Identification of Genes Involved in the Assembly and Biosynthesis of the $N$-linked Flagellin Glycan in the Archaeon, *Methanococcus maripaludis*

by

John Wu

A thesis submitted to the Department of Microbiology and Immunology

In conformity with the requirements for

the degree of Master of Science

Queen’s University

Kingston, Ontario, Canada

(July, 2009)

Copyright ©John Wu, 2009
Abstract

*N*-glycosylation is a metabolic process found in all three domains of life. It is the attachment of a polysaccharide glycan to asparagine (Asn) residues within the amino acid motif, Asn-Xaa-Ser/Thr. In the archaean, *Methanococcus maripaludis*, a tetrasaccharide glycan was isolated from purified flagella and its structure determined by mass spectrometry analysis. The linking sugar to the protein is surprisingly, N-acetylgalactosamine (β-GalNAc), with the next proximal sugar a derivative of N-acetylgulosamine (β-GlcNAc), being named β-GlcNAc3Ac, and the third sugar a derivative of N-acetylmannosamine (β-ManNAc), with an attached threonine residue on the C6 carbon (β-ManNAc3NAm). The terminal sugar is an unusual diglycoside of aldulose ((5S)-2-acetamido-2,4-dideoxy-5-O-methyl-α-L-erythro-hexos-5-ulo-1,5-pyranose). Previous genetic analyses identified the glycosyltransferases (GTs) responsible for the transfer of the second and third sugars of the glycan, as well as the oligosaccharyltransferase (OST) which attaches the glycan to protein. Left unidentified were the first and fourth GTs, the flippase as well as any genes involved in glycan sugar biosynthesis and modification. In this work, genes suspected to be involved in the biosynthesis of *N*-linked sugars, as well as those that might encode the missing GTs and flippase were targeted for in-frame deletion. Mutants with a deleted annotated GT gene (*MMP1088*) had a small decrease in flagellin molecular weight as determined by immunoblotting. Mass spectrometry (MS) analysis confirmed that the *N*-linked glycan was missing the terminal sugar as well as the threonine found on the third sugar of wildtype cells. Mutants with a deleted gene annotated to be involved in acetamidino synthesis (a functional group that is present on the third sugar), also had a decrease in flagellin molecular weight. MS analysis determined that the *N*-linked glycan was missing the acetamidino group on the third sugar as well as its attached threonine, along with the terminal sugar. Both mutants were able to assemble functional flagella but had impaired motility compared to wildtype cells in mini-
swarm agar. Deletions were also constructed in four other GT genes considered candidates in assembly of the linking sugar. However, none of these mutants had the expected decrease in flagellin molecular weight.

With the work done in this study, the glycosyl transferase that attaches the last sugar of the *M. maripaludis* N-linked assembly pathway has been identified as well as a gene involved in the biosynthesis and modification of the glycan sugars.
Acknowledgements

I would like to thank my supervisor, Dr. Ken Jarrell, for his guidance, generosity and kindness. I am grateful that he shared his knowledge and experience with me. I also appreciate the many opportunities provided for me to develop my research and intellectual skills. The following ‘work’ simply did not feel like such during my time with him and it was always a pleasure to go into the lab.

I would also like to thank my committee members, Dr. Keith Poole and Dr. Andrew Daugulis, for providing valuable advice throughout my degree and for accommodating my needs within their own time schedules and research.

In addition, many thanks to Susan Logan, who provided mass spectrometry analysis in this study (Figures 3.7 and 3.19) and to Dr. S.I. Aizawa, for providing electron micrographs of mutant strains (Figure 3.4). Deletions mutants in the glycosyl transferases, MMP1079 and MMP1080, and oligosaccharyl transferase, MMP1424, used in Figures 3.3 and 3.5 were created by David VanDyke.

I am grateful for past and present Jarrell lab members for their teaching, advice and assistance in the lab. I thank Dr. Dorothy Agnew for donating many lab supplies and equipment in my work and Jerry Dering for glassware and keeping me on my toes. I also extend my appreciation to the rest of the Microbiology and Immunology department for their tolerance (of me) during “incubation” times.

Finally, I thank my parents, who have advised and supported me throughout my education.
List of Figures

Figure 1.1: Tree of life for the Bacteria, Archaea and Eukarya domains ........................................ 5
Figure 1.2: Arrangements of flagella on a bacterial cell................................................................. 7
Figure 1.3: Assembly of the bacterial flagellum ............................................................................. 9
Figure 1.4: Twitching motility using the Type IV pilus ................................................................. 10
Figure 1.5: Organization of flagella associated genes for selected archaeal species .................... 13
Figure 1.6: N-linked glycosylation pathway for eukaryotes .......................................................... 16
Figure 1.7: Structure of the eukaryotic N-linked glycan ............................................................... 18
Figure 1.8: Structure of the bacterial N-linked glycan ................................................................. 20
Figure 1.9: N-linked glycosylation pathway in Bacteria ............................................................... 21
Figure 1.10: M. voltae and M. maripaludis N-linked glycan structure ........................................ 24
Figure 2.1: Vector map of pCRPrtNeo, pHW40 and pPAC60. ..................................................... 32
Figure 2.2: Schematic of an in-frame deletion and recombination when using pCRPrtNeo .......... 43
Figure 3.1: Creation of the in-frame deletion plasmid for MMP1088 ........................................ 51
Figure 3.2: Confirmation of MMP1088 deletion by PCR and Southern blotting ......................... 52
Figure 3.3: Immunoblot of whole cell lysates from the MMP1088 mutant .................................. 54
Figure 3.4: Electron micrograph of M. maripaludis ΔMMP1088 .................................................. 55
Figure 3.5: Motility assay for Δ1088 cells with semi-swarm agar plates ....................................... 56
Figure 3.6: Immunoblot of whole cell lysates complemented MMP1088 .................................... 58
Figure 3.7: MS data of Δ1088 N-linked glycan .......................................................................... 59
Figure 3.8: Creation of the in-frame deletion plasmid for MMP359 ............................................ 62
Figure 3.9: Confirmation of MMP359, MMP590, MMP1293 and MMP1294 deletion ............... 63
Figure 3.10: Immunoblots of M. maripaludis flagellins from potential 1st GT mutants ................ 64
Figure 3.11: Creation of deletion and disruption vectors of MMP1089 ...................................... 68
Figure 3.12: Whole cell PCR screening for multiple MMP1089 deletion attempts ................. 69
Figure 3.13: Immunoblot of whole cell lysates from putative ΔMMP1088/MMP1089 mutant .... 70
Figure 3.14: PCR of the putative mutant across MMP1088 and MMP1089 ......................... 71
Figure 3.15: Determination of co-transcription for MMP1081, MMP1082 and MMP1083 .... 73
Figure 3.16: Creation of deletion plasmid for MMP1081, MMP1082 and MMP1083 ......... 75
Figure 3.17: PCR & Southern blot of MMP1081, MMP1082 and MMP1083 mutants .............. 76
Figure 3.18: Immunoblot of Δ1081, Δ1082 and Δ1083 flagellins .............................................. 77
Figure 3.19: MS data of Δ1081 N-linked glycan ....................................................................... 78
Figure 3.20: Electron micrograph of ∆MMP1081 cell ......................................................... 79
Figure 3.21: Motility assay for ∆1081 cells with semi-swarm agar plates............................. 80
Figure 3.22: Immunoblot of whole cell lystates of complemented ∆1081 cells ..................... 81
Figure 4.1: Biosynthesis pathway of the acetamido sugar in eukaryotes and bacteria .......... 90
Figure 4.2: The N-linked glycosylation system and flagella assembly in M. maripaludis. ........ 93
Figure AppA.5.1: Creation of the in-frame deletion plasmid for flaJ .................................... 108
Figure AppA.5.2: PCR confirmation of flaJ deletion ............................................................. 109
Figure AppA.5.3: Electron micrograph of ∆FlaJ cell .......................................................... 110
Figure AppA.5.4: Immunoblot of ∆FlaI and ∆FlaJ membrane samples .............................. 111
### List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Strains used in this study</td>
<td>30</td>
</tr>
<tr>
<td>2.2</td>
<td>Plasmids used in this study</td>
<td>33</td>
</tr>
<tr>
<td>2.3</td>
<td>Deletion primers used in this study</td>
<td>36</td>
</tr>
<tr>
<td>2.4</td>
<td>Sequencing Primers</td>
<td>39</td>
</tr>
<tr>
<td>2.5</td>
<td>Complementation Primers</td>
<td>41</td>
</tr>
<tr>
<td>2.6</td>
<td>Primers for RT-PCR</td>
<td>45</td>
</tr>
<tr>
<td>3.1</td>
<td>Annotated glycosyl transferases in the <em>M. maripaludis</em> genome.</td>
<td>49</td>
</tr>
</tbody>
</table>
List of Abbreviations

ATP  
adenosine triphosphate

Bac  
bacillosamine

Co-B  
co-factor B

Co-M  
co-enzyme M

DIG  
digoxigenin

Dol-P  
dolichol phosphate

EDTA  
ethylenediamine tetraacetic acid

ER  
endoplasmic reticulum

GalNAc  
N-acetylgalactosamine

Glc  
glucose

GlcNAc  
N-acetylglucosamine

GT  
glycosyl transferase

Hpt  
hypoxanthine phosphoribosyltransferase

Man  
mannose

ManNAc  
N-acetylmannosamine

Neo  
neomycin

NMR  
nuclear magnetic resonance

OST  
oligosaccharyl transferase

PEG  
polyethylene glycol

SDS-PAGE  
sodium dodecyl sulfate polyacrylamide gel electrophoresis

TB  
transformation buffer

TFPP  
type IV prepilin peptidase

UDP  
uridine diphosphate
Chapter 1

Introduction

Introduction to Archaea

Carl Woese’s analysis of small subunit ribosomal RNA sequences led to a revised classification system of all living organisms that now consists of the three domains of Eukarya, Bacteria and Archaea (Woese and Olsen, 1986). Although the majority of microbiology research has focused on bacteria and their impact on our environment and health, archaea could also play a role in these fields and beyond.

Archaea may have initially gained recognition and curiosity for their ability to thrive in extreme environments (e.g. acidic, anaerobic, hyperthermophilic and halophilic), but recent culture-independent studies have revealed that they are ubiquitous (Chaban et al., 2006). Archaea are especially abundant in the ocean (estimated to comprise one third of all prokaryotic cells) and can live within the mesopelagic realm of the ocean where minimal light is present (Karner et al., 2001). These archaea may play a significant role in the oceanic carbon cycle with their ability to metabolize inorganic carbon under anaerobic conditions (Herndl et al., 2005). Other archaea are also capable of metabolic processes that benefit the environment, including the ability to oxidize ammonia to nitrate (contributing to the nitrogen cycle) and oxidize methane anaerobically, which converts one of the most potent greenhouse gases into environmentally friendly bicarbonate (Könneke et al., 2005; Caldwell et al., 2008). A better understanding of biodegrading archaeal species can have great environmental impact by reducing the amount of organic pollutants across the globe (Le Borgne et al., 2008).

Methanogens, in particular, have an important ecological niche in anaerobic environments, where complex organic matter, such as sewage, is degraded to methane and carbon dioxide (Schink, 1997). They create methane as a by-product of their anaerobic respiration,
obtaining ATP and energy through the oxidation of substrates (such as acetate, carbon dioxide and methanol) and the use of hydrogen as an electron donor (Zeikus, 1977). Anaerobic biodegradation of complex organic substrates requires three groups of organisms, starting with primary fermenting bacteria, followed sometimes by secondary fermenting bacteria and always terminating with methanogens (McInerney, 1988). Briefly, the primary fermenting bacteria utilize their hydrolytic enzymes to convert organic polymers (such as polysaccharides, proteins, nucleic acids and lipids) into oligomers and monomers (sugars, amino acids, purines, pyrimidines, fatty acids and glycerol), which are further broken down into fatty acids, succinate, lactate, acetate and alcohols. Some of these products (acetate, one-carbon compounds, and hydrogen gas) can be used directly by methanogens to produce methane and CO₂, while others need to be processed further by the secondary fermenting bacteria. The intermediate products are often toxic to the primary and secondary bacteria, along with most of the other organisms living within the ecosystem (Schink, 1997). Methanogens regulate the entire process with methanogenesis being essential for reducing the amount of intermediate waste.

Scientific pursuits to explore space and encounter extraterrestrial life may also point to archaea since they are capable of living in desolate environments. Recent landings on Mars indicate no signs of surface water but report a landscape full of canyons and gullies that suggests the presence of water at some point in the planet’s history (Baker, 2001). Furthermore, a methane cycle on the planet has been observed, suggesting biological life, while leading others to question the source of the replenished methane (Formisano et al., 2004; Mumma et al., 2009). Since on Earth, methanogens produce methane as a by-product of their anaerobic respiration, it would not be surprising to find a similar archaeal-like habitant on Mars (Kral et al., 2004; Kendrick and Kral, 2006). Interestingly, the presence of such organisms on Earth has led to the awareness of forward contamination where NASA equipment and landing gear may introduce microorganisms to Mars (Moissl et al., 2008).
Archaea may also have an impact on the medical sciences (Conway de Macario and Macario, 2008). Recently, it has been proposed that gut methanogens play a role in obesity (Zhang et al., 2009) while some patients with diseases such as diverticulitis and colon cancer have elevated numbers of archaea within their digestive tract (Eckburg et al., 2003). The most well understood disease associated with archaea is periodontitis, where the severity of the disease is strongly correlated with the presence of archaea (Vianna et al., 2008). The use of archaeal polar lipids in liposomes has also been shown to elicit a more potent adjuvant immune response for vaccines compared to the use of bacterial liposomes (Krishnan et al., 2001). With interest increasing in these related fields, archaea may be shown to have a larger impact in our daily lives than previously expected.

The revised tree of life, as shown in Figure 1.1, demonstrates further potential for research in archaea. Although there are many archaeal specific traits, archaea share many similarities to their prokaryotic relative in bacteria, while also possessing more eukaryotic-like properties than bacteria (Yutin et al., 2008). This is especially apparent in archaeal transcription and translation systems (Grabowski and Kelman, 2003). In addition to discoveries specific to archaea, some of which can be used commercially such as the DNA polymerases (Saiki et al., 1988; Cariello et al., 1991 and Marsic et al., 2008), archaea can also provide an equivalent system in a simpler study model compared to those found in Eukaryotes and Bacteria (Ouhammouch, 2004). In addition, studies in archaea can provide novel insight into eukaryotic or bacterial systems (Duggin and Bell, 2006). Some examples include the identification of a new catalytic motif in oligosaccharyl transferases after the crystallization of this enzyme from the hyperthermophilic archaeon Pyrococcus and how the knowledge of an archaeal cysteine pathway led to the discovery of selenocysteine biosynthesis in humans (Igura et al., 2008; Su et al., 2009). In this study, the genetics behind the archaeal N-linked glycosylation system was investigated by monitoring glycosylated flagellin proteins that assemble into the motility apparatus, the archaeal flagellum.
Although there are many lineages of archaea, relatively few can be cultivated and grown routinely in the laboratory (Schleper et al., 2005). Some of these include *Pyrococcus*, *Haloferax*, *Sulfolobus* and *Methanococcus* species, where sometimes challenging growth conditions are required. Still, within these species the genetic tools are limited in comparison to those available for many bacteria. Fortunately, recent advancements in technology and molecular biology techniques have enabled gene deletion studies to be performed in some of these archaea (Rother and Metcalf, 2005). The archaean used in this study is the obligate anaerobe, *Methanococcus maripaludis*, isolated from the salt marsh sediments in South Carolina (Whitman et al., 1986).

**Prokaryotic Motility Structures**

Prokaryotes have developed a wide variety of methods for movement to seek optimal living environments (Jarrell and McBride, 2008). Some methods include the rotational motor of flagella, use of gas vesicles, expulsion of polysaccharide mucous, a novel motility process called twitching which uses type IV pili and many more where the mechanisms of action are unknown. Extracellular appendages, such as flagella and pili, are most common and readily observed via microscopy (Oda et al., 2007; Craig et al., 2006). However, motile cells that lack any obvious appendages have also been documented (Jarrell and McBride, 2008). Research conducted to unravel the mysteries in assembly and function of motility structures, especially bacterial flagella and type IV pili, has been quite successful and will be reviewed in this study. Meanwhile, many archaeal species are known to be flagellated with a structure that superficially resembles that found in bacteria (Jarrell et al., 1996). However, further investigation revealed that fundamental differences exist in regard to assembly and structure despite the similarities in appearance and function of the two appendages. This work will be discussed later in the study.

**The Bacterial Flagellum and its Assembly**

The bacterial flagellum is a propulsive structure found in variable numbers and spatial arrangements, depending on the bacteria as shown in Figure 1.2 (Macnab, 2004). For example, lophotrichous bacteria (e.g. *Spirillum vultans*) possess multiple flagella at the same cell location
Figure 1.1: Tree of life depicting the relationship between the three domains Bacteria, Archaea and Eukarya. Taken from http://www.learner.org/courses/envsci/visual/img_lrg/common_ancestor.jpg
while monotrichous bacteria (e.g. *Pseudomonas aeruginosa*) have a single flagellum.

Amphitrichous bacteria (e.g. *Aquaspirillum magnetotacticum*) contain two flagella at opposite poles of the cell and peritrichous bacteria (e.g. *E. coli*) possess flagella inserted around the entire cell surface. Although there are exceptions, flagella generally function such that a counter-clockwise rotation generates a thrusting force that propels the cell forward whereas a clockwise rotation induces a tumbling motion that permits a change in direction, which is important for chemotaxis. Energy for rotation is usually achieved through a proton motive force while in some cases Na\(^+\) ions are used (Macnab, 2004).

Bacterial flagella can run lengths of 15-20 µm (Ferris and Minamino, 2006). They are long, hollow tube structures (18-24 nm in diameter) that are composed of many types of proteins. Structurally, flagella are characterized by three regions called the basal body, hook and filament (Macnab, 2004). The basal body anchors the flagellum through the use of ring structures located in the cytoplasm (C ring), the cell membrane (MS ring), peptidoglycan (P ring) and outer membrane (L ring) for Gram negatives (Macnab, 2003). The rings are connected throughout the periplasm by rod proteins. Nearly all of the flagella proteins need to be exported across the cell membrane, with the exception of the proteins that form the C-ring and MS-ring (Macnab, 2003). The L and P-ring proteins are transported across the cytoplasmic membrane using the Sec pathway while the rest of the flagellar subunits, including those for the rod, hook and filament, are exported using a type III secretion apparatus located at the base of the flagellum within the C-ring (Homma et al., 1987; Hirano et al., 2003). The presence of capping proteins is important for the proper assembly of the different flagellar subunits after they are exported (Yonekura et al., 2000). Each region (rod, hook, and filament) is assembled with its respective capping protein. Rod proteins are exported first and assemble under the rod capping protein, FlgJ, followed by hook proteins with the aid of the hook cap, FlgD (Macnab, 2003). The rod capping protein has a unique muramidase activity that allows the rod to assemble past the peptidoglycan layer (Hirano et al., 2001). Synthesis of the filament occurs when subunits are exported through the central
<table>
<thead>
<tr>
<th>Structure</th>
<th>Flagella Type</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Monotrichous</td>
<td><em>Pseudomonas aeruginosa</em></td>
</tr>
<tr>
<td></td>
<td>Lophotrichous</td>
<td><em>Spirillum vultans</em></td>
</tr>
<tr>
<td></td>
<td>Amphitrichous</td>
<td><em>Aquaspirillum magnetotacticum</em></td>
</tr>
<tr>
<td></td>
<td>Peritrichous</td>
<td><em>Escherichia coli</em></td>
</tr>
</tbody>
</table>

Figure 1.2: Different possible arrangements of flagella on the bacterial cell. Taken from http://gsbs.utmb.edu/microbook/images/fig2_4.jpg.
canal created with the proper assembly of the rod and hook region and assembled beneath the capping protein (HAP2) at the distal end of the growing flagellum as shown in Figure 1.3 (Imada et al., 1998).

**The Bacterial Type IV Pilus and its Assembly**

Type IV pili are found in a variety of Gram negative and, more recently, Gram positive bacteria (Varga et al., 2006). Type IV pili are used by bacteria for many processes such as adhesion, DNA uptake, biofilm formation, phage transduction and twitching motility (Pelicic, 2008). Their ability to extend and retract is of particular interest to this study, since twitching motility requires extension and retraction of the pilus to drag the cell forward along a surface as shown in Figure 1.4 (Henrichsen, 1983; Burrows 2005). Unlike many other pili structures which require very few proteins for assembly, type IV pili use at least a dozen and these are usually encoded by genes within the same operon (Craig and Li, 2008). Interestingly, many of these genes are homologues to those found in type II secretion systems, including the ATPase that powers pilus extension and retraction (Russel, 1998).

Type IV pili are primarily composed of a highly conserved 15-20 kDa pilin subunit that is synthesized as a preprotein (Craig and Li, 2008). Many bacterial proteins are synthesized as preproteins, containing N-terminal signal peptide sequences that are ultimately cleaved and this allows for protein recognition and trafficking (Pohlschroder et al., 1997). Signal peptide sequences contain a classical motif at the N-terminus with a basic n-region, followed by a hydrophobic core (h-region) and ending in a carboxyl-terminal region (c-region) that includes the cleavage site recognized by signal peptidases (von Heijne, 1983). There are several classes of signal peptides, with each class associated with its corresponding signal peptidase that recognizes and cleaves the peptide sequence (Paetzel et al., 2002). Type IV pilins require their own peptidase (type IV prepilin peptidase (TFPP)) and contain an unique class III signal peptide in which the cleavage site is found after the n-region, leaving the hydrophobic core with the mature protein
Figure 1.3: Assembly of the bacterial flagella showing the export of subunits through the central channel via a type III secretion apparatus, before assembling at the distal end. Hook-associated proteins (HAP) provide a junction between the hook and filament (HAP1 and HAP3) or act as a cap protein that induces polymerization of the filament at the growing distal tip (HAP2). Taken from http://www.millerandlevine.com/km/evol/design2/fig-1.jpg.
Figure 1.4: Twitching motility used by bacteria with the aid of type IV pili. The pilus extends away from the bacterial cell and attaches on the surface. Retraction of the pilus by removing pilin monomers pulls the cell forwards. Taken from http://www.uni-muenster.de/imperia/md/images/biologie_allgmzoo/maier/research/twitch2.png
The hydrophobic region makes up the core of the pilus structure and is important for the assembly and interaction of the pilin subunits (Albers and Pohlschröder, 2009).

Assembly of type IV pili occurs at the cell membrane (Crowther et al., 2004). It is proposed that the minimum assembly apparatus requires an inner membrane protein, responsible for recruiting the remaining accessory assembly proteins, and an energy providing ATPase (Craig and Li, 2008; Turner et al., 1993; Nunn et al., 1990). Pilin subunits pool in the inner membrane where their signal peptides are cleaved by a membrane bound signal peptidase before being inserted at the base of the pilus filament. ATP-hydrolysis is needed for incorporation of the pilin subunit causing the filament to extrude outwards (Wolfgang et al., 2000). Accessory secretin proteins homopolymerize within the outer membrane to form a hole that allows the growing pilus to pass through (Collins et al., 2005). Similarly for retraction, ATP-hydrolysis provided by a different ATPase, removes the pilin subunits at the base of the structure and returns them to the cell membrane (Burrows, 2005).

**Comparison of the Archaeal and Bacterial Flagella**

The archaeal flagellum is a unique motility structure that superficially looks and functions as its bacterial counterpart. It is characterized by three regions similar to those in the bacterial flagellum; a cell anchoring portion (basal body equivalent), a hook and a filament (VanDyke et al., 2008b). However, notable differences throughout each substructure of the archaeal and bacterial flagella can be seen. Starting at the basal body, the composition of the archaeal anchoring structure remains unknown, with the only observation pertaining to knob-like structures seen in electron micrographs of extracted archaeal flagella, which bear little resemblance to the rod and ring complexes seen in bacterial flagella (Kalmokoff et al., 1988; Kupper et al., 1994). Archaeal anchoring structure genes and proteins have not yet been identified in any species and homologues to bacterial basal body proteins cannot be found (Faguy et al., 1992). This suggests that the archaeal flagellum is most likely attached to the cell wall and membrane by a novel mechanism. Differences also occur in the hook region where a single
“flagellin” protein has been shown to compose the hook in *M. maripaludis* and likely in *M. voltae* (Chaban et al., 2007; Bardy et al., 2002). A “flagellin” was also identified in *Halobacterium salinarum*, as forming the hook region (Tarasov et al., 2000; Beznosov et al., 2007). In bacteria, the hook protein is very different in sequence from the flagellins. Electron microscopy reveals that the length of the archaeal hook (72-265 nm) is also more variable than in bacteria (consistently around 55 nm) (Bardy et al., 2002). Studies on the filament reveal that the diameter is much smaller for archaea in comparison to bacteria (10-14 nm vs 18-24 nm) (Trachtenberg and Cohen-Krausz, 2006). Variations in the composition of the filament are also observed where multiple copies of a single bacterial flagellin are typically assembled rather than the multiple types of flagellins (variations in the flagellin proteins FlaB and FlaA) found in all archaea, with the exception of *Sulfolobus* species (Jarrell et al., 1996). Another fundamental difference between the two organelles is that the rotation of archaeal flagella is driven by ATP hydrolysis rather than the proton motive force that drives bacterial flagella (Streif et al., 2008). Finally, and very significantly, the archaeal filament lacks the 2 nm central channel observed in bacterial flagella, which is essential for bacterial flagella assembly via the type III secretion system (Trachtenberg and Cohen-Krausz, 2006), suggesting an alternative method of assembly must occur in archaeal flagella (Jarrell et al., 1996).

The gene region encoding flagella structural proteins or proteins involved in flagella assembly is depicted in Figure 1.5 (VanDyke et al., 2008b). Assembly of the archaeal flagella is predicted to resemble the mechanism seen for type IV pili (Bardy et al., 2004). This is supported by the conservation of the *flaHIJ* gene cluster within known flagellated archaeal species, where *flaI* and *flaJ* are homologues to the type IV pilin ATPase (PilT) and a conserved membrane protein that interacts with the ATPase (PilC), respectively (Thomas and Jarrell, 2001). Furthermore, archaeal flagellins are synthesized as preflagellins that contain a signal peptide sequence similar to those for prepilin proteins (Ng et al., 2007). Not surprisingly, the signal
Figure 1.5: Organization of flagella genes for selected archaeal species. Matching colours correspond to gene homologues while arrows indicate direction of transcription. In *H. salinarum*, B flagellin genes are adjacent to accessory flagellin genes while A flagellins are located elsewhere in the genome (Modified from VanDyke et al., 2008b).
peptidase proteins (FlaK and PibD) found in *M. voltae* and *Sulfolobus solfataricus* species, are homologues to the type IV prepilin peptidase (Bardy and Jarrell, 2002; Albers et al., 2003). Inactivation of *flaK* in *M. voltae* and *M. maripaludis* resulted in nonflagellated cells demonstrating that peptidase processing is required for flagellins to assemble into flagella (Bardy and Jarrell, 2003). Homology is also observed when comparing the sequences of the archaeal flagellin and type IV pilin, where the sequence of the N-terminal 40 amino acids is highly conserved (Ng et al., 2006). Finally, the lack of a central channel in the archaeal flagellum also supports a type IV pili method of assembly, where subunits are incorporated at the base of the structure (Trachtenberg and Cohen-Krausz, 2006).

While sharing similarities to type IV pili, assembly of the archaeal flagella also involves unique proteins that are not found in any other domain (Thomas and Jarrell, 2001). In-frame deletion studies revealed that within the encoded region of flagellar accessory proteins, FlaCDEFG, all the genes were essential for flagellation, with the possible exception of *flaD* and *flaE*, which could not be deleted (Chaban et al., 2007). However, attempts to identify the structure, localization and function of these proteins have only recently begun.

**Protein Glycosylation**

Glycosylation is the attachment of polysaccharides (or glycans) to specific amino acids (sometimes within conserved motifs) within proteins and is likely an important factor for proper assembly, folding, stability and function (Helenius and Aebi, 2001; Mitra et al., 2006; Hanson et al., 2009). There are several types of glycosylation, with *O*^-^linked and *N*^-^linked systems found in prokaryotes. *O*^-^linked glycosylation is characterized by the attachment of the glycan to the hydroxyl group found on the side chains of serine and threonine (Peter-Katalinic, 2005). The majority of bacteria use *O*^-^linkage for protein glycosylation although there are several exceptions, with *Campylobacter* being the best studied due to the presence of both *O*^-^linked and *N*^-^linked glycosylation systems (Szymanski et al., 2003). *N*^-^linked glycosylation is the attachment of glycan through a β-glycosylamide linkage to the NH\(_2\) group found on the side chain of asparagine.
residues within the sequon Asn-Xaa-Ser/Thr, where Xaa is any amino acid other than proline (Weerapana and Imperiali 2006). N-linked glycosylation was once believed to be a unique eukaryotic trait until the discovery of N-linked glycans in the S-layer of the archaeon *Halobacterium salinarum* and later in the bacterium *Campylobacter jejuni*, which demonstrated the prevalence of N-linked glycosylation in all three domains of life (Mescher and Strominger, 1976; Szymanski et al., 2003).

**N-linked glycosylation in Eukaryotes**

N-linked glycosylation is essential in eukaryotes and has been most studied in the yeast, *Saccharomyces cerevisiae* (Burda and Aebi, 1999). Recent studies have demonstrated the role of N-linked glycoproteins in cell signaling and cell adhesion in cancer, as well as in accelerating protein folding and enhancing protein stability in the secretory pathway (Zhao et al., 2008; Hanson et al., 2009). Furthermore, deficiencies in glycosylation lead to a variety of disorders that are classified as congenital disorders of glycosylation. These disorders include deficiencies in growth and mental development, seizures and physical abnormalities (Leroy, 2006), and underline the importance of understanding the N-linked glycosylation pathway.

Eukaryote N-linked glycosylation begins on the cytoplasmic face of the rough endoplasmic reticulum (ER) membrane and terminates in the lumen of the ER with the transfer of a tetradecasaccharide (Glc$_3$Man$_9$GlcNAc$_2$) glycan to the N-linked motif in the protein (Burda and Aebi, 1999). The assembly of the glycan requires the sequential addition of activated nucleotide monosaccharides to a dolichol pyrophosphate lipid carrier embedded in the ER membrane, with the aid of glycosyltransferases or Alg (asparagine-linked glycosylation) proteins (entire pathway shown in Figure 1.6). The dolichol lipid is a polyisoprenoid composed of 80-100 carbons and differs from the undecaprenyl lipid carrier used in the bacterial N-linked pathway due to its terminating $\alpha$-saturated isoprenoid containing an alcohol functional group (Zhang et al., 2008) It is not present in the final glycan linkage and is used only for glycan assembly and transport.
Figure 1.6: N-linked glycosylation pathway for eukaryotes occurring at the endoplasmic reticulum (ER) membrane (Modified from Weerapana and Imperiali, 2006).
In eukaryotes, the linking sugar to the dolichol lipid is \( N \)-acetylglucosamine (GlcNAc) and it is attached by the \( N \)-acetylglucosamine-1-phosphate transferase called Alg7 (Kukuruzinska and Robbins, 1987). The second GlcNAc is attached by a protein complex consisting of Alg13/14, where Alg13 is the catalytic glycosyltransferase unit while Alg14 recruits Alg13 to the ER membrane (Gao et al., 2005; Gao et al., 2008). Finally, five mannose monosaccharides are added to the growing glycan through the activities of a series of glycosyltransferase in the order of Alg1, Alg2 and Alg11 (Gao et al., 2004). Alg1 attaches the first mannose in an unique \( \beta-1,4 \) linkage while Alg2 and Alg11 are each responsible for attaching two successive mannoses (Takeuchi et al., 1999; Cipollo et al., 2001).

The Man\(_5\)GlcNAc\(_2\)-dol intermediate is then translocated or flipped from the cytoplasmic side to the luminal side of the ER. Until recently, it had been widely accepted that this process was independent of ATP and required the protein Rft1, which is a transmembrane protein found only in the ER of eukaryotic cells that are capable of \( N \)-linked glycosylation (Helenius et al., 2002). However, the role of Rft1 as the flippase was recently questioned by Frank et al (2008) who suggested that a novel ER protein is performing the flipping function. The identity of this new flippase protein remains unknown and with early experiments only performed \textit{in vitro} (Sanyal et al., 2008; Sanyal and Mennon, 2009). Much debate on the true identity of the \( N \)-linked flippase remains, with its resolution being highly anticipated (Frank et al., 2008).

Once inside the lumen of the ER, sugar donors are no longer nucleotide-activated and instead are dolichol-phosphate linked (Dol-P-Man and Dol-P-Glc). Four additional mannose sugars are attached inside the lumen of the ER through the activity of Alg3, followed by Alg9, then Alg12 and finally Alg9 again (Verostek et al., 1993; Burda et al., 1999; Frank and Aebi, 2005). Alg6, Alg8 and Alg10 complete the glycan with the sequential addition of three glucose sugars (Reiss et al., 1996; Stagljar et al., 1994; Burda and Aebi, 1998). The final structure of the attached glycan is shown in Figure 1.7. In the final step of the \( N \)-linked process, the completed glycan is recognized by the oligosaccharyl transferase (OST) complex, and transferred to select
Figure 1.7: The N-linked glycan structure (\(\text{Glc}_3\text{Man}_9\text{GlcNAc}_2\); with the type of linkage occurring between each sugar) attached to a dolichol lipid used in eukaryotic N-glycosylation. Figure was taken from http://media.wiley.com/Lux/44/30644.1.gif.
asparagine residues within the N-linked sequon (Knauer and Lehle, 1999). Interestingly, N-linked glycans are not attached to all possible sequons, which may be a result of a variety of factors including protein folding and access for glycan attachment (Jones et al., 2005). The OST complex is composed of multiple proteins that span across the ER membrane. The largest subunit, Stt3p, contains the catalytic domain of the complex which is characterized by the amino acid sequence, WWDYG (Zufferey et al., 1995). This sequence is highly conserved in oligosaccharyl transferases across all three domains of life. The glycan can be processed further within the ER and Golgi complex with additional glycosylation, trimming, sulfation and epimerization (where sugars have the same molecular formula but enzymatic activity alters the stereochemistry at one of its stereogenic centers) (Szymanski and Wren, 2005).

**N-linked glycosylation in Bacteria**

The basic concepts of the eukaryotic N-linked glycosylation system are found in the equivalent bacterial system where a heptasaccharide (GlcNAc₂GalNAc₅, shown in Figure 1.8) is assembled on a lipid carrier embedded in the cytoplasmic face of the inner membrane before being flipped outwards to the periplasmic side. Unlike the eukaryotic pathway, no additional sugars are added to the glycan after flipping has occurred while an undecaprenyl pyrophosphate lipid is utilized as the lipid carrier rather than a dolichol pyrophosphate lipid (Chen et al., 2007). Genes residing in the model *Campylobacter* N-glycosylation gene cluster were given the prefix *pgl* (protein glycosylation). The functional transfer of the entire *pgl* gene cluster into *E. coli* was significant for further investigation of the N-linked system due to the molecular tools available in *E. coli* (Wacker et al., 2002).

Steps in the assembly of the bacterial N-linked glycan in *Campylobacter* are well studied and are shown in Figure 1.9 (Szymanski et al., 2003). The initial sugar, GlcNAc, is attached to the undecaprenyl lipid on the cytoplasmic face of the cell membrane by the glycosyltransferase, PglC (Glover et al 2006). The GlcNAc is then modified into a bacillosamine sugar by the enzymes PglF, PglE and PglD which function as a dehydratase, aminotransferase and
Figure 1.8: Structure of the bacterial N-linked glycan (GlcNAcGalNAc₅Bac; where Bac is a bacillosamine sugar) attached to the side chain of an asparagines residue (Weerapana and Imperiali, 2006).
Figure 1.9: Model N-linked glycosylation pathway located at the cytoplasmic membrane of bacterial cells (Modified from Weerapana and Imperiali, 2006).
acetyltransferase, respectively (Olivier et al., 2006). Next, five N-acetylgalactosamines (GalNAc) are assembled onto the glycan through the ordered actions of PglA, PglJ and PglH (Glover et al., 2005). Deficient pglH mutants accumulate a trisaccharide glycan and additional studies demonstrated that PglH is responsible for attaching the final three GalNAc sugars (Weerapana et al., 2005; Glover et al., 2005).

The completed heptasaccharide glycan is then flipped across the cell membrane to the periplasmic side. In contrast to the ATP-independent method of Rft1 in eukaryotes, bacteria require the use of ATP to translocate the glycan across membranes (Alaimo et al., 2006). A single ATP Binding Cassette (ABC)-transporter protein, PglK, is sufficient for glycan flipping as proven using the E. coli strain that possessed the Campylobacter pgl gene cluster (minus pglK) but was deficient in all known flippases (Kelly et al., 2006; Alaimo et al., 2006). Restoration of N-linked glycoproteins was observed when pglK was introduced via a plasmid (Alaimo et al., 2006). In contrast to the oligosaccharyl transferase complex used in eukaryotes, a single protein transfers the glycan to proteins with N-linked asparagine residues. The bacterial transferase is PglB, which contains the conserved catalytic domain of WWDYG, and is a homologue of the Stt3p subunit found in the eukaryotic OST complex (Glover et al., 2005).

**N-linked glycosylation in Archaea**

Until recently, knowledge of the archaeal N-linked system was fragmented although it is believed to follow a similar process to the eukaryotic and bacterial systems. Early studies in the archaeon Halobacterium salinarum, reported S-layer glycoproteins that contained both O- and N-linked glycans (Mescher and Strominger, 1976). Localization of the N-linked system was determined to be at the cell membrane and periplasm, as in bacteria, due to the observation that the cell-impermeable antibiotic bacitracin was capable of disrupting N-glycosylation (Mescher et al., 1976). Furthermore, peptides containing the N-linked sequon, and incapable of entering the cell, became glycosylated when introduced to growing cultures (Lechner et al., 1985). The archaeal N-linked system also shares similarities with the eukaryotic equivalent by using dolichol
as its lipid carrier (Burda and Aebi, 1999). The most recent archaeal N-linked glycosylation genetic studies have been fairly successful in species such as *Haloferax volcanii*, *M. voltae* and *M. maripaludis* thanks in part to the advancements of genetic tools for these organisms.

Mass spectrometry studies determined that the methanogens, *M. voltae* and *M. maripaludis*, possessed a trisaccharide and tetrasaccharide glycan, respectively, attached to flagellins and S-layer (Figure 1.10). The structure of the *M. voltae* glycan is $\beta$-ManNAcA6Thr-(1-4)-$\beta$-GlcNAc3NAcA-(1-3)-$\beta$-GlcNAc, with the attachment to the protein occurring through GlcNAc (Voisin et al., 2005). The *M. maripaludis* glycan is Sug-4-$\beta$-ManNAc3NAmA6Thr-4-$\beta$-GlcNAc3NAcA-3-$\beta$-GalNAc, with the terminal Sug being a novel monosaccharide, a diglycoside of aldulose or 2NAc-2,4 dideoxy-hex-5-uloze (Kelly et al., 2009). Several differences in the *M. voltae* and *M. maripaludis* glycans are observed including the nature of the initial sugar attached to dolichol, where GlcNAc and GalNAc are used, respectively (Chaban et al., 2006; VanDyke et al., 2009). The second sugar, GlcNAc3NAcA, is identical in both glycans while the third sugar is similar with the exception of an additional acetamidino group in *M. maripaludis*. The glycan of *M. maripaludis* includes a fourth sugar, a diglycoside of aldulose, which is the first reported natural occurrence of this sugar (VanDyke et al., 2009). Meanwhile, a pentasaccharide glycan was determined to be attached on *H. volcanii* glycoproteins (Yurist-Doutsh and Eichler, 2009). Although the specific structure of the pentasaccharide has not been determined, it is known to be composed of two hexoses, two hexuronic acids and an additional 190 Da saccharide (Abu-Qarn et al., 2007). The variable length of glycan and its composition demonstrate the diversity of the N-linked system among archaea (Yurist-Doutsh et al., 2008). Genetic studies in *M. voltae*, *M. maripaludis* and *H. volcanii* have identified several genes involved in the N-linked process and these have been annotated as *agl* (archaeal glycosylation) genes.

The identification of archaeal glycosyl transferases has only recently begun. In *M. voltae*, Chaban et al (2006) demonstrated that AglA was responsible for attaching the terminal sugar of
Figure 1.10: A) N-linked glycan structure (from c to a; β-ManpNAcA6Thr-(1-4)-β- GlcNAc3NacA-(1-3)-β-GlcNac) for the archaeal organism, *M. voltae* attached to an Asn residue (Chaban et al., 2006). B) N-linked glycan structure (from A to D; Sug-4-β-ManNAc3NAmA6Thr-4-β- GlcNAc3NacA-3-β-GalNac, where Sug is a novel diglycoside of aldulose) for the archaeal organism, *M. maripaludis* attached to an Asn residue (Kelly et al., 2009).
the glycan, while AglC and AglK were both required for second sugar attachment. In *M. voltae*, the glycosyl transferase AglH is a homologue of the yeast eukaryotic glycosyl transferase Alg7 (*N*-acetylglucosamine-1-phosphate transferase) and is believed to be responsible for the addition of the initial sugar to the dolichol lipid. Attempts to link this function to AglH through genetic knockout studies were unsuccessful likely due to the gene’s presumed involvement in the synthesis of co-factor B (7-mercaptoheptanoylthreoninephosphate), which is essential for methanogens (Noll et al., 1986). Co-factor B performs the redox reaction $\text{CH}_3\text{-S-CoM+ HS-CoB} \rightarrow \text{CH}_4 + \text{CoB-S-S-CoM}$ (Sauer, 1986). It contains a glycan that may require the activity of AglH for attachment (White, 2001; Chaban et al., 2006). The role of AglH was indirectly supported by the observation that AglH was capable of restoring the function of Alg7 in a conditional lethal *S. cervisiae* strain when *alg7* was deactivated (Shams-Eldin et al., 2008). While this provides strong evidence for the role of AglH in glycan assembly for *M. voltae* because the GlcNAc-dolichol linkage is consistent in both species, there is debate on whether the homologue found in *M. maripaludis, MMP1423*, can perform the same first step in assembly since the initial sugar of the glycan, GalNAc, is different. In the case of the *M. maripaludis* tetrasaccharide glycan, two of the predicted four glycosyl transferases (MMP1079 and MMP1080, responsible for attachment of the 2nd and 3rd sugar) have been found (Van Dyke et al., 2009). The fourth glycosyl transferase of *M. maripaludis* has not been identified. Similar advances have also been made in studies of glycan assembly in *H. volcanii*, where glycosyl transferases (AglD, AglE, AglF, AglG and AglI) have been identified (Abu-Qarn et al., 2007; Abu-Qarn et al., 2008 and Yourist-Doutsch et al., 2008).

Currently, the archaeal flippase remains unidentified with only a handful of genes considered strong candidates. However in these cases, either the gene could not be deleted or the gene knockout did not have the expected phenotype of a flippase mutant (Chaban et al., 2006; Van Dyke et al., 2009). On the other hand, the oligosaccharyl transferase, AglB, was easily identified in all organisms since the archaeal transferase is an obvious homologue of the Stt3p protein used by both eukaryotes and bacteria (Chaban et al., 2006; Van Dyke et al., 2009).
Glycosylation and Archaeal Protein Function

In eukaryotes, deficiencies in N-glycosylation are lethal and the glycan likely plays a role in protein function, transport, folding, sorting and for localization. In bacteria where N-linked glycosylation is rare, O-linked glycans are frequently observed on their structures (Messner, 2004). In archaea, N-glycosylation likely occurs throughout the domain although thus far only a few species including H. volcanii, H. salinarum, M. voltae and M. maripaludis, have been definitely shown to contain N-glycosylated proteins. The latter three organisms are flagellated and N-linked glycans can be found on their flagellin proteins as well as the S-layer protein that comprise their cell wall (Logan, 2006). H. volcanii is not flagellated but possesses N-linked glycoproteins as its S-layer cell wall (Abu-Qarn et al., 2008b). A function of the N-linked glycan was demonstrated in a H. volcanii non-glycosylated mutant strain. These cells grew significantly less in elevated salt conditions due to a disrupted S-layer caused by the lack N-linked glycan (Abu-Qarn et al., 2007). Similarly, in both M. voltae and M. maripaludis glycosyl transferase mutants, a partial loss of glycan resulted in deficient flagella motility or loss of assembled flagella, indicating that a truncated glycan leads to defective motility and a minimum two sugar glycan is needed for flagella assembly to occur (Chaban et al., 2006; VanDyke et al., 2009). In another M. maripaludis study, an acetyl transferase gene needed for the complete biosynthesis of the second monosaccharide in the glycan was deleted. In addition to the detrimental effects on flagella assembly, this mutant had an unexpected pilus phenotype. Unattached pili were observed in the growth media indicating that cells were able to export and assemble pili, but were defective in the attachment of the pili to the cell (VanDyke et al., 2008). This suggests a different role for glycosylation (either N-linked or O-linked, since the second monosaccharide could potentially be used in both pathways) where the glycan is involved in the attachment and anchoring of extracellular structures.

Aim of this Study
The purpose of this study was to complete the identification of the proteins required for assembly of the \( N \)-linked glycan in \( M. \) maripaludis. Currently, only the Stt3p transferase homologue and two of the four potential glycosyl transferases have been discovered, while the identification of the flippase remains elusive.

The first goal was to identify the glycosyl transferase responsible for attaching the fourth sugar on the tetrasaccharide glycan by using in-frame deletion techniques that were previously established (Moore and Leigh, 2005). In-frame deletions have advantages over more traditional methods, such as insertional inactivation, because any single gene can be studied, regardless of its location within an operon. Furthermore, mutant strains can be complemented by returning the deleted gene via plasmid. Immunoblotting against flagellin proteins was used to determine if any reduction in molecular weight occurred due to interference in assembly of the complete glycan caused by the deletion of the targeted gene. Motility of the mutant was assessed by using semi-swarm plates and effects on flagella assembly detected through electron microscopy.

The second objective was to identify the glycosyl transferase attaching GalNAc to the dolichol lipid. As previously mentioned, \( MMP1423 \) cannot be deleted because of its presumed involvement in co-factor B production. This gene’s function is to attach GlcNAc to a dolichol lipid, as reported in \( M. \) voltae and \( S. \) cerevisiae, and was initially believed to do the same in \( M. \) maripaludis. However with recent NMR and mass spectrometry data, the linking sugar in the \( M. \) maripaludis glycan was identified as GalNAc, not GlcNAc, so the role of \( MMP1423 \) in the \( M. \) maripaludis \( N \)-linked pathway can be questioned. Deletion of the remaining annotated glycosyl transferases was attempted in hopes of identifying a novel \( N \)-linked glycosyl transferase or to deduce the function of \( MMP1423 \) by eliminating all other possibilities.

The third goal of this study was to identify the elusive flippase. With a limited number of candidates, several deletion techniques, including in-frame and insertional inactivation, were attempted on a gene believed to be a strong flippase candidate due to its proximity to known \( N \)-linked genes and its annotation as a surface polysaccharide biosynthesis protein. A flippase
mutant would be expected to contain non-glycosylated flagellin proteins which can readily be verified by immunoblotting techniques and mass spectrometry.

The last goal of the study was to begin to identify genes believed to be involved in the biosynthesis and modification of the unusual sugars of the N-linked glycan as this aspect of the glycan assembly has thus far been overlooked in archaea.

This study aimed to further the understanding of the archaeal N-glycosylation system by reaching these four goals.
Chapter 2

Materials and Methods

2.1 Strains

*Escherichia coli* and *Methanococcus maripaludis* strains used in this study are shown in Table 2.1. *E. coli* DH5α was used for plasmid cloning while *M. maripaludis* MM900 was used for markerless in-frame deletion of targeted genes. *M. maripaludis* MM900 originated from the wild-type *Methanococcus maripaludis* S2, but is missing the MMP0145 gene, which encodes for the hypoxanthine phosphoribosyltransferase (Hpt) protein. As a result, *M. maripaludis* MM900 is resistant to the toxic nucleotide base analogue, 8-azahypoxanthine (Moore and Leigh, 2005).

2.2 *Methanococcus maripaludis* Plasmids

pCRPrtNeo is a 6.54 kb plasmid with encoded ampicillin and neomycin resistance and is shown in Figure 2.1 (Moore and Leigh, 2005). The vector also contains the *M. maripaludis* hpt gene, MMP0145. pCRPrtNEO cannot replicate in *M. maripaludis* and is used to introduce the in-frame deletion fragment to initiate the marker-less deletion process (Moore and Leigh, 2005).

pHW40 is a 10.4 kb vector with encoded ampicillin and puromycin selection and is used for gene complementation (obtained from John Leigh and shown in Figure 2.1). Transcription of the desired gene cloned in pHW40 is controlled through the nitrogen fixation (*nif*) promoter, where the presence of alanine, as the sole nitrogen source, induces gene expression and ammonia, as the sole nitrogen source, represses transcription (Lie and Leigh, 2005).

pPAC60 is a non-replicating plasmid with encoded ampicillin and puromycin resistance and is shown in Figure 2.1 (Thomas et al., 2001). Therefore, puromycin resistance of cells transformed with this plasmid only occurs when the plasmid integrates in the chromosome. pPAC60 is used to disrupt and inactivate desired genes (along with any co-transcribed downstream genes) by homologous recombination into the genome via a truncated version of the targeted gene cloned into pPAC60.
<table>
<thead>
<tr>
<th>Strains</th>
<th>Description</th>
<th>Source or Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>M. maripaludis</em> strains</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S2</td>
<td>wildtype</td>
<td>Whitman et al., 1986</td>
</tr>
<tr>
<td>MM900</td>
<td><em>M. maripaludis</em> S2 Δhpt</td>
<td>Moore and Leigh, 2005</td>
</tr>
<tr>
<td>MM359</td>
<td>MM900 ΔMMP359</td>
<td>This study</td>
</tr>
<tr>
<td>MM590</td>
<td>MM900 ΔMMP590</td>
<td>This study</td>
</tr>
<tr>
<td>MM1293</td>
<td>MM900 ΔMMP1293</td>
<td>This study</td>
</tr>
<tr>
<td>MM1294</td>
<td>MM900 ΔMMP1294</td>
<td>This study</td>
</tr>
<tr>
<td>MM1088</td>
<td>MM900 ΔMMP1088</td>
<td>This study</td>
</tr>
<tr>
<td>MM1081</td>
<td>MM900 ΔMMP1081</td>
<td>This study</td>
</tr>
<tr>
<td>MM1082</td>
<td>MM900 ΔMMP1082</td>
<td>This study</td>
</tr>
<tr>
<td>MM1083</td>
<td>MM900 ΔMMP1083</td>
<td>This study</td>
</tr>
<tr>
<td>FlaJ</td>
<td>MM900 ΔFlaJ</td>
<td>This study</td>
</tr>
<tr>
<td><em>E. coli</em> strain</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DH5α</td>
<td>Φ80lacZ ΔM15 Δ(lacZYA-argF) endA1 recA1</td>
<td>Laboratory strain</td>
</tr>
</tbody>
</table>
In addition to the *M. maripaludis* plasmids, pCR2.1 TOPO vector was sometimes used for cloning of PCR products. All *M. maripaludis* plasmids used in this study are shown in Table 2.2.

### 2.3 Media and Growth Conditions

*E. coli* cultures were grown in liquid (with shaking at 120 rpm) and solid Luria-Bertani (LB) media at 37°C. Solid plates require the addition of 15 g/L of agar. Ampicillin (100 µg/ml) was used for selection, when required.

*M. maripaludis* cultures were grown anaerobically, overnight with gentle shaking (100 rpm) at 30°C in Balch media III under an atmosphere of CO₂/H₂ (20:80). See Appendix B for Balch medium, III composition.

The transformation protocol for markerless gene deletion in MM900 requires the use of McCas media rather than Balch III medium (See Appendix B for McCas composition). In McCas medium, yeast extract and peptase that are present in Balch III are replaced with casamino acids which, is necessary for the 8-azahypoxanthine selection step to work. Plates of McCas agar media (standard McCas with 20 g of Noble agar) are used for single colony isolation and 8-azahypoxanthine selection (240 µg/ml). Selection using 8-azahypoxanthine does not occur when using less purified types of agar. Neomycin is used at a concentration of 1 mg/ml when required. Complementation of deleted genes in *M. maripaludis* required nitrogen-free media (Blank et al., 1995) (See Appendix B for recipe). Puromycin (2.5 µg/ml) is used for complementation plasmid selection.

### 2.4 *E. coli* Cloning and Transformation:

All plasmid cloning was performed with *E. coli* subcloning efficiency DH5α chemically competent cells (Invitrogen). Occasionally TOPO cloning kits (Invitrogen) were used to clone PCR products prior to subcloning into the appropriate *M. maripaludis* vectors. Transformation of plasmids into DH5α cells used the method recommended by the manufacturer (Invitrogen).
Figure 2.1: A) pCRPrtNeo vector used to generate in-frame deletions in *M. maripaludis* cells. The vector is driven by the *hmv* promoter and includes the *hpt* gene and encoded neomycin, ampicillin and kanamycin resistance (Moore and Leigh, 2005). B) pHW40 vector with ampicillin and puromycin resistance used for complementation of deleted genes. The *lac* gene is removed from the vector and desired genes are cloned into the same vector location. The *nif* promoter drives gene expression when alanine is used as the sole nitrogen source, while ammonia does not (Lie and Leigh, 2005). C) The pPAC60 vector used for insertional deactivation of targeted genes. The vector has encoded ampicillin and puromycin resistance and contains two multi-cloning sites (MCS) driven by the *phmvA* promoter (Thomas et al., 2001).
<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Description</th>
<th>Source or Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pCRPrtNeo</td>
<td><em>hmv</em> promoter-<em>hpt</em> fusion + Neo&lt;sup&gt;+&lt;/sup&gt; cassette in pCR2.1Topo, Amp&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Moore and Leigh, 2005</td>
</tr>
<tr>
<td>pKJ741</td>
<td>pCRPrtNeo with in-frame deletion of <em>MMP90</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKJ703</td>
<td>pCRPrtNeo with in-frame deletion of <em>MMP356</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKJ709</td>
<td>pCRPrtNeo with in-frame deletion of <em>MMP359</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKJ750</td>
<td>pCRPrtNeo with in-frame deletion of <em>MMP590</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKJ737</td>
<td>pCRPrtNeo with in-frame deletion of <em>MMP1293</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKJ739</td>
<td>pCRPrtNeo with in-frame deletion of <em>MMP1294</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKJ665</td>
<td>pCRPrtNeo with in-frame deletion of <em>MMP1088</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKJ685</td>
<td>pCRPrtNeo with in-frame deletion of <em>MMP1089</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKJ713</td>
<td>pCRPrtNeo with in-frame deletion of <em>MMP1088/MMP1089</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKJ735</td>
<td>pCRPrtNeo with in-frame deletion of <em>MMP1081</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKJ751</td>
<td>pCRPrtNeo with in-frame deletion of <em>MMP1082</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKJ757</td>
<td>pCRPrtNeo with in-frame deletion of <em>MMP1083</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKJ721</td>
<td>pCRPrtNeo with in-frame deletion of <em>FlaJ</em></td>
<td>This study</td>
</tr>
<tr>
<td>pPAC60</td>
<td><em>phmvA</em> promoter + Pur&lt;sup&gt;R&lt;/sup&gt; cassette</td>
<td>Thomas et al., 2001</td>
</tr>
<tr>
<td>pKJ693</td>
<td>pPAC60 with truncated <em>MMP1089</em></td>
<td>This study</td>
</tr>
<tr>
<td>pHW40</td>
<td><em>nif</em> promoter-<em>lacZ</em> fusion + Pur&lt;sup&gt;r&lt;/sup&gt; cassette, Amp&lt;sup&gt;r&lt;/sup&gt;</td>
<td>Obtained from J. Leigh</td>
</tr>
<tr>
<td>pKJ679</td>
<td>pHW40 with C-terminal his-tagged <em>MMP1088</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKJ752</td>
<td>pHW40 with C-terminal his-tagged <em>MMP1081</em></td>
<td>This study</td>
</tr>
</tbody>
</table>
Briefly, 50 uL of subcloning DH5α cells were incubated on ice with plasmid DNA for 30 min. The cells were then heat shocked at 42°C for 30 seconds, returned to ice for 2 minutes, followed by the addition of 950 µL of LB. After a 1 hour recovery at 37°C with shaking (200 rpm), cells were plated onto LB plates with ampicillin (100µg/mL) and grown overnight. For TOPO cloning, the plates also contained X-gal (40 µg/mL) to distinguish transformants containing vector with insert (white colonies) from vector alone (blue colonies). Single colonies are picked and grown overnight for screening. Plasmid DNA was isolated from E. coli with QIAprep® Spin Miniprep Kit (Qiagen) and digested with appropriate restriction enzymes.

Agarose (0.8%) gels containing ethidium bromide (0.00003%) were ran in TAE Buffer (40 mM Tris-HCl, pH 8.0, 0.1% glacial acetic acid, 1 mM EDTA) (Maniatis et al., 1982) to visualize DNA in the presence of UV light.

2.5 Restriction Enzyme Digestions and Ligations

Most restriction enzymes were obtained from New England Biolabs (NEB) or Fermentas and DNA was digested using their recommended buffers.

Rapid Ligation Kits (Fermentas) and protocol were used for all DNA ligations and vector cloning.

2.6 M. maripaludis Chromosomal DNA Isolation

Chromosomal DNA was isolated from desired M. maripaludis strains using a previously described protocol (Gernhardt et al., 1990). Briefly, 10mL of culture were centrifuged for 10 minutes at 5000 rpm to recover cells. Cells are then resuspended in 600 µL of TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) which causes the cells to partially lyse. Further cell lysis is induced by the addition of 6 µL of 10% sodium dodecyl sulphate (SDS) stock. Incubation with 4 µL of RNase (60 µg/mL, Fermentas) at 37°C for 1 hr, followed by 8 µl of Proteinase K (230 µg/mL, Fermentas) digestion at 55°C for 45 minutes allows for degradation of the RNA and proteins, respectively. DNA is initially extracted with phenol:chloroform (1:1 ratio) twice,
followed by a final extraction with only chloroform. DNA was precipitated with 95% ethanol and stored in distilled water (dH₂O).

2.7 Polymerase Chain Reaction (PCR)

Primers used for polymerase chain reactions (PCR) are shown in Tables 2.3, 2.4 and 2.5. A standard PCR 50 µl reaction requires: 40.5 µl of dH₂O, 5 µl of 10X Polymerase Thermobuffer (New England Biolabs), 1 µL each of forward and reverse primer (50 pmol), 1 µl dNTP mix (0.8 mM), 0.5 µl Vent Polymerase (New England Biolabs) and 1 µl of DNA template. PCR reactions tubes are then placed in a Mastercycler epgradient S (Eppendorf) and generally undergo the following cycle: 94°C for 5 minutes, followed by 30 cycles of 94°C for 45 seconds, primer annealing temperature of 55°C for 45 seconds and 72°C extension time based on 1 kb extension per minute. Primer annealing temperatures varied in some reactions because a gradient was used to optimize the amount and specificity of the desired PCR product. The PCR reaction then finishes with a final extension time of 10 minutes at 72°C to ensure completion of all reactions.

2.8 Transformation of M. maripaludis

Transformation of pCRPrtNeo deletion plasmids and pHW40 complementation plasmids into M. maripaludis was performed under anaerobic conditions according to the methodologies reported in previous studies (Tumbula et al., 1994). Briefly, 5 mL of M. maripaludis overnight cultures were inoculated into fresh media and grown for 1-2 hours at 37°C to allow cells to enter log phase. Cells were then harvested anaerobically by centrifuging 5 mL of the log phase culture at 5000 rpm for 10 minutes in sealed tubes. The pellets were resuspended in Transformation Buffer (TB; 50 mM Tris-HCl, pH 7.5, 0.35 M sucrose, 0.38 M NaCl, 1 mM MgCl₂·6H₂O, 0.02% cysteine-HCl, 1 mM dithiothreitol, and 0.00001% resazurin) and centrifuged at 5000 rpm for 20 minutes. Re-suspension of cells in 375 µL of TB with 5 µg of plasmid DNA in dH₂O and 225 µL PEG solution (40% PEG8000 solution with TB) along with an incubation of 1 hour (37°C without shaking), permits plasmid uptake. Following addition of McCas media (5 mL), the cells were centrifuged (5000 rpm for 20 minutes) and the pellet resuspended and subcultured in fresh
### Table 2.3: Deletion primers used in this study

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence (5’ to 3’)</th>
<th>Restriction Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMP090_For_Up</td>
<td>CCGATCCGAGTTTGAGATTACGCATACG</td>
<td>BamHI</td>
</tr>
<tr>
<td>MMP090_Rev_Up</td>
<td>TGGCGCGGCCAGGTGCTGCATAGTGCGTGG</td>
<td>Ascl</td>
</tr>
<tr>
<td>MMP090_For_Down</td>
<td>TGGCGCGGCCGAATCGATAGTGGCTTG</td>
<td>Ascl</td>
</tr>
<tr>
<td>MMP090_Rev_Down</td>
<td>CCGATCCAGTTCTCTGCTCTGTCC</td>
<td>BamHI</td>
</tr>
<tr>
<td>MMP356_For_Up</td>
<td>CCGATCCCATACCGGAGCGACTATGAC</td>
<td>BamHI</td>
</tr>
<tr>
<td>MMP356_Rev_Up</td>
<td>TGGCGCGGCCAGGTGCTGCATAGTGCGG</td>
<td>Ascl</td>
</tr>
<tr>
<td>MMP356_For_Down</td>
<td>TGGCGCGGCCGAATCGATAGTGGCTTG</td>
<td>Ascl</td>
</tr>
<tr>
<td>MMP356_Rev_Down</td>
<td>CCGATCGCAAAACAAACCGGATAATACGAG</td>
<td>BamHI</td>
</tr>
<tr>
<td>MMP359_For_Up</td>
<td>CCGATCCATATGTACTGGCCGATTATCC</td>
<td>BamHI</td>
</tr>
<tr>
<td>MMP359_Rev_Up</td>
<td>TGGCGCGGCCAGGTGCTGCATAGTGCGG</td>
<td>Ascl</td>
</tr>
<tr>
<td>MMP359_For_Down</td>
<td>TGGCGCGGCCGAATCGATAGTGGCTTG</td>
<td>Ascl</td>
</tr>
<tr>
<td>MMP359_Rev_Down</td>
<td>CCGATCCACCATAATGACGTACTATCC</td>
<td>BamHI</td>
</tr>
<tr>
<td>MMP590_For_Up</td>
<td>CCGATCCAGATTAGATCCTCATAAGC</td>
<td>BamHI</td>
</tr>
<tr>
<td>MMP590_Rev_Up</td>
<td>TGGCGCGGCCATCGTCTACGACAATACCTAC</td>
<td>Ascl</td>
</tr>
<tr>
<td>MMP590_For_Down</td>
<td>TGGCGCGGCCAGGTGCTGCATAGTGCGG</td>
<td>Ascl</td>
</tr>
<tr>
<td>MMP590_Rev_Down</td>
<td>CCGATCCCTGAAGGCGATTTAGATGCTG</td>
<td>BamHI</td>
</tr>
<tr>
<td>MMP1081_For_Up</td>
<td>CCGGATCCAGATACGGGAGATTAGAAG</td>
<td>BamHI</td>
</tr>
<tr>
<td>MMP1081_Rev_Up</td>
<td>ATGGCGCGGCCGTCAGTCTACGATAATATGC</td>
<td>Ascl</td>
</tr>
<tr>
<td>MMP1081_For_Down</td>
<td>ATGGCGCGGCCAGCGGAAGCAGTAGGACAG</td>
<td>Ascl</td>
</tr>
<tr>
<td>MMP1081_Rev_Down</td>
<td>CCGGATCCGAACCTCCATATGATGCTG</td>
<td>BamHI</td>
</tr>
<tr>
<td>MMP1082_For_Up</td>
<td>ATCGGGATCCCTGGCGTAGACAGTACTTATGTAGC</td>
<td>BamHI</td>
</tr>
<tr>
<td>MMP1082_Rev_Up</td>
<td>ATCGGCGCGCCCCATCCATAATCAATTATGC</td>
<td>Ascl</td>
</tr>
<tr>
<td>MMP1082_For_Down</td>
<td>ATCGGCGCGCCCTATGGAGATTCAATTCATCTG</td>
<td>Ascl</td>
</tr>
<tr>
<td>MMP1082_Rev_Down</td>
<td>ATCGGATCCGGAGCAATATTAATCTCCTTCGG</td>
<td>BamHI</td>
</tr>
<tr>
<td>MMP1083_For_Up</td>
<td>ATCGGGATCCCTAGAGCGGAAGCATTAGGACAG</td>
<td>BamHI</td>
</tr>
<tr>
<td>MMP1083_Rev_Up</td>
<td>ATCGGCGCGCGGATCACCTATATAATTTGGATC</td>
<td>Ascl</td>
</tr>
<tr>
<td>MMP1083_For_Down</td>
<td>ATCGGCGCGCCATCCGAGAGGAGATTAATTTGCTCC</td>
<td>Ascl</td>
</tr>
<tr>
<td>MMP1083_Rev_Down</td>
<td>ATCGGGATCCCGTGGTTATGCTGGCGATGCAGTGG</td>
<td>BamHI</td>
</tr>
<tr>
<td>MMP1088_For_Up</td>
<td>CCGATCCGCGTAGGTATGAAATATGTGTC</td>
<td>BamHI</td>
</tr>
<tr>
<td>MMP1088_Rev_Up</td>
<td>TGCGCGCGCCCTTTGTGGTACGTTGCAGCA</td>
<td>BamHI</td>
</tr>
<tr>
<td>MMP1088_For_Down</td>
<td>CGATCCGCGCCAACTAGCACCCTTTATCATTCAGC</td>
<td>Ascl</td>
</tr>
<tr>
<td>MMP1088_Rev_Down</td>
<td>CGAGATCCCGTAGGTATGAAATATGTGTC</td>
<td>BamHI</td>
</tr>
<tr>
<td>MMP1089_For_Up</td>
<td>GCCGTAGGTATGAAATATGTG</td>
<td>Xbal</td>
</tr>
<tr>
<td>MMP1089_Rev_Up</td>
<td>TGGCGCGGCCAAACAAATGCGGCGCCAA</td>
<td>Ascl</td>
</tr>
<tr>
<td>MMP1089_For_Down</td>
<td>TGGCGCGGCCATAAATCAGATGATTCCTC</td>
<td>Ascl</td>
</tr>
<tr>
<td>MMP1089_Rev_Down</td>
<td>GCTCTAGACCTGTGGATACGTTGCAGCA</td>
<td>Xbal</td>
</tr>
<tr>
<td>MMP1088/1089_For_Up</td>
<td>GCTCTAGAGCGTAGGTATGAAATATGTGTC</td>
<td>Xbal</td>
</tr>
<tr>
<td>MMP1088/1089_Rev_Up</td>
<td>TGCGCGCGCCCTTTGTGGTACGTTGCAGCA</td>
<td>BamHI</td>
</tr>
<tr>
<td>MMP1088/1089_For_Down</td>
<td>TGGCGCGGCCAACTAGCACCCTTTATCATTCAGC</td>
<td>Ascl</td>
</tr>
<tr>
<td>MMP1088/1089_Rev_Down</td>
<td>GCTCTAGACCTGTGGATACGTTGCAGCA</td>
<td>Xbal</td>
</tr>
<tr>
<td>MMP1089_For_Dis</td>
<td>GCAGATCCCGTAGGTATGAAATATGTGTC</td>
<td>EcoRI</td>
</tr>
<tr>
<td>MMP1089_Rev_Dis</td>
<td>GCAGATCCCGTAGGTATGAAATATGTGTC</td>
<td>EcoRI</td>
</tr>
<tr>
<td>MMP1293_For_Up</td>
<td>ATAGGATCCAGCTATCAAGCTG</td>
<td>BamHI</td>
</tr>
<tr>
<td>MMP1293_Rev_Up</td>
<td>TGGCGCGGCCCTTTGTGGTACGTTGCAGCA</td>
<td>Ascl</td>
</tr>
<tr>
<td>MMP1293_For_Down</td>
<td>TGGCGCGGCCATACGAAATACAGCTGGGATGC</td>
<td>Ascl</td>
</tr>
<tr>
<td>MMP1293_Rev_Down</td>
<td>CGGATCCAGCTATCAAGCTG</td>
<td>BamHI</td>
</tr>
<tr>
<td>MMP1294_For_Up</td>
<td>CGGATCCCTTGCAACAACAGGAGGTTC</td>
<td>BamHI</td>
</tr>
<tr>
<td>MMP1294_Rev_Up</td>
<td>TGGCGCGGCCCTTTGTGGTACGTTGCAGCA</td>
<td>Ascl</td>
</tr>
<tr>
<td>MMP1294_For_Down</td>
<td>TGGCGCGGCCAGACAGTTCTGAACGATACAG</td>
<td>Ascl</td>
</tr>
<tr>
<td>MMP1294_Rev_Down</td>
<td>CGGATCCCTTGCAACAACAGGAGGTTC</td>
<td>BamHI</td>
</tr>
<tr>
<td></td>
<td>Sequence</td>
<td>Restriction Site</td>
</tr>
<tr>
<td>------------</td>
<td>---------------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>FlaJ_For_Up</td>
<td>CGGATCCCTGTAGATGGTTCCCTTGCTG</td>
<td>BamHI</td>
</tr>
<tr>
<td>FlaJ_Rev_Up</td>
<td>TGCGCGCCGGACCAAACCGAGTGTAGCCTG</td>
<td>AscI</td>
</tr>
<tr>
<td>FlaJ_For_Down</td>
<td>TGGCGCGCCGGATGGAGGACACTACCCAGG</td>
<td>AscI</td>
</tr>
<tr>
<td>FlaJ_Rev_Down</td>
<td>CGGATCCCAAGCTCCATTGCATCAAGC</td>
<td>BamHI</td>
</tr>
<tr>
<td>Name</td>
<td>Sequence (5' to 3')</td>
<td></td>
</tr>
<tr>
<td>-----------------</td>
<td>------------------------------</td>
<td></td>
</tr>
<tr>
<td>MMP090_For_Seq</td>
<td>AATATAGTGATCCGTTAGC</td>
<td></td>
</tr>
<tr>
<td>MMP090_Rev_Seq</td>
<td>CCTCACAGCATGCAGGAACTATC</td>
<td></td>
</tr>
<tr>
<td>MMP356_For_Seq</td>
<td>TCTCCTGTCGGATACCTAG</td>
<td></td>
</tr>
<tr>
<td>MMP356_Rev_Seq</td>
<td>GGTTCCTGTTGTAGTTATGG</td>
<td></td>
</tr>
<tr>
<td>MMP359_For_Seq</td>
<td>CACTATTGCAATACTGAGTCC</td>
<td></td>
</tr>
<tr>
<td>MMP359_Rev_Seq</td>
<td>CAATACATCTGAAGGTAGTG</td>
<td></td>
</tr>
<tr>
<td>MMP590_For_Seq</td>
<td>CAACTTCAACAGAAGCACTAGCGG</td>
<td></td>
</tr>
<tr>
<td>MMP590_Rev_Seq</td>
<td>AAGGATGCTCCCAACAGAAACG</td>
<td></td>
</tr>
<tr>
<td>MMP1081_For_Seq</td>
<td>GATGCAGTATGCCATTTGAG</td>
<td></td>
</tr>
<tr>
<td>MMP1081_Rev_Seq</td>
<td>TCATTGGTTATAATAGCGTC</td>
<td></td>
</tr>
<tr>
<td>MMP1082_For_Seq</td>
<td>CTAGAGCGGAAGCATTAGGACAG</td>
<td></td>
</tr>
<tr>
<td>MMP1082_Rev_Seq</td>
<td>TATCATTGAATATCCTGACTGC</td>
<td></td>
</tr>
<tr>
<td>MMP1083_For_Seq</td>
<td>TCGAAGGATAACATCTATGGAG</td>
<td></td>
</tr>
<tr>
<td>Sequence</td>
<td>Sequence</td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>---------------------</td>
<td></td>
</tr>
<tr>
<td>MMP1083_Rev_Seq</td>
<td>TGGATTTATGGATAAACACATTG</td>
<td></td>
</tr>
<tr>
<td>MMP1088_For_Seq</td>
<td>TGGACATATTACACCCAGCT</td>
<td></td>
</tr>
<tr>
<td>MMP1088_Rev_Seq</td>
<td>ATGCCCAATTGACCCTATCG</td>
<td></td>
</tr>
<tr>
<td>MMP1089_For_Seq</td>
<td>TCGTAAATATTGGTCGGGA</td>
<td></td>
</tr>
<tr>
<td>MMP1089_Rev_Seq</td>
<td>TCGATAGGGTCAATTGGGCA</td>
<td></td>
</tr>
<tr>
<td>MMP1293_For_Seq</td>
<td>GCTAATGAAAAATATATGTAAAATAGAG</td>
<td></td>
</tr>
<tr>
<td>MMP1293_Rev_Seq</td>
<td>ACATGTACTCAATAGACCCTG</td>
<td></td>
</tr>
<tr>
<td>MMP1294_For_Seq</td>
<td>CTGCCCTCACGAGCGCTTCAG</td>
<td></td>
</tr>
<tr>
<td>MMP1294_Rev_Seq</td>
<td>GAAGATTCATTGTTAATGATTC</td>
<td></td>
</tr>
<tr>
<td>FlaJ_For_Seq</td>
<td>GATGGAAGTTATTACAAGAGG</td>
<td></td>
</tr>
<tr>
<td>FlaJ_Rev_Seq</td>
<td>CGAAGACATCGAAGGAGACC</td>
<td></td>
</tr>
</tbody>
</table>
Table 2.5: Primers used to create the complementation plasmids

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence (5' to 3')</th>
<th>Restriction Site (Underlined)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMP1081_For_Comp</td>
<td>CCAATGCATGGTTCAAAAATATGTAAGATGATTAT</td>
<td>NsiI</td>
</tr>
<tr>
<td>MMP1081_Rev_Comp</td>
<td>GC\text{TCTAGA}TTATGATGATGATGATGATGATGAAACCCTT</td>
<td>XbaI</td>
</tr>
<tr>
<td>MMP1088_For_Comp</td>
<td>CCAATGCATGAATGATTATTAAAAACGCTGTT</td>
<td>Nsi</td>
</tr>
<tr>
<td>MMP1088_Rev_Comp</td>
<td>GC\text{TCTAGA}TTATGATGATGATGATGATGATGGAGATATCA</td>
<td>XbaI</td>
</tr>
</tbody>
</table>

41
McCas media and subcultured overnight without selection. Cells were then grown in the presence of either neomycin (pCRPrtNeo) or puromycin (pHW40) to select for transformants with the plasmid.

2.9 Markerless In-frame Deletion

Markerless in-frame deletion constructs were created for the double recombination deletion methodology as previously developed by Moore and Leigh (2005). Primers amplify two 1 kb DNA fragments, upstream and downstream of the targeted gene. The upstream fragment still contains the start codon of the gene while the downstream fragment includes the stop codon. The two fragments are ligated through the introduced AscI restriction site contained within the primers. The ligated product has the DNA sequence that reads in-frame from the start codon to the stop (with the majority of the gene missing) and when cloned in the vector, pCRTprNeo, integrates into the genome to initiate the in-frame deletion following transformation.

pCRPrtNeo is a non-replicating plasmid that must integrate in the chromosomal DNA of *M. maripaludis* to render the cells neomycin resistant (NeoR). The pCRPrtNeo vector also encodes the *hpt* gene, which will allow for selection of desired genotypes during the deletion process. After transformation of the vector into *M. maripaludis*, neomycin selects for the first recombination, where the vector integrates into the chromosome. The neomycin resistant cells are then allowed to grow without any antibiotic. This provides the opportunity for the vector to recombine out of the chromosome where the recombination event can either return the wildtype genotype or the deletion sequence that was introduced with pCRPrtNeo (Figure 2.2). To select for this recombination, cells were plated on solid McCas media with the toxic base analog, 8-azahypoxanthine, since cells that still contain pCRPrtNeo have the *hpt* gene and are susceptible to 8-azahypoxanthine. Single colonies were picked and screened by PCR for deletion mutants. Sequencing primers were used to amplify over the targeted gene with deletion mutants producing a smaller PCR product than wildtype cells because of the missing gene.
Figure 2.2: Schematic of an in-frame deletion using pCRPrtNeo containing the up and down deletion fragment. After transformation of the plasmid into *M. maripaludis* MM900 cells, neomycin is used to select for transformants (1\(^{st}\) recombination). These cells are then grown without any selection, which allows for the vector to recombine out (2\(^{nd}\) recombination), either returning to the wild-type sequence or recombining to the introduced deletion sequence. Growth on 8-azahypoxanthine plates select for cells that have removed pCRPrtNeo, along with its encoded *hpt* gene. Single colonies undergo PCR screening (primers amplify over the gene region) to determine whether the second recombination event produced a mutant.
2.10 Complementation Vector Design

Deletion strains were complemented with the appropriate wildtype gene, which was cloned into the vector, pHW40, containing the nif promoter. The cloned gene is amplified using a forward primer that must include the restriction enzyme NsiI cleavage site (5' ATGCAT 3'), where the second AT begins at the ATG of the start codon of the gene. The restriction enzyme cleavage site incorporated into the reverse primer may be that of MluI, XbaI or BglII. However in this study only XbaI (5' TCTAGA 3') was used.

Cloned genes in pHW40 were transformed back into their respective deletion strains. Puromycin (2.5 ug/mL) was used to select transformants and these were subsequently cultured in nitrogen free media, supplemented with either alanine or NH4Cl. Under alanine conditions, the nif-promoter is turned on and results in gene expression. Alternatively, using NH4Cl as the sole nitrogen gene does not induce gene expression (Lie and Leigh, 2002).

2.11 Southern Blotting

Southern blotting protocols were performed using previously established methods (Thomas et al., 2001). Briefly, genomic DNA was digested overnight with the desired restriction enzyme and electrophoresed through a 0.8% agarose gel for 1.5 hours at 100 V. After denaturing and DNA transfer to positively charged nylon membrane by capillary transfer, the blot was blocked with prehybridization solution for 1 hour before hybridization overnight with a DNA probe (55°C). DNA probes were 400-500 bp PCR products that were amplified across the mutant gene region using sequencing primers and subsequently DIG labeled using a DIG labeled kit according to the manufacturer’s protocol (Roche). Bound probe were detected with Anti-digoxigenin-AP (Roche) and developed using CSPD Star (Roche).

2.12 RT-PCR

RNA was extracted from M. maripaludis MM900 cells using an RNeasy Mini Kit (Qiagen). RT-primers (Table 2.6) designed to amplify across the intergenic region were used with the One-Step RT-PCR kit (Qiagen). RT-PCR reaction mixes were prepared according to
<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
<th>Gene(s) of Interest</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMP1081 For RT</td>
<td>ACTAGAGCGGAAGCATTAGGACAG</td>
<td>MMP1081-MMP1082</td>
</tr>
<tr>
<td>MMP1082 Rev RT</td>
<td>CCTGAATCGAAGATCCCACTCCCGG</td>
<td>MMP1081-MMP1082</td>
</tr>
<tr>
<td>MMP1082 For RT</td>
<td>TATACTAATGACTACGGAATACG</td>
<td>MMP1082-MMP1083</td>
</tr>
<tr>
<td>MMP1083 Rev RT</td>
<td>TAATGCCGCCAGCATAACCCAGAGGCATG</td>
<td>MMP1082-MMP1083</td>
</tr>
<tr>
<td>MMP1083 For Int</td>
<td>TTCAATGATAAAAGAAGCTGATGAAC</td>
<td>MMP1083</td>
</tr>
<tr>
<td>MMP1083 Rev Int</td>
<td>CTACCTGCACCTGCAGACTTGCTCCAC</td>
<td>MMP1083</td>
</tr>
</tbody>
</table>
manufacturer, and were placed in a Mastercycler epgradient S (Eppendorf) to undergo the following program: initial incubations at 50°C for 30 minutes and 95°C for 15 minutes, followed by 30 cycles of 94°C for 30 seconds, 55°C for 30 seconds and 72°C for 50 seconds. An additional extension time of 2 minutes at 72°C completed the RT-PCR reaction. Controls PCR reactions on the RNA template lacking the reverse transcriptase step served as controls for potential DNA contamination.

2.13 SDS-PAGE and Immunoblotting

Whole cell or crude-membrane protein samples were boiled with 2xESB (125 mM Tris-HCl, pH 6.8, 2% SDS, 20% glycerol, 4% 2-mercaptoethanol, 0.002% bromophenol blue) for 10 minutes and loaded to a 15% acrylamide gel (Laemmli, 1970). A semi dry blotter (Trans Blot-SD, Biorad) was used to transfer the proteins to an Immobilon-P membrane (Millipore) as previously described (Towbin et al., 1979). Polyclonal chicken antibody to *M. voltae* FlaB2, shown to cross react with *M. maripaludis* flagellins (Van Dyke et al., 2008), and horseradish peroxidase-conjugated rabbit anti-chicken immunoglobulin Y (Jackson Immunoresearch Laboratories) were used as primary and secondary antibodies, respectively. Chemiluminescence kits (Roche) were used to develop the blots.

2.14 Flagella Isolation

*M. maripaludis* flagella were extracted according to protocol previously established by Bardy et al., (2002). 8 L of *M. maripaludis* cells were harvested via centrifugation (6000 x g for 15 min) and resuspended in 10mM Tris-HCl; pH 7 + 2% NaCl. OP-10 detergent (1% final concentration; Nikko Chemicals Co. Ltd., Tokyo, Japan) was used to lyse cells. DNase and RNase were added with an incubation time of 30 minutes at room temperature, with occasional inverting to reduce viscosity. After centrifugation (6000 x g for 15 min), the supernatant was incubated on ice (1 hr) with the addition of 1M NaCl + 20% polyethylene glycol (PEG8000) solution (1/10 volume). After another centrifugation (7000 x g for 20 minutes), pellets were resuspended in distilled water, loaded to a KBr gradient (0.5 g/mL) and centrifuged (35K rpm for 46
18 hours using an SW-41 rotor). A white band, corresponding to isolated flagella was removed, diluted with dH₂O and re-centrifuged (30K for 90 min using an SW-41 rotor). The final flagella pellet is resuspended in dH₂O. Samples were sent for mass spectrometry (MS) analysis (Dr. Susan Logan, NRC, Ottawa).

2.15 Mass Spectrometry

Purified flagellins (Bardy et al., 2002) were digested overnight with trypsin (Promega Madison, WI) at a ratio of 30:1 (protein: enzyme, v/v) in 50 mM ammonium bicarbonate at 37°C. The protein digest was analysed by nano-LC-MS/MS using a Q Tof Ultima hybrid quadropole time of flight mass spectrometer as previously described (Twine et al., 2008; Kelly et al., 2009). Glycopeptide spectra were examined manually to determine the glycan structure present.

2.16 Electron Microscopy

*M. maripaludis* cells were centrifuged and washed with 1X PBS and resuspended in 50 mM MgSO₄. Cells were transferred on Formvar-coated gold grids and stained with 2% phosphotungstic acid. Grids were viewed using a Hitachi H-700 electron microscope (at 75 kV).
Chapter 3

Results

3.1 Identification of potential glycosyl transferases:

Previous work had reported the role of two glycosyl transferase genes, MMP1079 and MMP1080, which attach the second and third sugars of the tetrasaccharide glycan in M. maripaludis (VanDyke et al., 2009). The first sugar in the M. maripaludis glycan is GalNAc, which differs from the GlcNAc seen in M. voltae (Kelly et al., 2009). While the glycosyl transferase for the first sugar in M. voltae is AglH, it may not carry out this function in M. maripaludis. However, since it is not possible to knock out this gene in either M. voltae or M. maripaludis, due to its proposed role in biosynthesis of the essential methanogenic cofactor B, the role of the AglH homologue in M. maripaludis (MMP1423) cannot be genetically tested directly.

The activity of AglH in M. voltae was indirectly proven when it was demonstrated to complement the eukaryote, S. cerevisiae, carrying a conditional lethal mutation in alg7, an N-acetylglucosamine-1-phosphate transferase responsible for attaching GlcNAc in the N-linked system (Shams-Eldin et al., 2008).

Potentially, the first glycosyl transferase may not be MMP1423. Therefore, in-frame deletions were attempted on the remaining annotated glycosyl transferases that were not yet investigated to identify the potential first glycosyl transferases (Table 3.1). Deletion of MMP1170 was previously unsuccessfully attempted by VanDyke et al. (2009).

Finally, the gene encoding the fourth glycosyl transferase has yet to be identified. An annotated glycosyl transferase gene, MMP1088, was a good candidate for an in-frame deletion since it is located in the same gene region as MMP1079 and MMP1080.

Creation of the in-frame deletion construct for the fourth glycosyl transferase:

Up and down deletion arms for MMP1088 were created via two PCR reactions using primer pairs: MMP1088_For_UP/MMP1088_Rev_Up and
Table 3.1: Annotated glycosyl transferases in the *M. maripaludis* genome.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Common Name</th>
<th>5' End</th>
<th>3' End</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMP0090</td>
<td>Glycosyl transferase</td>
<td>101817</td>
<td>100750</td>
</tr>
<tr>
<td>MMP0356</td>
<td>Glycosyl transferase, group 1</td>
<td>353708</td>
<td>352524</td>
</tr>
<tr>
<td>MMP0359</td>
<td>Glycosyl transferase, family 2</td>
<td>356598</td>
<td>355921</td>
</tr>
<tr>
<td>MMP0590</td>
<td>Glycosyl transferase, family 2</td>
<td>588034</td>
<td>587336</td>
</tr>
<tr>
<td>MMP1079</td>
<td>Glycosyl transferase, family 2</td>
<td>1069551</td>
<td>1070360</td>
</tr>
<tr>
<td>MMP1080</td>
<td>Glycosyl transferase, group 1</td>
<td>1070534</td>
<td>1071628</td>
</tr>
<tr>
<td>MMP1088</td>
<td>Glycosyl transferase, group 1</td>
<td>1079591</td>
<td>1080823</td>
</tr>
<tr>
<td>MMP1170</td>
<td>Glycosyl transferase, family 2</td>
<td>1153909</td>
<td>1154601</td>
</tr>
<tr>
<td>MMP1293</td>
<td>Glycosyl transferase, group 1</td>
<td>1278161</td>
<td>1276986</td>
</tr>
<tr>
<td>MMP1294</td>
<td>Glycosyl transferase, group 1</td>
<td>1278266</td>
<td>1279834</td>
</tr>
<tr>
<td>MMP1423</td>
<td>Putative glycosyl transferase, family 4</td>
<td>1395509</td>
<td>1394610</td>
</tr>
</tbody>
</table>

Note: Previous studies had investigated the roles of *MMP1079*, *MMP1080*, *MMP1170* and *MMP1423* (VanDyke et al., 2009).
MMP1088_For_Down/MMP1088_Rev_Down (Figure 3.1A). Up and down fragments were digested with AscI and ligated to create the deletion construct. Amplification of the desired construct was achieved through a PCR reaction consisting of the primers MMP1088_For_Up and MMP1088_Rev_Down (Figure 3.1B). The deletion PCR construct was digested with BamHI and cloned into pCRPrtNeo, creating the MMP1088 deletion vector, pKJ665 (Figure 3.1C).

**PCR screening and Southern blot confirmation of MMP1088 deletion:**

pKJ665 was transformed into the *M. maripaludis* strain, MM900, according to the established protocol (Moore and Leigh, 2005) and screened for potential mutants. The sequencing primers MMP1088_For_Seq and MMP1088_Rev_Seq were designed to amplify over the MMP1088 gene region to give a wildtype product of 1.4 kb versus a mutant product of 150 bp (Figure 3.2A). Single colonies were picked off agar plates containing 8-azahypoxanthine, grown in Balch III media and screened using whole cell PCR. Putative deletion mutants were also screened by Southern blotting to confirm the gene deletion. Genomic DNA from wildtype and ΔMMP1088 were extracted and digested with the restriction enzymes, EcoRI and PvuII. The PCR product, obtained with genomic ΔMMP1088 as DNA template and the sequencing primers, MMP1088_For_Seq and MMP1088_Rev_Seq, was DIG labeled and used as a probe to detect the digested DNA of interest. Under these conditions, the DIG-labeled probe binds a fragment of 2088 bp in the wildtype and 856 bp in the mutant (Figure 3.2B).

**Analysis of ΔMMP1088 using immunoblotting:**

To determine the effects of the MMP1088 deletion, western immunoblotting with anti-flagellin antibody was performed on ΔMMP1088 whole cell lysates alongside lysates of wildtype and previously identified glycosyl transferase deletion strains and an oligosaccharyl transferase deletion strain (VanDyke et al., 2009). *MMP1079* and *MMP1080* are responsible for attaching the second and third sugars, respectively, while *MMP1424* transfers the N-linked glycan to the side chain of asparagine residues within the N-linked sequon. Therefore, wildtype, ΔMMP1080,
Figure 3.1: Creation of the in-frame deletion plasmid for MMP1088. A) PCR amplification of the ~1kb upstream and downstream arms of the deletion construct. B) PCR amplification of 2 kb DNA fragment after ligation of upstream and downstream pieces. C) Restriction enzyme digest (BamH1) of pCRPrtNeo containing the 2 kb deletion construct. M represents lanes with the 1 kb ladder (NEB).
Figure 3.2: Confirmation of MMP1088 deletion. A) Whole cell PCR screens of wildtype (wt) and MMP1088 mutant (Δ). PCR of wildtype cells amplifies a 1.4 kb PCR product. PCR of mutant cells amplifies a 150 bp product. B) Southern blot analysis of wildtype (wt) and ΔMMP1088 (Δ) genomic DNA. DNA was digested with EcoR1 and PvuII and hybridized with a DIG labelled probe that binds in the targeted gene region. For wildtype DNA, the probe binds to the predicted 2088 bp sized DNA fragment whereas in the mutant, the predicted 856 bp fragment is bound. M represents lanes with the 1 kb ladder (NEB).
ΔMMP1079 and ΔMMP1424 flagellins have glycans of 4 sugars, 2 sugars, 1 sugar and no sugars. Western blot analysis revealed that ΔMMP1088 flagellins migrated at a lower molecular weight than wildtype and higher molecular weight than flagellins from any of the other deletion strains. In comparison to the known glycosyl transferases, there is a stepwise decrease in molecular weight from wildtype, starting with ΔMMP1088 flagellins and progressing to the completely non-glycosylated ΔMMP1423 flagellins (Figure 3.3).

**Analysis of ΔMMP1088 using Electron Microscopy:**

ΔMMP1088 and wildtype cells were viewed under the electron microscope to determine if flagella assembly had been affected. The ΔMMP1088 cells were fully flagellated and showed no striking differences from wildtype cells (Figure 3.4).

**Analysis of ΔMMP1088 using a Motility Assay:**

Motility of the ΔMMP1088 mutant was assessed with respect to wildtype and other glycosyl transferase deletion strains using 0.25% agar semi-swarm plates. Cells were inoculated at the center of 0.25% agar semi swarm plates and grown for 9 days at 30°C. The diameter of the swarm area correlates to effectiveness in motility. It was found that ΔMMP1088 was consistently less motile in this assay than wildtype cells, but more so than ΔMMP1080 cells when the diameters of the swarm were measured. ΔMMP1079 and ΔMMP1424 cells showed no motility because they lack flagella since a minimum 2 sugar glycan is needed for proper assembly of flagella (VanDyke et al., 2009) (Figure 3.5).

**Complementation of MMP1088:**

The *MMP1088* gene was complemented back into ΔMMP1088 to confirm the effects of the deletion. Genes cloned into the complementation vector, pHW40, are controlled through a nif promoter. This promoter is repressed in the presence of NH₄Cl but moderately induced in N-free media where alanine is supplemented as the sole nitrogen source (Lie and Leigh, 2002). To begin cloning into pHW40, *MMP1088* (1232 bp) was amplified using the primers
Figure 3.3: Immunoblot of whole cell lysates from the *MMP1088* mutant and previously created in-frame deletion strains that also affect glycosylation. The blot was developed with anti-flagellin antibody and bars indicates molecular weight marker. Figure modified from VanDyke et al., 2009.
Figure 3.4: Electron micrograph of *M. maripaludis* wildtype (left) and ∆MMP1088 (right) cells, showing that both are capable of assembling flagella. Cells were negatively stained with 2 % phosphotungstic acid. The arrows point to flagella and the bar represents a 0.5 µm.
Figure 3.5: Semi-swarm agar (0.25%) plates were used to determine the motility of Δ1088 in comparison to wildtype (wt) and previously created deletion strains that affect N-linked glycosylation. Plates were incubated at 30°C for 9 days (Figure modified from VanDyke et al., 2009).
MM1088_Rev_Comp and MM1088_Rev_Comp in a PCR reaction. After purification, MMP1088 was digested with Nsi and Xba and cloned into pHW40. The MMP1088 complementation vector was designated pKJ679. pKJ679 was transformed into ΔMMP1088 cells and desired transformants were selected in Balch III media with puromycin. These cells were then transferred to nitrogen free media, supplemented with either alanine (induced) or NH₄Cl (non-induced) for several transfers. Under inducing conditions, Western blotting revealed that the flagellin molecular weight returned to wildtype whereas in non-inducing conditions, flagellins remained the same size as in ΔMMP1088 cells (Figure 3.6).

**MS Analysis of ΔMMP1088 glycan:**

Flagella were purified from ΔMMP1088 cells and sent for MS analysis of the attached N-linked glycan in the laboratory of Dr. Susan Logan (NRC Ottawa, ON). It was determined that the glycan was missing the last sugar and threonine usually found on the third sugar (Figure 3.7). Analysis of the N-linked glycan from complemented ΔMMP1088 flagella saw a return of the glycan structure back to wildtype.

This set of data identified MMP1088 as the glycosyl transferase responsible for the transfer of the fourth sugar to the *M. maripaludis* N-linked glycan. This left only the first glycosyl transferase unaccounted for and the remaining genes annotated as glycosyl transferases were targeted for in-frame deletion in an attempt to identify this enzyme.
Figure 3.6: Immunoblot of whole cell lysates from *M. maripaludis* wildtype, Δ1088, Δ1088 complemented with pKJ679 grown in nitrogen free media supplemented with NH₄Cl (non-inducing; off) and Δ1088 complemented with pKJ679 grown in nitrogen free media supplemented with alanine (inducing; on) as the sole nitrogen source. Blot was developed with anti-flagellin antibody and bars indicates molecular weight marker (M).
Figure 3.7: MS/MS spectra of the triply protonated ion corresponding to the T^{41-469} glycopeptide of the mature FlaB2 protein at (A) m/z 1321.3 in the WT and (B) m/z 1215.3 in the ΔMMP1088 mutant. The amino acid sequence of this peptide is presented in the inset in (A) and the site of N-glycosylation is indicated. Some of the b and y fragment ions arising from fragmentation of the peptide bonds are indicated in the spectrum. The major carbohydrate oxonium ions are identified in the spectra using symbols to indicate the sugar residues present (GalNAc: ■ 203 Da; (5S)-2-acetamido-2,4-dideoxy-α-L-erythro-hexos-5-ulo-1,5-pyranose □ 217 Da; ManNAc3NamA6: ▲ 257 Da; GlcNAc3NAcA: ● 258 Da and ManNAc3NAc3NamA6Thr: ▲ 358 Da). The MS/MS spectrum for the wild-type glycopeptide (panel a) is dominated by the glycan oxonium ions at m/z 617.2 (● + ▲) and m/z 359.1 (□), although a weak ion signal for the intact WT glycan is observed at m/z 1037.4. MS/MS analysis of the equivalent glycopeptide from the 1088 mutant indicated the presence of a trimeric glycan composed of the GlcNAc (■), GlcNAc3NAcA (●) and ManNAc3NAc3NamA6 (▲) without the threonine modification. Figure modified from VanDyke et al, 2009.
Creation of the in-frame deletion constructs to identify the first glycosyl transferase:

For simplicity, complete data from only MMP359 will be shown in this section although MMP090, MMP356, MMP590, MMP1293 and MMP1294 were all tested as candidates for the first glycosyl transferase. Primers used for this study are listed in Table 2.1. The up and down arms of the in-frame deletion construct were obtained through two separate PCR reactions. The primers used in the up reaction were MMP359_For_Up and MMP359_Rev_Up while the down reaction had amplifying primers of MMP359_For_Down and MMP359_Rev_Down (Figure 3.8A). After purification of the up and down arms, PCR products were digested with the restriction enzyme, AscI, overnight. The two arms were then ligated to create the 2kb deletion construct and amplified using the primers MMP359_For_Up and MMP359_Rev_Down (Figure 3.8B). The 2kb deletion product was digested with the restriction enzyme, BamHI, and cloned into the M. maripaludis vector, pCRPrtNeo (Figure 3.8C). The deletion pCRPrtNeo vector was subsequently named pKJ709.

Screening for deletions:

pKJ709 was transformed into MM900 cells and cells incorporating the vector were selected with neomycin. These cells were then grown without any antibiotic, enabling the vector to recombine out. Single colonies were picked from agar plates containing 8-azahypoxanthine, grown in Balch III media and screened using whole cell PCR. The sequencing primers, MMP359_For_Seq and MMP359_Rev_Seq, were used to determine if cultures were either wildtype (1264 bp PCR product) or mutant (730 bp PCR product), while the same was determined for MMP590, MMP1293 and MMP1294 (Figure 3.9). With the exception of MMP090 and MMP356, all deletions could be created. Whole cell PCR screens were performed at two steps within the transformation protocol to demonstrate that a) the pCRPrtNeo vector had integrated in the genome (cells taken from neomycin cultures) and b) recombination after 8-azahypoxanthine selection occurred so as to generate gene deletion mutants, before plating and single colony isolation. In the case of MMP090, a PCR band corresponding to the deletion
construct could be detected in cells grown with neomycin, but not after 8-azahypoxanthine selection, indicating that recombination into the mutant genotype is extremely rare or not possible (Figure 3.9E) and thus providing a possible explanation for why screens of individual 8-azahypoxanthine colonies were always wildtype.

**Flagellin Immunoblotting:**

Western blotting revealed that there was no change in the apparent molecular weight of the flagellins of ΔMMP359 in comparison to wildtype, indicating that the ΔMMP359 flagellins appear fully glycosylated. Similarly, no changes in molecular weight of the flagellins in strains deleted for MMP590, MMP1293 and MMP1294 were observed (Figure 3.10). These results would appear to eliminate the MMP359, MMP590, MMP1293 and MMP1294 genes as potential candidates for encoding the first glycosyl transferase.
Figure 3.8: Creation of the in-frame deletion plasmid for *MMP359*. A) PCR amplification of the ~1 kb upstream and downstream arms of the deletion construct. B) PCR amplification of ~2 kb DNA fragment after ligation of upstream and downstream pieces. An unspecific PCR product at 1 kb was also observed leading to the use of TOPO cloning kit to isolate the 2 kb piece for eventual cloning into pCRPrtNeo. C) Restriction enzyme digest (BamH1) of pCR2.1 TOPO (T) with 2 kb insert and pCRPrtNeo (N) containing the 2 kb deletion construct.
Figure 3.9: Confirmation of MMP359, MMP590, MMP1293 and MMP1294 deletion. A) Whole cell PCR screens of wildtype (wt) and ΔMMP359 (Δ). PCR of wild-type cells amplifies a 1264 bp PCR product. PCR of mutant cells amplifies a 730 bp product. B) Whole cell PCR screens of wildtype (wt) and ΔMMP590 (Δ). A wild-type cell amplifies a 1066 bp PCR product while PCR of mutant cells amplifies a 574 bp product. C) Screens of wildtype (wt) and ΔMMP1293 (Δ). PCR of wild-type cells amplifies a 1415 kb PCR product whereas PCR of mutant cells amplifies a 467 bp product. D) PCR screen of wildtype (wt) and ΔMMP1294 (Δ). PCR of wild-type cells amplifies a 1905 kb PCR product while PCR of mutant cells amplifies a 438 bp product. E) PCR of M. maripaludis cells taken from neomycin selection and after 8-azahypoxanthine selection during MMP090 transformation (top arrow for MMP090 deletion band; bottom for MMP1293 and MMP1294). In all examples, M is the 1 kb ladder (NEB) used as a marker.
Figure 3.10: Immunoblots of *M. maripaludis* whole cell lysates comparing wildtype (wt) to A) ∆MMP359 B) ∆MMP590 C) ∆MMP1293 and D) ∆MMP1294. No apparent changes in molecular weight were observed. Blots developed with anti-flagellin antibodies and bars indicates molecular weight marker (M).
3.2 Identification of potential flippase:

The flippase protein has yet to be identified in any archaeal N-linked system. BLAST searches for homologues of the possible eukaryotic flippase, Rft1, in the *M. maripaludis* genome did not reveal any protein that was considered related based on the Expected (E) values given by the BLAST engine. On the other hand, similar searches for homologues of the bacterial ATP-dependent flippase gene in the *M. maripaludis* genome produced too many strong hits (some of which are clearly not flippase candidates) because of the conserved nature of ABC-type transporters. This makes the search for the actual flippase difficult. A deletion of the flippase gene should be possible given that a mutant for the oligosaccharyl transferase, which processes the glycan after the flippase in the N-linked pathway, has been constructed. A lethal flippase deletion would indicate that the protein is involved in another cellular process that is essential. In this work, the *MMP1089* gene was targeted for in-frame deletion as an N-linked flippase candidate, because it is located immediately downstream of the fourth glycosyl transferase, *MMP1088*, and is annotated as a polysaccharide biosynthesis gene that has BLAST hits to ABC-transporters.

Creation of the in-frame deletion construct for *MMP1089*:

A deletion construct for MMP1089 was created and cloned in pCRPrtNeo as for the other genes. The pCRPrtNeo deletion vector was named pKJ685 (Figure 3.11A).

Screening for *MMP1089* mutants:

After transformation of pKJ685 into MM900, single colonies on hypoxanthine plates were picked and screened by PCR using primers, 1089_For_Seq and 1089_Rev_Seq. Only wildtype transformants were obtained in screens of over 80 single colonies (Figure 3.12A). Whole cell PCR of *M. maripaludis* grown in neomycin indicates that pKJ685 had integrated in the chromosome (Figure 3.12B).

Since the conventional in-frame deletion attempt to delete *MMP1089* was unsuccessful, a second approach was taken, where the integration vector, pPAC60, was used. This method was
considered a plausible alternative because \textit{MMP1089} is the last gene in its operon therefore, a disruption will not have any downstream effects.

\textbf{Creation of the disruption vector, pPAC60, for \textit{MMP1089} inactivation:}

N- and C- terminally truncated \textit{MMP1089} (900 bp) was obtained though PCR using the primers, \textit{MMP1089\_For\_Dis} and \textit{MMP1089\_Rev\_Dis}. The truncated \textit{MMP1089} product was digested with EcoRI and cloned into the pPAC60 vector (Figure 3.11B). The vector was named pKJ693.

\textbf{Screening for disrupted \textit{MMP1089} mutants:}

Transformed MM900 cells with incorporated pKJ693 were selected with puromycin. PCR screening (\textit{MMP1089\_For\_Seq} and \textit{MMP1089\_Rev\_Seq}) of single colonies revealed cells that only had a wildtype sized \textit{MMP1089} (Figure 3.12C). No PCR product is expected for cells carrying the disrupted gene since the primers would be amplifying across a very large region that includes the pPAC60 vector and its insert. Over 50 colonies were screened indicating that the vector must have inserted elsewhere within the genome and the gene is likely essential. No attempt was made to identify the location of the inserted plasmid.

However, for a final attempt, the conventional in-frame deletion protocol was performed again. Unlike the previous work where \textit{MMP1089} was the only gene targeted for deletion, this attempt included the deletion of the \textit{MMP1088} gene. In several other previous gene deletions, the final recombination event was found to be heavily favoured towards wildtype, where frequencies such as one mutant in every fifty colonies are found. This may be a result of primer design and recombination hot spots throughout the \textit{M. maripaludis} genome. Therefore, the previously deleted region of \textit{MMP1088} is included in a final attempt to delete \textit{MMP1089} since the role of \textit{MMP1088} was now known and would not complicate interpretation of the desired gene deletion.

\textbf{Creation of the in-frame deletion construct for \textit{MMP1088}/\textit{MMP1089}:}

The deletion construct for \textit{MMP1088}/\textit{MMP1089} was created and cloned into pCRPrtNeo. The vector is called pKJ713 (Figure 3.11C).
Screening for MMP1088/MMP1089 mutants:

pKJ713 was transformed into MM900 cells and single colonies were screened by PCR after antibiotic selection was used according to the standard protocol. Initial screening with primers, MMP1088_For_Seq and MMP1089_For_Seq, revealed a potential mutant (Figure 3.12D).

Analysis of ΔMMP1088/MMP1089 using immunoblotting:

Western blotting analysis of the putative ΔMMP1088/MMP1089 strain showed that the flagellin molecular weight did not change from wild type (Figure 3.13). It is expected that even if MMP1089 was not the flippase, flagellins should migrate at a lower molecular weight than wildtype because of the deletion of MMP1088. This unexpected result led to further validation of the ΔMMP1088/MMP1089 mutant.

Are MMP1088 and MMP1089 truly deleted?:

To verify the deletions of MMP1088 and MMP1089, two PCR reactions were performed where each gene was amplified individually. Primers, MMP1088_For_Seq and MMP1088_Rev_Seq, were used for MMP1088 while 1089_For_Seq and 1089_Rev_Seq were used for MMP1089. The results indicate that both MMP1088 and MMP1089 are fully intact although the deletion screen using MMP1088_For_Seq and 1089_For_Seq still provided a mutant product (Figure 3.14).
Figure 3.11: Creation of deletion and disruption vectors of *MMP1089*. A) Restriction enzyme digestion (Xba1) of pCRPrtNeo containing the in-frame deletion construct for *MMP1089*. B) Restriction enzyme digestion (EcoR1) of pPAC60 containing the truncated version of *MMP1089* (900 bp). C) Restriction enzyme digestion (Xba1) of pCRPrtNeo containing the in-frame deletion construct for *MMP1088* and *MMP1089*. 
Figure 3.12: Whole cell PCR screening of MMP1089 deletion attempts. A) Screens of in-frame deletion of MMP1089 only produced wildtype (wt) PCR product (1.8 kb). Only 11 of the 80+ screens are shown. B) Whole cell PCR of *M. maripaludis* cells grown with neomycin selection produce wildtype and deletion PCR product (size), as shown by arrows. C) Screens for MMP1089 mutants disrupted by pPAC60 carrying the truncated MMP1089 insert revealed only wildtype products (only 11 of 50+ screens shown). D) Screens of in-frame deletion of MMP1088 and MMP1089 produced wildtype and a putative mutant (Δ).
Figure 3.13: Immunoblot of whole cell lysates from *M. maripaludis* wildtype (wt) and the putative ΔMMP1088/MMP1089 (Δ) cells. Blots were developed with anti-flagellin antibody and bars indicates molecular weight marker (M).
Figure 3.14: Whole cell PCR of the putative mutant with the usual sequencing primers (Δ), primers that amplify across *MMP1088* (1.4 kb) and primers that amplify across *MMP1089* (1.7 kb) individually.
3.3 Identification of genes involved in N-linked monosaccharide biosynthesis:

As previously mentioned, the structure of the N-linked tetrasaccharide in *M. maripaludis* has recently been published (Kelly et al., 2009). The linking sugar to the protein is N-acetylgalactosamine to which is attached a di-acetylated glucuronic acid followed by an acetylated and acetamidino modified mannuronic acid, which also includes threonine residue attachment. The final sugar is a very unusual diglycoside of aldulose, the first reported natural occurrence of this sugar. In previous work by VanDyke and work in this study, three glycosyl transferases, MMP1079, MMP1080 and MMP1088, were identified to be involved in the assembly of the *M. maripaludis* N-linked glycan (VanDyke et al., 2009). Lying between MMP1080 and MMP1088 were additional genes annotated as possibly involved in sugar biosynthesis and these were targeted for deletion in this section of the thesis. BLAST searches of MMP1081 reveal hits to enzymes responsible for attachment of acetamidino groups. Since the third sugar in the *M. maripaludis* glycan is a mannuronic acid with such a group, the defect in a potential MMP1081 mutant may be in the attachment of this acetamidino group. The final two genes in the operon, MMP1082 and MMP1083, have unusual annotations and their function cannot be easily predicted. Nevertheless, these two genes were also targeted for in-frame deletions because of their likelihood of co-transcription with MMP1081. RT-PCR was used to determine if in fact, these three genes are co-transcribed.

**Determination of co-transcription of MMP1081, MMP1082 and MMP1083:**

Co–transcription was likely given that the start codon of MMP1082 was immediately adjacent to the stop codon of MMP1081 while the length of the intergenic region between MMP1082 and MMP1083 was only 48 bp. RT-PCR analysis confirmed that MMP1081, MMP1082 and MMP1083 were in fact co-transcribed (Figure 3.15). Primers used in this study are shown in Table 2.6.
Figure 3.15: Determination of co-transcription for the genes MMP1081, MMP1082 and MMP1083 using RT-PCR. Primers designed to amplify across the intergenic regions were used for MMP1081/MMP1082 and MMP1082/MMP1083. Internal primers of MMP1083 were used to ensure that the entire RNA transcript is present. Classical PCR was performed with isolated RNA, but no reverse transcriptase, as template to show sample was not contaminated with DNA (DNA Ctrl). To validate the RT-PCR products sizes, classical PCR was also performed with chromosomal MM900 DNA.
Summary for biosynthesis gene deletions:

In-frame deletion PCR products of MMP1081, MMP1082 and MMP1083 were cloned in pCRPrtNeo (Figure 3.16) and transformed into MM900 cells. PCR screening and Southern blot analysis revealed mutants for all three genes were obtained (Figure 3.17). Isolated genomic DNA was digested with Nde1, Ssp1 or EcoRI. Western blot analysis revealed that ΔMMP1081 flagellins migrate to a lower molecular weight than wildtype and possibly ΔMMP1088 flagellins. On the other hand, ΔMMP1082 and ΔMMP1083 flagellins did not have a noticeable difference in molecular weight compared to those of wildtype cells (Figure 3.18). To identify the structure of the altered N-linked glycan, purified flagella from ΔMMP1081 were sent to S. Logan (Ottawa) for MS analysis. S. Logan et al (unpublished) determined that the acetamidino group on the third sugar was missing from the glycan, along with the fourth sugar and threonine usually found on the third sugar (Figure 3.19).

Electron microscopy revealed that ΔMMP1081 cells were still fully flagellated (Figure 3.20) while semi-swarm plates were used to determine if motility was affected as a result of the MMP1081 deletion. It was found that ΔMMP1081 was consistently less motile than both wildtype and ΔMMP1088 cells, but more so than ΔMMP1080 when the diameter of the swarm was measured and compared to wildtype (Figure 3.21).

Complementation of MMP1081 into ΔMMP1081 cells partially returned flagellins back to wildtype size under inducing conditions with alanine as the nitrogen source (Figure 3.22).
Figure 3.16: Creation of the in-frame deletion plasmid for MMP1081, MMP1082 and MMP1083. Restriction enzyme digestions (BamH1) of pCRPrtNeo containing the in-frame deletion construct for the genes A) MMP1081, B) MMP1082 and C) MMP1083.
Figure 3.17: Confirmation of MMP1081 deletion. A) Whole cell PCR screens of wildtype (wt) and MMP1081 (Δ). PCR of wild-type cells amplify a 1340 bp product. PCR of mutant cells amplifies a 489 bp product. B) PCR screen for potential MMP1082 mutants produces a 1002 bp wildtype product and a 499 bp deletion. C) PCR screen for potential MMP1083 deletions produces a 1048 bp wildtype product vs a 432 bp deletion. D) Southern blot confirmation of the putative MMP1081 deletion. Genomic DNA from wildtype and ΔMMP1081 were digested with NdeI and DIG-labelled probe binds to a 2494 bp wildtype fragment vs a 1651 bp DNA fragment in ΔMMP1081 cells. E) Southern blot of the putative MMP1082 deletion. Genomic DNA from wildtype and ΔMMP1082 were digested with SspI and DIG-labelled probe binds to a 1534 bp wildtype fragment vs a 1034 bp DNA fragment in ΔMMP1082 cells. F) Southern blot analysis of putative MMP1083 deletion. Genomic DNA from wildtype and ΔMMP1083 were digested with EcoRI and DIG-labelled probe binds to a 1962 bp wildtype fragment vs a 1339 bp DNA fragment in ΔMMP1083 cells.
Figure 3.18: Immunoblot of *M. maripaludis* mutant whole cell lysates for wildtype (wt), Δ1081, Δ1082 and Δ1083. Only Δ1081 cells had a reduction in molecular weight. Blots were developed with anti-flagellin antibody and bars indicates molecular weight marker.
Figure 3.19: NanoLC-MS/MS of flagellin glycopeptide aa 51-71 (DSTEQVASGLQIMGISGYQAGTANANITK) from Flagellin B2. a) M. maripaludis Mm900 and b) Δ1081. This peptide has one site of N-linked glycosylation (NIT). The mass spectra are from the triply protonated glycopeptide at m/z 1321.93, 1215.95 and 1316.66, respectively.
Figure 3.20: Electron micrograph of ∆MMP1081 cell showing that it can assemble flagella (arrows).
Figure 3.21: Semi-swarm agar (0.25%) plates were used to determine the motility of Δ1081 in comparison to wildtype (wt) and previously created deletion strains that affect N-linked glycosylation. Plates incubated at 30°C for 9 days.
Figure 3.22: Immunoblot of whole cell lysates from *M. maripaludis* wildtype, Δ1081, Δ1081 complemented with pKJ752 grown in nitrogen free media supplemented with alanine (inducing; on) as the sole nitrogen source and Δ1081 complemented with pKJ752 grown in nitrogen free media supplemented with NH₄Cl as the sole nitrogen source (non-inducing; off). Blots were developed with anti-flagellin antibody and bars indicate molecular weight marker.
Chapter 4

Discussion

The study of N-linked glycosylation in Archaea has not been well studied until the last few years, although significant progress has been made in a few model organisms. In this study a number of genes were deleted in an attempt to identify those involved in the biosynthesis and assembly of the four sugar glycan, important for flagella assembly and function in the archaeon, *Methanococcus maripaludis*. Two genes involved in the biosynthesis of the *M. maripaludis* N-linked glycan were identified. One of the genes plays a role in glycan assembly while the other is involved in the biosynthesis of one of the sugars that compose the glycan. In-frame deletions in both genes resulted in an altered glycan, as determined by immunoblotting and MS analysis, and consequently led to impairment of swimming motility, as shown by semi-swarm plate analysis.

Although there are many exceptions, several Gram negative bacteria have been shown to glycosylate their flagellin proteins (*e.g.* *Neisseria spp.*, *Aeromonas spp.*, *Helicobacter spp.*, *Campylobacter spp.* and *Pseudomonas spp.*), while only three Gram positive genera have been reported to do so (Logan, 2006). Interestingly, there is some heterogeneity in the composition of the glycan within an individual organism for some species such as *Pseudomonas aeruginosa*, while in others a single glycan is found on their flagellins, such as in *Helicobacter pylori* (Logan, 2006). The requirement of the glycan for assembly of flagella and motility also appears variable from species to species. Mutant strains in *Campylobacter jejuni* and *Campylobacter coli* that had disrupted glycosylation pathways were non-flagellated and accumulated intracellular flagellins proteins with reduced molecular weight (Goon et al., 2003). Meanwhile within the *Pseudomonas spp.*, glycosylation does not appear to be required for flagella assembly (Takeuchi et al., 2003; Schirm et al., 2004). With recent advancements in the sensitivity of mass spectrometers and NMR technology, detailed structures, including composition and linkage, from very small amounts of glycan can be obtained (Voisin et al., 2005).
In the case of bacterial flagellins and their glycans, thus far only $O$-linked sugars have been reported (Logan, 2006) but in Archaea only $N$-linked glycans have been found attached to flagellins. Previous studies demonstrated the requirement of glycosylation to archaeal flagellin proteins in *M. voltae* and *M. maripaludis* (Chaban et al., 2006; VanDyke et al., 2008; VanDyke et al., 2009). In these studies, reducing the wildtype glycan to two sugars or less disrupted flagella assembly causing the cells to become non-flagellated and subsequently non-motile (Chaban et al., 2006; VanDyke et al., 2009). Other mutations that left more than the two sugar minimum also affected motility, even though these cells could still assemble functional flagella (VanDyke et al., 2009).

The $N$-linked glycan structures for *M. voltae*, *M. maripaludis* and *H. volcanii* were reported to be a trisaccharide, tetrasaccharide and pentasaccharide, respectively (Chaban et al., 2006; Kelly et al., 2009; Yurist-Doutsch and Eichler, 2009). The glycans are assembled in the cytoplasm on a dolichol lipid embedded in the cytoplasmic membrane by glycosyl transferases. The glycan is then flipped across the cytoplasmic membrane and attached to selective asparagine residues. MS analysis of the *M. voltae* glycan determined the structure to be $\beta$-ManpNAcA6Thr-(1-4)-$\beta$-GlcNAc3NAcA-(1-3)-$\beta$-GlcNAc (where GlcNAc is attached to the dolichol lipid or Asn residue), while *M. maripaludis* structure is Sug-4-$\beta$-ManNAc3NAmA6Thr-4-$\beta$-GlcNAc3NAcA-3-$\beta$-GalNAc (where the terminal Sug is a diglycoside of aldulose) (Voisin et al., 2005; Kelly et al., 2009). The specific glycan structure of *H. volcanii* is not known however, MS revealed that the pentasaccharide is composed of two hexoses, two hexuronic acids and an additional 190 Da saccharide (Abu-Qarn et al., 2007). In *M. voltae*, all the genes involved in the assembly of the trisaccharide glycan have been found. The first sugar, GlcNAc, is attached to the dolichol lipid by AglH. The function for this gene could not be proven directly through inactivation or deletion methods because of its involvement in the biosynthesis of cofactor B. In order to demonstrate the function of AglH, the gene was complemented into a conditional lethal strain of the eukaryote, *S. cerevisiae*, whose expression of the homologue of AglH, Alg7, can be controlled. Alg7 activity
initiates the assembly of the eukaryotic N-linked glycan by attaching GlcNAc to the dolichol lipid embedded in the endoplasmic reticulum. Turning off the expression of Alg7 causes the cells to die because N-linked glycosylation is essential in eukaryotes. However, these cells were still viable when Alg7 expression was off if aglH was complemented into the cells, indicating that AglH could perform the same function as Alg7 and was most likely doing so in the M. voltae pathway. The remaining glycosyl transferases in M. voltae that attach the second and third sugar were identified through insertional inactivation experiments. In M. maripaludis, the development of in-frame deletion techniques has allowed for a more precise genetic study where any gene can be specifically targeted for deletion (rather than the multiple experiments that would be performed if inactivating a gene that is not the first or last gene of its operon) and returned through complementation. This has led to a similar characterization of the N-linked pathway in M. maripaludis, where the glycosyl transferases responsible for attaching the second and third sugar have been identified, while the first and last glycosyl transferases remained missing at the onset of this work. The flippase component is unidentified for the entire Archaeal domain.

The gene MMP1088 was targeted for in-frame deletion because of its potential role as one of the missing glycosyl transferase in the M. maripaludis N-linked pathway. Annotation and proximity to the second and third N-linked glycosyl transferases (MMP1079 and MMP1080, respectively) made MMP1088 an ideal candidate. Deletion of the MMP1088 gene resulted in a decrease in flagellin molecular weight that was between that observed for flagellins of wildtype and ΔMMP1080, indicating that the N-linked glycan had been altered but not to the extent of ΔMMP1080 (whose glycan only contains the full 2 sugars). Subsequently, MS analysis of purified ΔMMP1088 flagella confirmed that the N-linked glycan was missing the last sugar and unexpectedly, the threonine located on the third sugar. Complementation of the MMP1088 deletion returned the glycan structure to wildtype. Since the annotation of MMP1088 strongly suggests that the translated protein is a glycosyl transferase (based on conserved glycosyl transferase motifs), the role of MMP1088 is likely the addition of the fourth sugar of the
tetrasaccharide N-linked glycan. This step would be required before the attachment of the threonine amino acid. On the other hand, without direct evidence of the function of MMP1088, such as in vitro biochemical analysis, the small possibility also remains that MMP1088 may be adding the threonine, which would then allow for the last sugar to be attached, most likely by another enzyme.

In this study, the six remaining glycosyl transferases (MMP090, MMP356, MMP359, MMP590, MMP1293 and MMP1294) annotated in the M. maripaludis genome that have not been the target of previous researchers, were targeted for in-frame deletion in hopes of identifying the first glycosyl transferase. Deletions were created in four of these genes (ΔMMP359, ΔMMP590, ΔMMP1293 and ΔMMP1294), however western blot analysis revealed no apparent changes in flagellin molecular weight that would be expected of a deleted first glycosyl transferase. The remaining two genes (MMP090 and MMP356) could not be deleted because either recombination for the deletion construct is not favourable or these genes are involved in another pathway that is essential for the cells, such as an essential O-linked pathway or the biosynthesis of glycolipids. General PCR screening of cell cultures grown in 8-aza-hypoxanthine before plating and single colony selection revealed no traces of a mutant sized product, suggesting that isolating a deletion mutant in these genes is unlikely. Given these results, there are now only four possible genes encoding the first glycosyl transferase, all of which appear to be lethal or unfavourable for deletion: MMP090, MMP356, the previously attempted MMP1170 and the MMP1423 gene that is likely involved in the biosynthesis of the essential cofactor B. Homologues of MMP1423 have been demonstrated to play the role of the first glycosyl transferase using M. voltae and the eukaryote S. cerevisiae, which is to add a GlcNAc monosaccharide to a dolichol lipid. Experimental confirmation on whether this protein has relaxed substrate specificity for both GlcNAc and GalNAc is needed. Eukaryotic glycosyl transferases are known for their strict substrate specificity, despite only subtle differences between the monosaccharide sugars. BLAST searches against MMP1088 revealed strong hits (E value of approximately 10^{-74}) to other glycosyl
transferases in archaea such as *Pyrococcus furiosus* and other methanogens, as well as to other genes found within its own genome (MMP356). It is noteworthy that despite BLAST results indicating that MMP1088 and MMP356 are very similar glycosyl transferases, MMP356 must be performing a different function as it is unable to complement the MMP1088 deletion. However, there are rare examples of relaxed specificity among glycosyl transferases as well. It is reported that the β1, 4-N-acetylglucosaminyltransferase, which is capable of transferring GalNAc moieties in addition to GlcNAc, although at much less efficient rates (Ikeda et al., 2000).

Investigation of the relative concentrations of GlcNAc and GalNAc in the intracellular pool of *M. voltae* and *M. maripaludis* would also be useful, and may explain the differences observed in the linking sugar of their N-linked glycan if GalNAc is the major sugar available for MMP1423 to transfer in *M. maripaludis*. Unfortunately, no such studies have been performed in any methanogen.

A key component missing from the *M. maripaludis* and archaeal N-glycosylation pathway in general is the flippase protein that transfers the glycan from the cytoplasmic side of the inner membrane to the periplasmic side. Earlier studies have targeted several potential candidates based on BLAST searches, gene location and annotation in *M. maripaludis*, *M. voltae* and *H. volcanii* without any success (VanDyke et al., 2009; Chaban et al., 2006; J. Eichler, personal communication to K. Jarrell). Although deletion of MMP1089 had already been previously attempted, this gene was the focus for deletion by a variety of methods in this study. Three different attempts, classical in-frame deletion, insertional inactivation and the deletion of MMP1089 along with a neighbouring gene that was known to be susceptible for deletion (MMP1088) were performed. However, these three attempts were unsuccessful and may indicate that the MMP1089 gene is essential for the cells. It is expected that a deletion of the gene encoding the flippase protein should be possible. The N-linked system is not essential in any of the archaeal species currently studied, since the Stt3p protein, whose action occurs at the last step in the pathway, has been deleted in *M. voltae*, *M. maripaludis* and *H. volcanii* (Chaban et al., 2006).
The flippase would only be essential if it was involved in another process that is required by the *M. maripaludis* cells, such as in the case for *MMP1423* and co-factor B. In both eukaryotes and bacteria, flippase proteins, Rft1 and PglK respectively, have been identified. The flippases are quite different, where PglK is an ABC transporter that utilizes ATP for its function while Rft1 appears to flip the glycan independently of ATP. Recently, the function of Rft1 has been disputed by Frank et al., (2008), who suggested that Rft1 merely acts as a critical accessory protein, but is not the flippase itself. However, that group was unable to identify the endoplasmic protein(s) that would be the flippase. The lack of data on any archaeal flippase and the difficulty in identifying obvious homologues to the N-linked flippases found in Bacteria and Eukarya may indicate that there is a novel archaeal protein that makes the N-linked glycosylation pathway in Archaea unique.

Glycomics is the study of an organism’s entire array of oligosaccharides and is considered an emerging field (Nahalka et al., 2007). With research similar to work done in this study, we are beginning to realize the diverse informative role that oligosaccharides and other carbohydrates can have on proteins. Whether oligosaccharides are involved in every day cellular processes such as cell signalling, protein sorting and assembly or in health and disease, such as autoimmune disorders and bacterial pathogenesis, the need to understand them is apparent (Alavi and Axford, 2008). Recent advancements in carbohydrate structural characterization with mass spectrometry, as well as the vast amount of data on biosynthetic enzymes such as glycosyltransferases, have led to the development of exciting techniques that allow for the synthesis of many carbohydrates. These tools will be particularly valuable for research, such as in the study of deoxysugars from plants, fungi and bacteria, which are now recognized to possess bioactivity and holds great potential for drug therapies through the glycosylation of antibiotics and novel anti-cancer drugs (Salas et al., 2005; Trefzer et al., 1999). Future pharmaceutical and vaccine related studies will have the opportunity to investigate the direct immune response to these oligosaccharides by obtaining a pure sample through its *in vitro* synthesis (Seeberger and
Werz, 2007; Nalhalka et al., 2007). In another application, the interaction between two carbohydrates can also be tested through the use of carbohydrate arrays which can have implications for cell signalling, pathogenesis and cancer (Wu et al., 2009). At this time, only bacterial glycosyltransferases are being applied for *in vivo* and *in vitro* engineering of these natural carbohydrate moieties (Luzhetskyy and Bechthold, 2008). Many sugar modifications and glycosyl transferases are still unknown in Archaea, however, with their unique ability to survive in extreme environments, they most likely possess a diverse spectrum of oligosaccharides (Abu-Qarn et al., 2008). Identification of these novel sugars and enzymes (both biosynthetic and glycosyl transferases) will significantly enhance the oligosaccharide synthesis toolbox for future research and their applications.

Currently, our knowledge of the biosynthetic pathways of archaeal sugars is lacking. Sugar biosynthesis can require many different types of enzymes such as methyltransferases, acetyltransferases and epimerases. While most of these enzymes and their role in biosynthetic pathways are known in eukaryotes and bacteria, their homologues in Archaea can utilize different substrates, making it difficult to assign gene function based on annotation alone (Namboori and Graham, 2008). Eukaryotic and bacterial studies have focused on the biosynthesis of acetamido sugars, especially GlcNAc and its derivatives, due to their prevalence in cell structure and function. Bacterial peptidoglycan possesses many cross-linked polysaccharides composed of GlcNAc and one of its derivatives, *N*-acetylmuramic acid (Ghuysen et al., 1984). Gram positive and Gram negative bacteria also contain *N*-acetylmannosaminuronate (ManNAcA) sugars in their capsule or lipopolysaccharides, respectively (Kawamura et al., 1979). In eukaryotes, the *N*-linked glycan contains several GlcNAc sugars with many proteins glycosylated with these sugars (Burda and Aebi, 1999; Zachara and Hart, 2006). Interestingly, bacteria and eukaryotes use a different biosynthetic pathway to synthesize UDP-GlcNAc (a precursor to GlcNAc), where only the initial step, glucosamine synthase converting fructose-6-phosphate (Fru-6-P) to *α*-D-glucosamine-6-phosphate (GlcN-6-P), is conserved in both pathways (Namboori and Graham, 2008). In
eukaryotes, acetyl-coenzyme A is used to acetylate GlcN-6-P before the phosphate group is transferred to the anomeric position of the sugar to form GlcNAc-1-P. The sugar is activated to form UDP-GlcNAc through the actions of a nucleotidyltransferase enzyme using UTP. Alternatively in bacteria, GlcN-6-P is converted to glucosamine-1-phosphate (GlcN-1-P) and uses a bifunctional enzyme to transfer the sugar to UTP before acetylation. In addition to the order of acetylation seen in bacteria, another notable difference occurs at the end of the bacterial pathway where UDP-GlcNAc has the potential to be further processed into UDP-\(N\)-acetylmannosaminuronate (UDP-ManNAcA) (Namboori and Graham, 2008). The two biosynthetic pathways and their enzymes are summarized in Figure 4.1.

Archaea can also use GlcNAc sugars and its derivatives in their flagellin and S-layer proteins, and most likely in the biosynthesis of co-factor B (Chaban et al., 2006; Abu-Qarn and Eichler, 2006; White, 2001). Recently, the biosynthetic pathway of acetamido sugars in Methanococcales (M. maripaludis and M. jannaschii) was found to follow that of bacteria with many orthologues present in the archaeal genomes (Namboori and Graham, 2008). These genes were identified and characterized in both species through in vitro analysis where the universally conserved glucosamine synthase was encoded by MMP1160, MMP1077 converted GlcN-6-P to GleN-1-P, while MJ1101 possessed both acetyltransferase and nucleotidyltransferase ability and MMP705 and MMP706 were responsible for the epimerization of UDP-GlcNAc into UDP-ManNAc and oxidation of UDP-ManNAc into UDP-ManNAcA, respectively. The identification and characterization of more biosynthetic genes in archaea will lead to a better understanding of the evolutionary history between the three domains, provide more tools for glycoengineering and development of vaccines, as well as offer insight into the functional role of oligosaccharides, such as the \(N\)-linked glycan where GlcNAc, ManNAc and GalNAc are present, in assembly and stabilization of archaeal proteins. In M. maripaludis, located between the known \(N\)-linked glycosyl transferases (MMP1079, MMP1080 and MMP1088) were a series of genes, some of which have annotations likely involved in the biosynthesis and modifications of sugars.
Figure 4.1: Biosynthesis pathway of the acetamido sugar GlcNAc in Eukaryotes and Bacteria (Namboori and Graham, 2008).
The gene \textit{MMP1081} was targeted for in-frame deletion because of its location to known \textit{N}-linked glycosyl transferases and BLAST search hits to acetamidino biosynthesis enzymes. Given that the third sugar of the \textit{M. maripaludis} glycan contains such a group, MMP1081 was suspected to be involved in the acetamidino modification of this \textit{N}-linked sugar. Deletion of \textit{MMP1081} led to a decrease in flagellin molecular weight when immunoblot analysis was performed. The molecular weight of ∆MMP1081 flagellins appeared to be slightly lower than ∆MMP1088 flagellins and much higher than MMP1080 (which only has 2 sugars as its glycan). MS analysis confirmed differences in the \textit{N}-linked glycan compared to wildtype. Specifically, it was missing the acetamidino group on the third sugar and entire fourth sugar and in addition, the threonine usually found on the third sugar was also missing in the ∆MMP1081 glycan. This finding provides further insight into the biosynthesis of the sugars in the \textit{N}-linked glycan by suggesting that the acetamidino group plays an important role for the function of the fourth glycosyl transferase and that the final sugar must also be assembled before the addition of the threonine. In support of this, two different mutants (deletion of the fourth glycosyl transferase and deletion of the acetamidino biosynthesis gene) that results in the absence of the terminal sugar are also missing the threonine in the third sugar. This also suggests that the threonine modification may be added after glycan assembly and is not part of the third sugar biosynthetic pathway. Complementation of the MMP1081 gene returned flagellin molecular weights back to wildtype.

In \textit{M. maripaludis}, a recent study has demonstrated the importance of the \textit{N}-linked glycan for the assembly and function of flagella (VanDyke et al., 2009). They report that mutants with the third glycosyl transferase (∆MMP1080) deleted, resulting in only a 2 sugar glycan, could assemble flagella while mutants carrying a deletion of the second glycosyl transferase (∆MMP1079), with a single sugar glycan, could not. When flagella motility was compared between ∆MMP1080 and wildtype cells, ∆MMP1080 cells showed impairment in swimming, demonstrating the importance of a full glycan for optimal flagella function while a minimum size of glycan is needed for its assembly (VanDyke et al., 2009). Due to these observations, motility
and flagella assembly were determined for ∆MMP1081 and ∆MMP1088 cells using semi-swarm agar plates and electron microscopy. Electron microscopy analysis revealed that both ∆MMP1081 and ∆MMP1088 cells had no apparent deficits in the amount of flagella on the cell surface. However, when motility was compared to the wildtype strain, deficiencies in motility were observed in both mutants. Interestingly, the completeness of the glycan is highly correlated to motility and is highlighted when comparing the motility of ∆MMP1081 and ∆MMP1088 cells. The ∆MMP1081 glycan is nearly identical with the ∆MMP1088 glycan with the exception of a missing acetamidino functional group on the third sugar. However, ∆MMP1081 cells have a noticeable deficit in motility. In general, motility and completeness of glycan was found to occur in the following order (most motile to less motile; most complete glycan to least complete glycan): wildtype > ∆MMP1088 > ∆MMP1081 > ∆MMP1081 > ∆MMP1080 > ∆MMP1079. This study provides further insight into the relationship between the N-linked glycan and its role in the function of the flagella with the newly created ∆MMP1088 and ∆MMP1081 mutants.

With the recent work demonstrating the importance of N-glycosylation for archaeal flagellins and their assembly, the N-linked process is believed to be closely associated with the assembly of the flagellum. Figure 4.2 demonstrates how these two processes are hypothesized to interact at the cytoplasmic membrane of M. maripaludis cells. The N-linked process begins with the assembly of the glycan on a dolichol lipid embedded in the cytoplasmic membrane. While the identity of the first glycosyl transferase attaching the linking sugar remains unknown, the data presented in this work indicates that it is likely encoded by one of MMP090, MMP356, MMP1170 or MMP1423. The remaining glycosyl transferases assemble the glycan in the following order after the initial sugar has been attached: MMP1079, MMP1080 and the newly identified MMP1088. The tetrasaccharide glycan is then flipped across the cytoplasmic membrane, by an unidentified protein, with the glycan still attached to dolichol but it is now facing the periplasm. Meanwhile, at the cytoplasmic membrane, pre-flagellin proteins are cleaved
Figure 4.2: Topology and interaction between the N-linked glycosylation system and flagella assembly in *M. maripaludis*. 
by their signal peptidase, FlaK, and transported across the cytoplasmic membrane. Once the flagellin subunits are exposed in the periplasm, the Stt3p protein, MMP1424, transfers the \( N \)-linked glycan from the dolichol carrier to select \( N \)-linked sequons (Asn-Xaa-Ser/Thr, where Xaa is any amino acid other than proline) within the flagellins before they are assembled at the base of the growing flagellum. The mechanism for targeting Asn residues within the \( N \)-linked sequon is currently unknown, but would provide valuable insight as to why some proteins known to be associated with the cytoplasmic membrane are glycosylated while others are not, even though they contain the required Asn-Xaa-Ser/Thr sequon (Jones et al., 2005). In the case of \( M. maripaludis \), even within a single flagellin protein, all potential \( N \)-linked sites are not glycosylated, where the \( N^{26} \) position in FlaB1 and FlaB2 are not glycosylated (Kelly et al., 2009).

Future work is needed to identify the archaeal flippase, although this task is proving to be quite challenging. Classical methods that use homology to bacterial and eukaryotes, gene annotation and function appear to be inadequate and new screening methods may be needed. If the flippase gene happens to be essential, due to its involvement in a separate essential pathway, more sophisticated molecular tools rather than direct deletion can be performed, similar to those that contributed to the identification of the first glycosyl transferase in \( M. voltae \), where cross species complementation was performed. The first glycosyl transferase in \( M. maripaludis \) will also need to be validated since the three remaining candidates appear to be essential. \textit{In vitro} analysis demonstrating sugar substrate specificity and perhaps inter-species complementation will reveal the identity of this protein. Finally, genes that are involved with the biosynthesis of the \( N \)-linked sugars can be investigated fairly easily using MS with the newly identified \( N \)-linked glycan structure. Given the novel diglycoside of aldulose seen in \( M. maripaludis \), archaea likely possess more unique sugars and any insight into the synthesis of these sugars will deepen our growing knowledge in Archaea, as well as to provide an increased repertoire of enzymes for possible glycoengineering studies.
Chapter 5

Conclusion

This study has provided data for the identification of two genes involved in the assembly and biosynthesis of the \( N \)-linked glycan attached to \textit{Methanococcus maripaludis} flagellins. The glycosyl transferase responsible for attaching the fourth sugar to the glycan was found along with a biosynthetic protein responsible for modifying the third sugar in the glycan with the addition of an acetamidino functional group. It was found that both deletion mutants of these proteins were missing the last sugar of the glycan and the threonine amino acid that is located at the C6 position of the third sugar, while the mutant of the biosynthetic gene was also lacking the acetamidino group. The study also limits the potential number of gene candidates that encode for the glycosyl transferase responsible for attaching the first sugar, to only four candidates. To summarize, the data presented in this work provides further insight into the \( N \)-linked glycosylation process as well as the biosynthesis and modification of these \( N \)-linked sugars in archaea.
References


Imperiali, B. and J. Troutman. 2009. Campylobacter jejuni PglH is a single active site processive polymerase that utilizes product inhibition to limit sequential glycosyl transfer reactions. Biochemistry


101


bacteria, and acetogenic bacteria, within plaque biofilms associated with human periodontal

Voisin, S., R. S. Houliston, J. Kelly, J. R. Brisson, D. Watson, S. L. Bardy, K. F. Jarrell, and
S. M. Logan. 2005. Identification and characterization of the unique N-linked glycan common to
the flagellins and S-layer glycoprotein of Methanococcus voltae. J. Biol. Chem. 280:16586-
16593.

Biochem. 133:17-21.


Whitman, W.B., J. Shieh, S. Sohn, D.S. Caras, and U. Premachandran. 1986. Isolation and

Components and dynamics of fiber formation define a ubiquitous biogenesis pathway for
bacterial pili. EMBO J. 19:6408-6418.

Biomol. Chem. 7:2247-2254.

Yonekura, K., S. Maki, D. G. Morgan, D. J. DeRosier, F. Vonderviszt, K. Imada, and K.
Science 290:2148-2152.

Yurist-Doutsch, S., B. Chaban, D. J. VanDyke, K. F. Jarrell, and J. Eichler. 2008. Sweet to

protein expression studies reveal novel genes in the agl cluster responsible for N glycosylation in


Zachara, N. E. and G. W. Hart. 2006. Cell signaling, the essential role of O-GlcNAc! Biochim.


Zhang, H., J. K. DiBaise, A. Zuccolo, D. Kudrna, M. Braidotti, Y. Yu, P. Parameswaran, M.


Appendix A: FlaJ and its function in the *M. maripaludis* flagellum

**FlaJ as a recruiting platform for the flagellar assembly ATPase, FlaI:**

FlaI is an essential ATPase protein for archaeal flagella biogenesis and is likely providing the energy for assembly. Interestingly, FlaI is located in the membrane in wildtype cells despite the lack of any transmembrane domains. It is proposed that FlaI interacts with an integral membrane recruiting protein. The identification of this protein has not yet been found.

FlaJ is also an essential accessory protein of the archaeal flagella with multiple transmembrane domains, although its function is yet to be characterized. Previous FlaJ studies in *M. voltae* had inactivated the *flaJ* gene however the equivalent gene has not yet been deleted in *M. maripaludis*. In order to investigate the role of FlaJ as a possible membrane interacting protein for recruitment of FlaI, a deletion of *flaJ* was constructed.

**Creation of ∆FlaJ and verification of ∆FlaI using PCR:**

The ∆FlaJ strain was constructed in the same manner as previous mutants (Figure AppA.1) and the *flaJ* deletion was confirmed by PCR (Figure AppA.2).

Electron microscopy confirmed that ∆FlaJ cells were unable to assemble flagella (Figure AppA.3), in agreement with earlier reports of its disruption in *M. voltae* (Thomas et al., 2002).

**Analysis of ∆FlaJ using anti-FlaI antibody for Immunoblotting:**

To test if FlaJ acts as a membrane recruiting protein by attaching FlaI to the cytoplasmic membrane, western blotting using anti-FlaI antibody was performed on membrane preparations. Results revealed that FlaI is present in ∆flaJ membranes while positive and negative controls of MM900 and ∆FlaI membranes were used to validate the effectiveness of the antibody and preparation of each membrane (Figure AppA.4).
Figure AppA.5.1: Creation of the in-frame deletion plasmid for *flaJ*. A) PCR amplification of the ~1 kb upstream and downstream arms of the deletion construct. B) PCR amplification of 2 kb DNA fragment (arrow) after ligation of upstream and downstream pieces. C) Restriction enzyme digest (BamHI) of pCRPrtNeo containing the 2 kb deletion construct.
Figure AppA.5.2: Confirmation of flaJ deletion. Whole cell PCR produces a wildtype (wt) size product of 2280 bp, while ∆FlaJ cells produce an 865 bp sized product.
Figure AppA.5.3: Electron micrograph of *M. maripaludis* strains A) MM900 (flagellated) B) ΔflaJ (non-flagellated).
Figure AppA.5.4: Immunoblot using anti-FlaI antibody against *M. maripaludis* wildtype (wt), ΔflaI and ΔflaJ crude membrane extracts, with the corresponding band for FlaI (indicated by arrows). A) and B) are two separate immunoblots. Blots were developed with antibodies raised in chickens against *M. voltae* FlaI.
Appendix B: Media Recipes

Balch III medium was prepared by boiling under an atmosphere of CO₂/H₂ (20:80) and reducing with cysteine-sulfide. The medium contained resazurin (final concentration of 1 mg/L) as an oxygen indicator. To prepare solid plates, 15 g of agar was added to the medium. Balch III medium contains, per litre dissolved in water:

- 0.335 g KCl
- 2.75 g MgCl₂·6H₂O
- 3.45 g MgSO₄·7H₂O
- 0.25 g NH₄Cl
- 0.07 g CaCl₂·2H₂O
- 0.18 g K₂HPO₄
- 18 g NaCl
- 10 ml Mineral elixir
- 10 ml Vitamin mixture
- 1 g NaAcetate·3H₂O
- 2 g yeast extract
- 2 g pepticase
- 6 ml FeSO₄ solution
- 0.1 ml NiCl solution
- 4.8 ml 8% NaCO₃
Mineral Elixer (per litre of water):

- 1.5 g Nitrilotriacetate
- 3 g MgSO₄·7H₂O
- 0.5 g MnSO₄·2H₂O
- 1 g NaCl
- 0.1 g CoCl₂·6H₂O
- 0.1 g CaCl₂·2H₂O
- 0.1 g ZnSO₄·7H₂O
- 0.01 g CuSO₄·5H₂O
- 0.01 g AlK(SO₄)₂·12 H₂O
- 0.01 g H₃BO₃
- 0.01 g Na₂MoO₄·2H₂O

Vitamin mixture (per litre of water):

- 2 mg biotin
- 2 mg folic acid
- 10 mg pyridoxine HCl
- 5 mg thiamine HCl
- 5 mg riboflavin
- 5 mg nicotinic acid
- 0.5 mg vitamin B₁₂
- 5 mg p-aminobenzoic acid
- 5 mg lipoic acid

FeSO₄ solution:

- 4 mg FeSO₄/L H₂O
McCas medium was prepared using the same method as Balch III media. It also contained resazurin (1 mg/L) as an oxygen indicator. To prepare solid plates, 20 g of noble agar (Difco) was added. McCas medium contains, per litre:

- 500 ml H$_2$O
- 500 ml General Salts solution
- 5 g NaHCO$_3$ (liquid) or 2 g (agar plates)
- 22 g NaCl
- 10 ml K$_2$HPO$_4$ solution
- 5 ml FeSO$_4$ solution
- 10 ml Trace Mineral solution
- 10 ml Vitamin mixture
- 1.4 g NaAcetate·3H$_2$O
- 2 g Casamino acids

General Salts solution (per litre of water):

- 0.67 g KCl
- 5.5 g MgCl$_2$·6H$_2$O
- 6.9 g MgSO$_4$·7H$_2$O
- 0.28 g CaCl$_2$·2H$_2$O
- 1 g NH$_4$Cl (omit from nitrogen-free stock)

K$_2$HPO$_4$ solution:

- 14 g K$_2$HPO$_4$ per litre of H$_2$O
Trace Mineral solution (per litre of water):

- 2.1 g Na$_3$Citrate·2H$_2$O
- Adjust pH to 6.5
- 0.5 g MnSO$_4$·2H$_2$O
- 0.1 g CoCl$_2$·6H$_2$O
- 0.1 g ZnSO$_4$·7H$_2$O
- 0.01 g CuSO$_4$·5H$_2$O
- 0.01 g AlK(SO$_4$)$_2$
- 0.01 g H$_3$BO$_4$
- 0.1 g Na$_2$MoO$_4$·2H$_2$O
- 0.025 g NiCl$_2$·6H$_2$O
- 0.2 g Na$_2$SeO$_3$

The nitrogen-free media used in the complementation studies was prepared the same as McCas media, however, NH$_4$Cl and Casamino acids were omitted from the recipe. Also, the media was reduced by the addition of dithiothreitol (DTT) (0.5 g) as opposed to cysteine-sulfide.