmed as done with the rOMT (data not

shown). Because a plateau was not reached during the methylation assays above, using bacterial and host cell lysates, the reaction was allowed to proceed for 4 hours, so that fluorescence measurements would yield a typical enzyme activity curve. Four recombinant *A. phagocytophilum* proteins were tested using this methylation assay: Msp4, APH_0406, TypA, and P44-16b. I selected these proteins from eight candidates that yielded expression ratios of <0.60 (Table 5) because they presented the strongest reductions in the expression ratios. Production of recombinant preprotein translocase subunit SecA was unsuccessful with any of the *E. coli* strains used. Of the four proteins produced, the only protein that resulted in a significant increase in resorufin activation was Msp4 (Figure 7). APH_0406, TypA, and p44-16b produced high background fluorescence that resulted in high initial readings (~1500 – 2000 RU), but did not continue to accumulate a significant number of activated resorufin units, and only reached values of ~2100 RU (Figure 7). Also, the readings did not yield a typical enzyme activity curve like that obtained using Msp4 as the substrate. By contrast, we Msp4 was used as the substrate the fluorescence started at a lower reading (~1300 RU) but reached higher values (~3400 RU) and reached to present a plateau at around 210 min after the reaction was initiated (Figure 7).

Putative positioning of methylated residues in Msp4

The tertiary structure of Msp4 was predicted using Phyre2 that compares conserved residues of a query protein to the sequence of proteins with known crystallized structures. The predicted tertiary and secondary structures were used to predict the probable positioning of the methylated residues. Msp4 was predicted to form a beta-

barrel typical of porins and the glutamic acid residues that are modified by the OMT are positioned at the start of one of the beta-strands forming the beta-barrel (Figure 8A and B). Furthermore, transmembrane and signal peptide prediction software suggested that the first ~30 aa residues represented a signal peptide to direct transport of the protein from the cytoplasm to the outer membrane (Figure 8C). These residues correspond to the alpha helix at the N-terminus (dark blue) that is probably cleaved before the protein is positioned in the outer membrane (Figure 8B and C). The protein does not contain predicted transmembrane domains, but it is very likely that its positioning in the outer membrane is similar to that reported in other porins in that the beta- barrel spans the membrane, and the portion of the protein with the longest loops is exposed on the outside of the bacteria.

Discussion

Post-translational modification of molecules that function as adhesins and other bacterial virulence factors that respond to environmental conditions or interactions with host cells has been reported [115,118,166,167]. One such modification is the addition of methyl groups by methyltransferases that functionally modify a range of diverse acceptor molecules, such as DNA, RNA, or proteins [110]. Methylation of outer membrane proteins and virulence factors has been increasingly recognized as an essential process during host invasion and infection by several obligate and facultative intracellular bacteria [117,138,148,152]. For example, outer membrane protein OmpB is an immune-dominant rickettsial antigen in several members of the genus *Rickettsia* that is known to be important for binding to and entry into mammalian cells in vitro [168], and is methylated by protein lysine methylases [115]. Methylation of proteins commonly occurs

by two types of methyltransferases: N-methyltransferases that modify lysine (K), arginine (R), or glutamine (Q), and *O*-methyltransferases that usually modify glutamic acid (E) by transferring a methyl group to the carboxyl residue, forming a methyl ester (as reviewed in [109]). Fifty four percent of the methyltransferases included in the subclass EC.2.1.1 comprise enzymes that modify an oxygen (O) atom in acceptor molecules, making *O*-methyltransferases (OMTs) the most common group of methyltransferases with the greatest genetic diversity and functional expansion [110]. Although, this group of enzymes has mostly been studied in plants and mammals [110], its importance in bacterial pathogenesis has been a subject of increasing interest [117]. *O*-methylation of outer membrane proteins has most commonly been linked to bacterial chemotaxis, signal transduction and adaptation to the environment [109,169,170]. More recent reports have, however, expanded the role of this modification in bacterial pathogenesis to include methylation of an outer membrane protein of *Leptospira interrogans* that has been shown to be important for bacterial binding to and infection of kidney cells [118].

Here, I report methylation of a major surface protein, Msp4, by an *O*-methyltranferase (OMT) in *A. phagocytophilum*. In the previous chapter, I described the effects of the mutation of the gene *aph_0584*, which encodes an OMT family 3 member, on the ability of *A. phagocytophilum* to bind to and infect the embryonic *I. scapularis* cell line ISE6. In this chapter, I used IFAs to confirm the colocalization of the enzyme only with bacteria that are binding and interacting with ISE6 cells, as extracellular bacteria that were not bound to the tick cells were not recognized by the anti-rOMT serum (Figure 1). This suggested involvement of the enzyme in the methylation of a molecule important for adhesion to and infection of ISE6 cells.

Proteomic analyses were then used to identify molecules that could be the potential substrates of the enzyme. iTRAQ, which can be used to quantify abundance of certain peptides in samples subject to different treatment conditions, revealed differences in peptide methylation of *A. phagocytophilum* and *I. scapularis* proteins. This technique identified several potential OMT substrates, two of which included *A. phagocytophilum* proteins that were previously shown to be highly expressed during infection of ISE6 cells, i.e., Msp4 and APH_0406 [54]. Both of these proteins presented a substantial decrease in their methylation due to the mutation of the OMT (Table 5 and Figure 3). Methylation of Msp4 was further confirmed using *in vitro* methylation assays (Figure 7). Msp4 is an antigenic protein encoded by a single copy gene that is highly conserved between different strains of *A. phagocytophilum* [171], as well as other members of the genus *Anaplasma* [165]. Furthermore, Msp4 is a member of the Msp2 superfamily of proteins [172] and is homologous to another major surface protein from *A. phagocytophilum*, Msp2 (P44), which has been shown to also be post-translationally modified [173], has been demonstrated to be a porin [51], and to facilitate binding to mammalian cells [53]. Thus, it is very likely that Msp4 carries out similar functions and has a similar structure to those reported for Msp2 (P44), but that it completes this function in the arthropod vector.

Several bacterial outer membrane proteins have been discovered to act as porins of various specificities and structures [174]. The most common "general" porins form 16-strand beta barrels configured as trimeric peptide subunits and in most cases are non-specific [174,175]. However, other less common porins, which have been shown to be more substrate specific, have been described and are comprised by 18-, 14-, 12-, and even

8- strands and in some cases are present as monomers (e.g., the 14 beta-stranded porins OmpG and CymA in *Escherichia coli*) [174]. Some of these porins have been shown to have double functionality since they can also act as adhesins [174]. The expression of porins depends on the environmental conditions experienced by bacteria as many porins only transport specific molecules (reviewed by Nikaido, 2003 [175]). Moreover, specific residues present on *E. coli* OmpC and OmpF affect the conformation and function of these porins [175]. Some porins have been shown to respond to cues in the environment by opening or closing, which regulates the entry or exit of specific molecules. One such porin is the 14-stranded porin OmpG from *E. coli,* that reacts to the pH in the environment and changes the conformation of its beta-strands, resulting in modifications of the arrangement of the outside loops, which leads to closure of the gate [176]. This activity is mediated by two single residues, His231 and His261, that serve as pH sensors and trigger the changes in beta-strand 11 – 13 [176,177].

Like OmpG [174], Msp4 possesses a signal peptide within the first 30 aa in the N-terminus that probably directs transport to the membrane (alpha-helix shown in Figure 8B). Analysis using Phyre2 additionally suggests that Msp4 is also a 14-beta stranded porin (Figure 8A and B), and the Msp4 homolog from the closely related bovine pathogen, *A. marginale,* was previously shown to exist as a monomer in the membrane of this bacterium [178]. Thus, is very likely that Msp4 is indeed a monomeric 14-stranded porin, similar to OmpG. It is interesting to note that the glutamic acid residues (E) of Msp4 that appear to be important for binding and host cell invasion by *A. phagocytophilum* are close to one of the loops on the outside of the channel (Figure 3A and Figure 8B). Similarly, *L. interrogants* OmpL32 contains methylated glutamic acid

residues that are important for infection and colonization of kidney and liver cells [118]. It is possible that modification of glutamic acid residues in Msp4 is involved in the specificity of the porin's channel during the binding of the bacteria to ISE6 cells and thus essential for the infection of this cell type. Additionally, methylation of these glutamic acid (E) residues could aid in chemotaxis of the bacteria towards host cells, as previously described for other bacterial taxa [109]. A lack in the ability of the *omt*-mutant to respond to environmental changes when transferred from mammalian to tick cell cultures could explain the continued up-regulation of a set of proteins that are known to be important during infection of mammalian cells. Under these conditions, wild-type *A. phagocytophilum* bacteria presented a subset of proteins previously shown to be highly expressed in tick cell culture, or which have been reported to be involved in infection of ticks [54,55]. It is possible that in addition to a role in binding to host cells, methylation of Msp4 serves to sense changes in the environment including temperature and pH, such as occur when *A. phagocytophilum* is transferred from mammalian to tick hosts, or *vice versa.* These assumptions are supported by the finding that *E. coli* OmpG porin responds to changes in pH [176]. In this context, it is of interest to note that the gene encoding the OMT, APH_0584, is in close proximity to a sensor histidine kinase encoding gene (APH_0582), which is thought to be part of the two component systems of *A. phagocytophilum* [179]. However, more research is needed to determine the exact function of these methyl modifications of Msp4. This is the first report of the methylation of an outer membrane protein in *A. phagocytophilum* by an *A. phagocytophilum O*-methylase, and its importance in infection of tick cells.

Conclusion

As reported in the previous chapter, the OMT is important during binding of *A. phagocytophilum* in ISE6 cells. The binding process of bacteria to tick host cells is related to the methylation of the outer membrane protein, Msp4. Proteomic experiments done in this study showed that Msp4 is methylated at two glutamic acid (E) residues near position 130, which according to the prediction analyses corresponds to one of the beta-strands that forms part of a beta-barrel porin structure. Thus, it is likely that the methylation of these residues is in response to the interaction with the ISE6 cells and affects the conformation and function of this protein, which in turns affects the ability of the bacteria to infect and develop in this cell type.

Figures and Tables

Table 1. Primers used for the amplification of *A. phagocytophilum* and *I. scapularis* genes for the production of recombinant proteins.

Primer name	Primer sequence	Protein target	Organism
rOMT Fw	5' GCT AGA GTC GAC TCA TGT GAG CTT TA '3	GI:88598384	*A. phagocytophilum*
rOMT Rv	5' GCT CAG ATA TCG TGC GCA ATG TCT CT '3	GI:88598384	*A. phagocytophilum*
rOMTns Fw	5' GCC AGA TCA TAT GGT GCG CAA TGT CT '3	GI:88598384	*A. phagocytophilum*
rOMTns Rv	5' CTT TAG CGA ATT CCA TGT GAG CTT TAT '3	GI:88598384	*A. phagocytophilum*
TypA Fw	5' TGA TCC TCG AGT GCA GCA TAT CCA TGG AAC AAT C '3	GI:88607727	*A. phagocytophilum*
TypA Rv	5' GGA TCC ATA TGA TGT CGA GTG CCT ACG ATT CCA '3	GI:88607727	*A. phagocytophilum*
p4416b Fw	5' ATA GCT CGA GCC TAA CAC CAA ATT CCC CAC CGA C '3	GI:88607043	*A. phagocytophilum*
p4416b Rv	5' CAT ATG GTC ATG GCT GGG ACT GAT GTC AGG GCT '3	GI:88607043	*A. phagocytophilum*
msp4 Fw	5' GTA GTG CAT ATG TAC AGA GAA TTG CTG GT '3	GI:88607879	*A. phagocytophilum*
msp4 Rv	5' CGT AGC CTC GAG TCT TGC TCC TAT GTT GAA GCC G '3	GI:88607879	*A. phagocytophilum*
aph0406 Fw	5' GAT AGC ATA TGA TAG TAG CGG TAA AAC GAG TA '3	GI:88607117	*A. phagocytophilum*
aph0406 Rv	5' GTA TGT CTC GAG TAT AAC TTG ACC TCT ATT TAC AC '3	GI:88607117	*A. phagocytophilum*

Table 2. Annealing temperatures used for amplification of *A. phagocytophilum* genes for the production of recombinant proteins.

	Annealing Temperature (°C)	
Primer set	**First 10 cycles**	**Next 20 cycles**
pet29a-p44-16b	50	60
pet29a - typA	43	53
pet29a- msp4	45	52
pet29a - aph0406	43	55
pet29a-0906	46	56

Figure 1. OMT co-localization with bacteria during binding of A. phagocytophilum to ISE6 cells.

(A-B) Bacteria are labeled with FITC-labeled antibodies in the green channel. *A. phagocytophilum* expressing the OMT were labeled with Rhodamine-labeled antibodies in the red channel. Bacteria expressing the OMT were co-labeled and visualed with a yellow color. C) Pre-immune mice serum was used as a negative control. D) Uninfected ISE6 cells were also used as negative control.

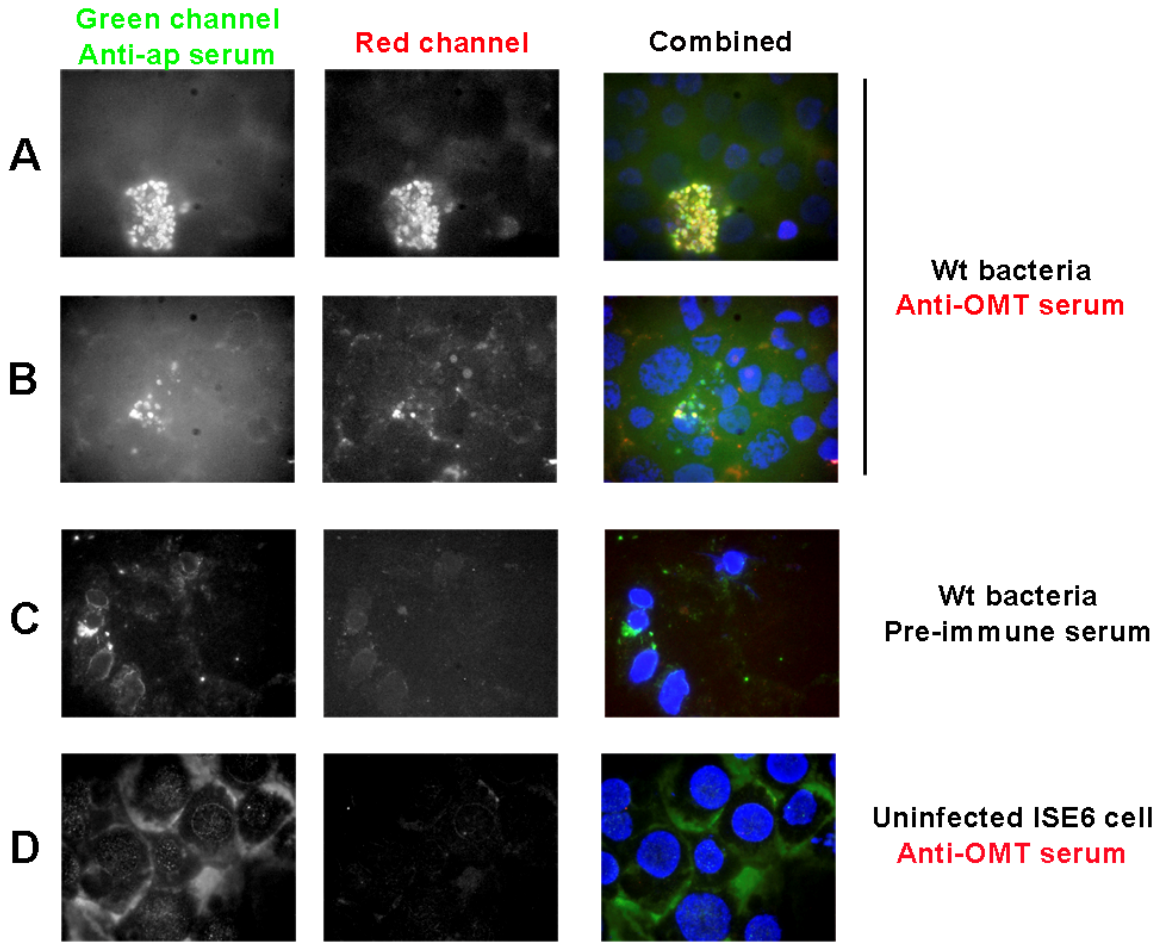


Table 3. *A. phagocytophilum* differentially expressed proteins in the OMT-mutant

Accession Number	Protein ID	Function	Expression in K/O	Ratio KO/Wt
GI88607117	Hypothetical protein APH_0406	Hypothetical	Downregulated	0.1773
GI88607879	Major Surface Protein 4	Hypothetical porin	Downregulated	0.1190
GI88607666	Glutamine synthetase, type 1	Amino acid metabolism	Downregulated	0.0997
GI88607183	Anti-oxidant AhpCTSA family protein	Signal transduction	Downregulated	0.8069
GI88607707	Ankyrin	Membrane interactions	Downregulated	0.5898
GI88606723	Chaperonin GroEl	Stress response	Upregulated	1.5489
GI88607549	Chaperone DnaK	Stress response	Upregulated	1.1503
GI88607105	DNA-directed RNA polymerase beta subunit	Transcription	Upregulated	1.3672
GI88607442	Bifunctional proline dehydrogenase/pyrroline-5-carboxylate dehydrogenase	Amino acid metabolism	Upregulated	1.2042
GI88606872	DNA-directed RNA polymerase beta subunit	Transcription	Upregulated	1.3873
GI88607267	Hypothetical protein APH_0404	Unknown	Upregulated	1.3778
GI88607778	Polynucleotide phosphorylase/polyadenylase	RNA metabolism	Upregulated	1.5950
GI88607654	Hypothetical protein APH_0405	Membrane	Upregulated	1.3533
GI88607567	OMP85 family outer membrane protein	Membrane	Upregulated	1.2934
GI88607014	Leucyl Aminopeptidase	Protein metabolism	Upregulated	1.3933
GI88606911	Hypothetical protein APH_0906	Unknown	Upregulated	1.2636
GI88607774	F0F1 ATP synthase subunit beta	Energy metabolism	Upregulated	1.2689
GI88607426	P44-1 Outer membrane protein	Hypothetical porin	Upregulated	1.1646
GI88607319	Translation initiation factor IF-2	Protein metabolism	Upregulated	1.2237
GI88606885	Hypothetical protein APH_1235	Unknown	Upregulated	1.3020
GI88607566	Co-chaperone GrpE	Stress response	Upregulated	1.3726
GI88607299	OmpA family protein	Membrane	Upregulated	1.3939
GI88607721	Cytochrome C oxidase, subunit II	Energy	Upregulated	1.6799
GI88607259	P44-18ES, expression locus with P44-18	Hypothetical porin	Upregulated	3.1262

Table 4. Pathways that involve proteins that showed a difference in expression in the OMT-mutant during binding and internalization into ISE6 cells.

Pathway	**Pathway ID**	**# proteins downregulated**
Pyrimidine metabolism	Aph00240	1
Purine metabolism	Aph00230	1
Nicotinate and nicotamide metabolism	Aph00760	1
Pathway	**Pathway ID**	**# proteins upregulated**
RNA degradation	Aph03018	3
Pyrimidine metabolism	Aph00240	3
Purine metabolism	Aph00230	2
Nitrogen metabolism	Aph00910	2
Oxidative phosphorylation	Aph00190	2
Alanine, aspartate, and glutamate metabolism	Aph00250	2
Arginine and proline metabolism	Aph00330	2
Glyoxylate and dicarboxylate metabolism	Aph00630	1
Two-component system	Aph02020	1
Glutathione metabolism	Aph00480	1

Figure 2. Spectrum intensity in methyl modified peptides with ratios above 0.7 in A. phagocytophilum proteins.

The numbers under the 115:114 columns and the lines labeled 114.1108 and 115.1079 represent the ratios and intensities of the peptides in the wild-type bacteria. The numbers under the 116:114 and 117:114 fields represent the ratios of the peptides in the mutant bacteria compared to the wild-type and the lines labeled 116.1112 and 117.1146 represent intensities of the peptides in the mutant samples.

Figure 3. Spectrum intensity differences of the methyl modified peptides in MSP4 and APH_0406 in the OMT-mutant versus the wild-type A. phagocytophilum

The numbers under the 115:114 columns and the lines labeled 114.1108 and 115.1079 represent the ratios and intensities of the peptides in the wild-type bacteria. The numbers under the 116:114 and 117:114 represent the ratios of the peptides in the mutant bacteria compared to the wild-type and the lines labeled 116.1112 and 117.1146 represent intensities of the peptides in the mutant samples in A) Msp4 and B) APH_0406.

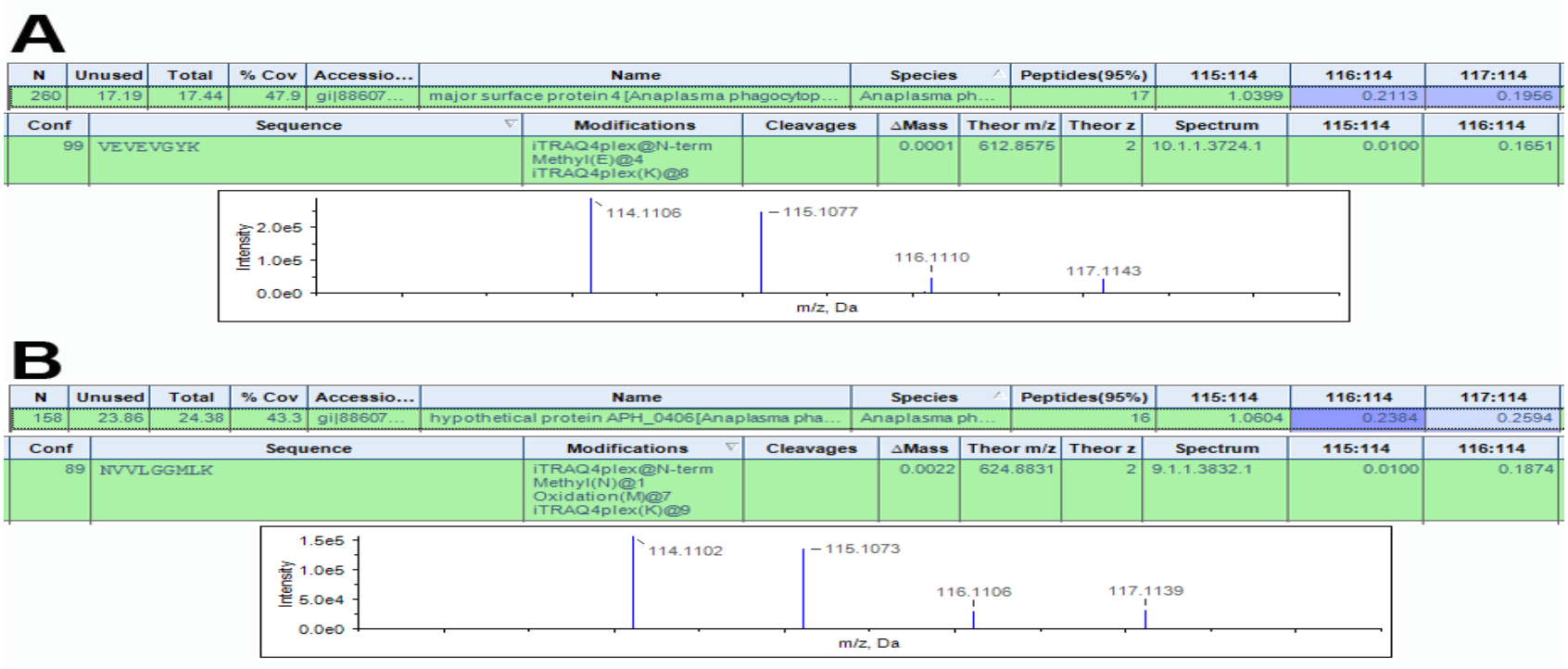


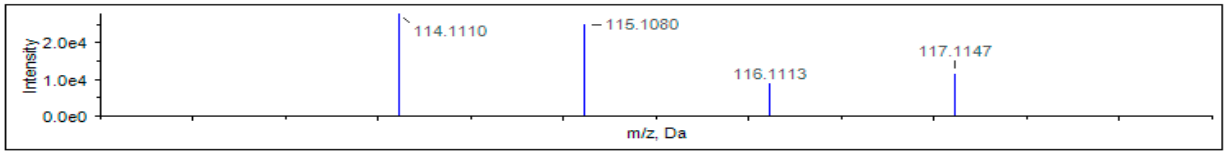
Figure 4. Spectrum intensity differences of the non-methyl modified peptides in MSP4 and APH_0406 in the OMT-mutant versus the wild-type A. phagocytophilum

The numbers under the 115:114 columns and the lines labeled 114.1108 and 115.1079 represent the ratios and intensities of the peptides in the wild-type bacteria. The numbers under the 116:114 and 117:114 represent the ratios of the peptides in the mutant bacteria compared to the wild-type and the lines labeled 116.1112 and 117.1146 represent intensities of the peptides in the mutant samples in A) Msp4 and B) APH_0406.

A

N	Unused	Total	% Cov	Accessio...	Name	Species	Peptides(95%)	115:114	116:114	117:114
260	17.19	17.44	47.9	gi|88607...	major surface protein 4 [Anaplasma phagocytop...	Anaplasma ph...	17	1.0399	0.2113	0.1956

Conf	Sequence	Modifications	Cleavages	ΔMass	Theor m/z	Theor z	Spectrum	115:114	116:114	117:114
77	VEVEVGYKK	iTRAQ4plex@N-term iTRAQ4plex(K)@6 iTRAQ4plex(K)@9	missed K-K...	0.0022	494.9679	3	15.1.1.2918.1	0.9688	0.2546	0.3433



B

N	Unused	Total	% Cov	Accessio...	Name	Species	Peptides(95%)	115:114	116:114	117:114
158	23.86	24.38	43.3	gi|88607...	hypothetical protein APH_0406 [Anaplasma pha...	Anaplasma ph...	16	1.0604	0.2384	0.2594

Conf	Sequence	Modifications	Cleavages	ΔMass	Theor m/z	Theor z	Spectrum	115:114	116:114	117:114
< 1	NVVLGGMLK	iTRAQ4plex@N-term iTRAQ4plex(K)@9		-0.0031	609.8777	2	11.1.1.5223.1	1.4368	0.8801	0.7348

Table 5. *A. phagocytophilum* proteins with reduced peptide methylation in the OMT mutant
* Proteins analyzed with the methylation assay.

Accession Number	Protein ID	Peptides with modifications	Ratio Wt:KO
GI88607441	Branched-chain alpha-keto acid dehydrogenase subunit E2	TLSELSK Methyl(S)@6	0.6651
GI88607727*	GTP-binding protein TypA	INSQVK Methyl (N)@2	0.5106
GI88607117*	Hypothetical protein APH_0406	NVVLGGMLK Methyl(N)@1	0.1773
GI88607879*	Major Surface Protein 4	VEVEVGYK Methyl(E)@4 VEVEVGYK Methyl(E)@2	0.109 0.1178
GI88607043*	P44-16B Outer membrane protein	TKDTAIANFSME Methyl(S)@6	0.4831
GI88607849	Peprotein translocase subunit SecA	RIDNQLR Methyl (D)@3	0.4966
GI88607473	Phosphoribosylamine-glycine ligase	VLVIGSGGR Methyl(I)@4	0.6483
GI88607510	RNA polymerase sigma factor RpoD	AVLADLR Methyl(D)@5	0.6772

Table 6. *I. scapularis* proteins that present a reduction in peptide methylation in the omt-mutant infected cells

Accession Number	Protein ID	Peptides with modifications	Ratio Wt:KO
GI241044082	3-phosphoglycerate kinase, putative	ALDNPSRPFLAILGGAK Dimethyl(R)@7	0.4780
GI241157545	Actin, putative	YPIEHGIVTNWDDMEK Methyl(H)@6	0.7378
GI241326700	Cell division protein, putative	LYEFPCDDEEENKR Methyl(D)@8	0.6880
GI241104748	Fasciclin	SFFNNMLLQTAEGDDKIR Methyl(D)@14	0.6323
GI241703753	Flavonol reductase/cinnamoyl-CoA reductase, putative	EVLEIEPR Methyl(E)@6 Methyl(E)@4 LLLEDGQLR Methyl(E)@4	0.6495 0.6523
GI241558809	Glycoprotein gC1qBP, putative	IEGFDVK Methyl(D)@5	0.6467
GI241263225	Heatshock protein 20.6, putative	DGILTIEAPLPALEAPNR Methyl(E)@14	0.2999
GI241830514	Hsp90 protein, putative	DQVANSAFVER Methyl(D)@1 LMKDILDIL Methyl(D)@7	0.7276 0.5119
GI241600204	Phosphatidylinositol-4-phosphate 5-kinase type II, putative	AEQEAVER Methyl(E)@7	0.5439
GI241586870	Plexin domain-containing protein, putative	DLPVPVTEIPDK Methyl(D)@1	0.6624
GI240974259	Prolyl 4-hydroxylase alpha subunit, putative	GDDGDVPMDEAAVGK Methyl(E)@10	0.6978
GI241627626	RNA-Binding protein musashi	RGGGGGASGGGGGYHPYSR Dimethyl(R)@1	0.4190
GI241720809	TraB domain-containing protein, putative	AVQEAEK Methyl(E)@4	0.4417
GI242004214	Ubiquitin-activating enzyme E1, putative	ITAHENR Methyl(N)@6	0.6694
GI241061134	Voltage-dependent anion-selective channel, putative	VNASLETK Methyl(E)@6	0.6737

Figure 5. Eluted recombinant OMT visualized in stained protein gel.

The eluted recombinant version produced in *E. coli* BL21(DE3) cells visualized by Coomasie blue staining after running in 4-16% gel for 1 hour. Protein standards were loaded in the first lane on the left, the second lane (U) contains lysate from uninduced *E. coli* cells and the last lane (P) contains the eluted protein after His-tag purification with nickel columns. The black arrow head points at the purified protein and the two small gray arrows show proteins from unspecific binding.

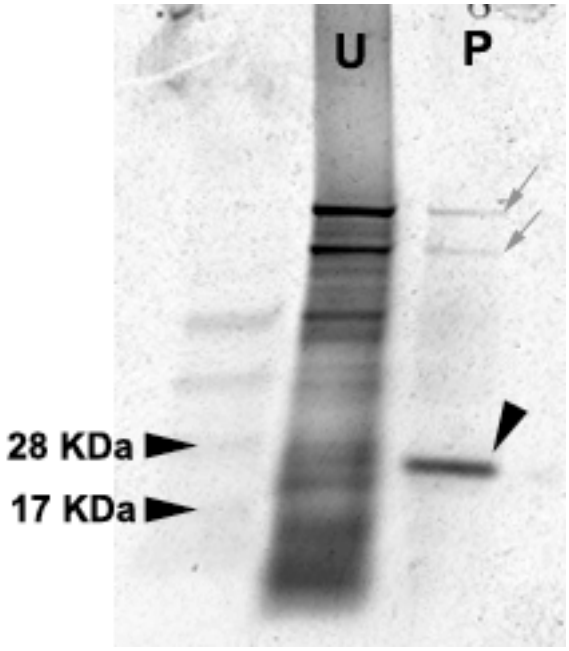


Figure 6. Methylation activity of the rOMT during incubation with A. phagocytophilum and I. scapularis lysates.

Enzymatic activity of the rOMT was measured by the increments in fluorescence in each sample. The lines represent the average fluorescence in each sample for that given time point from three replicates. The red line represents the change in fluorescence when the tick lysate was used in combination with the rOMT. The back line represents the activity of the rOMT using Ap lysate as substrate.

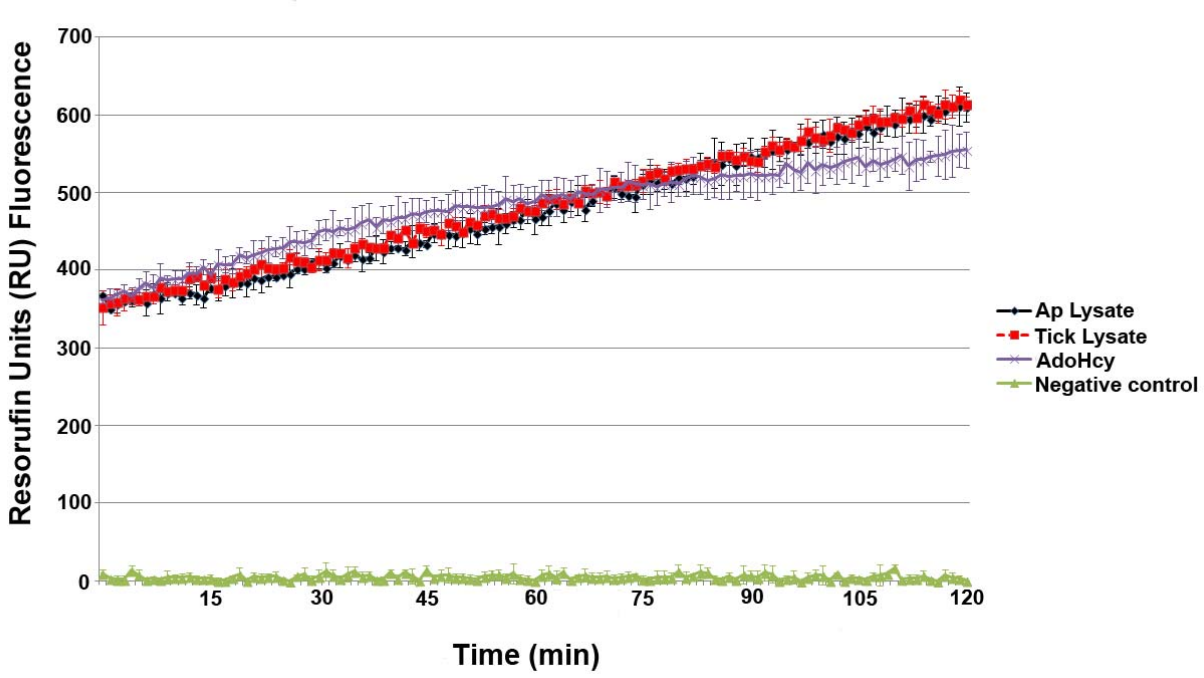


Figure 7. In vitro methylation of recombinant A. phagocytophilum proteins by rOMT

Enzymatic activity of the rOMT was measured by the increase in fluorescence (Resorufin units) in each sample. The lines represent the average fluorescence in each sample at that given time point from three replicates. The blue dotted line represents the reaction using Msp4 as the substrate, which is the only reaction that shows an exponential phase that reached a plateau at around 1 hr 30 mins, typical of an enzymatic reaction. The reactions with APH_0406, TypA, and p44-16b as substrates are represented by the black line, a dotted red line, and dotted gray line, respectively. The control is represented by a green dotted line, but it did not pass background levels.

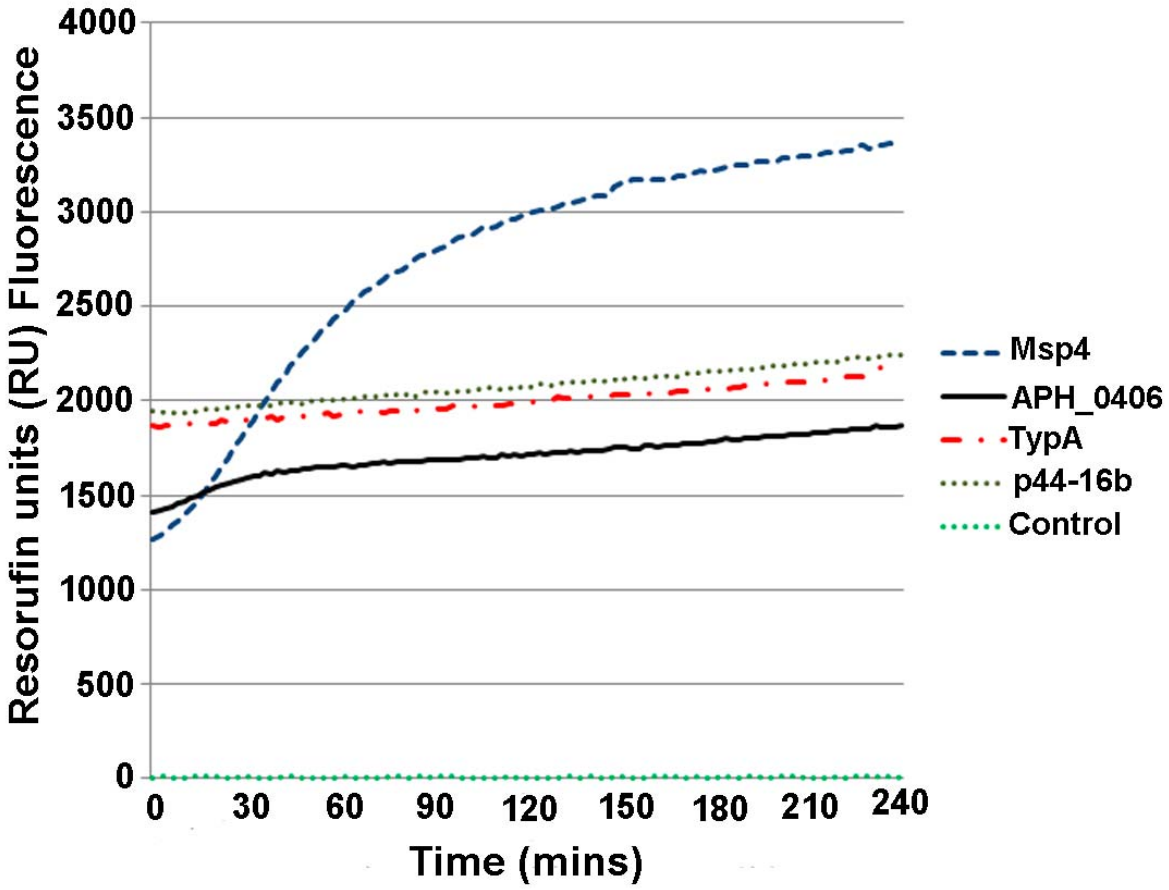
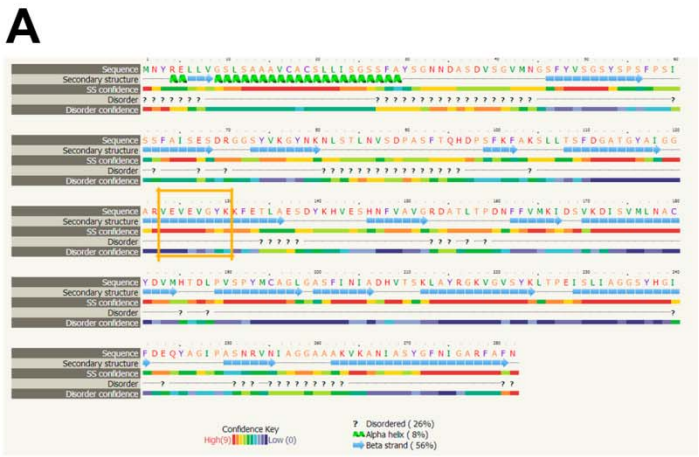
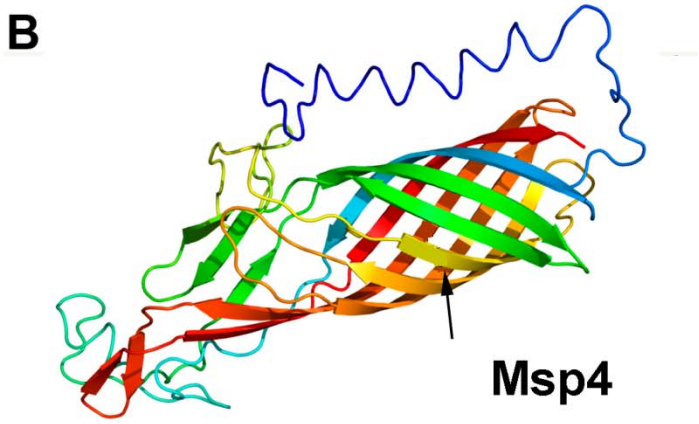
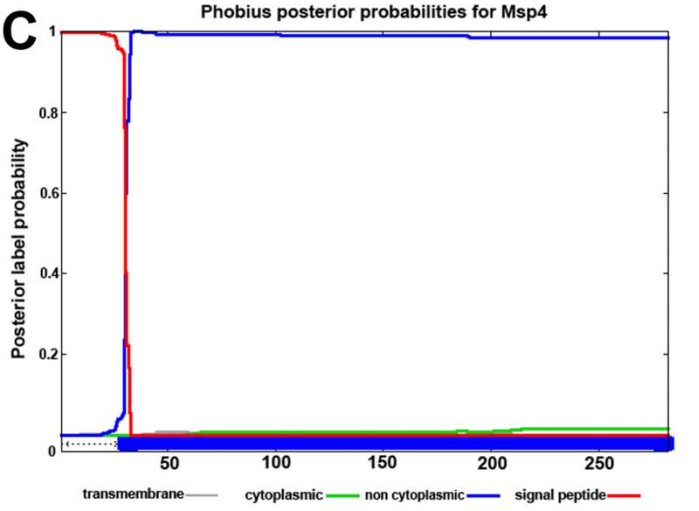


Figure 8. Msp4 putative tertiary structure and potential positioning of methylated glutamic acid (E) residues

A) Putative secondary structure of Msp4 protein sequence according to Phyre2. The yellow square shows the position at the beginning of the 7th beta strand. B) Msp4 tertiary structure showing the typical porin-like structure form by several beta-strands. The black arrow shows the potential positioning of the modified glutamic acid residues. C) Msp4 transmembrane and signal peptide prediction. The red line represents the positioning of a signal peptide according to the protein sequence from GenBank and its probability. The blue line represents the portion of the protein that is non-cytoplasmic and the probability of a right prediction.


A
Sequence
Secondary structure
SS confidence
Disorder
Disorder confidence
Confidence Key
High(9)
Low (0)
? Disordered (26%)
Alpha helix (8%)
Beta strand (56%)


B
Msp4


C
Phobius posterior probabilities for Msp4
Posterior label probability
1
0.8
0.6
0.4
0.2
0
50
100
150
200
250
transmembrane
cytoplasmic
non cytoplasmic
signal peptide

Chapter 4

The *Anaplasma phagocytophilum* hypotethical protein APH_0906 is secreted into the cytoplasm of mammalian host cells

Introduction

Anaplasma phagocytophilum is an obligately intracellular organism and a member of the Anaplasmataceae family along with several pathogens of human and veterinary importance [5]. *A. phagocytophilum* has been recognized as an important tick-borne pathogen of ruminants for over 200 years in Europe, where the infection with this organism has been known as "tick-borne fever", causing important economic losses in many countries [25,27,180]. However, in 1994 three cases of human infection with this pathogen, human granulocytic anaplasmosis (HGA), were reported for the first time [4], and since then the number of human cases have been increasing steadily, making HGA the second most commonly diagnosed tick-borne illness in the US [123]. The human version of the disease is also gaining importance in Europe [181] and Asia [182]. *A. phagocytophilum* is a zoonotic pathogen known to also infect horses, dogs, cats, rodents, domestic and wild ruminants, wild boars, foxes, and thus has a broad vertebrate host range [158]. A characteristic of this organism is its development inside a vacuole to form a "morula" [5] and it is one of the few bacteria that are capable of infecting polymorphonuclear leucocytes, also known as neutrophils, and its progenitors [183] as well as endothelial cells [85].

Because of the intracellular nature of *A. phagocytophilum* and the harsh environment where it develops, within neutrophils, the survival of this bacterium relies on its ability to influence and hijack the host's defensive machinery and metabolism in order to avoid being killed and to obtain nutrients necessary for its growth [62]. Several studies have shown that infection with *A. phagocytophilum* affects the gene expression [74], cytokine production [184], and the ability of neutrophils to adhere to endothelial

cells [185]. It also inhibits apoptosis in neutrophils and tick cells [186,187], and co-opts the host cell's transport system to obtain cholesterol and nutrients [188,189]. These activities, which are shared among pathogens in the family Anaplasmataceae [190], rely on effectors in order to manipulate host cell survival [191]. Although only two proteins have been experimentally shown to be translocated into the cytoplasm and nucleus of host cells during *A. phagocytophilum* infection [191], it is very likely that additional effectors are encoded within the bacterium's genome.

A. phagocytophilum is an organism with a small genome (~1.5 Mbps), but over 70% of its genome encodes to hypothetical genes [128]. Over 40% of *A. phagocytophilum* genome encodes hypothetical proteins and conserved hypothetical proteins. These hypothetical genes comprise the majority of the genes that are unique to this bacterium [128]. One such unique *A. phagocytophilum* gene encodes the hypothetical protein APH_0906 predicted to be an outer membrane protein that is highly expressed during infection of HL-60 and HMEC-1 cells, as demonstrated by whole genome tiling-array analysis [54]. A mutant resulting from transposition of a GFPuv marker gene and spectinomycyn resistant gene [108], into the APH_0906 locus of *A. phagocytophilum* is unable to grow in HL-60 cells as reported in Chapter 2. This mutant demonstrated a slight decrease in its ability to bind to these cells, and was further defective for development in RF/6A and HMEC-1 cells, but able to infect hamsters. In this chapter, I present evidence for translocation of the protein into the cytoplasm and nucleus of infected HL-60 cells as well as in HL-60 and HMEC-1 cells transfected with a construct encoding the entire APH_0906 region. I also determined the relationship and similarities between the homologs of this gene from several strains of *A. phagocytophilum.*

Materials and Methods

Recombinant protein production

Due to the size (4587 bp) of the APH_0906 locus, recombinant versions of the two halves of the APH_0906 protein were produced using the pET29a expression vector (Novagen, Germany). The portion between nucleotides 965395 – 967663 was amplified using primers APH_0906 fw1 (5'- ATG GTA GCA TAT GAT GAC TCT GCT GCT TAA GCC AAA C – 3') and APH_0906 rv1 (5' - GTG ATG CCT CGA GAT CGA TCA GAG TGT CAC CGA GCA T – 3') to obtain the first half of the protein (r1APH_0906), and the second half (r2APH_0906), located between nucleotides 967516 – 969978 was generated using primers APH_0906 fw2 (5'- GTG CGA TCC ATG GAA GAA GAT GGT ACC GCG T – 3') and APH_0906 rv2 (5' - ATT GCT TAG CTC GAG ATG CTG CTG CTG TGA TAC G – 3'). PCR employed *pfu* DNA polymerase (Promega, Madison, WI) under the following conditions: one denaturing cycle at 94 °C for 3 min, 10 cycles with a denaturing step at 94 °C for 1 min, 48 °C for 1 min for annealing, and extension at 72 °C for 2 min, then 20 additional cycles with a denaturing step at 94 °C for 1 min, 60 °C for 1 min for annealing, and extension at 72 °C for 2 min, and a final extension step of 7 min at 72 °C. The amplified products were digested with the restriction enzymes *XhoI* and *NdeI* (r1APH_0906) and *XhoI* and *NcoI* (r2APH_0906), followed by ligation into the vector (pET29a) at 15°C overnight. Plasmids were cloned into One Shot® TOP10 competent cells (Invitrogen, Grand Island, NY) for replication, and purified using the High pure plasmid isolation kit (Roche, Indianapolis, IN). The integrity of the plasmid was checked by sequencing with the T7 promoter (5'- TAA TAC GAC TCA CTA TAG

GG – 3') and the T7 terminator (5'- GCT AGT TAT TGC TCA GCG G – 3') primers at the Biomedical Genomics Center (the University of Minnesota). Sequences were reviewed using MacVector 12.5 (MacVector, Inc., Cary, NC) and when the integrity was confirmed plasmids were transfected into BL21(D3) *E. coli* (New England Biolabs, Ipswich, MA) for expression. BL21(D3) *E. coli* were inoculated into 100 ml of Superior Broth (AthenaES, Baltimore, MD) induced with 200 μM IPTG, final concentration, and incubated at 37°C overnight with constant shaking. Proteins were purified using Ni-NTA Fast Start Kit columns (Qiagen, Germantown, MD) and concentrations measured using the BCA protein assay kit (Pierce, Rockford, IL). The correct molecular weight of recombinant proteins was verified by electrophoresis on a 4 – 15% Mini protean TGX gels (Biorad, Hercules, CA) and bands visualized using silver stain (Biorad, Hercules, CA).

Production of anti-APH_0906 antibodies

Both recombinant r1APH_0906 and r2APH_0906 were dialyzed against tris-buffered saline (TBS) in 3 ml Slide-A-Lyzer Dialysis Cassettes with a 10KDa molecular weight cut-off (Thermo Scientific, Rockford, IL). TBS was changed after 2 hr and dialysis continued overnight. Three 6 – 8 weeks old C57BL/6J mice (Jackson Laboratories, Bar Harbor, ME) were injected subcutaneously at the base of the tail with 100 μg of each recombinant protein emulsified in TiterMax Research adjuvant (CytRx Co., Norcross, GA). Mice were given a second and third booster injection with 100 μg of each recombinant protein 14 days and 24 days later. To obtain antiserum, blood was collected 10 days after the third booster by heart puncture after CO_2 euthanasia. Serum

from mice of the same cohort but that were not immunized with the recombinant protein was used as a control. Serum was prepared by incubating the blood at 37 °C for 60 min, the clot was separated from the wall of the centrifuge tube, and clotted blood was held at 4 °C overnight. Serum was collected after centrifugation at 1610 x g for 10 min and stored frozen at -20°C.

Plasmid design for mammalian cell transfection

pIMGreen was used as the parent plasmid for the construction of pIM 0906. pIMGreen was made by cloning the *Not*I/*Nhe*I fragment from pVITRO4-nGFPLacZ_TDS (Invitrogen, San Diego, CA) into a plasmid between two IR/DR repeats recognized by the Sleeping beauty transposase. The IR/DR(L) repeats (Figure 1) were digested from pT-HB whereas IR/DR(R) was digested from pT-BH (Figure 1) [192]. The complete pIM-0906 was constructed by PCR amplification of the *aph_0906* coding region in two fragments, one half was digested with *NcoI*, whereas the second half was digested with *NcoI* and *AvrII*. The two halves were ligated together and cloned into pIM GFP by replacing the GFP gene in the plasmid with the entire coding sequence of *aph_0906*. This was done to keep the size of the construct small enough to not affect transposition efficiency. Thus, the two IR/DR repeats flanked the entire *aph_0906* and a Neomycin resistance gene (Figure 1), referred to as PIM-0906core.

Transfection of RF/6A, HMEC-1, and HL-60 cells with PIM-0906core containing complete APH_0906

RF/6A and HMEC-1 cells were routinely grown in 12.5 cm^2 flasks with vented caps using 2.5 ml of RPMI medium supplemented with 10% Fetal Bovine Serum (FBS) and 2 mM of L-glutamine and then incubated at 37 °C in a humidified atmosphere of 5% CO_2 in air. Confluent cells layers were subcultured by trypsinization (Gibco, New York) as follows: cell layers were exposed to 1 ml of trypsin with EDTA for 1 minute at 37°C, trypsin was aspirated, and cell layers incubated for an additional 4 min at 37°C. Cells were resuspended in 3 ml of growth medium, and 0.5 ml of the cell suspension was inoculated into a new flask containing 2 ml fo fresh supplemented RPMI medium. This was repeated every two weeks.

For transfection, RF/6A and HMEC-1 cells were subcultured as described, except that 1 ml of cell suspension was inoculated into a 12.5 cm^2 flask containing 1.5 ml of medium. Cells were incubated at 37°C as described above for 24 hr and then transfected at 50 – 57% confluency using Lipoafectamine 2000 (Invitrogen, Grand Island, NY) as follows: 2.5 μg of PIM-0906core and 2.5 μg of SB11 ([193] and Figure 2) plasmids, which is the plasmid containing the transposase, were mixed with 100 μl OPTI-MEM I (Gibco, Grand Island, NY) in 0.5 ml microcentrifuge tubes and combined with 100 μl OPTI-MEM I plus 12.5 μl of the Lipoafectamine reagent. Plasmids and Lipoafectamine were mixed by flicking the tubes gently for 20 sec followed by 5 min incubation at room temperature. Lipoafectamine-plasmid complexes were then inoculated onto RF/6A or HMEC-1 layers in 2 ml of OPTI-MEM I. Cell layers were incubated at 37°C for 1 – 3 days for transfection to occur, after which time the transfection medium was replaced

with RPMI medium containing G418 (100 μg/ml for RF/6A and or 5 μg/ml for HMEC-1 cells) for selection of transformed cells.

To transfect HL-60 cells, the procedure was modified to accommodate non-adherent cells (growing in suspension) using a transfection protocol from Invitrogen. HL-60 cells were counted with a hemacytometer and ~$1x10^5$ cells resuspended in 0.5 ml of RPMI supplemented with 10% FBS and 2 mM of L-glutamine were seeded into a 48 well cell culture plate. SB11 and PIM-0906core (0.5 μg each) were diluted in 100 μl of OPTI-MEM I, 1.75 μl of Lipofectamine 200 was added, and the mixture was incubated for 25 min at room temperature to allow for the formation of complexes. Plasmid-Lipoafectamine complexes (100 μl) were added to each well containing HL-60 cells. Plates were incubated at 37°C in a humid atmosphere of 5% CO_2 in air for 3 days for transfection to occur and 0.5 mg/ml of neomycin was then added for selection. After 1 week, cells were expanded in 25 cm^2 flasks with 20 ml of growth medium containing selection agent at 0.5 mg/ml.

Expression of the APH_0906 protein was verified by western blot. Three weeks after transfection of endothelial and HL-60 cells, protein was extracted from whole cells in 2x sample buffer (6mM Tris-HCL, 10% glycerol, 2% SDS, 5% 2-mercaethanol, and 0.1% bromophenol blue) heated to >95°C for 5 min and centrifugated at 21,000 x g for 10 min to remove insoluble debris. Protein concentrations were measured using the BCA protein assay kit (Pierce, Rockford, IL) and 40 ηg of protein was loaded into protein gels as described in Chapter 3. MultiMark (Invitrogen, Grand Island, NY) protein standard was run in the same gel to verify the size of the protein. Proteins were electrophoresed for 1.5 hr at 100 V and transferred to an Immobilon-P membrane (EMD Millipore Billerica,

MA) at 30 V overnight at 4 – 6°C. The membrane was blocked for 1 hr using 5% no-fat dry milk in Phosphate buffered Saline (PBS) and then incubated with anti-r1APH_0906 or anti-r2APH_0906 antiserum followed by labeling with anti-mouse horseradish peroxidase conjugated IgG. Protein bands were visualized using the Metal enhanced DAB substrate kit (Thermo Scientific, Waltham, MA) as specified by the manufacturer.

APH_0906 localization assay

HGE2, a strain isolated from a patient from MN [92], wild-type bacteria were purified from fully infected ISE6 cell layers, after cells had lifted from the bottom of 25 cm^2 flasks, by vortexing the infected cells with 60/90 grit silicon carbide (Lortone, Inc., Mukilteo, WA) for 30 s followed by filtration through a 2.0 µm pore size filter and centrifugation at 700 x g for 5 min to remove any remaining cell debris. Bacteria were inoculated into 25 cm^2 vented cap flasks containing HL-60 cells. To examine localization of APH_0906 during infection, 5 days p.i., 300 µl of an infected HL-60 culture was diluted with 700 µl PBS, and 40 – 60 µl of the cell suspension was spotted onto microscope slides by centrifugation with a Cytospin 4 centrifuge (Thermo Shandon, Asheville, NC). Cells were fixed with methanol and baked for 1 hr at 50 °C. To further determine localization of the protein at specific times during bacterial replication, infection of HL-60 cells was synchronized as follows: wild-type bacteria from one 25 cm^2 flask containing fully infected cultures of HL-60 cells as described above for infected ISE6 cells. The cell-free bacteria were then combined with 2.5 $x10^5$ cells in a 1.5 ml microcentrifuge tube containing 200 µl of supplemented RPMI medium and incubated for 30 min at 37 °C with mixing every 10 min and then transferred to 4 °C for an

additional 30 min to synchronize the bacterial entry. Cells were then centrifuged down at 700 x g for 5 min at room temperature and unbound bacteria were removed by 3 washes with supplemented RPMI medium. Cells were added to a 15 ml tube containing 5 ml of supplemented RPMI and incubated at 37 °C. Samples were taken 1.5 hr, 12 hr, and 24 hr p.i. and centrifuged onto microscope slides as described above.

To observe localization of APH_0906 within transformed cells, RF/6A and HMEC-1 cells expressing the transgene were trypsinized, diluted 10-fold in PBS, and 60 µl of the cell suspension spotted onto a microscope slide as described above. Transformed HL-60 cells were diluted and then spotted onto microscope slides as described for infected HL-60 cells. All cell spots of transformed cells were fixed in methanol for 10 min and dried at 50°C for 1 hr.

APH_0906 protein was detected by incubation with r1APH_0906 or r2APH_0906 antiserum diluted 1:250 in PBS with 3% bovine serum albumin (BSA). Cell spots were incubated with antibody in a humid environment for 2.5 - 3 hr at room temperature. The slides were washed 3 times in PBS and blocked in 3% BSA in PBS for 10 min. The protein was then labeled with an anti-mouse antibody conjugated with FITC (for RF/6A cells transformed with PIM-0906core) or Cy3 (for infected and transfected HL-60 cells) at 1:500 dilution by incubating spots for 1 - 2 hrs at room temperature. All *A. phagocytophilum* bacteria were labeled with dog anti- *A. phagocytophilum* antibody and FITC labeled anti-dog IgG to aid in identification of the bacteria, as described previously but using a 1:500 dilution for the dog anti-serum. Cell nuclei were labeled using DAPI present in VectaShield mounting medium (Vector laboratories, Burlingame, CA). Samples were viewed with a Nikon Eclipse E400 microscope (Nikon Instruments,

Melville, NY) and imaged with a Nikon DXM1200 digital camera (Nikon Inc., Melville, NY), using Nikon ACT-1 software.

Sub-cellular localization and identification of a putative Nuclear Localization Signal

The APH_0906 (GI:88597974) protein sequence was used to predict the putative subcellular localization of the protein within the bacteria and within the eukaryotic host cell in addition to other information. Subcellular localization, positioning of binding sites, amino acid composition, presence of transmembrane helices, and exposed residues were predicted using PredictProtein (https://www.predictprotein.org/). Since analyses using CELLO and P-sortb in Chapter 2 suggested that APH_0906 was an outer membrane protein which differed from the analysis done herein, Nuc-PLoc (http://www.csbio.sjtu.edu.cn/bioinf/Nuc-PLoc/) was used to verify the results obtained with PredictProtein. The presence of Nuclear Localization Signal (NLS) residues within the protein was predicted using NLS_Mapper (http://nls-mapper.iab.keio.ac.jp/cgi-bin/NLS_Mapper_form.cgi), NucPred (http://www.sbc.su.se/~maccallr/nucpred/cgi-bin/single.cgi), and NLStradamus (http://www.moseslab.csb.utoronto.ca/NLStradamus/).

A region similar to BRO (chromatin-associated), anti-repressor motifs, and predicted NLSs was also predicted in the protein homolog from four additional *A. phagocytophilum* isolates (HGE1, Dog2, JM, and CRT38) from a human patient, a dog, a jumping mouse (*Zapus hudsonius*), and a black-legged tick collected from deer [32,194-196]. A phylogenetic tree showing the relationship of these protein sequences was produced using MEGA 4.0 [136] after generating a sequence alignment of all the proteins in MacVector 12.0 (MacVector Inc, Cary, NC). Additionally, PredictProtein was used to

determine if the binding regions of the protein were conserved among the different homologs.

Results

APH_0906 is translocated into the cytoplasm and nucleus of HL-60 cells during infection with A. phagocytophilum

In Chapter 2, I reported the upregulation of the expression of APH_0906 during infection of HL-60 cells starting 72 hr (3 d) post-inoculation (late phases of infection). Because of those results, I performed Immunofluorescence Assays (IFAs) to determine the localization of APH_0906 during late phases of infection in HL-60 cells, which showed that the protein is secreted into the cytoplasm and translocated into the nucleus of infected cells (Figure 3). APH_0906 appears to be transported to the cytoplasm of the host cell from the morulae as the protein is observed surrounding the morulae, but not in association with bacteria inside morulae (Figure 3A and C). The protein appeared as granules that accumulated in the cytoplasm of infected cells surrounding large morulae and throughout the cytoplasm of infected cells. In some cases, it was not labeled by the anti-*A. phagocytophilum* but that can be visualized in the DAPI channel serum (Figure 3A), probably due to inefficient labeling of the bacteria. In the nucleus the protein localized to diffused structures and was not evenly distributed as seen in the cytoplasm in most cases (Figure 3A and C). APH_0906 co-localized with bacteria released from cells that had lysed due to heavy infection (Figure 3B), and less so with extracellular bacteria that were seen binding to cells (Figure 3C). Uninfected cells incubated with the anti-

APH_0906 serum were not labeled, suggesting that the serum raised against the recombinant protein was specific to APH_0906 (Figure 3D).

To determine if a similar localization was observed at specific time points early and late during infection, infection of HL-60 cells infection was synchronized as described above. Synchronized cultures were sampled at 1.5, 12, and 24 hr p.i. and the localization of the protein was determined using IFAs, as described above. These experiments confirmed results obtained with unsynchronized cultures. Cells incubated for 24 hr with bacteria contained several morulae and APH_0906 was observed both in the cytoplasm and in the nucleus (Figure 4A and B). In the nucleus, APH_0906 appeared in diffused islands and localized to parts of the nucleus (Figure 4A), whereas in the cytoplasm the protein appeared more granular and distributed randomly (Figure 4B). At 12 hr p.i (Figure 4C), the protein was found in the nucleus and cytoplasm of the host cells as well as localized to the morulae. At 1.5 hr, the protein mostly co-localized with individual bacteria and bacteria forming morulae (Figure 4D). APH_0906 protein was not detected in uninfected cells (Figure 4E), as reported above, and pre-immune serum did not label any protein or structure of *A. phagocytophilum* infected cells (Figure 4F).

APH_0906 protein translocated into the nucleus of transformed HL-60 and RF/6A cells

In order to test whether or not the cells could express the APH_0906 protein, RF/6A cells were transfected with only the PIM-0906core plasmid (Figure 1) without the SB11 plasmid ([193] and Figure 2). This is referred as “first trasnfection” from herein. RF/6A cells transfected with PIM-0906core plasmid (first transfection) as well as those

transfected using the PIM-0906core plasmid and SB11 (second transfection) grew more slowly than wild-type cells (data not shown). After 3 weeks of continuous culture under G418 selection, RF/6A cells from both transfections were subcultured and samples were taken for protein extraction. Transfected RF/6A cells were subcultured every 3 – 4 weeks thereafter using a 1:2 dilution, reflecting their reduced growth rate compared to wild-type cells which were subcultured by 1:5 dilution every 2 weeks. Transfected cells expressed a protein of approximately ~160 kDa, migrating just above the 148 kDa marker (Figure 5A and B), which correlates well with the 166.8 kDa mass of APH_0906 predicted by Expasy (http://web.expasy.org/cgi-bin/compute_pi/pi_tool) and Science Gateway (http://www.sciencegateway.org/tools/proteinmw.htm). Protein samples from tranfected and wild-type cells were analyzed by Western blot to determine expression of APH_0906 protein. The APH_0906 protein was detected in RF/6A cells after both transfection attempts (Figure 5A and B). Immunofluorescence assays (IFAs) identified the localization of the protein within transfected cells. Only a few (<1%) transfected cells from the "first transfection" showed expression of the protein after subculturing (data not shown). However, the cells that harbored the PIM-0906core plasmid showed strong expression of the protein localized to the nucleus of the cells (Figure 6). Furthermore, APH_0906 appeared to accumulate to some areas of the nucleus whereas other parts of the nucleus showed no protein presence at all (Figure 6). The maintenance of these cells was discontinued and cells were stored in liquid nitrogen. In the second transfection attempt using SB11 in RF/6A, the proportion of cells that expressed the protein increased to almost 80 % (data not shown) and this population of cells is considered to be transformed and is being continuously grown.

By comparison, transfected HL-60 cells could be subcultured under selection with neomycin by passing around 6.4 x 10^4 cells to a new 25 cm^2 flask containing fresh medium once a week. This subculture schedule was similar to that for wild-type HL-60 cells. Samples for protein analysis were taken from transfected HL-60 cells after the second subculture (3 weeks after transfection). Like transformed RF/6A cells, transformed HL/60 cells expressed the APH_0906 protein (Figure 5B). Transfection of HMEC-1 cells was not achieved and APH_0906 was not detected by Western blotting (data not shown). IFAs with transfected HL-60 cells showed that the protein was translocated to and distributed in the nucleus of these cells similarly to what was observed in RF/6A cells (Figure 7). Transfection of cells with PIM-0906core in the presence of SB11 yielded a population of cells that exhibited greater variability in transgene expression; however, most of the cells expressed the protein in relatively high amounts (Figure 7). Thus, expression of the protein by two different cell lines resulted in the translocation of APH_0906 to the nucleus of the cells (Figures 6 and 7).

Nuclear localization signals (NLSs) are conserved between human pathogenic *A. phagocytophilum* strains but differs in non-human pathogenic strains

The sequences of several *A. phagocytophilum* strains are available in GenBank, including one isolated from a human (HGE1 accession numbers: NZ_APHH01000002.1 and NZ_APHH01000001.1), one isolated from a dog (Dog2, accession number: CP006618.1), a jumping mouse isolate (JM, accession number: CP006617.1). There is also a sequence from a deer tick isolate Ap-Variant 1 found in white-tailed deer that has been shown to not infect mice and is suspected to not infect humans ([33]; CRT38

accession number: NZ_APHI01000001.1 and NZ_APHI01000002.1). The putative protein sequences from all the homologs of APH_0906 in these genomes were analyzed to identify putative DNA binding domains, the BRO-domain, predicted subcellular localization, and the presence of putative Nuclear Localization Signals (NLSs). The BRO (antirepressor) domain was detected in APH_0906 homologs from all strains except the Ap-Variant 1 strain (CRT38) (Table 1), whereas the presence of several NLSs was detected in all the strains. Two of these NLSs were shared between all of the sequences (YRTAVKKVKLA at position 202 or 165 and KKPSTTVPKKPPRPARGAK at position 1451) (Table 1). The NLS DLRKRHRWA identified by NLSMAPPER at position 388 was detected only in the protein sequence of APH_0906 from the jumping mouse (Table 1), while a putative NLS (ENAVKERMLPALTLYAKRIAKR at position 610) was present in all homologs except the one from Ap-Variant 1 (CRT38) homolog Table 1).

APH_0906 from all sources was predicted to be an outer membrane protein (Psort), secreted (ProteinPredict) and translocated to the nuclei of the host cells (Nuc Ploc) (Table 1). Overall, the APH_0906 homolog from CRT38 presented the most differences in its protein sequence, only 82% identity with the other homologs, and was distinct from the cluster formed by the protein from the jumping mouse, dog, and human strains in a Neighbor Joining tree (Figure 8). Interestingly, the APH_0906 homolog from HGE1 was more closely related to that from dog strain than to the homolog from another human isolate, HZ (Figure 8). An analysis of all putative binding sites in APH_0906 homologs revealed significant differences in the number and localization of binding sites throughout the different homologs (Figure 9). However, all except the HZ isolate contained a central and a C-terminal cluster binding site (Figure 9). APH_0906 from HZ

presented the most differences in the binding sites, which appeared to be more dispersed throughout the protein. The secondary structures of the different homologs and their exposed and hidden residues were highly conserved and almost identical (Figure 9).

Discussion

Several human pathogenic viruses, bacteria, and protozoa, transport proteins and other molecules, known as effectors, to the different compartments of eukaryotic host cells, including the cytoplasm and nucleus to manipulate the cell's machinery to their advantage [197]. The mechanisms that bacteria use to survive within host cells include hijacking of cell signaling pathways (MAPK, G-proteins, and lipid signaling), the manipulation of actin and tubulin, targeting of ubiquitination, and possibly other unknown mechanisms [197,198]. One example of proteins that translocate into the nucleus of cells is the IpaH protein from *Shigella flexneri*, which is expressed by the bacteria only after entry into host cells, delivered into the cell cytoplasm and then translocated into the nucleus using the actin and tubulin machinery of mammalian cells [199]. This protein has been shown to reduce the activation of the nuclear factor KB (NF-KB) by promoting ubiquitination of NEMO/IKKγ and leading to its degradation, thus affecting the inflammatory response of the host [200]. Like this protein, most effector proteins identified and described to date are secreted by Type Secretion Systems (TSS) I, II, III, and IV [201]. However, not all effectors and secreted proteins present TSS signal peptides and some are secreted using other systems, e.g., through their own transport channels or *via* outer membrane vesicles [202].

Like most intracellular pathogens, *A. phagocytophilum* has been shown to secrete effector proteins into the cytoplasm and nucleus of its mammalian host cell [73,188]. *A. phagocytophilum* is an obligately intracellular bacterium that develops in neutrophil granulocytes and their bone marrow progenitors [183]. Neutrophils are highly efficient phagocytes and play an important role in the innate immune system. To survive and replicate in this normally hostile environment, *A. phagocytophilum* has evolved several mechanisms, including the injection of effectors [48]. The *A. phagocytophilum* genome encodes homologs of the T1SS and T4SS found in other bacteria [48]. But the two effectors described to date in *A. phagocytophilum*, i.e., AnkA and Ats-1 [48,73,188], are secreted by the T4SS. AnkA is a 160 kDA protein that is secreted into the cytoplasm of neutrophils and then translocated into their nucleus [73] where it binds to DNA and nuclear proteins [73]. In the nucleus AnkA binds to AT-rich regions of the *CYBB* promoter causing a decrease in the acetylation of histone 3 (H3), thus modulating the expression of the promoter and several immune genes including *RAC2, MPO, BPI,* and *MYC* [104].

Even though AnkA is the only *A. phagocytophilum* protein that has been shown to be translocated into the nucleus and affect host cell gene expression, it is likely that its genome encodes additional nuclear effector proteins. In the present study, I used IFAs to show that the hypothetical protein APH_0906 was expressed early during infection in bacteria adherent to host cells. As infection of the cells progressed, secretion of the protein into the cytoplasm became obvious (Figures 3 and 4), and in heavily infected cells, the protein was subsequently translocated into the nucleus (Figure 2 and 3). The distribution of the protein in the nucleus appeared to follow diffused structures perhaps

representing chromatin or nucleoprotein (Figure 3 and 4). In RF/6A and HL-60 cells that had been transfected to express the entire Aph_0906 protein (Figure 6 and 7), the same pattern was observed, suggesting that the translocation of the protein was dependent on the host cell machinery and was independent of the presence of the pathogen in the cells.

In eukaryotes, proteins with molecular mass of > 50 kDa cannot pass unaided through the nuclear membrane and require NLSs in order to be translocated into the nucleus [201]. NLSs consisting of one (monopartite) or two (bipartite) short amino acid sequences rich in positively charged amino acids (K and R) are recognized by carriers of the importin family (α or β) which effect translocation into the nucleus [201]. Intracellular pathogens exploit the host cell machinery by including NLSs within bacterial effector amino acid sequences [201]. However, the NLSs present in many bacterial effector proteins can differ from the standard NLSs found in eukaryotes, making it difficult for software to correctly predict the localization of many of these effector proteins [203]. This is true for 49 effector proteins in *Helicobacter pylori*, which according to Psortb, a bacterial protein localization predictor, are assigned to the cytoplasm, inner membrane, or outer membrane of the bacteria and not predicted to be secreted into the host cell [203]. These proteins were then analyzed for the presence of NLSs, based on amino acid composition, and then experimentally shown to localize to the nucleus of COS-7 cells transfected to express these effectors [203]. The Artemis genome browser used for tiling array analysis [54] and my research presented in Chapter 2, predicted APH_0906 to be an outer membrane protein and a signal peptide for its secretion was not identified. However, the IFA experiments employing infected cells as well as APH_0906 transfected cells demonstrated that the protein was being secreted and

translocated into the cytoplasm and nucleus of host cells. Therefore, I used ProteinPredict, the program that had successfully identified NLSs in *H. pylori* to further examine the status of this protein and its homologs in several strains of *A. phagocytophilum.* Unlike CELLO and Psortb, ProteinPredict and a suite of different bioinformatics software programs suggested that all APH_0906 homologs were secreted and contained several NLSs (Table 1).

The variations in the NLSs present in APH_0906 homologs from *A. phagocytophilum* isolates derived from different hosts prompted an analysis of any anti-repressor (BRO-like) sequences and the positioning of putative protein binding sites within the protein. All APH_0906 homologs except that in CRT38 (Ap-Variant 1) possessed the BRO-like anti-repressor sequence (Table 1 and Figure 9). Research examining the phylogenetic relationship of *ankA* genes in *A. phagocytophilum* strains from different hosts showed that the *ankA* genes from isolates of human origin clustered in a group together with dog, horse, deer, cat, and sheep strains. By contrast, *ankA* genes from roe-deer associated strains not known to be pathogenic to humans clustered in groups distant from human pathogenic strains [204]. Moreover, AnkA proteins from the divergent clusters contained different numbers of ankyrin repeats, which are postulated to mediate interactions with other proteins [204]. My analysis of APH_0906 homologs similarly indicated that the Ap-Variant 1 protein from a strain infectious for white-tailed deer (a close relative to European roe deer) diverged from APH_0906 cluster representing isolates from humans, a dog, and jumping mouse (Figure 8). When the putative protein binding sites from each of the homologs were analyzed, the dog, jumping mouse, and HGE1 (isolated from a human in WI) showed the most similarity, with only

one small region at the beginning of the HGE1 protein being different from the other two homologs (Figure 9). While the pathogenicity of the Dog2 and JM (jumping mouse) strains for humans has not been proven, they are clearly related. By contrast, the Ap-Variant 1 has not been associated with human cases of anaplasmosis and cannot infect mice or HL-60 cells experimentally [205], It is possible that similarly to AnkA proteins, the sequence and structures of the different APH_0906 homologs depend on the host cells which the strains infect. It was surprising that the APH_0906 homolog from the human isolate HZ differed from all others with respect to protein binding sites. Ap-Variant 1 showed the same central and C-terminal binding regions in the sequence as did isolates JM, Dog2 and HGE1 (although minor differences were seen at the N-terminus), whereas the protein binding sites in HZ APH_0906 were distributed throughout the whole sequence with no pattern evident (Figure 9). Since the sequence of the APH_0906 homolog of *A. phagocytophilum* HGE (the transposon mutant with the disrupted *aph_0906* gene used here) has not been determined, the exact structure, localization, and presence or absence of a BRO-domain and NLSs in the protein is unknown. Although it is tempting to speculate that the BRO-domain and other features are conserved in the APH_0906 of HGE2, the differences in the homolog from isolate HZ demonstrate that human-derived isolates may not all conform to each other.

The results presented here and in Chapter 2 suggest that APH_0906 is an *A. phagocytophilum* effector protein that is secreted into the cytoplasm of neutrophils and possibly endothelial cells and is translocated into the nucleus using the host cell machinery. Once in the nucleus, it is possible that APH_0906 binds to specific DNA sequences that are repeated throughout the genome or to proteins that recognize specific

DNA sequences in the host nucleus in a similar manner as that shown for AnkA [206]. Whether APH_0906 plays a role in host-specificity of *A. phagocytophilum* remains to be determined. Additional sequences from different strains that have shown human or non-human specificity will need to be analyzed to draw a conclusion.

Conclusion

In conclusion, APH_0906 is a protein secreted into the nucleus of host cells where it appears to interact with specific regions of the DNA or with specific host nuclear proteins. The sequence, structure, and presence of known DNA binding-motifs appears to change depending on the host species that the strain infects and anti-repressor (BRO-domain) was lacking from a strain not known to infect humans. Using random mutagenesis, we have identified a hypothetical protein predicted to localize to the bacterial outer membrane, which is consistent with experimental evidence that it is secreted and translocated into the nucleus. These studies support the validity of functional genomic analyses based on random mutagenesis to identify genes involved in bacterial pathogenesis of human anaplasmosis.

Figures and Tables

Figure 1. PIM-0906core plasmid.

Plasmid used for the transfection of HL-60 and HMEC-1 cells. The plasmid contains the entire *aph_0906* coding region (lighter blue arrow) driven by the CAGGS promoter (red arrow). The EM7 (yellow arrow) promoter drives the kanamycin/neomycin (dark blue arrow) resistance gene and a polyA signal (light blue) at the end of the resistance gene. All these genes were flanked by two IR/DR repeats (pink blocks) recognize by the transposase.

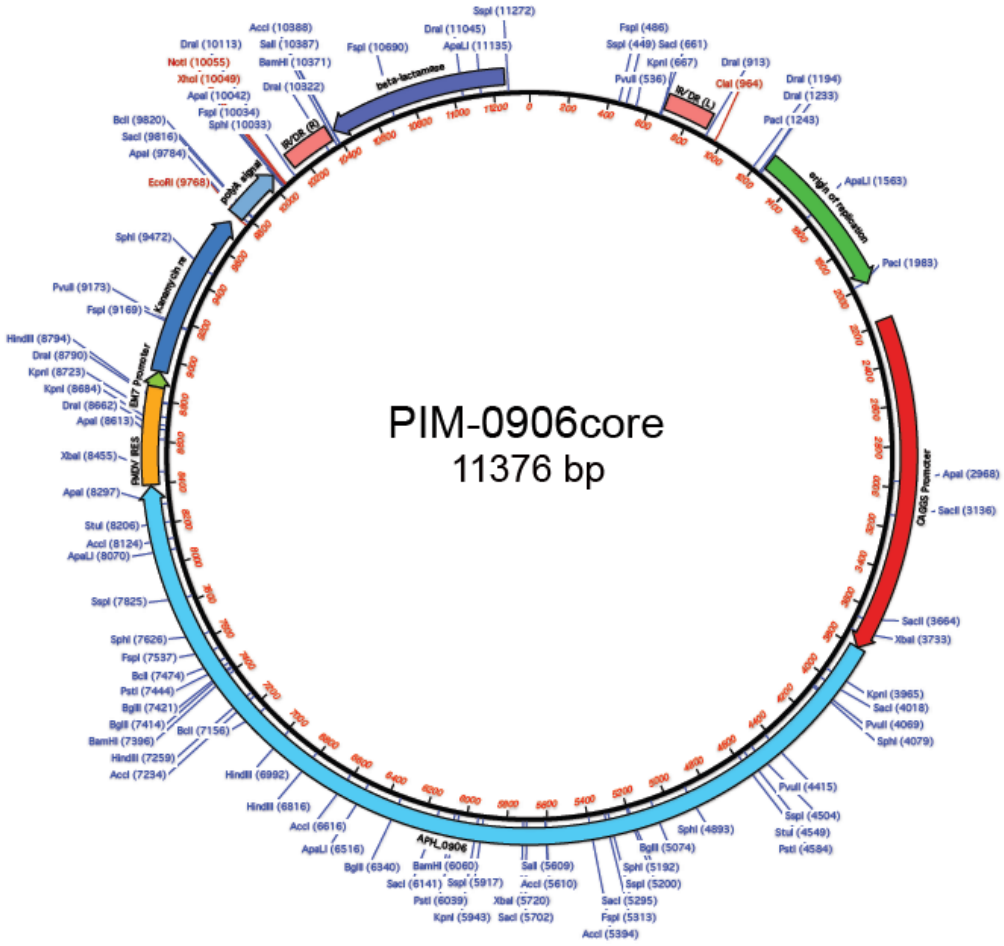


Figure 2. SB11 Plasmid.

Map available in the addgene website (http://www.addgene.org/26552/). The blue arrow (ORF frame 2) represents the full coding sequence for the optimized Sleeping Beauty version 11 [193]. The green arrow (CMV) represents the position of the CMV2 promoter which contains an intron (red box) for optimal expression of the protein. The red arrow represents the ampicillim resistance gene for selection in *E. coli.*

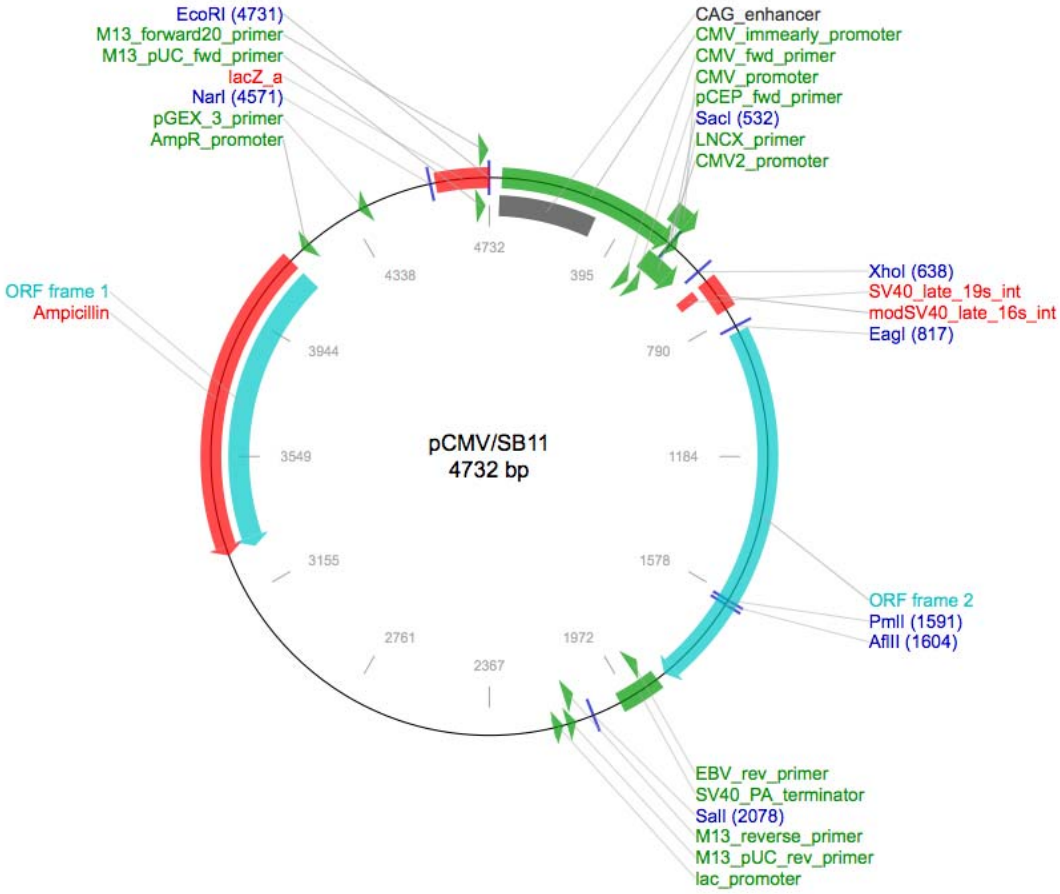


Figure 3. Immunofluorescence Assays showing the localization of the hypothetical protein APH_0906 in HL-60 cells during infection with A. phagocytophilum 5 days after inoculation.

Hypothetical protein *Aph_0906* labeled using serum raised against the first half of the protein (r1APH0906) is shown in the first column with the red channel (first column). *A. phagocytophilum* was labeled using serum against whole bacteria and is shown in the green in the second column). Bacteria are pointed at in all infected cells by green arrow heads. The nuclei of the cells and bacteria were labeled with DAPI and are shown in the blue channel (third column). The fourth column shows all channels combined. All cells presented in the figure represent cells from unsynchronized heavily infected cultures at different levels of infection. A) Cell with several internalized bacteria forming morulae. B) and C) cells with several morulae forming bacteria and some extracellular (bound and unbound) bacteria. D) Uninfected HL-60 cells (control).

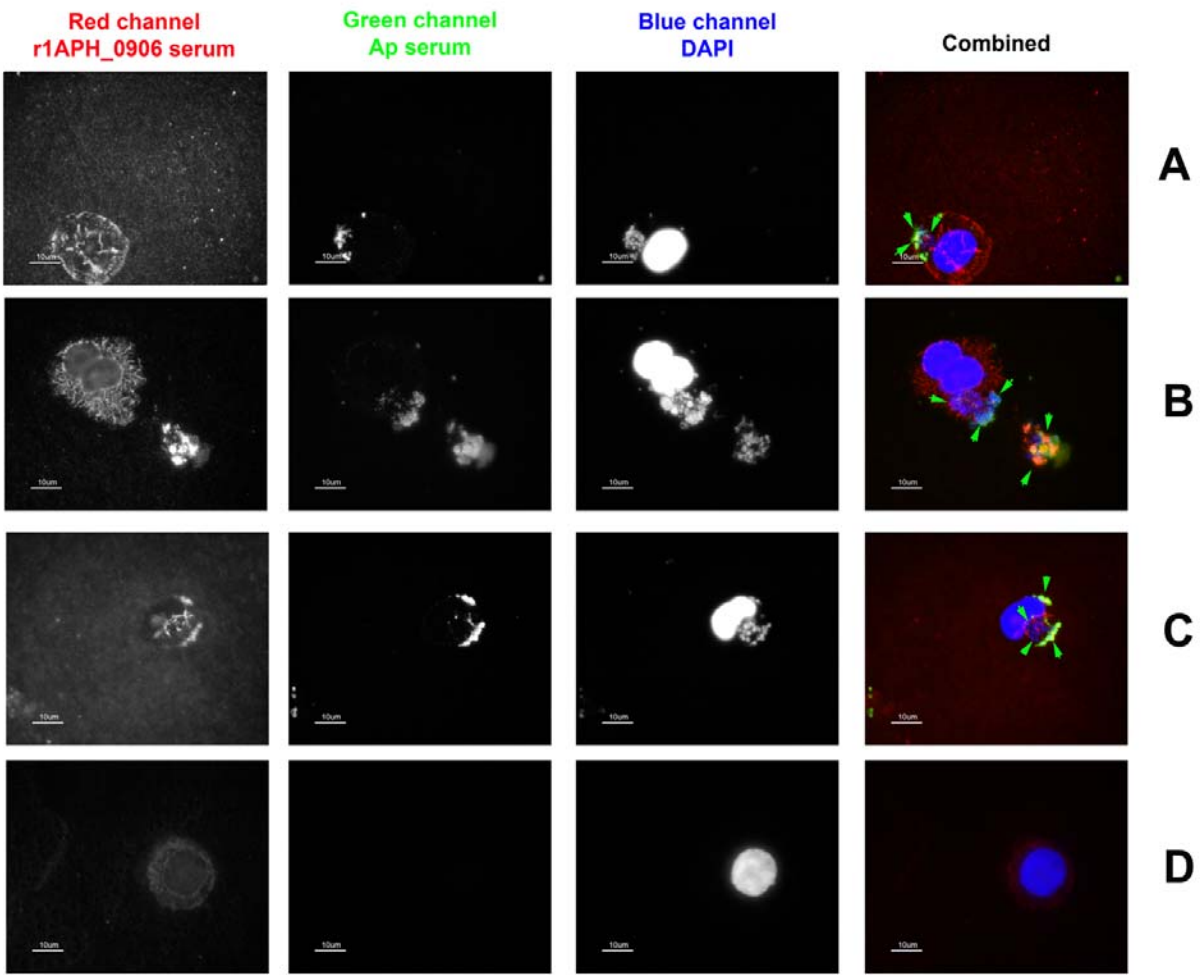

Red channel
r1APH_0906 serum
Green channel
Ap serum
Blue channel
DAPI
Combined
A
B
C
D

Figure 4. Immunofluorescence Assays showing the localization of the hypothetical protein APH_0906 in HL-60 cells with A. phagocytophilum synchronized infection.

Hypothetical protein *Aph_0906* labeled using serum raised against the first half of the protein (r1APH_0906) is shown in the first column with the red channel (first column) in most panels. Pre-immunized serum was used on infected cells as negative control. *A. phagocytophilum* was labeled using serum against whole bacteria and is shown in the green channel (the second column). Bacteria are pointed at in all infected cells by green arrow heads. The nuclei of the cells and bacteria were labeled with DAPI and are shown in the blue channel (third column). The fourth column shows all channels combined. HL-60 cells infected with cell-free *A. phagocytophilum* for synchronization of the cultures. A and B) 24 hr, C) 12 hr, and D) 1.5 hr. E) Shows the negative control consisting of uninfected HL-60 cells incubated with r1APH_0906 serum. F) Infected cells incubated with pre-immune mouse serum.

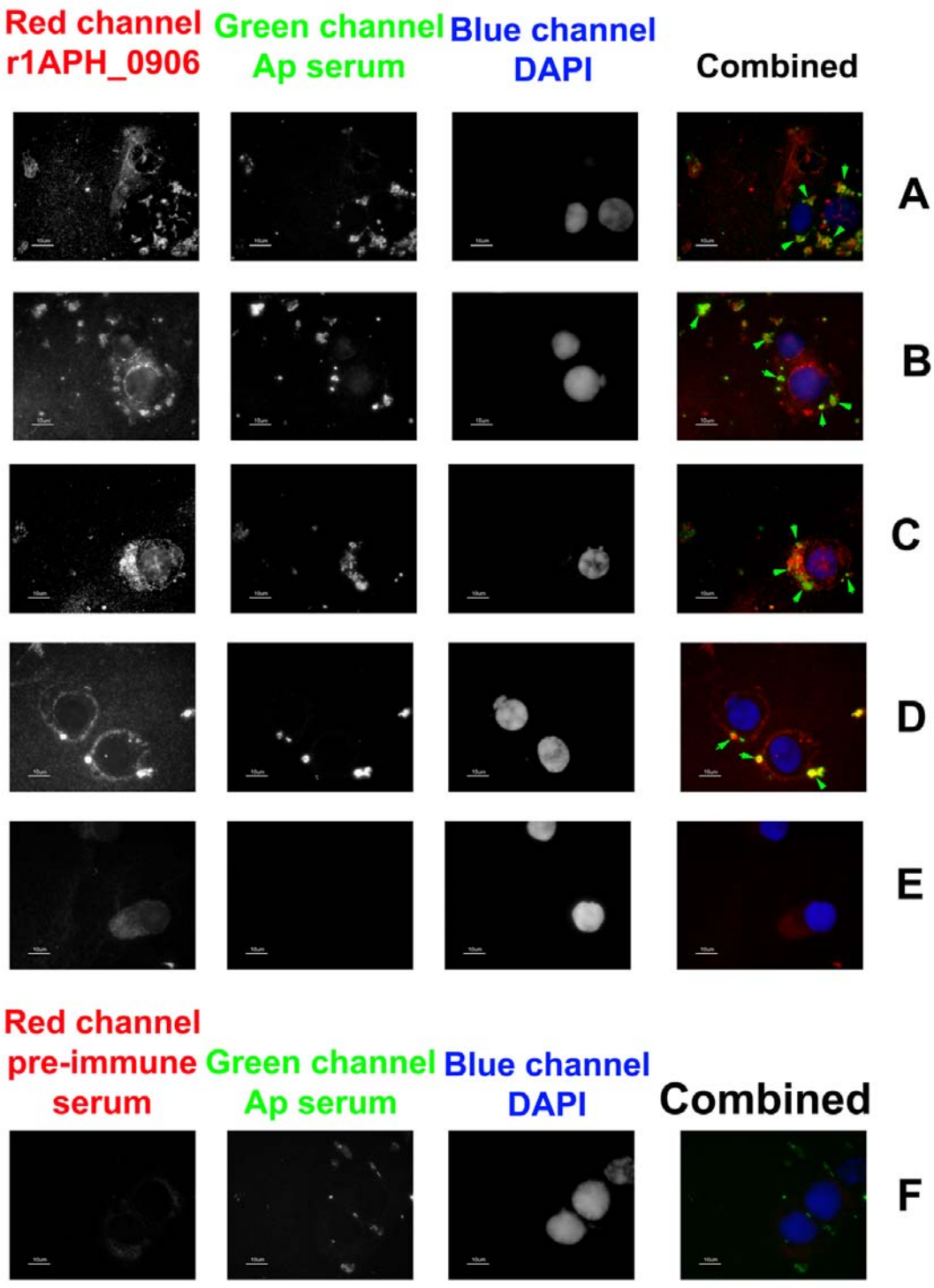

Red channel
r1APH_0906
Green channel
Ap serum
Blue channel
DAPI
Combined
A
B
C
D
E
Red channel
pre-immune
serum
Green channel
Ap serum
Blue channel
DAPI
Combined
F

Figure 5. Western blots of transfected cells expressing APH_0906 detected with r1APH0906 serum against the first half of the protein.

A) First transfection event WB shows RF/6A cells transfected only using the PIM0906 core plasmid without the transposase. B) The second transfection event WB shows the protein band in RF/6A and HL-60 cells transfected with both the PIM0906 core and SB11 plasmid containing the Sleeping Beauty transposase. The red arrow points at the APH_0906 band.

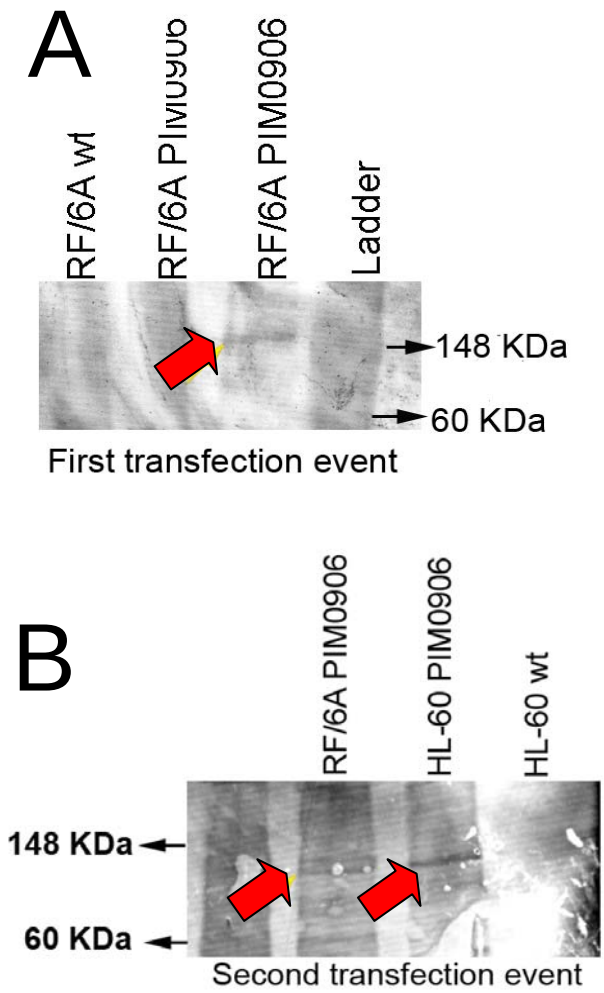


Figure 6. Translocation of APH_0906 to the nucleus of transfected RF/6A cells.

APH_0906 within transfected cells was labeled using antibodies against the first half of the protein (r1APH_0906) and then detected using FITC-labeled anti-mouse antibodies. The localization of the protein is shown in green fluorescence (First column), whereas the nucleus is shown by DAPI staining in the blue channel (Second column). The third column shows the combined channels. The fifth row shows wild-type RF/6A (control) cells incubated with r1APH_0906 antibodies.

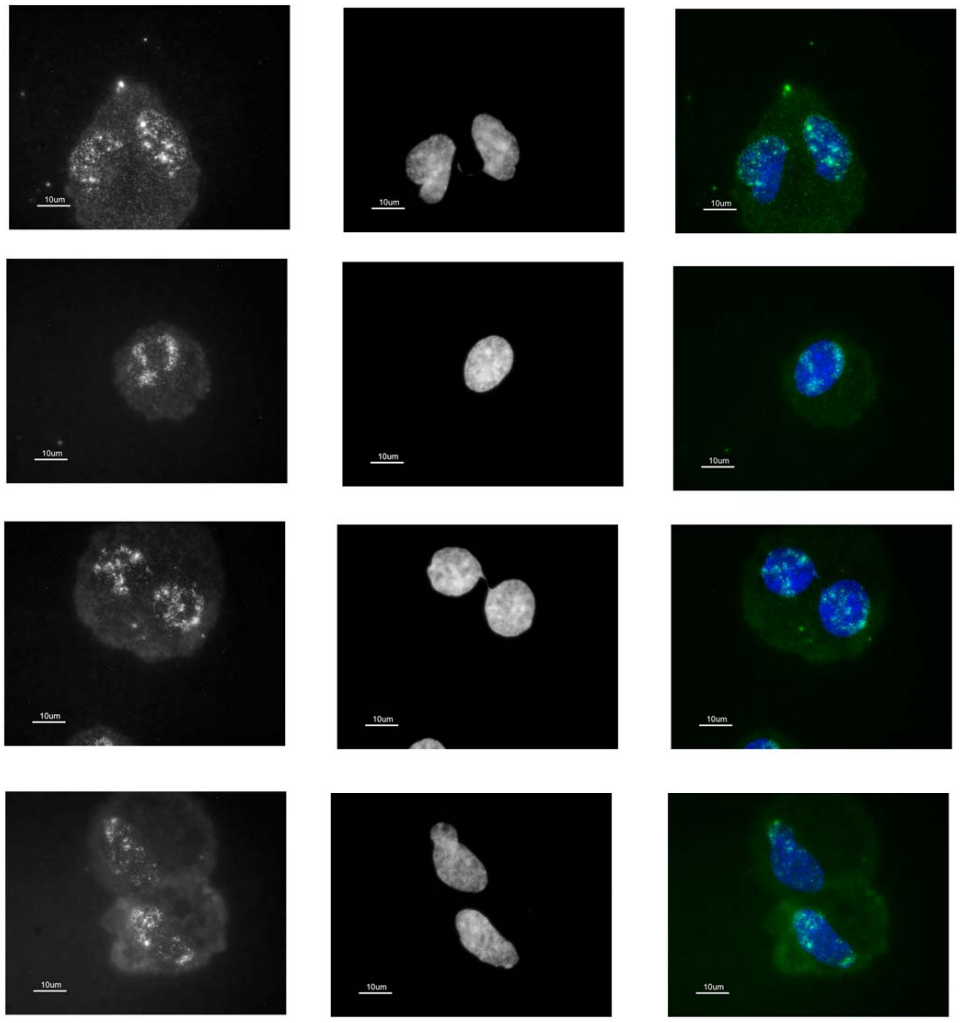

Anti-r1APH0906
(FITC)
Host nuclei
(DAPI)
Combined
Transfected RF/6A
Anti-r1APH0906
(FITC)
Host nuclei
(DAPI)
Combined
Wild-type RF/6A

Figure 7. Translocation of APH_0906 to the nucleus of transfected HL-60 cells

APH_0906 within transfected cells was detected using antibodies against the first half of the protein (r1APH_0906) and then labeled using Cy3-labeled anti-mouse antibodies. The localization of the protein is shown in red fluorescence (First column), whereas the nucleus is shown by DAPI staining in the blue channel (Second column). The third column shows the combined channels.

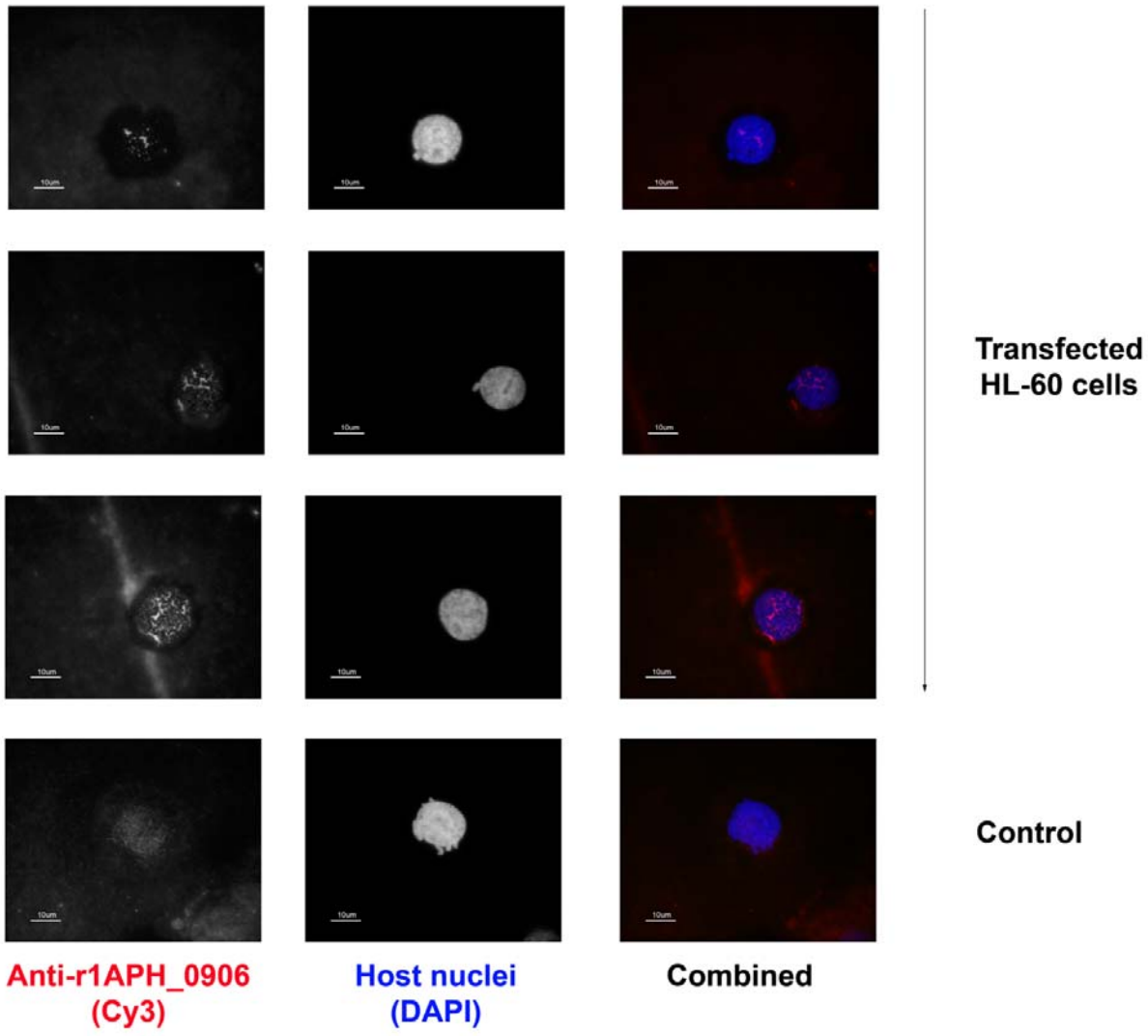


Table 1. Differences in sizes and Nuclear Localization Signals (NLS)s between strains of A. phagocytophilum.

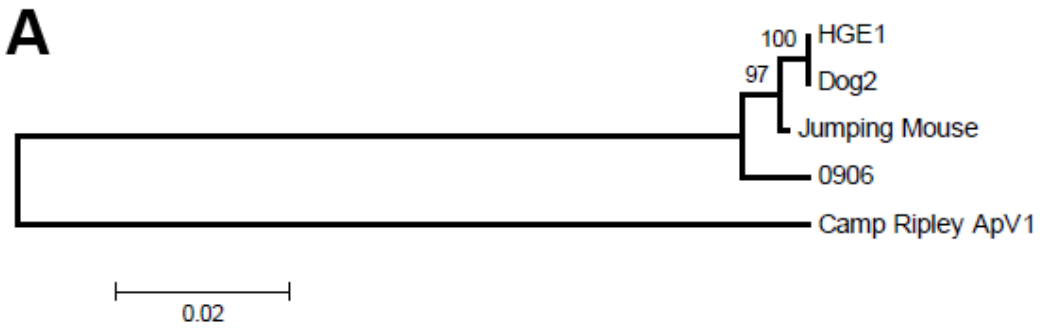
Strain	Host	Size	BRO Anti-repressor	Sequences NLS	Predictor	Position	P-sort	PP** (Bacteria)	PP** (Eukaryotic)	Nuc-Ploc
HZ	Human	1528	yes	YRTAVKKVKLA ENAVKERMLPALTLYAKRIAKR KGTDGKKPSTTVPKKPPRPARGAK	NLSMAPPER NLSMAPPER NLStradamus	202 647 1451	OM*	Secreted	Cytoplasm	Nuclei
HGE1	Human	1528	yes	YRTAVKKVKLA ENAVKERMLPALTLYAKRIAKR DGKKPSTTVPKKPPRPARGAK	NLSMAPPER NLSMAPPER NLStradamus	202 647 1454	OM*	Secreted	Cytoplasm	Nuclei
JM	Jumping Mouse	1491	yes	YRTAVKKVKLA DLRKRHRWA ENAVKERMLPALTLYAKRIAKR DGKKPSTTVPKKPPRPARGAK	NLSMAPPER NLSMAPPER NLSMAPPER NLStradamus	165 388 610 1417	OM*	Secreted	Cytoplasm	Nuclei
Dog2	Dog	1491	yes	YRTAVKKVKLA ENAVKERMLPALTLYAKRIAKR DGKKPSTTVPKKPPRPARGAK	NLSMAPPER NLSMAPPER NLStradamus	165 610 1417	OM*	Secreted	Cytoplasm	Nuclei
ApV1	Deer	1552	no	YRTAVKKVKLA KKPSTTVPHKKPPRPARGAK	NLSMAPPER NLStradamus	205 1459	OM*	Secreted	Cytoplasm	Nuclei

* OM = Outer Membrane

** PP = Protein Predict prediction

Figure 8. Similarities and relationship between Aph_0906 homologs in different A. phagocytophilum strains.

A) Phylogenetic tree showing the relationship between the homologs of the protein APH_0906 from 5 different strains of *A. phagocytophilum.* 0906 represents the protein present in the HZ strain. Phylogenetic tree was constructed using alignment of the predicted amino acid sequences from *A. phagocytophilum* strains, which have been sequenced and are available in GenBank. The relationship of the protein sequences was inferred using Neighbor Joining and the total number of differences to determine the length of the branches. The numbers at the nodes represents the bootstrap value from 1000 replications. B) Similarity matrix showing the identity scores of the different amino acid sequences from the APH_0906 homolog present in each of the *A. phagocytophilum* strains.



B

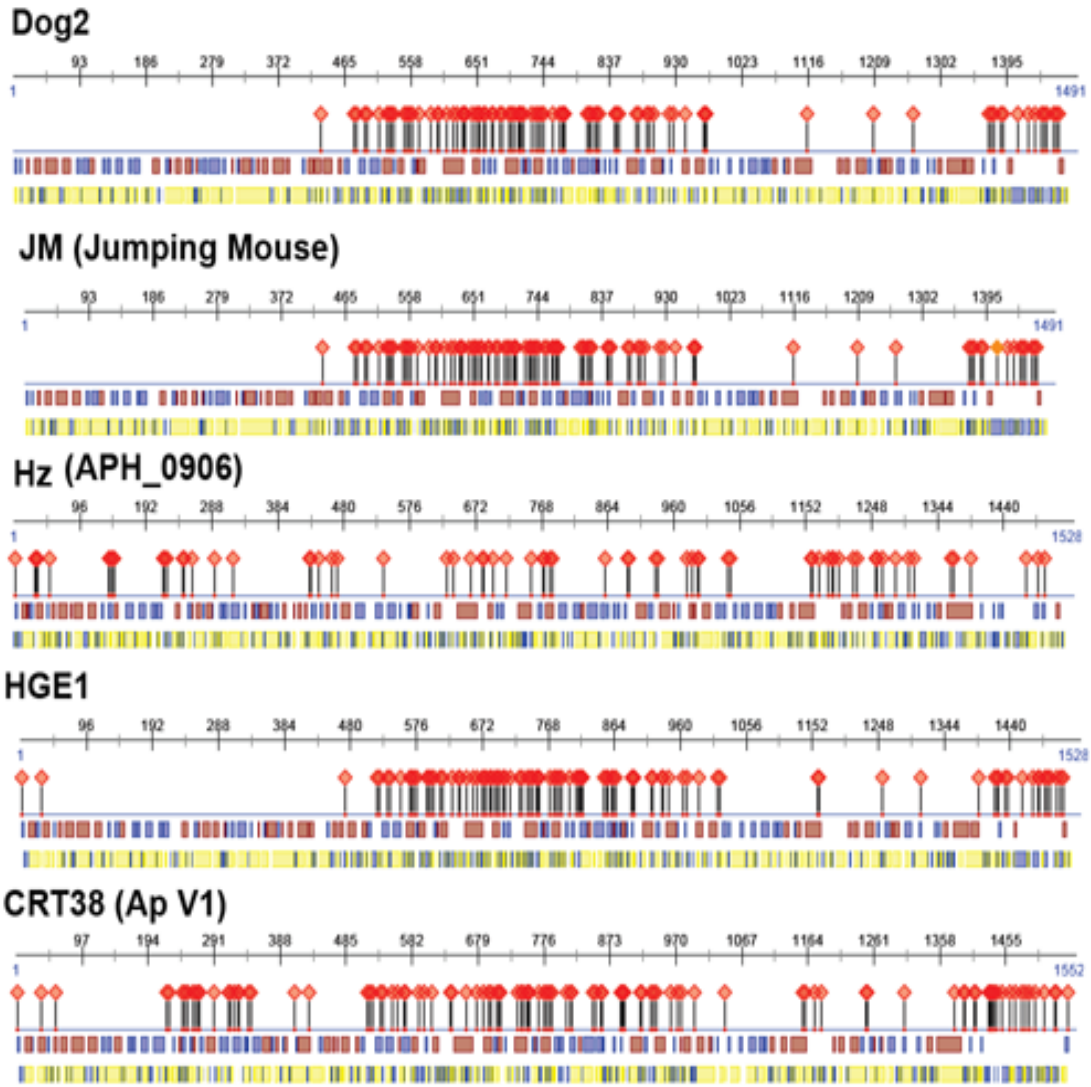
Identity Scores (%)

	HGE1	0906	Camp Ripley ApV1	Dog2	Jumping Mouse
HGE1	100.0	98.5	81.4	100.0	99.6
0906	98.9	100.0	81.4	98.5	98.7
Camp Ripley ApV1	86.7	86.8	100.0	81.4	81.6
Dog2	100.0	98.9	86.7	100.0	99.6
Jumping Mouse	99.8	99.0	86.8	99.8	100.0

Similarity Scores (%)

Figure 9. Binding sites, secondary structure, and solvent accessibility of the Aph_0906 homologs present in 5 strains of A. phagocytophilum.

The black line and the numbers represent the number of the residue in the linear form of the amino acid sequence. The positions of binding sites in the different homologs are shown by the red pins. The red squares, below the binding sites, represent the residues that form α-helices whereas the blue squares represent β-strands. The third row shows the exposed (yellow squares) and the hidden (blue squares) residues on the proteins.



Summary

Two *A. phagocytophilum* mutants that present phenotypical changes in their ability to infect tick or mammalian cells were produced using the HIMAR transponson-transposase system. The transposition sites were determined to have occurred in a gene that encodes a *O*-methyltransferase, referred to the OMT-mutant, and in the gene *aph_0906* that encodes a hypothetical protein, referred to APH0906-mutant. The OMT-mutant failed to grow in ticks cells and the bacteria were rapidly degraded, showing a decrease in bacterial numbers at 24 hr p.i. In contrast, the APH0906-mutant was not able to grow in HL-60 cells but the bacterial numbers remained the same throughout the 5 days of the experiment. PCR showed that the APH0906-mutant was able to infect and persist in hamsters for up to 21 days.

Gene expression analyses of the OMT coding gene showed that the gene is upregulated during binding and internalization at 1 – 2 hr p.i. The OMT-mutant failed to bind to ISE6 cells (tick cells). Further analyses showed that inhibition of methyltransferases in wild-type *A.phagocytophilum*, using an S-adenosyl-homocysteine hydrolase inhibitor (Adenosine Periodate (Adox), resulted in a similar phenotype. Immuno Fluorescent Assays (IFAs) determined that the OMT was expressed by *A. phagocytophilum* when the bacteria were invading tick cells. Proteomic analyses identified several proteins, known to be expressed during tick infection, that were downregulated in the OMT-mutant during binding to tick cells whereas that several proteins suspected to be important for infection of mammalian cells were upregulated. This proteomic analysis also identified several proteins that were not methylated in the OMT-mutant. *In vitro* methylation assays identified the Major Surface Protein 4 (Msp4) as the substrate for the OMT. Thus, suggesting that methylation in Msp4 is important for *A. phagocytohilum* binding to and infection of tick cells.

Analyses of *aph_0906* gene expression showed that the gene is upregulated during later phases of infection and the effect on binding was not as pronounced. Bioinformatic analyses of the gene showed that the protein presents several DNA binding sites. Localization experiments of the protein using IFAs of infected cells demonstrated that the protein localizes to the bacteria when *A. phagocytophilum* was outside of the cell and during binding. However, when the bacteria were inside of HL-60 cells, the protein was secreted into the cytoplasm and as the infection progressed the protein was found in the nucleus of the host cells. Likewise, RF/6A and HL-60 cells transformed to express the APH_0906 protein showed a similar translocation of the protein to their nuclei. Bioinformatic analyses of the protein showed several Nuclear Localization Signal (NLS) sequences within APH_0906 and also identified several differences in these NLSs in the homologs of this protein within strains of *A. phagocytophilum* that present different host-tropisms. These results suggest that APH_0906 may be an effector molecule in *A. phagocytophilum* important for infection of human cells.

References

1. CDC (2010) Anaplasmosis: Symptoms, Diagnosis, and Treatment.

2. Johnson RC, Kodner C, Jarnefeld J, Eck DK, Xu Y (2011) Agents of Human Anaplasmosis and Lyme Disease at Camp Ripley, Minnesota. Vector Borne Zoonotic Dis 2011: 25.

3. Ravyn MD, Kodner CB, Carter SE, Jarnefeld JL, Johnson RC (2001) Isolation of the etiologic agent of human granulocytic ehrlichiosis from the white-footed mouse (*Peromyscus leucopus*). J Clin Microbiol 39: 335-338.

4. Chen SM, Dumler JS, Bakken JS, Walker DH (1994) Identification of a granulocytotropic *Ehrlichia* species as the etiologic agent of human disease. J Clin Microbiol 32: 589-595.

5. Dumler JS, Barbet AF, Bekker CP, Dasch GA, Palmer GH, et al. (2001) Reorganization of genera in the families *Rickettsiaceae* and *Anaplasmataceae* in the order *Rickettsiales*: unification of some species of *Ehrlichia* with *Anaplasma*, *Cowdria* with *Ehrlichia* and *Ehrlichia* with *Neorickettsia*, descriptions of six new species combinations and designation of *Ehrlichia equi* and 'HGE agent' as subjective synonyms of *Ehrlichia phagocytophila*. Int J Syst Evol Microbiol 51: 2145-2165.

6. Dumler JS, Choi KS, Garcia-Garcia JC, Barat NS, Scorpio DG, et al. (2005) Human granulocytic anaplasmosis and *Anaplasma phagocytophilum*. Emerg Infect Dis 11: 1828-1834.

7. Dumler JS (2005) *Anaplasma* and *Ehrlichia* infection. Ann N Y Acad Sci 1063: 361-373.

8. Thomas RJ, Dumler JS, Carlyon JA (2009) Current management of human granulocytic anaplasmosis, human monocytic ehrlichiosis and *Ehrlichia ewingii* ehrlichiosis. Expert Rev Anti Infect Ther 7: 709-722.

9. Ismail N, Bloch KC, McBride JW (2010) Human ehrlichiosis and anaplasmosis. Clin Lab Med 30: 261-292.

10. Chen G, Severo MS, Sakhon OS, Choy A, Herron MJ, et al. (2012) *Anaplasma phagocytophilum* dihydrolipoamide dehydrogenase 1 affects host-derived immunopathology during microbial colonization. Infect Immun 80: 3194-3205.

11. Choi KS, Webb T, Oelke M, Scorpio DG, Dumler JS (2007) Differential innate immune cell activation and proinflammatory response in *Anaplasma phagocytophilum* infection. Infect Immun 75: 3124-3130.

12. Choi KS, Dumler JS (2007) Mitogenic component in polar lipid-enriched *Anaplasma phagocytophilum membranes.* Clin Vaccine Immunol 14: 1260-1265.

13. Granick JL, Reneer DV, Carlyon JA, Borjesson DL (2008) *Anaplasma phagocytophilum* infects cells of the megakaryocytic lineage through sialylated ligands but fails to alter platelet production. J Med Microbiol 57: 416-423.

14. Parkins MD, Church DL, Jiang XY, Gregson DB (2009) Human granulocytic anaplasmosis: First reported case in Canada. Can J Infect Dis Med Microbiol 20: e100-102.

15. Dahlgren FS, Mandel EJ, Krebs JW, Massung RF, McQuiston JH (2011) Increasing incidence of *Ehrlichia chaffeensis* and *Anaplasma phagocytophilum* in the United States, 2000-2007. Am J Trop Med Hyg 85: 124-131.

16. CDC (2010) Statistics and Epidemiology: Annual Cases of Anaplasmosis in the United States.

17. Health MDo (2012) Reported Cases of Human Anaplasmosis in Minnesota by Year.

18. Petrovec M, Lotric Furlan S, Zupanc TA, Strle F, Brouqui P, et al. (1997) Human disease in Europe caused by a granulocytic *Ehrlichia* species. J Clin Microbiol 35: 1556-1559.

19. Doudier B, Olano J, Parola P, Brouqui P (2009) Factors contributing to emergence of *Ehrlichia* and *Anaplasma spp.* as human pathogens. Vet Parasitol 167: 149-154.

20. Sreter T, Sreter-Lancz Z, Szell Z, Kalman D (2004) *Anaplasma phagocytophilum*: an emerging tick-borne pathogen in Hungary and Central Eastern Europe. Ann Trop Med Parasitol 98: 401-405.

21. Zhang L, Cui F, Wang L, Zhang L, Zhang J, et al. (2011) Investigation of anaplasmosis in Yiyuan County, Shandong Province, China. Asian Pac J Trop Med 4: 568-572.

22. Zhang L, Liu Y, Ni D, Li Q, Yu Y, et al. (2008) Nosocomial transmission of human granulocytic anaplasmosis in China. Jama 300: 2263-2270.

23. Zhang S, Hai R, Li W, Li G, Lin G, et al. (2009) Seroprevalence of human granulocytotropic anaplasmosis in central and southeastern China. Am J Trop Med Hyg 81: 293-295.

24. Foley JE, Nieto NC, Massung R, Barbet A, Madigan J, et al. (2009) Distinct ecologically relevant strains of *Anaplasma phagocytophilum*. Emerg Infect Dis 15: 842-843.

25. Stuen S (2007) *Anaplasma phagocytophilum* - the most widespread tick-borne infection in animals in Europe. Vet Res Commun 31 Suppl 1: 79-84.

26. Woldehiwet Z (2006) *Anaplasma phagocytophilum* in ruminants in Europe. Ann N Y Acad Sci 1078: 446-460.

27. Grova L, Olesen I, Steinshamn H, Stuen S (2011) Prevalence of *Anaplasma phagocytophilum* infection and effect on lamb growth. Acta Vet Scand 53: 30.

28. Garcia-Perez AL, Barandika J, Oporto B, Povedano I, Juste RA (2003) *Anaplasma phagocytophila* as an abortifacient agent in sheep farms from northern Spain. Ann N Y Acad Sci 990: 429-432.

29. Kawahara M, Rikihisa Y, Lin Q, Isogai E, Tahara K, et al. (2006) Novel genetic variants of *Anaplasma phagocytophilum*, *Anaplasma bovis*, *Anaplasma centrale*, and a novel *Ehrlichia sp.* in wild deer and ticks on two major islands in Japan. Appl Environ Microbiol 72: 1102-1109.

30. Portillo A, Perez-Martinez L, Santibanez S, Santibanez P, Palomar AM, et al. (2011) *Anaplasma spp.* in wild mammals and Ixodes ricinus from the north of Spain. Vector Borne Zoonotic Dis 11: 3-8.

31. Massung RF, Courtney JW, Hiratzka SL, Pitzer VE, Smith G, et al. (2005) *Anaplasma phagocytophilum* in white-tailed deer. Emerg Infect Dis 11: 1604-1606.

32. Massung RF, Levin ML, Munderloh UG, Silverman DJ, Lynch MJ, et al. (2007) Isolation and propagation of the Ap-Variant 1 strain of *Anaplasma phagocytophilum* in a tick cell line. J Clin Microbiol 45: 2138-2143.

33. Reichard MV, Roman RM, Kocan KM, Blouin EF, de la Fuente J, et al. (2009) Inoculation of white-tailed deer (*Odocoileus virginianus*) with Ap-V1 Or NY-18 strains of *Anaplasma phagocytophilum* and microscopic demonstration of Ap-V1 In *Ixodes scapularis* adults that acquired infection from deer as nymphs. Vector Borne Zoonotic Dis 9: 565-568.

34. Woldehiwet Z (2009) The natural history of *Anaplasma phagocytophilum*. Vet Parasitol 167: 108-122.

35. Woldehiwet Z, Horrocks BK (2005) Antigenicity of ovine strains of *Anaplasma phagocytophilum* grown in tick cells and ovine granulocytes. J Comp Pathol 132: 322-328.

36. Castellaw AH, Chenney EF, Varela-Stokes AS (2011) Tick-borne disease agents in various wildlife from Mississippi. Vector Borne Zoonotic Dis 11: 439-442.

37. Bowman D, Little SE, Lorentzen L, Shields J, Sullivan MP, et al. (2009) Prevalence and geographic distribution of *Dirofilaria immitis*, *Borrelia burgdorferi*, *Ehrlichia canis*, and *Anaplasma phagocytophilum* in dogs in the United States: results of a national clinic-based serologic survey. Vet Parasitol 160: 138-148.

38. Domingos MC, Trotta M, Briend-Marchal A, Medaille C (2011) Anaplasmosis in two dogs in France and molecular and phylogenetic characterization of *Anaplasma phagocytophilum*. Vet Clin Pathol 40: 215-221.

39. Kohn B, Silaghi C, Galke D, Arndt G, Pfister K (2011) Infections with *Anaplasma phagocytophilum* in dogs in Germany. Res Vet Sci 91: 71-76.

40. Lim S, Irwin PJ, Lee S, Oh M, Ahn K, et al. (2010) Comparison of selected canine vector-borne diseases between urban animal shelter and rural hunting dogs in Korea. Parasit Vectors 3: 32.

41. Majlathova V, Majlath I, Vichova B, Gul'ova I, Derdakova M, et al. (2011) Polymerase chain reaction Confirmation of *Babesia canis canis* and *Anaplasma phagocytophilum* in Dogs Suspected of Babesiosis in Slovakia. Vector Borne Zoonotic Dis: 7.

42. Santos HA, Pires MS, Vilela JA, Santos TM, Faccini JL, et al. (2011) Detection of *Anaplasma phagocytophilum* in Brazilian dogs by real-time polymerase chain reaction. J Vet Diagn Invest 23: 770-774.

43. Carrade DD, Foley JE, Borjesson DL, Sykes JE (2009) Canine granulocytic anaplasmosis: a review. J Vet Intern Med 23: 1129-1141.

44. Magnarelli LA, Bushmich SL, JW IJ, Fikrig E (2005) Seroprevalence of antibodies against *Borrelia burgdorferi* and *Anaplasma phagocytophilum* in cats. Am J Vet Res 66: 1895-1899.

45. Billeter SA, Spencer JA, Griffin B, Dykstra CC, Blagburn BL (2007) Prevalence of Anaplasma phagocytophilum in domestic felines in the United States. Vet Parasitol 147: 194-198.

46. Foley JE, Leutenegger CM, Dumler JS, Pedersen NC, Madigan JE (2003) Evidence for modulated immune response to Anaplasma phagocytophila sensu lato in cats with FIV-induced immunosuppression. Comp Immunol Microbiol Infect Dis 26: 103-113.

47. Silaghi C, Liebisch G, Pfister K (2011) Genetic variants of *Anaplasma phagocytophilum* from 14 equine granulocytic anaplasmosis cases. Parasit Vectors 4: 161.

48. Rikihisa Y (2011) Mechanisms of obligatory intracellular infection with *Anaplasma phagocytophilum*. Clin Microbiol Rev 24: 469-489.

49. Rikihisa Y (2009) Molecular events involved in cellular invasion by Ehrlichia chaffeensis and *Anaplasma phagocytophilum*. Vet Parasitol 19: 19.

50. Caspersen K, Park JH, Patil S, Dumler JS (2002) Genetic variability and stability of *Anaplasma phagocytophila* msp2 (p44). Infect Immun 70: 1230-1234.

51. Huang H, Wang X, Kikuchi T, Kumagai Y, Rikihisa Y (2007) Porin activity of *Anaplasma phagocytophilum* outer membrane fraction and purified P44. J Bacteriol 189: 1998-2006.

52. Park J, Kim KJ, Grab DJ, Dumler JS (2003) *Anaplasma phagocytophilum* major surface protein-2 (Msp2) forms multimeric complexes in the bacterial membrane. FEMS Microbiol Lett 227: 243-247.

53. Park J, Choi KS, Dumler JS (2003) Major surface protein 2 of *Anaplasma phagocytophilum* facilitates adherence to granulocytes. Infect Immun 71: 4018-4025.

54. Nelson CM, Herron MJ, Felsheim RF, Schloeder BR, Grindle SM, et al. (2008) Whole genome transcription profiling of *Anaplasma phagocytophilum* in human and tick host cells by tiling array analysis. BMC Genomics 9: 364.

55. Ge Y, Rikihisa Y (2007) Identification of novel surface proteins of *Anaplasma phagocytophilum* by affinity purification and proteomics. J Bacteriol 189: 7819-7828.

56. Lin M, Rikihisa Y (2003) *Ehrlichia chaffeensis* and *Anaplasma phagocytophilum* lack genes for lipid A biosynthesis and incorporate cholesterol for their survival. Infect Immun 71: 5324-5331.

57. Andersson SG, Kurland CG (1998) Reductive evolution of resident genomes. Trends Microbiol 6: 263-268.

58. Herron MJ, Nelson CM, Larson J, Snapp KR, Kansas GS, et al. (2000) Intracellular parasitism by the human granulocytic ehrlichiosis bacterium through the P-selectin ligand, PSGL-1. Science 288: 1653-1656.

59. Reneer DV, Kearns SA, Yago T, Sims J, Cummings RD, et al. (2006) Characterization of a sialic acid- and P-selectin glycoprotein ligand-1-independent adhesin activity in the granulocytotropic bacterium *Anaplasma phagocytophilum*. Cell Microbiol 8: 1972-1984.

60. Ojogun N, Kahlon A, Ragland SA, Troese MJ, Mastronunzio JE, et al. (2012) *Anaplasma phagocytophilum* outer membrane protein A interacts with sialylated glycoproteins to promote infection of mammalian host cells. Infect Immun 80: 3748-3760.

61. Kahlon A, Ojogun N, Ragland SA, Seidman D, Troese MJ, et al. (2012) *Anaplasma phagocytophilum* Asp14 is an invasin that interacts with mammalian host cells via its C terminus to facilitate infection. Infect Immun 81: 65-79.

62. Truchan HK, Seidman D, Carlyon JA (2013) Breaking in and grabbing a meal: *Anaplasma phagocytophilum* cellular invasion, nutrient acquisition, and promising tools for their study. Microbes Infect 4579: 00210-00214.

63. Carlyon JA, Abdel-Latif D, Pypaert M, Lacy P, Fikrig E (2004) *Anaplasma phagocytophilum* utilizes multiple host evasion mechanisms to thwart NADPH oxidase-mediated killing during neutrophil infection. Infect Immun 72: 4772-4783.

64. Thomas V, Fikrig E (2007) *Anaplasma phagocytophilum* specifically induces tyrosine phosphorylation of ROCK1 during infection. Cell Microbiol 9: 1730-1737.

65. Troese MJ, Carlyon JA (2009) *Anaplasma phagocytophilum* dense-cored organisms mediate cellular adherence through recognition of human P-selectin glycoprotein ligand 1. Infect Immun 77: 4018-4027.

66. Huang B, Hubber A, McDonough JA, Roy CR, Scidmore MA, et al. (2010) The *Anaplasma phagocytophilum*-occupied vacuole selectively recruits Rab-GTPases that are predominantly associated with recycling endosomes. Cell Microbiol 2010: 25.

67. Niu H, Yamaguchi M, Rikihisa Y (2008) Subversion of cellular autophagy by *Anaplasma phagocytophilum*. Cell Microbiol 10: 593-605.

68. Huang B, Troese MJ, Howe D, Ye S, Sims JT, et al. (2010) *Anaplasma phagocytophilum* APH_0032 is expressed late during infection and localizes to the pathogen-occupied vacuolar membrane. Microb Pathog 49: 273-284.

69. Huang B, Troese MJ, Ye S, Sims JT, Galloway NL, et al. (2010) *Anaplasma phagocytophilum* APH_1387 is expressed throughout bacterial intracellular development and localizes to the pathogen-occupied vacuolar membrane. Infect Immun 78: 1864-1873.

70. Sukumaran B, Mastronunzio JE, Narasimhan S, Fankhauser S, Uchil PD, et al. (2011) *Anaplasma phagocytophilum* AptA modulates Erk1/2 signalling. Cell Microbiol 13: 47-61.

71. Huang B, Ojogun N, Ragland SA, Carlyon JA (2012) Monoubiquitinated proteins decorate the *Anaplasma phagocytophilum*-occupied vacuolar membrane. FEMS Immunol Med Microbiol 2011.

72. Lee HC, Kioi M, Han J, Puri RK, Goodman JL (2008) *Anaplasma phagocytophilum*-induced gene expression in both human neutrophils and HL-60 cells. Genomics 92: 144-151.

73. Park J, Kim KJ, Choi KS, Grab DJ, Dumler JS (2004) *Anaplasma phagocytophilum* AnkA binds to granulocyte DNA and nuclear proteins. Cell Microbiol 6: 743-751.

74. Thomas V, Samanta S, Fikrig E (2008) *Anaplasma phagocytophilum* increases cathepsin L activity, thereby globally influencing neutrophil function. Infect Immun 76: 4905-4912.

75. Hotopp JC, Lin M, Madupu R, Crabtree J, Angiuoli SV, et al. (2006) Comparative genomics of emerging human ehrlichiosis agents. PLoS Genet 2: e21.

76. Choi KS, Grab DJ, Dumler JS (2004) *Anaplasma phagocytophilum* infection induces protracted neutrophil degranulation. Infect Immun 72: 3680-3683.

77. Schaff UY, Trott KA, Chase S, Tam K, Johns JL, et al. (2010) Neutrophils exposed to *A. phagocytophilum* under shear stress fail to fully activate, polarize, and transmigrate across inflamed endothelium. Am J Physiol Cell Physiol 299: C87-96.

78. Sukumaran B, Carlyon JA, Cai JL, Berliner N, Fikrig E (2005) Early transcriptional response of human neutrophils to *Anaplasma phagocytophilum* infection. Infect Immun 73: 8089-8099.

79. Baldridge GD, Burkhardt N, Herron MJ, Kurtti TJ, Munderloh UG (2005) Analysis of fluorescent protein expression in transformants of *Rickettsia monacensis*, an obligate intracellular tick symbiont. Appl Environ Microbiol 71: 2095-2105.

80. Choi KS, Park JT, Dumler JS (2005) *Anaplasma phagocytophilum* delay of neutrophil apoptosis through the p38 mitogen-activated protein kinase signal pathway. Infect Immun 73: 8209-8218.

81. Lee HC, Goodman JL (2006) *Anaplasma phagocytophilum* causes global induction of antiapoptosis in human neutrophils. Genomics 88: 496-503.

82. Hodzic E, Feng S, Fish D, Leutenegger CM, Freet KJ, et al. (2001) Infection of mice with the agent of human granulocytic ehrlichiosis after different routes of inoculation. J Infect Dis 183: 1781-1786.

83. Granquist EG, Bardsen K, Bergstrom K, Stuen S Variant -and individual dependent nature of persistent *Anaplasma phagocytophilum* infection. Acta Vet Scand 52: 25.

84. Munderloh UG, Lynch MJ, Herron MJ, Palmer AT, Kurtti TJ, et al. (2004) Infection of endothelial cells with Anaplasma marginale and A. phagocytophilum. Vet Microbiol 101: 53-64.

85. Herron MJ, Ericson ME, Kurtti TJ, Munderloh UG (2005) The interactions of *Anaplasma phagocytophilum*, endothelial cells, and human neutrophils. Ann N Y Acad Sci 1063: 374-382.

86. Carreno AD, Alleman AR, Barbet AF, Palmer GH, Noh SM, et al. (2007) In vivo endothelial cell infection by *Anaplasma marginale*. Vet Pathol 44: 116-118.

87. Ohashi N, Inayoshi M, Kitamura K, Kawamori F, Kawaguchi D, et al. (2005) *Anaplasma phagocytophilum*-infected ticks, Japan. Emerg Infect Dis 11: 1780-1783.

88. Wirtgen M, Nahayo A, Linden A, Losson B, Garigliany M, et al. Detection of *Anaplasma phagocytophilum* in Dermacentor reticulatus ticks. Vet Rec 168: 195.

89. Baldridge GD, Scoles GA, Burkhardt NY, Schloeder B, Kurtti TJ, et al. (2009) Transovarial transmission of *Francisella*-like endosymbionts and *Anaplasma phagocytophilum* variants in *Dermacentor albipictus* (*Acari: Ixodidae*). J Med Entomol 46: 625-632.

90. Hodzic E, Fish D, Maretzki CM, De Silva AM, Feng S, et al. (1998) Acquisition and transmission of the agent of human granulocytic ehrlichiosis by *Ixodes scapularis* ticks. J Clin Microbiol 36: 3574-3578.

91. Liu L, Narasimhan S, Dai J, Zhang L, Cheng G, et al. (2011) *Ixodes scapularis* salivary gland protein P11 facilitates migration of *Anaplasma phagocytophilum* from the tick gut to salivary glands. EMBO Rep 12: 1196-1203.

92. Munderloh UG, Jauron SD, Fingerle V, Leitritz L, Hayes SF, et al. (1999) Invasion and intracellular development of the human granulocytic ehrlichiosis agent in tick cell culture. J Clin Microbiol 37: 2518-2524.

93. Carlyon JA, Akkoyunlu M, Xia L, Yago T, Wang T, et al. (2003) Murine neutrophils require alpha1,3-fucosylation but not PSGL-1 for productive infection with *Anaplasma phagocytophilum*. Blood 102: 3387-3395.

94. Pedra JH, Narasimhan S, Rendic D, Deponte K, Bell-Sakyi L, et al. (2010) Fucosylation enhances colonization of ticks by *Anaplasma phagocytophilum*. Cell Microbiol 2010: 19.

95. Stadtmann A, Germena G, Block H, Boras M, Rossaint J, et al. (2013) The PSGL-1-L-selectin signaling complex regulates neutrophil adhesion under flow. J Exp Med 210: 2171-2180.

96. de la Fuente J, Almazan C, Blouin EF, Naranjo V, Kocan KM (2006) Reduction of tick infections with *Anaplasma marginale* and *A. phagocytophilum* by targeting the tick protective antigen subolesin. Parasitol Res 100: 85-91.

97. Mastronunzio JE, Kurscheid S, Fikrig E (2012) Postgenomic analyses reveal development of infectious *Anaplasma phagocytophilum* during transmission from ticks to mice. J Bacteriol 194: 2238-2247.

98. Omsland A, Sager J, Nair V, Sturdevant DE, Hackstadt T (2012) Developmental stage-specific metabolic and transcriptional activity of *Chlamydia trachomatis* in an axenic medium. Proc Natl Acad Sci U S A 109: 19781-19785.

99. Luo ZQ (2012) *Legionella* secreted effectors and innate immune responses. Cell Microbiol 14: 19-27.

100. Lin M, den Dulk-Ras A, Hooykaas PJ, Rikihisa Y (2007) *Anaplasma phagocytophilum* AnkA secreted by type IV secretion system is tyrosine phosphorylated by Abl-1 to facilitate infection. Cell Microbiol 9: 2644-2657.

101. Kulp A, Kuehn MJ (2010) Biological functions and biogenesis of secreted bacterial outer membrane vesicles. Annu Rev Microbiol 64: 163-184.

102. Simone D, Bay DC, Leach T, Turner RJ (2013) Diversity and Evolution of Bacterial Twin Arginine Translocase Protein, TatC, Reveals a Protein Secretion System That Is Evolving to Fit Its Environmental Niche. PLoS One 8: e78742.

103. Reneer DV, Troese MJ, Huang B, Kearns SA, Carlyon JA (2008) A*naplasma phagocytophilum* PSGL-1-independent infection does not require Syk and leads to less efficient AnkA delivery. Cell Microbiol 10: 1827-1838.

104. Rennoll-Bankert KE, Dumler JS (2012) Lessons from *Anaplasma phagocytophilum*: chromatin remodeling by bacterial effectors. Infect Disord Drug Targets 12: 380-387.

105. Niu H, Kozjak-Pavlovic V, Rudel T, Rikihisa Y (2010) *Anaplasma phagocytophilum* Ats-1 is imported into host cell mitochondria and interferes with apoptosis induction. PLoS Pathog 6: e1000774.

106. Niu H, Xiong Q, Yamamoto A, Hayashi-Nishino M, Rikihisa Y (2012) Autophagosomes induced by a bacterial Beclin 1 binding protein facilitate obligatory intracellular infection. Proc Natl Acad Sci U S A 109: 20800-20807.

107. Rikihisa Y, Lin M, Niu H (2010) Type IV secretion in the obligatory intracellular bacterium *Anaplasma phagocytophilum*. Cell Microbiol 2010: 28.

108. Felsheim RF, Herron MJ, Nelson CM, Burkhardt NY, Barbet AF, et al. (2006) Transformation of *Anaplasma phagocytophilum*. BMC Biotechnol 6: 42.

109. Cain JA, Solis N, Cordwell SJ (2013) Beyond gene expression: The impact of protein post-translational modifications in bacteria. J Proteomics 3919: 00453-00453.

110. Liscombe DK, Louie GV, Noel JP (2012) Architectures, mechanisms and molecular evolution of natural product methyltransferases. Nat Prod Rep 29: 1238-1250.

111. Garcia-Fontana C, Reyes-Darias JA, Munoz-Martinez F, Alfonso C, Morel B, et al. (2013) High specificity in CheR methyltransferase function: CheR2 of *Pseudomonas putida* is essential for chemotaxis, whereas CheR1 is involved in biofilm formation. J Biol Chem 288: 18987-18999.

112. Kumar R, Mukhopadhyay AK, Ghosh P, Rao DN (2012) Comparative transcriptomics of *H. pylori* strains AM5, SS1 and their hpyAVIBM deletion mutants: possible roles of cytosine methylation. PLoS One 7: e42303.

113. Szretter KJ, Daniels BP, Cho H, Gainey MD, Yokoyama WM, et al. (2012) 2'-O methylation of the viral mRNA cap by West Nile virus evades ifit1-dependent and -independent mechanisms of host restriction in vivo. PLoS Pathog 8: e1002698.

114. Willemsen NM, Hitchen EM, Bodetti TJ, Apolloni A, Warrilow D, et al. (2006) Protein methylation is required to maintain optimal HIV-1 infectivity. Retrovirology 3: 92.

115. Abeykoon AH, Chao CC, Wang G, Gucek M, Yang DC, et al. (2012) Two protein lysine methyltransferases methylate outer membrane protein B from *rickettsia*. J Bacteriol 194: 6410-6418.

116. Staerkel C, Boenisch MJ, Kroger C, Bormann J, Schafer W, et al. (2013) CbCTB2, an O-methyltransferase is essential for biosynthesis of the phytotoxin cercosporin and infection of sugar beet by *Cercospora beticola*. BMC Plant Biol 13: 50.

117. Champion MD (2011) Host-pathogen o-methyltransferase similarity and its specific presence in highly virulent strains of *Francisella tularensis* suggests molecular mimicry. PLoS One 6: e20295.

118. Eshghi A, Pinne M, Haake DA, Zuerner RL, Frank A, et al. (2012) Methylation and in vivo expression of the surface-exposed *Leptospira interrogans* outer-membrane protein OmpL32. Microbiology 158: 622-635.

119. Chao CC, Chelius D, Zhang T, Daggle L, Ching WM (2004) Proteome analysis of Madrid E strain of *Rickettsia prowazekii*. Proteomics 4: 1280-1292.

120. Chao CC, Zhang Z, Wang H, Alkhalil A, Ching WM (2008) Serological reactivity and biochemical characterization of methylated and unmethylated forms of a recombinant protein fragment derived from outer membrane protein B of *Rickettsia typhi*. Clin Vaccine Immunol 15: 684-690.

121. Bakken JS, Dumler JS (2006) Clinical diagnosis and treatment of human granulocytotropic anaplasmosis. Ann N Y Acad Sci 1078: 236-247.

122. Dumler JS (2012) The biological basis of severe outcomes in *Anaplasma phagocytophilum* infection. FEMS Immunol Med Microbiol 64: 13-20.

123. CDC (2012) Anaplasmosis: Statistics and Epidemiology.

124. Jin H, Wei F, Liu Q, Qian J (2012) Epidemiology and Control of Human Granulocytic Anaplasmosis: A Systematic Review. Vector Borne Zoonotic Dis 2012: 4.

125. Rejmanek D, Bradburd G, Foley J (2012) Molecular characterization reveals distinct genospecies of *Anaplasma phagocytophilum* from diverse North American hosts. J Med Microbiol 61: 204-212.

126. Beugnet F, Marie JL (2009) Emerging arthropod-borne diseases of companion animals in Europe. Vet Parasitol 163: 298-305.

127. Bakken JS, Dumler S (2008) Human granulocytic anaplasmosis. Infect Dis Clin North Am 22: 433-448, viii.

128. Dunning Hotopp JC, Lin M, Madupu R, Crabtree J, Angiuoli SV, et al. (2006) Comparative genomics of emerging human ehrlichiosis agents. PLoS Genet 2: e21.

129. Troese MJ, Kahlon A, Ragland SA, Ottens AK, Ojogun N, et al. (2011) Proteomic analysis of *Anaplasma phagocytophilum* during infection of human myeloid cells identifies a protein that is pronouncedly upregulated on the infectious dense-cored cell. Infect Immun 79: 4696-4707.

130. Troese MJ, Kahlon A, Ragland SA, Ottens AK, Ojogun N, et al. (2011) Proteomic analysis of *Anaplasma phagocytophilum* during infection of human myeloid cells identifies a protein that is pronouncedly upregulated on the infectious dense-cored cell. Infect Immun 79: 4696-4707.

131. Shaner NC, Steinbach PA, Tsien RY (2005) A guide to choosing fluorescent proteins. Nat Methods 2: 905-909.

132. Rikihisa Y, Zhi N, Wormser GP, Wen B, Horowitz HW, et al. (1997) Ultrastructural and antigenic characterization of a granulocytic ehrlichiosis agent directly isolated and stably cultivated from a patient in New York state. J Infect Dis 175: 210-213.

133. Burkhardt NY, Baldridge GD, Williamson PC, Billingsley PM, Heu CC, et al. (2011) Development of shuttle vectors for transformation of diverse *Rickettsia* species. PLoS One 6: e29511.

134. Barlough JE, Madigan JE, DeRock E, Bigornia L (1996) Nested polymerase chain reaction for detection of *Ehrlichia equi* genomic DNA in horses and ticks (Ixodes pacificus). Vet Parasitol 63: 319-329.

135. Ades EW, Candal FJ, Swerlick RA, George VG, Summers S, et al. (1992) HMEC-1: establishment of an immortalized human microvascular endothelial cell line. J Invest Dermatol 99: 683-690.

136. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24: 1596-1599.

137. Rodionov AV, Eremeeva ME, Balayeva NM (1991) Isolation and partial characterization of the M(r) 100 kD protein from *Rickettsia prowazekii* strains of different virulence. Acta Virol 35: 557-565.

138. Bechah Y, El Karkouri K, Mediannikov O, Leroy Q, Pelletier N, et al. (2010) Genomic, proteomic, and transcriptomic analysis of virulent and avirulent *Rickettsia prowazekii* reveals its adaptive mutation capabilities. Genome Res 20: 655-663.

139. Paranjpye RN, Lara JC, Pepe JC, Pepe CM, Strom MS (1998) The type IV leader peptidase/N-methyltransferase of *Vibrio vulnificus* controls factors required for adherence to HEp-2 cells and virulence in iron-overloaded mice. Infect Immun 66: 5659-5668.

140. Montagna M, Sassera D, Epis S, Bazzocchi C, Vannini C, et al. (2013) "*Candidatus Midichloriaceae*" fam. nov. (*Rickettsiales*), an Ecologically Widespread Clade of Intracellular *Alphaproteobacteria*. Appl Environ Microbiol 79: 3241-3248.

141. Sockett RE (2009) Predatory lifestyle of *Bdellovibrio bacteriovorus*. Annu Rev Microbiol 63: 523-539.

142. Lambert C, Chang CY, Capeness MJ, Sockett RE (2010) The first bite--profiling the predatosome in the bacterial pathogen *Bdellovibrio*. PLoS One 5: e8599.

143. Kang W, Kurihara M, Matsumoto S (2006) The BRO proteins of *Bombyx mori* nucleopolyhedrovirus are nucleocytoplasmic shuttling proteins that utilize the CRM1-mediated nuclear export pathway. Virology 350: 184-191.

144. Bideshi DK, Renault S, Stasiak K, Federici BA, Bigot Y (2003) Phylogenetic analysis and possible function of bro-like genes, a multigene family widespread among large double-stranded DNA viruses of invertebrates and bacteria. J Gen Virol 84: 2531-2544.

145. Pustelny C, Brouwer S, Musken M, Bielecka A, Dotsch A, et al. (2013) The peptide chain release factor methyltransferase PrmC is essential for pathogenicity and environmental adaptation of Pseudomonas aeruginosa PA14. Environ Microbiol 15: 597-609.

146. Chao CC, Chelius D, Zhang T, Mutumanje E, Ching WM (2007) Insight into the virulence of *Rickettsia prowazekii* by proteomic analysis and comparison with an avirulent strain. Biochim Biophys Acta 1774: 373-381.

147. Viswanathan VK, Edelstein PH, Pope CD, Cianciotto NP (2000) The *Legionella pneumophila* iraAB locus is required for iron assimilation, intracellular infection, and virulence. Infect Immun 68: 1069-1079.

148. Dekkers KL, You BJ, Gowda VS, Liao HL, Lee MH, et al. (2007) The *Cercospora nicotianae* gene encoding dual O-methyltransferase and FAD-dependent monooxygenase domains mediates cercosporin toxin biosynthesis. Fungal Genet Biol 44: 444-454.

149. An D, Apidianakis Y, Boechat AL, Baldini RL, Goumnerov BC, et al. (2009) The pathogenic properties of a novel and conserved gene product, KerV, in proteobacteria. PLoS One 4: e7167.

150. Garbom S, Olofsson M, Bjornfot AC, Srivastava MK, Robinson VL, et al. (2007) Phenotypic characterization of a virulence-associated protein, VagH, of *Yersinia pseudotuberculosis* reveals a tight link between VagH and the type III secretion system. Microbiology 153: 1464-1473.

151. Beels D, Heyndrickx L, Vereecken K, Vermoesen T, Michiels L, et al. (2008) Production of human immunodeficiency virus type 1 (HIV-1) pseudoviruses using linear HIV-1 envelope expression cassettes. J Virol Methods 147: 99-107.

152. Krzywinska E, Bhatnagar S, Sweet L, Chatterjee D, Schorey JS (2005) *Mycobacterium avium* 104 deleted of the methyltransferase D gene by allelic replacement lacks serotype-specific glycopeptidolipids and shows attenuated virulence in mice. Mol Microbiol 56: 1262-1273.

153. Jabbouri S, Relic B, Hanin M, Kamalaprija P, Burger U, et al. (1998) nolO and noeI (HsnIII) of *Rhizobium sp.* NGR234 are involved in 3-O-carbamoylation and 2-O-methylation of Nod factors. J Biol Chem 273: 12047-12055.

154. Lovrich SD, Jobe DA, Kowalski TJ, Policepatil SM, Callister SM (2011) Expansion of the Midwestern focus for human granulocytic anaplasmosis into the region surrounding La Crosse, Wisconsin. J Clin Microbiol 49: 3855-3859.

155. Clark KL (2012) *Anaplasma phagocytophilum* in small mammals and ticks in northeast Florida. J Vector Ecol 37: 262-268.

156. Roellig DM, Fang QQ (2012) Detection of *Anaplasma phagocytophilum* in ixodid ticks from equine-inhabited sites in the Southeastern United States. Vector Borne Zoonotic Dis 12: 330-332.

157. Keesing F, Hersh MH, Tibbetts M, McHenry DJ, Duerr S, et al. (2012) Reservoir competence of vertebrate hosts for *Anaplasma phagocytophilum.* Emerg Infect Dis 18: 2013-2016.

158. Stuen S, Granquist EG, Silaghi C (2013) *Anaplasma phagocytophilum*--a widespread multi-host pathogen with highly adaptive strategies. Front Cell Infect Microbiol 3: 31.

159. Rikihisa Y (2010) Molecular events involved in cellular invasion by *Ehrlichia chaffeensis* and *Anaplasma phagocytophilum*. Vet Parasitol 167: 155-166.

160. Pedra JH, Narasimhan S, Rendic D, DePonte K, Bell-Sakyi L, et al. (2010) Fucosylation enhances colonization of ticks by *Anaplasma phagocytophilum*. Cell Microbiol 12: 1222-1234.

161. Sukumaran B, Narasimhan S, Anderson JF, DePonte K, Marcantonio N, et al. (2006) An Ixodes scapularis protein required for survival of *Anaplasma phagocytophilum* in tick salivary glands. J Exp Med 203: 1507-1517.

162. Zivkovic Z, Blouin EF, Manzano-Roman R, Almazan C, Naranjo V, et al. (2009) *Anaplasma phagocytophilum* and *Anaplasma marginale* elicit different gene expression responses in cultured tick cells. Comp Funct Genomics 705034: 705034.

163. Yang F, Shen Y, Camp DG, 2nd, Smith RD (2012) High-pH reversed-phase chromatography with fraction concatenation for 2D proteomic analysis. Expert Rev Proteomics 9: 129-134.

164. Lin-Moshier Y, Sebastian PJ, Higgins L, Sampson ND, Hewitt JE, et al. (2013) Re-evaluation of the role of calcium homeostasis endoplasmic reticulum protein (CHERP) in cellular calcium signaling. J Biol Chem 288: 355-367.

165. Oberle SM, Palmer GH, Barbet AF (1993) Expression and immune recognition of the conserved MSP4 outer membrane protein of *Anaplasma marginale.* Infect Immun 61: 5245-5251.

166. Chamot-Rooke J, Mikaty G, Malosse C, Soyer M, Dumont A, et al. (2011) Posttranslational modification of pili upon cell contact triggers *N. meningitidis* dissemination. Science 331: 778-782.

167. Narimatsu M, Noiri Y, Itoh S, Noguchi N, Kawahara T, et al. (2004) Essential role for the gtfA gene encoding a putative glycosyltransferase in the adherence of *Porphyromonas gingivalis.* Infect Immun 72: 2698-2702.

168. Uchiyama T, Kawano H, Kusuhara Y (2006) The major outer membrane protein rOmpB of spotted fever group *rickettsiae* functions in the rickettsial adherence to and invasion of Vero cells. Microbes Infect 8: 801-809.

169. Goldman DJ, Ordal GW (1984) In vitro methylation and demethylation of methyl-accepting chemotaxis proteins in *Bacillus subtilis*. Biochemistry 23: 2600-2606.

170. Perez E, Zheng H, Stock AM (2006) Identification of methylation sites in *Thermotoga maritima* chemotaxis receptors. J Bacteriol 188: 4093-4100.

171. de la Fuente J, Massung RF, Wong SJ, Chu FK, Lutz H, et al. (2005) Sequence analysis of the msp4 gene of *Anaplasma phagocytophilum* strains. J Clin Microbiol 43: 1309-1317.

172. Palmer GH, Eid G, Barbet AF, McGuire TC, McElwain TF (1994) The immunoprotective *Anaplasma marginale* major surface protein 2 is encoded by a polymorphic multigene family. Infect Immun 62: 3808-3816.

173. Sarkar M, Troese MJ, Kearns SA, Yang T, Reneer DV, et al. (2008) *Anaplasma phagocytophilum* MSP2(P44)-18 predominates and is modified into multiple isoforms in human myeloid cells. Infect Immun 76: 2090-2098.

174. Galdiero S, Falanga A, Cantisani M, Tarallo R, Della Pepa ME, et al. (2012) Microbe-host interactions: structure and role of Gram-negative bacterial porins. Curr Protein Pept Sci 13: 843-854.

175. Nikaido H (2003) Molecular basis of bacterial outer membrane permeability revisited. Microbiol Mol Biol Rev 67: 593-656.

176. Damaghi M, Bippes C, Koster S, Yildiz O, Mari SA, et al. (2010) pH-dependent interactions guide the folding and gate the transmembrane pore of the beta-barrel membrane protein OmpG. J Mol Biol 397: 878-882.

177. Korkmaz-Ozkan F, Koster S, Kuhlbrandt W, Mantele W, Yildiz O (2010) Correlation between the OmpG secondary structure and its pH-dependent alterations monitored by FTIR. J Mol Biol 401: 56-67.

178. Vidotto MC, McGuire TC, McElwain TF, Palmer GH, Knowles DP, Jr. (1994) Intermolecular relationships of major surface proteins of *Anaplasma marginale*. Infect Immun 62: 2940-2946.

179. Cheng Z, Kumagai Y, Lin M, Zhang C, Rikihisa Y (2006) Intra-leukocyte expression of two-component systems in *Ehrlichia chaffeensis* and *Anaplasma phagocytophilum* and effects of the histidine kinase inhibitor closantel. Cell Microbiol 8: 1241-1252.

180. Guyot H, Ramery E, O'Grady L, Sandersen C, Rollin F (2011) Emergence of bovine ehrlichiosis in Belgian cattle herds. Ticks Tick Borne Dis 2: 116-118.

181. Edouard S, Koebel C, Goehringer F, Socolovschi C, Jaulhac B, et al. (2012) Emergence of human granulocytic anaplasmosis in France. Ticks Tick Borne Dis 3: 403-405.

182. Ohashi N, Gaowa, Wuritu, Kawamori F, Wu D, et al. (2013) Human granulocytic Anaplasmosis, Japan. Emerg Infect Dis 19: 289-292.

183. Carlyon JA, Fikrig E (2003) Invasion and survival strategies of *Anaplasma phagocytophilum*. Cell Microbiol 5: 743-754.

184. Akkoyunlu M, Malawista SE, Anguita J, Fikrig E (2001) Exploitation of interleukin-8-induced neutrophil chemotaxis by the agent of human granulocytic ehrlichiosis. Infect Immun 69: 5577-5588.

185. Choi KS, Garyu J, Park J, Dumler JS (2003) Diminished adhesion of *Anaplasma phagocytophilum*-infected neutrophils to endothelial cells is associated with reduced expression of leukocyte surface selectin. Infect Immun 71: 4586-4594.

186. Ayllon N, Villar M, Busby AT, Kocan KM, Blouin EF, et al. (2013) *Anaplasma phagocytophilum* inhibits apoptosis and promotes cytoskeleton rearrangement for infection of tick cells. Infect Immun 81: 2415-2425.

187. Niu H, Kozjak-Pavlovic V, Rudel T, Rikihisa Y (2010) *Anaplasma phagocytophilum* Ats-1 is imported into host cell mitochondria and interferes with apoptosis induction. PLoS Pathog 6: e1000774.

188. Niu H, Rikihisa Y (2013) Ats-1: a novel bacterial molecule that links autophagy to bacterial nutrition. Autophagy 9: 787-788.

189. Xiong Q, Rikihisa Y (2012) Subversion of NPC1 pathway of cholesterol transport by *Anaplasma phagocytophilum*. Cell Microbiol 14: 560-576.

190. Rikihisa Y (2003) Mechanisms to create a safe haven by members of the family *Anaplasmataceae*. Ann N Y Acad Sci 990: 548-555.

191. Rikihisa Y, Lin M (2010) *Anaplasma phagocytophilum* and *Ehrlichia chaffeensis* type IV secretion and Ank proteins. Curr Opin Microbiol 13: 59-66.

192. Hackett PB, Largaespada DA, Cooper LJ (2010) A transposon and transposase system for human application. Mol Ther 18: 674-683.

193. Geurts AM, Yang Y, Clark KJ, Liu G, Cui Z, et al. (2003) Gene transfer into genomes of human cells by the sleeping beauty transposon system. Mol Ther 8: 108-117.

194. Al-Khedery B, Lundgren AM, Stuen S, Granquist EG, Munderloh UG, et al. (2012) Structure of the type IV secretion system in different strains of *Anaplasma phagocytophilum*. BMC Genomics 13: 678.

195. Goodman JL, Nelson C, Vitale B, Madigan JE, Dumler JS, et al. (1996) Direct cultivation of the causative agent of human granulocytic ehrlichiosis. N Engl J Med 334: 209-215.

196. Johnson RC, Kodner C, Jarnefeld J, Eck DK, Xu Y, et al. (2011) Agents of human anaplasmosis and Lyme disease at Camp Ripley, Minnesota Vector Borne Zoonotic Dis 11: 1529-1534.

197. Alto NM, Orth K (2012) Subversion of cell signaling by pathogens. Cold Spring Harb Perspect Biol 4: a006114.

198. Henry T, Gorvel JP, Meresse S (2006) Molecular motors hijacking by intracellular pathogens. Cell Microbiol 8: 23-32.

199. Toyotome T, Suzuki T, Kuwae A, Nonaka T, Fukuda H, et al. (2001) *Shigella* protein IpaH(9.8) is secreted from bacteria within mammalian cells and transported to the nucleus. J Biol Chem 276: 32071-32079.

200. Ashida H, Kim M, Schmidt-Supprian M, Ma A, Ogawa M, et al. (2010) A bacterial E3 ubiquitin ligase IpaH9.8 targets NEMO/IKKgamma to dampen the host NF-kappaB-mediated inflammatory response. Nat Cell Biol 12: 66-73; sup pp 61-69.

201. Hicks SW, Galan JE (2013) Exploitation of eukaryotic subcellular targeting mechanisms by bacterial effectors. Nat Rev Microbiol 11: 316-326.

202. Olofsson A, Vallstrom A, Petzold K, Tegtmeyer N, Schleucher J, et al. (2010) Biochemical and functional characterization of *Helicobacter pylori* vesicles. Mol Microbiol 77: 1539-1555.

203. Lee JH, Jun SH, Baik SC, Kim DR, Park JY, et al. (2012) Prediction and screening of nuclear targeting proteins with nuclear localization signals in *Helicobacter pylori.* J Microbiol Methods 91: 490-496.

204. Scharf W, Schauer S, Freyburger F, Petrovec M, Schaarschmidt-Kiener D, et al. (2011) Distinct host species correlate with *Anaplasma phagocytophilum* ankA gene clusters. J Clin Microbiol 49: 790-796.

205. Massung RF, Mather TN, Priestley RA, Levin ML (2003) Transmission efficiency of the AP-variant 1 strain of *Anaplasma phagocytophila*. Ann N Y Acad Sci 990: 75-79.

206. Garcia-Garcia JC, Rennoll-Bankert KE, Pelly S, Milstone AM, Dumler JS (2009) Silencing of host cell CYBB gene expression by the nuclear effector AnkA of the intracellular pathogen *Anaplasma phagocytophilum*. Infect Immun 77: 2385-2391.