GENETIC ANALYSIS AND POST TRANSLATIONAL MODIFICATIONS OF TYPE IV PILI IN
METHANOCOCCUS MARIPALUDIS

by

Divya Balachandran Nair

A thesis submitted to the Department of Biomedical and Molecular Sciences
in conformity with the requirements for the degree of
Doctor of Philosophy

Queen's University
Kingston, Ontario, Canada
(September, 2014)

Copyright©Divya Balachandran Nair, 2014
Abstract

*Methanococcus maripaludis* has two different surface appendages that are considered to be type IV pili-like structures: the well-studied archaella as well as type IV pili. Though prevalent among various archaea, studies on type IV pili in any archaeon are extremely limited. In this thesis, many novel essential components necessary for type IV pili formation in *M. maripaludis* have been identified by gene deletion analysis coupled with electron microscopy. Several of these essential components were found in a previously identified eleven gene locus (*mmp0231-mmp0241*) containing three previously identified pilin-like genes and a prepilin peptidase. The genes in this locus were shown in this work, by reverse transcriptase-polymerase chain reaction experiments, to be arranged in a single operon. The remaining unstudied genes, except for *mmp0231*, were deleted and, with the exception of *mmp0235* and *mmp0238*, found to be essential for piliation. Outside this locus, the conserved assembly ATPase (*mmp0040; epdL*) and, unusually, two tandem versions of the conserved type IV pilus platform protein (*mmp0038 and mmp0039; epdJ and epdK*) were identified. All three were shown to be essential for piliation. Furthermore, the gene encoding the major pilin structural protein (*mmp1685; epdE*) was identified at a separate locus. In addition, six other pilin-like genes (*mmp0528, mmp0600, mmp0601, mmp0709, mmp0903, mmp1283*) scattered around the genome were targeted for deletion, with electron microscopy of the individual deletion strains revealing that normal piliation was only affected when *mmp1283* (*epdD*) was missing, suggesting it was a fourth minor pilin. Lastly, studies were conducted to examine the order of the two posttranslational modifications of pilins and archaellins. While it was determined that signal peptide cleavage and N-linked glycosylation of archaellins could occur independently of each other, this was not the case for pilins. Detection of epitope-tagged pilins expressed in various mutant backgrounds unable to carry out one, the other, or both posttranslational modifications, strongly suggested that pilins cannot be N-glycosylated unless the signal peptide
is first removed. These studies reveal that the type IV pili of *M. maripaludis* are more complex than those of other studied archaea and also raise questions about how the cell is able to distinguish between the two types of type IV-pilin-like proteins, archaellins and pilins.
Co-Authorship

Dr. Ken Jarrell has provided the editorial input for this thesis that has greatly improved the quality of this thesis. Experiments in this thesis were designed by myself in collaboration with Ken Jarrell. All first drafts of published papers were written by me and edited by Ken Jarrell. Chapter 3 is a published manuscript co-authored by Kaoru Uchida and Dr. Shin-Ichi Aizawa (Prefectural University of Hiroshima), and they were responsible for the electron microscopy analysis. They did not contribute to the writing of the manuscript. Chapter 4 is co-authored by Daniel K.C. Chung and James Schneider (former MICR499 undergraduate thesis students), who contributed in making some of the pilin-like gene deletions and Kaoru Uchida and Dr. Shin-Ichi Aizawa (Prefectural University of Hiroshima) who did the electron microscopy analysis. None of these individuals contributed to the writing of the manuscript. Chapter 5 is a manuscript in preparation.
Acknowledgements

Firstly, with deep sense of gratitude, I would like to acknowledge my supervisor, Professor Ken F. Jarrell, for his excellent guidance, helpful discussions, and continuous encouragement throughout this journey. I will always be grateful for the time you have committed to my development. Prof. Jarrell has been instrumental in ensuring my academic and professional well being ever since I started my course by constant guidance and providing opportunities to attend international conferences. His enthusiasm and critical thinking came as a boon during my candidature. In offering me the opportunity to work and learn in his laboratory, he offered me the invaluable gifts of knowledge and critical thinking, which I will cherish for life.

I would like to thank my committee members, Dr. Virginia Walker, Dr. Keith Poole and Dr. Katrina Gee, for their continuous advice and constant direction all throughout the course. Thesis committee meetings helped me to plan and execute this work in time. Also, I would like to thank Dr. Martin for serving on my thesis committee and support through the formalities of the graduate program.

I am greatful to all my fellow research students for their support. I am always grateful to all my lab mates especially to John Wu, Gareth Jones, Yan Ding and Sarah Siu. Special thanks are due to the staff in the department including Chris Boer, Jerry Dering, Tammy Henry, Jackie Jones and Diane Sommerfeld. Special thanks to Dorothy Agnew for her help during TAing.

Most importantly I thank my parents and siblings whose love and supports carried me through good and bad times. I am very greatful to my mother for being here to support me by taking care of my son and all my family members for the support at that time. Last, but not least, I thank my husband, Dileep Nair, who constantly supported me throughout my studies and without whom I could not have accomplished what I did as well as my son Dev Nandan for
being extremely supportive during the thesis writing period by adapting very well to the circumstances.
Statement of Originality

I hereby certify that all of the work described within this thesis is the original work of the author. Any published (or unpublished) ideas and/or techniques from the work of others are fully acknowledged in accordance with the standard referencing practices.

(Divya Balachandran Nair)

(September, 2014)
Table of Contents

Abstract........................................................................................................................................... ii
Co-Authorship.................................................................................................................................. iv
Acknowledgements....................................................................................................................... v
Statement of Originality................................................................................................................... vii
List of Figures .................................................................................................................................... xi
List of Tables ....................................................................................................................................... xiii
List of Abbreviations....................................................................................................................... xiv

Chapter 1
General Introduction ........................................................................................................................ 1
Literature cited .................................................................................................................................. 4

Chapter 2
Literature Review .............................................................................................................................. 8
Archaeal Surface Appendages ......................................................................................................... 10
Non-type IV pili-like structures ...................................................................................................... 13
  (i) Hami ........................................................................................................................................ 13
  (ii) Cannulae ................................................................................................................................. 15
  (iii) Mth60 fimbriae ..................................................................................................................... 15
Type IV pili-like structures ............................................................................................................. 18
  Bacterial Type IV pili characteristics ......................................................................................... 18
  (i) Bindosome .............................................................................................................................. 20
  (ii) Iho670 fibers ........................................................................................................................ 21
  (iii) Archaella .............................................................................................................................. 23
    (a) Glycosylation ...................................................................................................................... 29
        N-Glycosylation in methanogens ........................................................................................ 30
  (iv) Type IV-like pili ................................................................................................................... 33
    (a) Adhesive (Aap) pili ............................................................................................................. 34
    (b) UV-induced (Ups) pili ........................................................................................................ 36
    (c) Pili in Haloferax volcanii .................................................................................................. 38
    (d) Pili in Methanococcus maripaludis .................................................................................. 39
Aims of the study ............................................................................................................................. 40
Chapter 3
Genetic analysis of a type IV pili-like locus in the archaeon Methanococcus maripaludis

Abstract
Introduction
Materials and Methods
Results
Discussion
Literature cited

Chapter 4
Identification of an additional minor pilin essential for piliation in the archaeon Methanococcus maripaludis

Abstract
Introduction
Materials and Methods
Results
Discussion
Literature cited

Addendum to chapter 4
Identification of MMP1685 as the Major Pilin in Methanococcus maripaludis

Literature cited

Chapter 5
Pilin processing follows a different temporal route than that of archaellins in the archaeon Methanococcus maripaludis

Abstract
Introduction
Materials and Methods
Results and Discussion
Literature cited
Chapter 6

General Discussion ........................................................................................................................................... 160

Genes involved in *M. maripaludis* piliation are scattered around the genome ................................................. 161
Deciphering the role of *M. maripaludis* epd genes of unknown function .......................................................... 163
*M. maripaludis* type IV pilus ATPase and conserved membrane protein ......................................................... 166
Minor pilins in *M. maripaludis* .......................................................................................................................... 168
Pilin posttranslational modifications ................................................................................................................ 171
Mode of assembly of *M. maripaludis* pili ........................................................................................................... 175
Conclusions and Future Directions .................................................................................................................. 177

Literature cited .................................................................................................................................................... 180
List of Figures

Chapter 2
Figure 2.1: Hami of SM1 euryarchaeon ................................................................. 14
Figure 2.2: Scanning electron micrograph showing a network of Pyrodictium cells and cannulae ................................................................. 16
Figure 2.3: Scanning electron micrograph of Mth60 fimbriae of M. thermotrophicus ......... 17
Figure 2.4: Numerous lho670 fibers emerging from the surface of Ignicoccus hospitalis cells.. ........................................................................................................ 22
Figure 2.5: Electron micrograph of M. maripaludis surface showing thin pili (arrows) with thicker and more numerous archaella ........................................................................................................ 24
Figure 2.6: Comparison of the fla operons of selected archaea ........................................ 25
Figure 2.7: A proposed model of archaellin glycosylation in M. maripaludis ..................... 31
Figure 2.8: Type IV pili-like loci found in three different Archaea .................................... 35

Chapter 3
Figure 3.1: Analysis of the major type IV pilus locus .................................................... 68
Figure 3.2: Confirmation of the deletions of the targeted pili locus genes by PCR ............. 72
Figure 3.3: Electron micrographs of strains carrying the indicated pili locus gene deletion showing its effect on piliation ........................................................................................................ 73
Figure 3.4: Electron micrographs of deletion strains complemented with a plasmid-borne wildtype copy of the deleted gene ........................................................................................................ 75
Figure 3.5: The five gene locus containing genes for the pili assembly ATPase and two copies of the platform (pilC-like) protein ........................................................................................................ 76
Figure 3.6: Electron micrographs of strains deleted for the type IV pili ATPase and platform protein genes and their complemented strains ........................................................................................................ 78

Chapter 4
Figure 4.1: Sequences of the six pilin-like proteins as well as the major pilin (MMP1685) of M. maripaludis studied in this report ........................................................................................................ 104
Figure 4.2: RT-PCR analysis reveals that all six pilin-like genes are transcribed under standard laboratory conditions ........................................................................................................ 105
Figure 4.3: Confirmation of the deletion of each of the six pilin-like genes ..................... 107
Figure 4.4. Electron micrographs of strains carrying deletions of each of the six pilin-like genes.................................................................................................................................108

Figure 4.5: Electron micrographs showing that complementation restores piliation to the M. maripaludis Δmmp1283 strain ........................................................................................................109

Addendum to Chapter 4
Addendum Figure 1: Posttranslational modifications of M. maripaludis pilins ....................123
Addendum Figure 2: PCR analysis of wildtype and the epdE deletion strain .........................125
Addendum Figure 3: Electron microscopy examination of M. maripaludis ΔflaKΔepdE mutant cells carrying the vector control (A, B) and carrying the complementation construct (C, D) ....126

Chapter 5
Figure 5.1: Alignment of the N-terminus region including the signal peptide of type IV pilin-like proteins of M. maripaludis .................................................................140
Figure 5.2: Western blot analysis of various M. maripaludis mutant strains expressing a plasmid borne C-terminal histagged version of EpdE .........................................................141
Figure 5.3: Western blot analysis of various M. maripaludis mutant strains expressing a plasmid borne C-terminal FLAG-tagged version of EpdE or EpdE with the +3 amino acid position (relative to the signal peptide cleavage site) changed to glycine ..................144
Figure 5.4: Coomassie stained SDS-PAGE gel following the purification of C-terminal Histagged EpdE by Ni-affinity chromatography .................................................................146
Figure 5.5: Western blot detection of native EpdE in various mutant backgrounds ................147
Figure 5.6: Confirmation of the double gene deletion strains by PCR ..................................149
Figure 5.7. Examination of posttranslational modifications of archeallins and pilins in various mutant backgrounds ..........................................................................................................150

Chapter 6
Figure 6.1. A hypothetical model of pili assembly in M. maripaludis ........................................176
List of Tables

Chapter 2
Table 2.1: Comparison of different archaeal surfacestructures...........................................11

Chapter 3
Supplemental Table 3.1: Primers used in this study for gene deletion, complementation, site-directed mutagenesis and screening mutants.................................................................61
Supplemental Table 3.2: Plasmids used in this study................................................................63
Supplemental Table 3.3: Primers used in this study for RT-PCR experiments.........................66
Table 3.1: Properties of the genes/proteins in the pili locus of Methanococcus maripaludis ......69
Table 3.2: Summary of the targeted genes effects on piliation....................................................80

Chapter 4
Table 4.1: Primers used in this study........................................................................................99
Table 4.2: Plasmids used in this study.......................................................................................101

Chapter 5
Table 5.1: Primers used in this study for PCR.........................................................................134
Table 5.2: Strains and plasmids used in this study.................................................................136
### List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aap</td>
<td>Archaeal-adhesive pilus</td>
</tr>
<tr>
<td>Da</td>
<td>Dalton</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediamine tetraacetic acid</td>
</tr>
<tr>
<td>Epd</td>
<td>EppA-dependent pili</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropylthio-beta-galactoside</td>
</tr>
<tr>
<td>Kb</td>
<td>kilobase</td>
</tr>
<tr>
<td>kDa</td>
<td>kiloDalton</td>
</tr>
<tr>
<td>ORF</td>
<td>open reading frame</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PEG</td>
<td>polyethylene glycol</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulfate</td>
</tr>
<tr>
<td>TMDs</td>
<td>Trans membrane domains</td>
</tr>
<tr>
<td>Ups</td>
<td>UV-inducible type IV pilus of <em>Sulfolobales</em></td>
</tr>
<tr>
<td>Lrp</td>
<td>Leucine-responsive regulatory protein</td>
</tr>
</tbody>
</table>
Chapter 1

General Introduction

Three recent reviews (Albers et al. 2013; Cavicchioli. 2011; Jarrell et al. 2011a) have summarized the significant contributions that work on Archaea has contributed to the scientific community. This growing awareness of the importance of Archaea stems from research in widely diverse areas ranging from evolution to biotechnology, human disease and the search for extra-terrestrial life forms. Archaea were initially thought to be restricted to several “extreme” environments including ones featuring high temperature (hyperthermophilic), low pH, saturating concentrations of NaCl and extreme anaerobiosis (Chaban et al. 2006; Kashefi and Lovley 2003; Tu et al. 1982; Zillig et al. 1981; Grant et al. 2001; Futterer et al. 2004; Schleper et al. 1995). With the advent of methodology that allowed the detection of organisms that were not readily cultivated from the environment using molecular means, it became clear that Archaea were found in basically all environments where serious efforts were made to look for them (Barns et al. 1996; Chaban et al. 2006). These habitats include ocean waters, freshwater sediments, soils, the rumen, sewage digestors, marshlands and in various locations in the human body (Chaban et al. 2006; Hoffman et al. 2013; Probst et al. 2013; Karner et al. 2001; Olsen 1994; Offre et al. 2013).

The important role of Archaea in the environment, especially in the carbon, nitrogen and sulfur cycles, is now being realized as the distribution and metabolic potential of Archaea becomes more understood (Thauer et al. 2008; Walker et al. 2010; Bruneel et al. 2008; Offre et al. 2013; Liu et al. 2012). The key terminal role of methanogens in the anaerobic degradation of complex carbohydrates has been long studied and methanogens are estimated to contribute 1
billion tonnes of methane, a significant greenhouse gas, per annum, of which almost half escapes into the atmosphere (Thauer 2011). However, more recently, a major archaeal involvement in two other systems has been elucidated; anaerobic methane oxidation and ammonia oxidation (Knittel and Boetius 2009; Schleper and Nicol 2010)). Anaerobic methane oxidation or “reverse methanogenesis” is carried out by a specific archaeal lineage called ANME (anaerobic methane-oxidizing archaea) in a symbiotic relationship with sulfate-reducing bacteria, resulting in the oxidization of methane to carbon dioxide with the reduction of sulfate (Thauer 2011), thus reducing the atmospheric emission of methane. Two distinct groups, ANME-1 and ANME-2 belonging to the methanotrophic archaea were known to be involved in AOM at methane seeps (Conrad 2009). ANME archaea contain gene homologues of the methyl-coenzyme M reductase, the key enzyme of methanogenesis, suggesting that a similar enzyme may be involved in anaerobic oxidation of CH\textsubscript{4} to CO\textsubscript{2} (Hallam et al. 2004). All methane-containing environments are believed to have methane oxidizing archaea but they are best studied in ocean environments (Liu et al. 2012). They are found in almost all environments and are among the most abundant micro-organisms on Earth (Schleper and Nicol 2010). Their discovery led to a complete re-evaluation of the nitrogen cycle, since it was previously believed that ammonia oxidation was carried out solely by bacteria.

In their various environments, the role of archaeal surface structures is paramount. For example both archaella and pili as well as other surface appendages including hami and cannulae have been shown to play key roles in such varied functions as adhesion, motility, cell to cell interactions, cell aggregation, DNA transfer and biofilm formation (Offre et al. 2013; Lassak et al 2012; Jarrell et al 2013; Perros et al. 2014; Nickell et al. 2003; Jarrell et al. 2011b; Frols et al. 2008; Ajon et al. 2011). Yet, despite their obvious importance in the interactions of archaea with their environment, with the possible exception of archaella (Shahapure et al. 2014; Jarrell and Albers 2012; Ng et al. 2006), the study of archaeal surface structures has been
relatively neglected. Indeed, many studies of archaeal surface appendages, such as hami, cannulae, Mth60 fimbriae, and Iho670 fibres are reported in organisms which have no genetic systems to fully analyze the assembly and function of the structure (Muller et al. 2009; Rieger et al. 1995; Moissl et al. 2005; Thoma et al. 2008).

The major goal of this study was to add to the very little information available on pili systems in Archaea. Before this study, Sulfolobus was the only known archaeal genus in which significant information was available (Albers et al. 1999; Frols et al. 2008; Henche et al. 2012; Ajon et al. 2011). M. maripaludis is one of the few Archaea for which genetic tools have been developed (Leigh et al. 2011). Using a variety of genetic tools available in M. maripaludis, such as the ability to create in-frame deletions (Moore and Leigh 2005), an efficient transformation system (Tumbula et al. 1994) and complementation technology (Kessler and Leigh in 1999), we were able to identify for the first time a number of essential genes, including novel ones, involved in the structure and assembly of the pilus as well as discover an unusual difference in the sequence of post-translational modifications of pilins compared to archaellins.
Literature cited


and autotrophy in globally distributed marine crenarchaea. Proc Natl Acad Sci U S A. 107, 8818-8823.

Chapter 2

Literature Review

Archaea were proposed as a third domain of life, distinct from Bacteria and Eukarya, in 1977 by Carl Woese based primarily on small subunit rRNA sequence analysis (Woese and Fox 1977; Woese et al. 1990). Early on, the notion of archaea as extremophiles was established, as they were known to thrive in extreme environmental conditions that defined the known limits for life (Stetter 1999). These included such hostile conditions as temperatures over 100°C (Kashefi and Lovley 2003; Tu et al. 1982; Zillig et al. 1981), concentrated NaCl (Grant et al. 2001) and a pH approaching 0 (Futterer et al. 2004; Schleper et al. 1995). Later, it was recognized through culture-independent methodologies that archaeal distribution in nature is cosmopolitan and that archaea were, in fact, significant and sometimes dominant components of a variety of “nonextreme” habitats including ocean waters, freshwater sediments and soils (Barns et al. 1994, 1996; DeLong 1992, 2005; Delong et al. 1994).

In the initial proposal of Woese et al. (1990), the Archaea were divided into two phyla. Euryarchaeota included methanogens, extreme halophiles, the sulfate-reducing genus Archaeoglobus as well as Thermoplasma and the hyperthermophiles Thermococcus and Pyrococcus. The Crenarchaeota contained the thermoacidophiles like Sulfolobus, as well as numerous hyperthermophilic genera such as Pyrodictium. Later, other phyla were proposed including Nanoarchaeota, comprised of very small cells (Nanoarchaeum equitans) that only grow when attached to a larger archaeal host (Ignicoccus hospitalis) (Huber et al. 2002) and Korarchaeota, based on environmental sequence data, although no pure cultures exist (Barns et al. 1996; Gupta and Shami 2011). Most recently, it has been proposed that several genera
previously considered to be mesophilic members of the Crenarchaeota actually should be grouped as the novel phylum Thaumarchaeota (Brochier-Armanet et al. 2008).

Traditionally, the cultured members of the archaea included methanogens, extreme halophiles and a variety of thermoacidophiles and later hyperthermophiles (Chaban et al. 2006a). Methanogens can be found in a wide variety of anaerobic habitats that include marine waters, freshwater sediments, swamps, anaerobic digestors, the rumen, termite guts and the human body (Chaban et al. 2006a; Liu and Whitman 2008). Extreme halophiles thrive in hypersaline environments like salterns, salt lakes and the Dead Sea (Andrei et al. 2012; Oren 2002, 2013). Thermoacidophiles, such as Sulfolobus and Thermoplasma, grow at elevated temperature and low pH, as found in environments like solfatara fields, acid mine drainage and self-heating coal spoil heaps (Lucheta et al. 2013; Volant et al. 2012). Hyperthermophiles, first isolated in the 1980s (Stetter et al. 1981), are defined as organisms having an optimal growth temperature above 80°C (Stetter 1996) and they are often capable of growth at temperatures in excess of 100°C (Adams 1998; Huber et al. 2000). They thrive in niches that include deep sea hydrothermal vents, hot springs, and in volcanic terrestrial environments (Stetter 1999). More recently, members of the novel phylum Thaumarchaeota (Pester et al. 2011) have been cultured in the laboratory, including the marine organism Nitrosopumilis maritimus (Konneke et al. 2005) and the terrestrial archaeon Nitrososphaera viennensis (Tourna et al. 2011). These organisms are known to play an important role in the oxidation of ammonia to nitrite in the nitrogen cycle (Offre et al. 2013).

While studies on archaeal systems have contributed significantly to our understanding of many basic processes in the other domains (Igura et al. 2008; Van den Berg et al. 2004; Yamagata and Tainer 2007), for many reasons, including difficulty in routine cultivation and general lack of genetic tools in most genera, many very basic aspects of archaeal physiology, ecology, biochemistry and genetics remain poorly understood (Rother and Metcalf 2005).
Archaeal Surface Appendages

Interactions of archaea with their environment are often due to a variety of unusual and often archaeal-specific surface appendages which are necessary for a variety of functions including motility, attachment, biofilm formation, DNA uptake and cell communication (Jarrell et al. 2013; Lassak et al. 2012a; Ng et al. 2008).

Archaea, like bacteria, are known to possess many kinds of cell surface appendages; some of them resemble bacterial appendages while many are unique to the archaeal domain. In archaea, appendages have been classified into two major subtypes (Albers and Pohlschroder 2009; Jarrell et al. 2011b): non type IV pili-like (Horn et al. 1999; Moissl et al. 2005; Thoma et al. 2008) and type IV pili-like (Ajon et al. 2011; Frols et al. 2008; Henche et al. 2012a; Ng et al. 2011). A comparison of these different surface appendages are presented in Table 2.1.

Important components in the type IV pilus-like assembly system distinguish these structures from non-type IV pili-like structures (Jarrell et al. 2013; Lassak et al. 2012). These components include structural proteins made initially with class 3 signal peptides (Pohlschroder et al. 2005), a prepilin peptidase that cleaves these signal peptides, a polymerizing ATPase that supplies energy to incorporate new subunits (sometimes a second, depolymerizing, ATPase is involved to remove structural proteins as in twitching motility in bacterial type IV pili (Burrows 2012)), and an integral membrane protein that acts as a scaffold for the assembly system (Peabody et al. 2003).
Table 2.1. Comparison of different archaeal surface structures

<table>
<thead>
<tr>
<th>Type IV pili-like</th>
<th>Diameter</th>
<th>Structural proteins</th>
<th>TFP-like ATPase</th>
<th>TFP-like conserved membrane protein</th>
<th>Prepilin peptidase</th>
<th>Glycosylation of structural proteins</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Archaella</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>M. maripaludis</em></td>
<td>12 nm</td>
<td>FlaB1, FlaB2, FlaB3</td>
<td>Flal</td>
<td>FlaJ</td>
<td>FlaK</td>
<td>N-linked Tetrasaccharide</td>
<td>Motility, adhesion</td>
</tr>
<tr>
<td><em>S. acidocaldarius</em></td>
<td>12-14 nm</td>
<td>FlaB</td>
<td>Flal</td>
<td>FlaJ</td>
<td>PibD</td>
<td>N-linked glycosylation of FlaB shown but attached glycan structure not determined</td>
<td>Motility, adhesion</td>
</tr>
<tr>
<td><em>Hfx. volcanii</em></td>
<td>ND*</td>
<td>FLGA1 and FLGA2</td>
<td>ND*</td>
<td>ND*</td>
<td>PibD</td>
<td>N-linked glycosylation but attached glycan structure not determined</td>
<td>Motility</td>
</tr>
<tr>
<td>Pili</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>M. maripaludis</em></td>
<td>6 nm</td>
<td>EpdA, EpdB, EpdC</td>
<td>EpdL</td>
<td>EpdJ, EpdK</td>
<td>EppA</td>
<td>N-linked Pentasaccharide</td>
<td>Adhesion</td>
</tr>
<tr>
<td><em>S. acidocaldarius</em></td>
<td>8-10 nm</td>
<td>AapA, AapB</td>
<td>AapE</td>
<td>AapF</td>
<td>PibD</td>
<td>ND*</td>
<td>Adhesion</td>
</tr>
<tr>
<td>Aap pili</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>S. solfatarius</em></td>
<td>10 nm</td>
<td>UpsA, UpsB</td>
<td>UpsE</td>
<td>UpsF</td>
<td>PibD</td>
<td>ND*</td>
<td>Cell aggregation</td>
</tr>
<tr>
<td>Ups pili</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Haloarcula marismortui</em></td>
<td>14 nm</td>
<td>Igni_0670</td>
<td></td>
<td>PibD</td>
<td></td>
<td>Glycoprotein staining reactions were negative</td>
<td>Adhesion</td>
</tr>
<tr>
<td><em>Non-Type IV pili-like</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Hami</em></td>
<td>7-8 nm</td>
<td>120 kDa protein</td>
<td>Not present</td>
<td>Not present</td>
<td>Not present</td>
<td>Glycoprotein staining reactions were negative</td>
<td>Cell to cell contact and biofilm formation</td>
</tr>
<tr>
<td>Cannulae</td>
<td>25 nm</td>
<td>CanA, CanB, CanC</td>
<td>Not present</td>
<td>Not present</td>
<td>Not present</td>
<td>Glycoproteins but no structures or linkages of glycans reported</td>
<td>Intercellular communication and transport, adhesion; Cell to cell contact</td>
</tr>
<tr>
<td>----------</td>
<td>-------</td>
<td>-----------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
<td>---------------------------------------------------------------</td>
<td>---------------------------------------------------------------------</td>
</tr>
<tr>
<td>Mth60 Fimbriae</td>
<td>5 nm</td>
<td>Mth60</td>
<td>Not present</td>
<td>Not present</td>
<td>Not present</td>
<td>Glycoprotein but no structures or linkages of glycans reported</td>
<td>Adhesion; Cell to cell contact</td>
</tr>
</tbody>
</table>

* ND - not determined
Non-type IV pili-like structures

Non type IV pili-like appendages in archaea include ones, like hami and cannulae that are unique to the Domain Archaea while also including the Mth60 fimbriae of *Methanothermobacter thermautotrophicus*, so far the only studied example of a non-type IV pilus in archaea.

(i) Hami

The SM1 euryarchaeon, discovered in unique strings-of-pearls communities in association with *Thiothrix* or IMB1 proteobacterium in low temperature sulfidic springs (Moissl et al. 2002) express a novel appendage with a highly complex structure, called the hamus (Moissl et al. 2005) (Figure 2.1). There are about 100 filamentous hami located peritrichously on each SM1 archaeal cell. Each hamus is 7-8 nm in diameter with a length ranging from 1-3 µm and each is composed of a 120 kDa protein (Moissl et al. 2005). Three so-called prickles emanate from the helical filament every 46 nm while the distal end terminates in a tripartite hook giving the appearance of a miniature grappling hook. Hami are known to be stable at high temperatures and over a wide range of pH and help the cells to adhere to chemically diverse surfaces, as well as to their bacterial partners (Moissl et al. 2005). Studies on a SM1 biofilm revealed that cells are regularly spaced at a distance of about 4 µm, believed to be the result of contact of the hami (average length near 2 µm) from neighboring cells (Henneberger et al. 2006). It has been suggested that the roles played by hami in cell to cell contact and biofilm formation are variations of roles typically played by pili and flagella in bacteria (Henneberger et al. 2006).
**Figure 2.1: Hami of SM1 euryarchaeon.** A) Electron micrograph of hami radiating from SM1 euryarchaeal cell, Bar = 500nm. (B) The hook and prickle region of a hamus filament Bar = 50 nm. (From Jarrell et al. 2013).
ii) Cannulae

Cells of the hyperthermophilic crenarchaeal genus, *Pyrodictium*, were shown to grow embedded in a dense network of unusual appendages called cannulae (Figure 2.2) (Horn et al. 1999). These hollow tubules have an outer diameter of 25 nm and are composed of at least three related small glycoproteins designated CanA, CanB and CanC. Cannulae connect neighbouring cells and three dimensional reconstructions of cannulae-cell interactions provided evidence that cannulae enter the periplasmic space (Nickell et al. 2003). While known functions for cannulae are lacking, the connection to the periplasm suggests a possible role in intercellular communication and transport, in addition to a likely role in adhesion (Albers and Meyer 2011). These structures are known to be highly heat-tolerant and insensitive to denaturing agents. They may also be essential for *Pyrodictium* cell growth, as cannulae-free mutants have never been observed (Nickell et al. 2003).

(iii) Mth60 fimbriae

The surface appendages of the thermophilic archaeon, *M. thermautotrophicus*, are known as Mth60 fimbriae (Figure 2.3) and they are the only archaeal pili studied thus far that are clearly not type IV pili (Thoma et al. 2008). These structures are extremely thin, at only 5 nm, and are composed of a 16 kDa glycoprotein that is predicted to be 143 amino acids long after signal peptide removal. Mth60 fimbriae were the first archaeal pili shown to be adhesins, enabling cells to bind to both abiotic surfaces and to other cells. Interestingly, the number of fimbriae per cell increases dramatically when cells are grown on a solid surface compared to planktonically (Thoma et al. 2008). As with all reported non-type IV appendages, their identification is in an organism which lacks genetic systems, severely limiting their study.
Figure 2.2: Scanning electron micrograph showing a network of *Pyrodictium* cells and cannulae.

Bar = 2 µm. (From Jarrell et al. 2013).
Figure 2.3: Scanning electron micrograph of Mth60 fimbriae of *M. thermautotrophicus*.

Cells grown on gold EM grids, shown to express many Mth60 fimbriae. Bar = 1µm. (From Jarrell et al. 2013).
Type IV pili-like structures

Apart from the unique appendages mentioned above, several other cell surface structures exist in archaea that are much more widely distributed in both major phyla (Crenarchaeota and Euryarchaeota) and assembled using the bacterial type IV pili model (Jarrell et al. 2009, 2013; Lassak et al. 2012a; Ng et al. 2008; Pohlschroder et al. 2011). These include archaella (formerly archaeal flagella: Jarrell and Albers 2012) (Ghosh and Albers 2011; Lassak et al. 2012b; Ng et al. 2006), type IV-like pili (Albers and Pohlschroder 2009; Frols et al. 2008; Henche et al. 2012a; Ng et al. 2011; Szabo et al. 2007), the bindosome identified in Sulfolobus solfataricus (Albers et al. 1999; Zolghadr et al. 2007, 2011) and the Iho670 fibers of Ignicoccus hospitalis (Yu et al. 2012).

Bacterial Type IV pili characteristics

Type IV pili are common appendages found on many Gram negative bacteria (Burrows 2012) and more recently Gram Positive bacteria (Melville and Craig 2013), including many pathogens where the structure is known to play a role in virulence (Burrows 2012; Salomonsson et al. 2012). Distinguishing features of type IV pili in bacteria include structural proteins that are synthesized initially with class 3 signal peptides (Pohlschroder et al. 2005), a prepilin peptidase (or class three signal peptidase) responsible for the cleaving of these signal peptides and generally two ATPases that supply the energy to incorporate new subunits into the base of the structure (polymerizing ATPase) or remove subunits from the base (depolymerizing ATPase). The alternating action of these two ATPases leads to pili extension and retraction and results in pulling the cell forward on solid surfaces in a translocation mechanism termed twitching motility (Burrows 2012; Bradley 1980). These ATPases interact with a conserved integral membrane protein that acts as a scaffold for the assembly system (Peabody et al. 2003). The prepilin peptidases are a novel type of aspartic acid protease (Lapointe and Taylor 2000), distinct from
the other two types of signal peptidases present in most cells (signal peptidase I and II). Their substrates are limited to proteins with class III signal peptides. Cleavage of the class III signal peptide by prepilin peptidases occurs on the cytoplasmic side of the cytoplasmic membrane (Strom and Lory 1991). In bacteria, the prepilin peptidase can be a bifunctional enzyme, responsible for the removal of the signal peptide from the prepilins as well as the N-methylation of the terminal amino acid (Phe) of the mature protein (Nunn and Lory. 1991; Strom et al. 1993). Class III signal peptides typically end with a glycine residue and are characterized by a prepilin peptidase cleavage site that occurs before the hydrophobic stretch of amino acids found in class I and II signal peptides (Pohlschroder et al. 2005; Giltner et al. 2012).

There are two major classes of type IV pili in bacteria, type IVa (T4a) and type IV b (T4b). T4a pili are a broadly distributed subtype found widely in many plant and animal pathogens including Pseudomonas and Neisseria as well as environmental organisms such as Thermus (Giltner et al. 2012) and are considered relatively homogeneous (Pellicic 2008) The T4b class is a more heterogeneous group best studied in enteric organisms like Vibrio and enteropathogenic E. coli. The T4a and T4b classes are distinguished on the basis of the lengths and the sequences of signal peptides in their major pilins (Giltner et al. 2012; Pohlschroder et al. 2011). The T4a pilins have a shorter signal peptide, typically only 6 or 7 residues and these pilins are known to have a methylated phenylalanine at the N-terminus of the mature protein. The T4b pilins have a longer signal peptide ranging from 15 to 30 residues and the N-terminus of the mature pilin can contain other hydrophobic residues. The size of the mature protein is also considered a feature for distinguishing the two types of pilins. T4a pilins typically have a smaller mature protein length (150-175 residues) compared to the T4b pili system (180-200 residues). The T4b pili class includes a subtype called the tight adherence pili or Tad pili, also called Flp and Fap pili, first identified in Aggregatibacter actinomycetemcomitans (Kachlany et al. 2001). The Tad mature pilins are significantly smaller than the others type IV pilins at only
50-80 residues (Giltner et al. 2012; Pelicic 2008). Most T4b and Tad pili are known to lack the retraction ATPase and consequently are not known to be associated with motility; the bundle forming pilus of enteropathogenic *E. coli* and the T4b system in *Pseudomonas aeruginosa* are exceptions (Milgotina and Donnenberg 2009; Carter et al. 2010). T4a pili are always associated with twitching motility and the more recently discovered other forms of motility using T4a pili includes swarming, walking, and slingshot motilities (Burrows 2012).

**(i) Bindosome**

Several sugar-binding proteins including GlcS and AraS in *S. solfataricus* are known to be cleaved by the type IV prepilin-like signal peptidase PibD and assembled into a cell surface structure that aids in sugar uptake (Albers et al. 1999, Elferink et al. 2001). The GlcS and the AraS are expressed on the cell surface by the bindosome assembly system (Bas), a type IV pilus-like assembly system. It consists of three pilin-like proteins with class 3 signal peptides (BasABC), a PilT-like ATPase (BasE) and a PilC-like integral membrane protein (BasF) (Zolghadr et al. 2007). Deletion of these genes leads to cells defective in growth on sugars transported by the sugar-binding proteins (Zolghadr et al. 2007). Initially, it was thought that the genes in the Bas system encoded for a pilus-like structure, due to the similarity of the Bas system to the type IV pilus assembly system. Later, it was found that the sugar-binding structure is essential for the normal appearance of the envelope in *S. solfataricus* (Zolghadr et al. 2011). Abnormal cell morphology and S-layer architecture was observed in a basEF deletion strain indicating a role for the sugar-binding proteins along with the S-layer in maintaining proper cell shape. Thus, it was established that the sugar-binding proteins are a functional part of the S-layer in *S. solfataricus* (Zolghadr et al. 2011). Actual electron microscopic visualization of the putative pilus-like structure composed of sugar-binding proteins is lacking, possibly due to its
short length. While studies have been conducted only in *S. solfataricus*, *in silico* analyses indicate the presence of bindosomes may be widespread in archaea (Szabo et al. 2007).

(ii) Iho670 fibers

Iho670 fibres are novel structures found in *I. hospitalis*, which function as adhesive filaments (Figure 2.4). Iho670 fibres are so called because they are comprised of the Igni_0670 gene product, a 33 kDa protein (Muller et al. 2009). The diameter of this structure is reported to be 14 nm with a length up to 20 µm (Muller et al. 2009). The hydrophobic amino terminus of the Iho670 protein resembles that of type IV pilin-like proteins. Apart from this, the Iho670 protein shows no primary sequence similarity to any other known structural protein in archaea, including archaellins, the Mth60 fimbrin of *M. thermautotrophicus*, the hamus protein or the three cannulae proteins of *Pyrodictium* (Muller et al. 2009). N-terminal sequencing of the mature protein revealed the absence of the first 7 amino acids predicted from the gene sequence, indicating that the Iho670 protein was initially made with a short, type IV pilin-like signal peptide (Muller et al. 2009). The fiber protein also has a potential site for N-glycosylation and a putative N-oligosaccharyltransferase (Igni_0016) was identified in the *I. hospitalis* genome (Magidovich and Eichler 2009). Structural analysis revealed that the conserved N-terminal type IV pilin-like α-helices form the core of the Iho670 fiber (Yu et al. 2012). Though the overall helical symmetry has similarity to that of the archaella filaments of *H. salinarum* (Trachtenberg et al. 2005) and *S. shibatae* (Cohen-Krausz and Trachtenberg 2008), the quaternary structure is unique (Yu et al. 2012).
Figure 2.4: Numerous lho670 fibers emerging from the surface of *Ignicoccus hospitalis* cells. The three cells are on the carbon support film showing the thin and long lho670 fibers emerging from them. Bar = 2 µm. (From Jarrell et al. 2013).
(iii) Archaella

To date, the best studied archaeal surface appendage is the archaellum. The proposal for the renaming of this appendage from archaeal flagellum sought to address possible confusion concerning the motility structures of the two prokaryotic domains (Jarrell and Albers 2012). While both are rotating structures (Alam and Oesterhelt 1984; Shahapure et al. 2014), they are not evolutionarily related and do not share any conserved proteins (Faguy and Jarrell 1999; Nutsch et al. 2005). Indeed, the archaellum bears many similarities instead to bacterial type IV pili in structure and likely assembly, including the presence of several conserved proteins involved in assembly (Ghosh and Albers 2011; Jarrell et al. 2009; Ng et al. 2006; Thomas et al. 2001). The major structural proteins, the archaellins, have signal peptides similar to that seen in bacterial type IV pilins (Faguy et al. 1994). The signal peptides are processed by a pre-archaellin peptidase (FlaK/PibD) that is homologous to the bacterial type IV prepilin peptidase (Bardy and Jarrell 2002; Albers et al. 2003). FlaI is a homologue of the polymerizing ATPase essential for the type IV pili assembly system (Bayley and Jarrell 1998), and FlaJ is a type IV pilus conserved membrane (platform) protein homologue (Ng et al. 2006; Peabody et al. 2003). Archaella are usually 10-14 nm in diameter, typically thicker than archaeal pili (7-8 nm) but thinner than most bacterial flagella (18-22 nm) (Jarrell et al. 2009; Jones and Aizawa 1991; Thomas et al. 2001). A M. maripaludis cell showing both archaella and pili is shown in Figure 2.5. The genes involved in archaellation are clustered in the fla operon which usually begins with several archaellin genes (Ng et al. 2006). The fla operon generally includes other fla-associated genes (flaC to flaJ) as a complete set (Figure 2.6). While several fla genes have similarities to ones in bacterial type IV pili systems (i.e., flaI, flaJ and flaK), the others are found uniquely in archaea (Ng et al. 2006). Deletion analysis of fla operon genes conducted in several different species shows that all fla genes so far successfully deleted are essential for archaellation (Chaban et al. 2007; Lassak et al. 2012b; Patenge et al. 2001; Thomas et al. 2001, 2002). Similar to the situation in the bacterial type IV pili systems (Giltner et al. 2012;
Figure 2.5: Electron micrograph of *M. maripaludis* surface showing thin pill (arrows) with thicker and more numerous archaella. Bar = 0.5 µm. (From Jarrell et al. 2013).
Figure 2.6. Comparison of the fla operons of selected archaea. Homologues are indicated by similar colors. The arrows indicate the direction of transcription.
other genes essential for archaellation are located outside the *fla* operon. Examples include *flaK* which codes for the signal peptidase responsible for processing archaellins (Vandyke et al. 2009; Bardy et al. 2002, 2003) as well as various *agl* (archaeal glycosylation) genes, essential for biosynthesis and attachment of N-glycans that are necessary for assembly of archaella (Chaban et al. 2006b; Jarrell et al. 2010; Vandyke et al. 2009; Tripepi et al. 2012).

The number of archaellin genes varies depending on the archaeal species and can range from 1-6 (Jarrell et al. 2009). The distribution of the archaellins in the genome differs in various archaea. In the case of *Thermococcus kodakarensis*, there are five archaellin genes co-transcribed with the *fla* accessory genes (Nagahisa et al. 1999). In *Halobacterium salinarum*, three archaellin genes are located adjacent to the *fla* accessory genes but orientated in the opposite direction while two other archaellin genes are located elsewhere on the chromosome (Gerl and Sumper 1988; Gerl et al. 1989). In *M. maripaludis*, there are three archaellins co-transcribed with the *fla* accessory genes (Thomas et al. 2001). The major archaellins are FlaB1 and FlaB2, while FlaB3 is responsible for the hook region of the archaella (Chaban et al. 2007). A *flaB3* deletion mutant has functional, but hookless, archaella (Chaban et al. 2007). In *M. maripaludis* a complete set of *fla* associated genes from *flaC* to *flaJ* is found and all the *fla* genes that could be successfully deleted were found to be essential for archaellation; however, *flaD* and *flaE* deletions could not be created (Chaban et al. 2007). Flal, the ATPase essential for archaellation has been particularly well studied in the crenarchaeote *S. acidocaldarius*. Flal was shown to exhibit ATP-dependent oligomerization into a hexamer (Ghosh et al. 2011) and deletion analysis demonstrated its essential role in archaella assembly in *S. acidocaldarius* (Reindl et al. 2013), as shown previously for the euryarchaeotes *M. maripaludis* (Chaban et al. 2007), *M. voltae* (Thomas et al. 2002) and *H. salinarum* (Patenge et al. 2001). Importantly, and uniquely, Flal was shown to have two crucial roles in archaella function. Not only is it required
for archaella assembly but it is also necessary for archaella rotation (Reindl et al. 2013), thus explaining the earlier observation that archaella rotation is driven by ATP hydrolysis and not the proton motive force (Streif et al. 2008). The N-terminal domain was shown to be essential for archaella assembly and its proper localization but not for ATPase activity. Furthermore, the N-terminal 29 amino acids were demonstrated to be necessary for motility but not archaellar assembly (Reindl et al. 2013). This was the first instance of a separation of two functions of FlaI.

Despite the fact that flaC was shown to be essential for archaellation in *M. maripaludis* (Chaban et al. 2007), the flaCDE genes are all missing in Crenarchaeota, while a novel gene, not found in Euryarchaeota, flaX, is present and shown to be essential for archaellum assembly in *S. acidocaldarius* (Lassak et al. 2012b). A recent report showed that FlaI and FlaH interact with FlaX and that FlaX may act as the priming protein for archaellum assembly. Together these proteins are proposed to form the motor complex of the *S. acidocaldarius* archaellum (Banerjee et al. 2013).

As mentioned above, FlaK and PibD are the prepilin peptidases initially identified as involved in class 3 signal peptide processing of type IV pilin-like proteins (Albers et al. 2003; Bardy and Jarrell 2002, 2003). In *M. maripaludis*, FlaK processes only archaellins and a second, novel, prepilin-like peptidase, EppA, processes type IV pilins (Szabo et al. 2007). This is not the case for other studied archaea, such as *Sulfolobus* species and *Hfx. volcanii* where a single enzyme, PibD, processes all type IV pilin-like proteins (Albers et al. 2003; Tripepi et al. 2010). Site-directed mutagenesis of both FlaK and PibD identified them as members of an unusual family of aspartic acid proteases that included bacterial prepilin peptidases (Lapointe and Taylor 2000). Two conserved aspartic acid residues were shown to be critical for the signal peptidase activity of both enzymes (Bardy and Jarrell 2003; Hu et al. 2011; Szabo et al. 2006). The processing of archaellins is an essential step in the assembly of the filament. Mutants of both *M. voltae* and *M. maripaludis* carrying deletions of flaK were nonarchaellated (Bardy and Jarrell
Archaellins generally have a short signal peptide with a length of only 6-12 amino acids (Ng et al. 2006). Various amino acid positions in and around the signal peptide were found to be important for the cleavage of the signal peptide by FlaK. As found in the case of bacterial type IV pilins, changes to the highly conserved -1 position (relative to the cleavage site) glycine and -2 basic amino acid were generally not tolerated while the conserved +3 glycine of the mature protein was also found to be essential (Thomas et al. 2001). Similar studies were conducted in S. solfataricus, but as might be expected by its broader substrate range, PibD was not as restrictive as FlaK in accepting changes at conserved amino acid positions in the signal peptide, although site directed changes to the conserved -1 glycine were the least tolerated (Albers et al. 2003).

Recently, Hu et al (2011) solved the crystal structure of FlaK. The structure demonstrated that FlaK has six transmembrane helices, as predicted earlier (Bardy and Jarrell 2003). The conserved GXGD motif and a short transmembrane helix (helix 4) were found positioned at the centre and surrounded by the remaining transmembrane helices. This positioning showed that a conformational change is essential to bring into close proximity the GXGD motif and an essential aspartic acid residue from transmembrane helix 1 for catalysis (Hu et al. 2011).

The archaellum is a widespread motility organelle found throughout the domain archaea and is known to be responsible for swimming and swarming (Lassak et al. 2012b). However, in addition to this function, there are several reports which indicate further roles for archaella in several archaea. These include adhesion (Jarrell et al. 2011a; Nather et al. 2006; Zolghadr et al. 2010) although at least in M. maripaludis this required the presence of pili as well (Jarrell et al. 2011a). In S. acidocaldarius, S. solfataricus and S. tokodaii, biofilm formation has been reported (Koerdt et al. 2010, 2011). Though not involved in the development of the biofilm, the archaella
help the cells to detach during the dispersal phase (Henche et al. 2012b). In *P. furiosus*, archaella were proposed to be swimming organelles, but also to aid in attachment of cells to abiotic surfaces and to form cell to cell contacts through aggregated archaella forming cable-like structures (Nather et al. 2006).

(a) Glycosylation

Until 1976, it was thought that glycosylation was a posttranslational modification unique to the proteins of eukaryotes (Eichler 2013). However, Mescher and Strominger in 1976 demonstrated that the S-layer protein of the halophilic archaeon, *Halobacterium salinarum*, was a glycoprotein (Mescher and Strominger 1976, 1978). It is now known that representatives of all three domains of life are able to glycosylate their proteins, although among the prokaryotes the N-linked glycosylation pathway appears most widespread in archaea (Abu-Qarn et al. 2007; Kaminski et al. 2013; Nothaft and Szymanski 2010). The N-glycosylation pathways in the three domains show a number of common features coupled with domain-specific modifications (Dell et al. 2010; Jarrell et al. 2010). The archaea share the glycosylation consensus sequon N-x-S/T with Eukarya, where (x ≠ P) and not the longer extended bacterial version (D/E-z-N-x-S/T, where z, x ≠ P). The lipid carrier upon which the glycan is initially assembled is undecaprenol di-phosphate in Bacteria (Linton et al. 2005) while both Eukarya and archaea utilize a dolichol derivative (Guan et al. 2010, 2011; Larkin et al. 2013). Uniquely, archaea employ a much greater diversity in the nature of the sugar that links the glycan to the protein (Jarrell et al. 2010; Jarrell et al. 2014).
Among the archaea, the N-linked glycosylation pathway has been well studied in *Hfx. volcanii*, *M. maripaludis* and to a lesser extent in *S. acidocaldarius* (Eichler 2013; Eichler et al. 2013; Jarrell et al. 2014; Meyer and Albers 2013).

The first archaeal genes proven to be involved in the N-glycosylation pathway were reported in *M. voltae* (Chaban et al. 2006b) where the *agl* (archaeal glycosylation) nomenclature since adopted by the community was first proposed. In this organism, a trisaccharide is N-linked to multiple sites in the four archaellins that make up the archaellar filament, as well as to the S-layer protein (Chaban et al. 2009; Kelly et al. 2009). The trisaccharide was shown by mass spectrometry studies to be comprised of the linking sugar N-acetyl-glucosamine with the second sugar being a di-acetylated glucuronic acid and a terminal sugar that was an acetylated mannuronic acid with an attached threonine. A tetrasaccharide was later reported in other strains of *M. voltae* that was identical to this trisaccharide but with an additional, unidentified, component of 220 or 260 Da (Chaban et al. 2009). However, advances in genetic systems for a closely related methanogen, *M. maripaludis* (Leigh et al. 2011; Moore and Leigh. 2005; Tumbula and Whitman 1999), led to a switch in the study of the N-glycosylation pathway of methanogens (Figure 2.7) to this model organism from *M. voltae*.

The structure of the N-linked archaellin glycan in *M. maripaludis* was found to be related to, but different from, that seen in *M. voltae* (Kelly et al. 2009). The N-linked glycan of *M. maripaludis* was determined to be a tetrasaccharide in which the linking sugar was N-acetyl-galactosamine, not N-acetyl-glucosamine as in *M. voltae*. While the second sugar was the same di-acetylated glucuronic acid found in the *M. voltae* glycan, the third sugar was slightly different. While both are modified N-acetylmannuronic acid with attached threonine, the *M. maripaludis* version has a 3-acetamidino modification not found in the *M. voltae* structure (Kelly et al. 2009).
Figure 2.7. A proposed model of archaellin glycosylation in *M. maripaludis*. Shown are the steps for which various known *agl* gene products act in the biosynthesis and assembly of the N-glycan and the steps for which Agl protein involvement remains to be identified.
Finally, the *M. maripaludis* glycan contained a fourth sugar that has not been identified elsewhere in nature: (5S)-2-acetamido-2,4-dideoxy-5-O-α-l-erythro-hexos-5-ulo-1,5-pyranose (Kelly et al. 2009). In both *M. voltae* and *M. maripaludis*, AglB is the oligosaccharyltransferase that transfers the assembled lipid-bound glycan to the target proteins and AglA is the glycosyltransferase that adds a similar third sugar to the di-N-acetyl-glucuronic acid located at position two of the glycan. The remaining glycosyltransferases are different in the two species reflecting the differences in glycan structures. In *M. maripaludis*, AglO and AglL have been shown to be the enzymes responsible for transfer of the second and fourth sugars onto the lipid carrier (VanDyke et al. 2009). In-frame deletion of these *agl* genes led to a truncated glycan which significantly affected the assembly and function of archaella (VanDyke et al. 2009). Electron microscopy and mass spectrometry analyses in various *agl* gene deletion strains led to the conclusion that a minimum glycan size of two sugars attached to the archaellins was necessary for their assembly into archaella (VanDyke et al. 2009).

Additional studies identified a number of genes involved in the biosynthesis pathways for the sugars involved in the glycan structure (Ding et al. 2013; Jones et al. 2012). Recently, the *aglXYZ* gene cluster was shown to be responsible for the acetamidino group of the third sugar (Jones et al. 2012). Another operon consisting of four genes (*aglU, aglV, mmp1086* and *mmp1087*), located between *aglZ* and *aglL*, contained the putative threonine transferase for the third sugar (*aglU*) and the methyltransferase (*aglV*) which transfers the methyl group to the second hydroxyl group on the C-5 position in the terminal sugar (Ding et al. 2013). Earlier studies had identified a putative acetyltransferase (MMP0350) responsible for attaching one of the acetyl groups to the di-acetylated second sugar (VanDyke et al. 2008). A mutant deleted for *mmp0350* possessed archaellins with a single sugar attached, resulting in non-archaellated cells. At the same time, the pili on the cell were also affected by the deletion of this gene. In ∆*mmp0350* mutants, pili were made but seemingly weakly attached to the cell surface, leading
to poorly piliated cells and resulting in pili being shed into the medium (VanDyke et al. 2008). Complementation of the mmp0350 deletion with a plasmid-borne wildtype copy of the gene led to the return of normally piliated cells. These results suggested the importance of glycosylation in the anchoring of the pili rather than in the assembly of the structure, as seen in archaella (VanDyke et al. 2008).

In vitro studies conducted by Namboori and Graham (2008) on heterologously expressed and purified enzymes have also contributed important information to the formation of acetamido sugars in *M. maripaludis*. Their results established that the euryarchaeal hexosamine biosynthesis pathway shows similarities to the bacterial biosynthesis pathway rather than to the eukaryotic pathway. In step 1, Fructose-6-phosphate (Fru-6-P) is isomerized into GlcN-6-P (α-D-glucosamine-6-phosphate) by an isomerizing glutamine-Fru-6-P transaminase MMP1680 (GlmS). The phosphate group of GlcN-6-P is transferred to C-1 position to form GlcN-1-P (α-D-glucosamine-1-phosphate) by the phosphoglucosamine mutase activity of MMP1077 (GlmM). MMP1076 (GlmU) catalyzes the next two steps, namely the acetylation of the 2-amino group of GlcN-1-P to form GlcNAc-1-P and its subsequent transfer to UTP to form UDP-GlcNAc. The UDP-GlcNAc 2-epimerase activity of MMP0705 converts UDP-GlcNAc into UDP-ManNAc (UDP-N-acetylmannosamine). In the final step, formation of UDP-ManNAcA (UDP-N-acetylmannosaminuronate) is catalyzed by MMP0706, a UDP-ManNAc 6-dehydrogenase (Namboori and Graham 2008).

(iv) Type IV-like Pili

Specific reports on archaeal pili are relatively recent in the literature, although these structures were likely observed decades ago in electron microscopic studies of various archaea (Doddema et al. 1979; Weiss 1973). Most of the pili reported to date in archaea are type IV pili-like, apart from the Mth60 fimbriae of *M. thermotrophicus* (see above) (Thoma et al. 2008).
Many different functions are associated with bacterial type IV pili, such as adhesion, twitching motility, DNA uptake, and biofilm formation (Burrows 2012; Pelicic 2008) as well as more specific roles as in the attachment of *Bdellovibrio bacteriovorus* to prey cells (Mahmoud and Koval 2010) and in electron transfer (the so-called nanowires of *Geobacter* species) (Malvankar and Lovley 2014). Reports indicate that archaeal type IV-like pili are also involved in various functions including adhesion, cell aggregation, biofilm formation, and DNA exchange (Ajon et al. 2011, Frols et al. 2008; Jarrell et al. 2013; Lassak et al. 2012). Recently, dedicated studies on the structure (Wang et al. 2008; Henche et al. 2012a), assembly (Henche et al. 2012a), genetics (Ng et al. 2011; Van Wolderen et al. 2013; Esquivel et al. 2013) and function (Jarrell et al. 2011a; Ajon et al. 2011) of archaeal type IV-like pili have been reported in both crenarchaeota and euryarchaeota.

**(a) Adhesive (Aap) pili**

The presence of Aap pili has only been reported in *S. acidocaldarius* where these organelles are known to play a role in the adherence of the cells to surfaces (Hench et al. 2012a). The Aap pili locus consists of five genes (Figure 2.8A), including two pilin-like genes *aapA* and *aapB*. The proteins encoded by these genes were shown to be processed by PibD. Genes encoding a type IV pilus polymerizing ATPase (*aapE*) and conserved integral membrane protein (*aapF*) have also been identified in the gene cluster. An iron-sulfur oxidoreductase-encoding gene, *aapX*, was also present in the locus. Analysis of deletion mutants revealed that all five genes were essential for piliation (Hench et al. 2012a). Aap pili are extremely stable, likely an adaptation of the cells to the harsh environment in which the *S. acidocaldarius* is found (Hench et al. 2012a). The pilus has a diameter of 11 nm and a unique structure not found in any bacterial or archaeal pilus with a three stranded helical structure with a rotation of 138° and a rise per subunit of 5.7 (Hench et al. 2012a).
Figure 2.8. Type IV pili-like loci found in three different Archaea. Functionally similar genes are given identical colors. (A) The adhesive pili locus of *S. acidocaldarius* showing the pilin-like genes *aapA/B*, the iron sulfur oxidoreductase (*aapX*), the ATPase (*aapE*) and the inner membrane protein (*aapF*). (B) The UV-inducible pili locus of *S. solfataricus* contains genes for two pilins (*upsA* and *upsB*), an ATPase (*upsE*) and a conserved membrane protein (*upsF*) as well as a protein of unknown function (*upsX*). (C) The adhesive pili locus found in *M. maripaludis* contains genes for three minor pilins (*epdA,B,C*), a prepilin peptidase (*eppA*) and a number of proteins of unknown function. Located outside this locus are genes for the ATPase (MMP0040), inner membrane proteins (MMP0038 and MMP0039) and the major pilin gene (MMP1685).
Recently, it was also observed that the expression of Aap pili in S. acidocaldarius is dependent on the growth phase of the cells (Henche et al. 2012a). In stationary phase, cells expressed less Aap pili compared to the exponential phase. This was consistent with the lower transcript level of aap genes in stationary phase compared to that seen in exponential phase. In contrast to Aap pili, archaella numbers are up-regulated in the stationary phase (Henche et al. 2012a). The expression of archaella was increased in all mutants carrying deletions in aap genes, particularly in the aapF gene deletion strain, which was hyper-archaellated (Hench et al. 2012a, Lassak et al. 2012b). It was also demonstrated that overexpression of an archaellation repressor (ArnA) led to increased production of Aap pili (Reimann et al. 2012). These data suggests that there is an interplay of regulation between the Aap pili and archaella.

In S. acidocaldarius, Aap pili are involved in adhesion while the presence of Ups pili has little effect (Henche et al. 2012b). Aap pili are also known to influence biofilm formation in S. acidocaldarius (Henche et al. 2012a; 2012b). Strains unable to make Aap pili had biofilms that were flat and denser compared to the wildtype cells and these strains were also known to lack the tower structures observed in the wildtype biofilms (Hench et al. 2012b).

b) UV-induced (Ups) pili

All Sulfolobus species studied so far (S. acidocaldarius, S. solfataricus and S. tokodaii) are known to express UV-inducible (Ups) pili (Frols et al. 2008; Szabo et al. 2007). Ups pili have a diameter of about 10 nm with variable lengths (Frols et al. 2008; Szabo et al. 2007). They are encoded by five ups genes: upsA, upsB, upsE, upsF and upsX (Figure 2.8B). Recently, a molecular analysis of the ups genes established that the genes were co-transcribed and that UV light strongly upregulated transcription of the ups operon (Frols et al. 2008; Henche et al. 2012a). Deletion analysis of the genes encoding the prepilin-like proteins UpsA and UpsB suggested that they were both major pilin proteins and that mutants deleted for either gene
could still produce pili although only in small numbers (Van Wolferen et al. 2013). Deletion of upsE, encoding the ATPase and of upsF, encoding the conserved integral membrane protein, resulted in nonpiliated cells, indicating their essential role in pili formation (Van Wolferen et al. 2013). Deletion of upsX did not have any effect on piliation or aggregation mediated by pili but the data indicated UpsX was needed for efficient DNA transfer following UV treatment and aggregation (Van Wolferen et al. 2013). UpsB was experimentally shown to be processed by PibD (Frols et al. 2008). Since PibD is the only prepilin-like peptidase found in the organism, it is likely that UpsA is also processed by PibD. Ups pili help the cells to aggregate which enhances DNA exchange and repair, thus overcoming the DNA damage caused by UV exposure or chemicals like bleomycin. In addition, Ups pili are known to be responsible for species-specific recognition in aggregation (Ajon et al. 2011). The structure of the Ups pilus consisted of three helices evenly spaced with a pitch of 15.5 nm (Ajon et al. 2011, Frols et al. 2008).

Recently, it was shown that Sa-Lrp, a transcriptional regulator that belongs to the leucine-responsive regulatory protein (Lrp)-like family, is involved in the transcriptional regulation of the Ups pili in S. acidocaldarius. A strain deleted for the Sa-Lrp gene was defective in the cell aggregation that normally occurs as a result of UV-irradiation. Results from qRT-PCR showed that after UV induction, the transcript level of upsA was significantly lower in a Sa-Lrp gene deletion strain compared to that of the wildtype cells, resulting in a lower production of the UV-induced pili (Vassart et al. 2013).

In Sulfolobus, the Ups pili result in cell aggregation and chromosomal DNA exchange that helps to overcome the UV-induced DNA damage (Frols et al. 2008; Frols et al. 2007). Mutants unable to produce Ups pili do not exchange chromosomal DNA and thus have a lower survival rate after UV treatment compared to wildtype cells. Though all known species of Sulfolobus are capable of making UV-induced pili, aggregation can only happen between
members of the same species, suggesting a specific recognition used by the pili to differentiate the cell surface (Frols et al. 2008). Ups pili are also known to aid in the surface attachment of the cells, but their involvement varies in the different species. In *S. solfataricus*, Ups pili are essential for attachment along with the archaellum, and the absence of either one of the appendages leads to inefficient attachment. Additionally, Ups pili might play a role in biofilm maturation (Koerdt et al. 2010, Zolghadr et al. 2010). The qRT-PCR analysis of adherent cells revealed an upregulation in transcription of the Ups pilin genes (*upsA* and *upsB*) and a decreased transcription of the archaellin gene, *flaB*, suggesting that the role of archaella here may be in the initiation of the attachment with pili being involved in the persistence of the attached cells on the surface.

**Pili in *Haloferax volcanii***

Previously, it was known that non-archaellated mutants in *Hfx. volcanii* adhere to glass cover slips as effectively as wildtype cells, indicating that archaella in this species are not essential, at least under laboratory conditions, for the attachment of cells (Tripepi et al. 2010). Meanwhile, deletion of the prepilin peptidase *pibD* left the cells unable to attach, suggesting that a PibD-dependent structure, likely pili, was responsible for the observed attachment, (Tripepi et al. 2010).

In support of this hypothesis, a novel adhesive pilus was recently identified in *Hfx. volcanii* (Esquivel et al. 2013). There were a total of six identified pilin proteins (PilA1, PilA2, PilA3, PilA4, PilA5, and PilA6). PibD was responsible for processing of the pilins, suggesting they are type IV-pilin-like. The deletion of any five of the six pilin genes resulted in the assembly of a functional pilus, though the ability of such mutant cells to adhere varied. It was also shown that the complementation of any one of the *pilA* genes to a strain deleted for all *pilA* genes restored piliation and adhesion function to at least some degree.
**Pili in *Methanococcus maripaludis***

Bioinformatic analysis of the genome of *M. maripaludis* identified a type IV pili-like locus consisting of 11 potential genes (*mmp0231-mmp0241*) (Szabo et al. 2007). The locus was shown to include three pilin-like genes (*epdA, epdB* and *epdC*) as well as a gene encoding for a novel prepilin peptidase (*eppA*) (Figure 2.8C). The other genes found in the locus do not have any homologues in either archaeal or bacterial type IV pili systems. Deletion analysis of the three pilin-like genes confirmed a role of these genes in the appearance of surface pili (Ng et al. 2011). Electron microscopic examination of ∆*epdB* and ∆*epdC* mutants showed them to be completely nonpiliated while ∆*epdA* mutant cells were typically underpiliated compared with wildtype cells, usually possessing only one or two pili per cell.

A unique characteristic of the *M. maripaludis* pili system is the presence of a separate prepilin peptidase (EppA) for the processing of prepilins (Szabo et al. 2007) that is distinct from the prepilin-like peptidase (FlaK) that cleaves class III signal peptides from the archaellins (Bardy and Jarrell 2003). EppA is a larger protein than FlaK with an additional four transmembrane segments. The specificity of each of the two signal peptidases was demonstrated by creating genetic hybrids of their substrates, where the archaellin FlaB2 cleavage site (KGAS, positions -2 to +2) was substituted with the cleavage site (RGQI) of the pilin EpdA and vice versa. This study showed that FlaK was able to process the pilin cleavage site substituted in the archaellin, but did not process the archaellin cleavage site substituted in the pilin. This indicated that FlaK substrate recognition and cleavage involved more amino acid positions than just those in the immediate vicinity of the cleavage site. Unlike the case with FlaK, EppA was able to process the hybrid substrates as long as the +1 position was a glutamine, suggesting that the +1 position is critical for substrate recognition and processing by EppA.
The first solved structure of an archaenal pilus was that of the *M. maripaludis* appendage (Wang et al. 2008). The pili in this organism are approximately 6 nm in diameter, peritrichously located and less numerous than archaella (Ng et al. 2011). Cryo-electron microscopy analysis of purified *M. maripaludis* pili showed a different structure from that of any bacterial pili with two different helical symmetries co-existing within the same pilus (Wang et al. 2008). Interestingly, the *M. maripaludis* pili have a central lumen with a diameter of 20Å (Wang et al. 2008) that is missing in archaella structures (Cohen-Krausz and Trachtenberg 2008; Trachtenberg and Cohen-Krausz 2006).

The only known function of *M. maripaludis* pili is in cellular attachment to solid surfaces. However, attachment is co-dependent on the presence on archaella since the absence of either one of the appendages prevents adhesion (Jarrell et al. 2011a). Since the attachment of cells seems to take place via bundles of archaella, it may be that pili only initiate attachment, and that archaella are responsible for prolonged attachment. This is the exact opposite of what is suggested to happen in the case of *S. solfataricus* attachment (Zolghadr et al. 2010). Based on these observations, one of the genes in the type IV pili operon could encode an adhesin protein that would be located at the tip of the pilus and be responsible for adherence of the cell onto surfaces. Twitching motility is a very common function of type IV pili in many bacterial species (Burrows 2012). It is known that twitching motility is facilitated by the presence of two different ATPases, one for extension and the other for retraction of pili (Burrows 2012). To date, there is no evidence in any archaeon, including *M. maripaludis*, of twitching motility.

**Aims of the study**

Analysis of pili systems in archaea is in its infancy, despite the obvious important roles that these structures are likely to play in the interactions of archaenal cells with their environment. The only well studied archaenal pili systems are found in the crenarchaeote *Sulfolobus* species...
while almost no data is available for such structures in any euryarchaeote. With genetic systems available in *M. maripaludis* (Moore and Leigh 2005), an organism considered to be a model organism for archaeal study (Leigh et al. 2011), and extensive data already available for another surface appendage (archaella) in this methanogen (Chaban et al. 2007; Jarrell et al. 2009; Jarrell and Albers 2012), we sought to study the type IV pili with the following specific aims.

1. To delineate the gene region in the type IV pili-like locus in *M. maripaludis* that was involved in pili formation. A type IV pilus locus was predicted for *M. maripaludis* based on bioinformatics analysis (Szabo et al. 2007). That study predicted three pilin-like genes and a prepilin peptidase among the 11 genes found in the locus but no genetic work was done to show that any of these genes were involved in the appearance of the surface pili. Subsequent work by the Jarrell group revealed the involvement of the three pilin-like genes in normal piliation (Ng et al. 2011). Since this gene locus included many genes predicted to code for hypothetical proteins with no homology to known type IV proteins it was necessary to analyze deletions in each gene to confirm a role in pili formation.

2. The same bioinformatical analysis identified numerous other potential type IV pilin-like proteins that could be processed by the pilin-specific peptidase EppA (Szabo et al. 2007). Whether any of these were essential for pili formation was unknown and untested. Therefore, deletion analysis coupled with electron microscopy was performed to identify other essential pili structural components.

3. Type IV pili systems typically contain a readily identifiable, conserved polymerizing ATPase and conserved membrane or platform component. Since these were not evident in the major type IV pilus locus identified by Szabo et al (2007), their identification elsewhere in the
genome was sought by bioinformatics and their involvement in piliation determined by deletion analysis and electron microscopy.

4. It has been well established that archaellins in *M. maripaludis* are glycoproteins with an N-linked tetrasaccharide attached at multiple sites to each of the three archaellins (Kelly et al. 2009) and that the archaellins must also have a signal peptide removed before they can assemble into the archaellar filament (Bardy and Jarrell 2003; Ng et al. 2011). Glycosylation of archaellins can still occur in the absence of signal peptide removal (Bardy and Jarrell 2003). However, there is no information on the order of posttranslational modifications for any archaeal pilin. This problem was addressed through a series of complementation experiments expressing FLAG-tagged pilins in various mutant backgrounds.
Literature cited


Jarrell KF, Stark M, Nair DB, Chong JP (2011a) Flagella and pili are both necessary for efficient attachment of Methanococcus maripaludis to surfaces. FEMS Microbiol Lett 319:44-50


Jarrell KF, Walters AD, Bochiwal C, Borgia JM, Dickinson T, Chong JP (2011b) Major players on the microbial stage: why archaea are important. Microbiology 157:919-936


Kaminski L, Guan Z, Yurist-Doutch S, Eichler J (2013) Two distinct N-glycosylation pathways process the Haloferax volcanii S-layer glycoprotein upon changes in environmental salinity. MBio 4:e00716-13


48


Thomas NA, Chao ED, Jarrell KF (2001) Identification of amino acids in the leader peptide of Methanococcus voltae preflagellin that are important in posttranslational processing. Arch Microbiol 175:263-269


Weiss RL (1973) Attachment of bacteria to sulfur in extreme environments J Gen Microbiol 77:501-507


Chapter 3

Genetic Analysis of a Type IV Pili-like Locus in the Archaeon *Methanococcus maripaludis*

Abstract

*Methanococcus maripaludis* is a stringently anaerobic archaeon with two studied surface structures, archaella and type IV pili. Previously, it was shown that three pilin genes (*mmp0233 [epdA], mmp0236 [epdB] and mmp0237 [epdC]*) located within an 11 gene cluster in the genome were necessary for normal piliation. This study focused on analysis of the remaining genes to determine their potential involvement in piliation. Reverse transcriptase-PCR experiments demonstrated the 11 genes formed a single transcriptional unit. Deletions were made in all the non-pilin genes except *mmp0231*. Electron microscopy revealed that all the genes in the locus except *mmp0235* and *mmp0238* were essential for piliation. Complementation with a plasmid borne wildtype copy of each examined deleted gene restored at least some piliation. We identified genes for an assembly ATPase and two versions of the conserved pilin platform forming protein necessary for pili assembly at a separate genetic locus.
Introduction

Archaea, like Bacteria, are known to possess a variety of cell surface structures (Jarrell et al. 2013, Lassak et al. 2012a, Ng et al. 2008, Pohlschroder et al. 2011). Many of thesearchaeal appendages appear to be made using a system like that employed by bacteria to assemble type IV pili (Jarrell et al. 2009, Jarrell et al. 2013, Lassak et al. 2012a, Ng et al. 2008, Pohlschroder et al. 2011). The core components of such a system include structural subunits with class 3 signal peptides (Pohlschroder et al. 2005) which are removed by a specific signal peptidase (type IV prepilin signal peptidase), an ATPase to incorporate new subunits into the base of the growing structure [sometimes a second ATPase to remove subunits from the structure], and a conserved membrane protein that is thought to interact with the ATPase(s) as an export complex for the structural proteins (Burrows 2012, Pohlschroder et al. 2011, Takhar et al. 2013). These core components of type IV pili systems have been observed in Archaea in the loci responsible for formation of archaella (formerly archaeal flagella; (Jarrell and Albers 2012) (Ghosh and Albers 2011, Lassak et al. 2012b, Ng et al. 2006), type IV-like pili (Albers and Pohlschroder 2009, Frols et al. 2008, Henche et al. 2012a, Ng et al. 2011, Szabo et al. 2006) and the bindosome involved in substrate uptake in Sulfolobus solfataricus (Albers et al. 1999, Zolghadr et al. 2011, Zolghadr et al. 2007). In addition, the Iho670 fibers of Ignicoccus hospitalis are also made from type IV pilin-like proteins (Yu et al. 2012). Presumably, there are genes that encode the other core components but they do not appear to be located in the immediate vicinity of the structural protein gene (Muller et al. 2009). While the archaellum is the best-studied of the archaeal surface structures (Ghosh and Albers 2011, Jarrell et al. 1996, Jarrell and McBride 2008, Thomas et al. 2001) more recently, various studies have targeted the structure, function and genetics of archaeal pili systems, especially ones that are type IV-like (Esquivel et al. 2013, Frols et al. 2008, Henche et al. 2012a, Lassak et al. 2012a, Ng et al.
Among the archaeal type IV-like pili, the majority of progress has been reported in *Sulfolobus* species (Henche et al. 2012a, Orell et al. 2013a, Orell et al. 2013b, van Wolferen et al. 2013) and, to a lesser degree, in *Methanococcus maripaludis* (Jarrell et al. 2011, Nair et al. 2013, Ng et al. 2011, Wang et al. 2008) and *Halofex volcanii* (Esquivel et al. 2013).

In *Sulfolobus acidocaldarius*, there are at least 2 different operons in the genome that are responsible for two unique pili types on the cell surface, namely Aap pili and Ups pili (Lassak et al. 2012a, Pohlschroder et al. 2011). Aap pili, observed on cells during growth under normal laboratory conditions, are involved in adhesion of cells (Henche et al. 2012a). The Ups pili, on the other hand, are observed only when the cells are subjected to DNA damaging conditions like UV exposure (Ajon et al. 2011, Frols et al. 2008). The upregulation of the *ups* operon led to cell aggregation and an enhanced ability of the cells to exchange DNA (Ajon et al. 2011). A recent study showed the importance of both Ups and Aap pili on the structure of the biofilm produced by *S. acidocaldarius* (Henche et al. 2012b, Orell et al. 2013a). A genetic locus containing five genes (encoding two predicted pilins, a type IV pilus polymerizing ATPase, a conserved pilus membrane protein and a putative iron-sulfur oxidoreductase) has been implicated in Aap pili formation (Henche et al. 2012a). Mutants carrying deletions in any one of the five genes do not assemble Aap pili. The *ups* locus consists also of five genes, encoding two prepilins, a hypothetical protein as well as homologues of a type IV pilus polymerization ATPase and conserved membrane protein. No pili were observed if the gene for the ATPase or the conserved membrane protein was deleted (van Wolferen et al. 2013). Mutants carrying deletions of either of the two prepilin genes still made pili but these cells were defective in aggregation (van Wolferen et al. 2013). Mutants carrying a deletion of the gene encoding the hypothetical protein UpsX still made pili but decreased DNA exchange was observed (van Wolferen et al. 2013). Very recently, insights into the regulation of the pili systems of *S.*
*Acidocaldarius* have also been presented (Orell et al. 2013b, Vassart et al. 2012), including studies that show that there is an intertwined regulation of archaella and Aap pili formation (Hence et al. 2012a, Orell et al. 2013b, Reiman et al. 2012).

The type IV pili-like locus in *M. maripaludis* genome was originally predicted to consist of 11 potential genes, with three of them, *mmp0233 (epdA)*, *mmp0236 (epdB)* and *mmp0237 (epdC)* thought to encode structural proteins (pilins) and another shown to be a type IV prepilin peptidase (*mmp0232, EppA*) required to cleave the class 3 signal peptides from the prepilins (Szabo et al. 2007). Subsequent genetic work demonstrated that all three of the genes encoding the predicted structural proteins were essential for normal piliation (Ng et al. 2011). However, mass spectrometry of purified pili revealed that the major structural pilin was MMP1685, encoded by a gene located outside the pilin locus. The deletion and complementation of *mmp1685* showed that it was indeed essential for piliation (Ng et al. 2011). More recently, another minor pilin, encoded by *mmp1283*, was also shown to be essential for pili formation (Nair et al. 2013). An important, and thus far unique, characteristic of the pili system in *M. maripaludis* is the presence of a second, apparently pilin-specific peptidase (EppA, (Szabo et al. 2007)), distinct from the prepilin peptidase, FlaK, (Bardy and Jarrell 2002, Bardy and Jarrell 2003) needed to process archaellins for archaella assembly.

In addition to the genetic studies, the structure of *M. maripaludis* pili was also determined and shown to be different from that of any known bacterial pili (Wang et al. 2008). Furthermore, at least one function for the pili of *M. maripaludis* was demonstrated, that of surface adhesion. However, this function is dependent on the co-expression of archaella (Jarrell et al. 2011).

In this study, we complete the in-frame deletion and complementation study on the genes in the type IV pili-like locus of *M. maripaludis*. In addition, we identify an ATPase (*mmp0040*) as well as two conserved type IV pili membrane component homologues (*mmp0038*...
and mmp0039) critical for pili assembly which are located adjacent to each other but separate from the known pili locus.

Materials and methods

Strains and growth conditions

*Methanococcus maripaludis* (Mm900) (Moore and Leigh 2005) and a Δ*flaK* mutant strain derived from Mm900 (Ng et al. 2009) were grown in Balch medium III (Balch et al. 1979) at 35°C under a headspace of CO₂/H₂ (20:80). McCas medium (Moore and Leigh 2005) was used for transformation experiments with the addition of neomycin (1 mg/ml) or 8-azahypoxanthine (240 μg/ml) for selection at various steps of the procedure. Puromycin (2.5 μg/ml) was used to select for transformants carrying the complementation vectors. For complementation studies, cells were grown in nitrogen free medium (Blank et al. 1995) supplemented with sterile anaerobic solutions of either NH₄Cl (10mM) or alanine (10mM). *Escherichia coli* strain DH5α (Novagen) or *E. coli* TOP10 cells (Invitrogen) used for various cloning steps was grown at 37°C in Luria-Bertani medium, with ampicillin (100 μg/ml) added for selection, when necessary.

Bioinformatic analysis. Each predicted pilus locus gene product was analyzed by a variety of online tools to gather information about its likely subcellular location and possible function. This included programs which predicted archaeal signal peptides (PRED-SIGNAL (http://bioinformatics.biol.uoa.gr/PRED-SIGNAL/ (Bagos et al. 2009)), transmembrane domains (TmPred, http://www.ch.embnet.org/software/TMPRED_form.html (Hofmann and Stoffel 1993)) and subcellular location (PSORTb version 3.0.2 trained to Archaea, http://www.psort.org/psortb/index.html (Yu et al. 2010) as well as ones that searched for conserved motifs (BLAST search (Basic Local Alignment Search Tool,
Plasmid construction to create gene deletions

Plasmids used for the generation of inframe deletions of pilus locus genes were generated as previously described (Moore and Leigh 2005, VanDyke et al. 2008). Briefly, P1 and P2 PCR primers for each gene (Supplemental Table 3.1) were selected to amplify approximately 1kb upstream and the P3 and P4 PCR primers to amplify approximately 1kb downstream of the targeted gene. The P2 and P3 primers were designed so that after ligation a small internal fragment of the targeted gene was left. The gene specific P1 and P4 primers had added BamH1 restriction sites while P2 and P3 had added Asc1 restriction sites. The upstream and downstream PCR products were ligated after digestion with Asc1 and this product was used as template for a further PCR using primers P1 and P4. This approximately 2 kb piece was digested with BamH1 and cloned into pCRPrtNeo to create the plasmids used for generating the deletion strains (Moore and Leigh 2005).

*M. maripaludis* mutant generation

The pCRPrtNeo derivatives carrying deletions of the pilus locus genes (listed in Supplemental Table 3.2) were transformed into *M. maripaludis ΔflaK* using a PEG precipitation method as described by Tumbula et al. (Tumbula et al. 1994). After recovery overnight, the cells were transferred to McCas medium containing neomycin (1mg/ml) to select for transformants that had integrated the vector. This culture was then used to inoculate McCas media without neomycin to allow for a second recombination event to remove the vector. At this stage, a recombination event that removes the inserted plasmid can either return the chromosome to its
Supplemental Table 3.1: Primers used in this study for gene deletion, complementation, site-directed mutagenesis and screening of potential mutants. The restriction sites are underlined (GGATCC – BamH1; GGCGCGCC – Asc1; ATGCAT – Nsi1; ACGGCT – Mu1).

<table>
<thead>
<tr>
<th>GENE</th>
<th>PRIMER</th>
<th>SEQUENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMP0231</td>
<td>P1</td>
<td>5'-CGGATCCGGCCGATCTAGCTTTGCCGCTATGAATTAGCG-3'</td>
</tr>
<tr>
<td></td>
<td>P2</td>
<td>5'-TTGGCCGGCGCTGCTCCGAGAGGATAGA-3'</td>
</tr>
<tr>
<td></td>
<td>P3</td>
<td>5'-TTGGCCGGCGCCATTGGAAGGATATAATGTAATTCACAAGCTTGCAGG-3'</td>
</tr>
<tr>
<td></td>
<td>P4</td>
<td>5'-GGATCCGGTAAGCGACAAAGCACAACGTGGTACG-3'</td>
</tr>
<tr>
<td></td>
<td>Seq-for</td>
<td>5'-TTCTATTATCCTCAGGCG-3'</td>
</tr>
<tr>
<td></td>
<td>Seq-rev</td>
<td>5'-CATCTTAAACTCTCCACCGGC-3'</td>
</tr>
<tr>
<td>MMP0232</td>
<td>P1</td>
<td>5'-GGCGGATCCCATTCTCTATTACTTGCTGTGCG-3'</td>
</tr>
<tr>
<td></td>
<td>P2</td>
<td>5'-TTGGCCGGCGCTGCTCCGAGAGGATAGA-3'</td>
</tr>
<tr>
<td></td>
<td>P3</td>
<td>5'-TTGGCCGGCGCCATTGGAAGGATATAATGTAATTCACAAGCTTGCAGG-3'</td>
</tr>
<tr>
<td></td>
<td>P4</td>
<td>5'-GGATCCGGTAAGCGACAAAGCACAACGTGGTACG-3'</td>
</tr>
<tr>
<td></td>
<td>Seq-for</td>
<td>5'-CTGGAGCTGTAGAATGCAAC-3'</td>
</tr>
<tr>
<td></td>
<td>Seq-rev</td>
<td>5'-GGATGACCTGGGATAATGCAACG-3'</td>
</tr>
<tr>
<td></td>
<td>Comp-for</td>
<td>5'-CCATGCGATTTGTTGAGTTGAATATTTAATT-3'</td>
</tr>
<tr>
<td></td>
<td>Comp-rev</td>
<td>5'-CCAGCTTCCGTTATATCCACGATAATTATA-3'</td>
</tr>
<tr>
<td></td>
<td>Sdm-for</td>
<td>5'-CCTTAATTTGCTCGTGATA-3'</td>
</tr>
<tr>
<td></td>
<td>Sdm-rev</td>
<td>5'-GAATATTTGACGACGGTATTAGG-3'</td>
</tr>
<tr>
<td>MMP0234</td>
<td>P1</td>
<td>5'-GGCGGATCCCATTCTCTATTACTTGCTGTGCG-3'</td>
</tr>
<tr>
<td></td>
<td>P2</td>
<td>5'-TTGGCCGGCGCTGCTCCGAGAGGATAGA-3'</td>
</tr>
<tr>
<td></td>
<td>P3</td>
<td>5'-TTGGCCGGCGCCATTGGAAGGATATAATGTAATTCACAAGCTTGCAGG-3'</td>
</tr>
<tr>
<td></td>
<td>P4</td>
<td>5'-GGATCCGGTAAGCGACAAAGCACAACGTGGTACG-3'</td>
</tr>
<tr>
<td></td>
<td>Seq-for</td>
<td>5'-CCATGCGATTTGTTGAGTTGAATATTTAATT-3'</td>
</tr>
<tr>
<td></td>
<td>Seq-rev</td>
<td>5'-GGATGACCTGGGATAATGCAACG-3'</td>
</tr>
<tr>
<td></td>
<td>Comp-for</td>
<td>5'-CCATGCGATTTGTTGAGTTGAATATTTAATT-3'</td>
</tr>
<tr>
<td></td>
<td>Comp-rev</td>
<td>5'-CCAGCTTCCGTTATATCCACGATAATTATA-3'</td>
</tr>
<tr>
<td></td>
<td>Sdm-for</td>
<td>5'-CCTTAATTTGCTCGTGATA-3'</td>
</tr>
<tr>
<td></td>
<td>Sdm-rev</td>
<td>5'-GAATATTTGACGACGGTATTAGG-3'</td>
</tr>
<tr>
<td>MMP0235</td>
<td>P1</td>
<td>5'-GGCGGATCCCATTCTCTATTACTTGCTGTGCG-3'</td>
</tr>
<tr>
<td></td>
<td>P2</td>
<td>5'-TTGGCCGGCGCTGCTCCGAGAGGATAGA-3'</td>
</tr>
<tr>
<td></td>
<td>P3</td>
<td>5'-TTGGCCGGCGCCATTGGAAGGATATAATGTAATTCACAAGCTTGCAGG-3'</td>
</tr>
<tr>
<td></td>
<td>P4</td>
<td>5'-GGATCCGGTAAGCGACAAAGCACAACGTGGTACG-3'</td>
</tr>
<tr>
<td></td>
<td>Seq-for</td>
<td>5'-GGATGACCTGGGATAATGCAACG-3'</td>
</tr>
<tr>
<td></td>
<td>Seq-rev</td>
<td>5'-GGATGACCTGGGATAATGCAACG-3'</td>
</tr>
<tr>
<td></td>
<td>Comp-for</td>
<td>5'-CCATGCGATTTGTTGAGTTGAATATTTAATT-3'</td>
</tr>
<tr>
<td></td>
<td>Comp-rev</td>
<td>5'-CCAGCTTCCGTTATATCCACGATAATTATA-3'</td>
</tr>
<tr>
<td></td>
<td>Sdm-for</td>
<td>5'-CCTTAATTTGCTCGTGATA-3'</td>
</tr>
<tr>
<td></td>
<td>Sdm-rev</td>
<td>5'-GAATATTTGACGACGGTATTAGG-3'</td>
</tr>
<tr>
<td>MMP0238</td>
<td>P1</td>
<td>5'-GGCGGATCCCATTCTCTATTACTTGCTGTGCG-3'</td>
</tr>
<tr>
<td></td>
<td>P2</td>
<td>5'-TTGGCCGGCGCTGCTCCGAGAGGATAGA-3'</td>
</tr>
<tr>
<td></td>
<td>P3</td>
<td>5'-TTGGCCGGCGCCATTGGAAGGATATAATGTAATTCACAAGCTTGCAGG-3'</td>
</tr>
<tr>
<td></td>
<td>P4</td>
<td>5'-GGATCCGGTAAGCGACAAAGCACAACGTGGTACG-3'</td>
</tr>
<tr>
<td></td>
<td>Seq-for</td>
<td>5'-GGATGACCTGGGATAATGCAACG-3'</td>
</tr>
<tr>
<td></td>
<td>Seq-rev</td>
<td>5'-CCCTTCCGAATCTGCTTCTTCAATACGG-3'</td>
</tr>
<tr>
<td></td>
<td>Comp-for</td>
<td>5'-CCATGCGATTTGTTGAGTTGAATATTTAATT-3'</td>
</tr>
<tr>
<td></td>
<td>Comp-rev</td>
<td>5'-CCAGCTTCCGTTATATCCACGATAATTATA-3'</td>
</tr>
<tr>
<td></td>
<td>Sdm-for</td>
<td>5'-CCTTAATTTGCTCGTGATA-3'</td>
</tr>
<tr>
<td></td>
<td>Sdm-rev</td>
<td>5'-GAATATTTGACGACGGTATTAGG-3'</td>
</tr>
<tr>
<td>MMP0239</td>
<td>P1</td>
<td>5'-GGCGGATCCCATTCTCTATTACTTGCTGTGCG-3'</td>
</tr>
<tr>
<td></td>
<td>P2</td>
<td>5'-TTGGCCGGCGCTGCTCCGAGAGGATAGA-3'</td>
</tr>
<tr>
<td></td>
<td>P3</td>
<td>5'-TTGGCCGGCGCCATTGGAAGGATATAATGTAATTCACAAGCTTGCAGG-3'</td>
</tr>
<tr>
<td></td>
<td>P4</td>
<td>5'-GGATCCGGTAAGCGACAAAGCACAACGTGGTACG-3'</td>
</tr>
<tr>
<td></td>
<td>Seq-for</td>
<td>5'-GGATGACCTGGGATAATGCAACG-3'</td>
</tr>
<tr>
<td></td>
<td>Seq-rev</td>
<td>5'-CCCTTCCGAATCTGCTTCTCAGGG-3'</td>
</tr>
<tr>
<td>MMP0240</td>
<td>P1</td>
<td>5'-GGGATCCCAAGTAGCTTGAATGATTCTG-3'</td>
</tr>
<tr>
<td></td>
<td>P2</td>
<td>5'-TTGGGCGGCGCCTACACTTCTTGCACC-3'</td>
</tr>
<tr>
<td></td>
<td>P3</td>
<td>5'-GGGATCCGACGAGATGATGATTCTTGCACC-3'</td>
</tr>
<tr>
<td></td>
<td>P4</td>
<td>5'-GGGATCCGACGAGATGATGATTCTTGCACC-3'</td>
</tr>
<tr>
<td></td>
<td>Seq-for</td>
<td>5'-CAATATGAGGCTGATGGA-3'</td>
</tr>
<tr>
<td></td>
<td>Seq-rev</td>
<td>5'-CAATATGAGGCTGATGGA-3'</td>
</tr>
<tr>
<td></td>
<td>Comp-for</td>
<td>5'-CAATATGAGGCTGATGGA-3'</td>
</tr>
<tr>
<td></td>
<td>Comp-rev</td>
<td>5'-CAATATGAGGCTGATGGA-3'</td>
</tr>
<tr>
<td>MMP0241</td>
<td>P1</td>
<td>5'-GGGATCCCAAGTAGCTTGAATGATTCTG-3'</td>
</tr>
<tr>
<td></td>
<td>P2</td>
<td>5'-TTGGGCGGCGCCTACACTTCTTGCACC-3'</td>
</tr>
<tr>
<td></td>
<td>P3</td>
<td>5'-GGGATCCGACGAGATGATGATTCTTGCACC-3'</td>
</tr>
<tr>
<td></td>
<td>P4</td>
<td>5'-GGGATCCGACGAGATGATGATTCTTGCACC-3'</td>
</tr>
<tr>
<td></td>
<td>Seq-for</td>
<td>5'-CAATATGAGGCTGATGGA-3'</td>
</tr>
<tr>
<td></td>
<td>Seq-rev</td>
<td>5'-CAATATGAGGCTGATGGA-3'</td>
</tr>
<tr>
<td></td>
<td>Comp-for</td>
<td>5'-CAATATGAGGCTGATGGA-3'</td>
</tr>
<tr>
<td></td>
<td>Comp-rev</td>
<td>5'-CAATATGAGGCTGATGGA-3'</td>
</tr>
<tr>
<td>MMP0040</td>
<td>P1</td>
<td>5'-GGGATCCCAAGTAGCTTGAATGATTCTG-3'</td>
</tr>
<tr>
<td></td>
<td>P2</td>
<td>5'-TTGGGCGGCGCCTACACTTCTTGCACC-3'</td>
</tr>
<tr>
<td></td>
<td>P3</td>
<td>5'-GGGATCCGACGAGATGATGATTCTTGCACC-3'</td>
</tr>
<tr>
<td></td>
<td>P4</td>
<td>5'-GGGATCCGACGAGATGATGATTCTTGCACC-3'</td>
</tr>
<tr>
<td></td>
<td>Seq-for</td>
<td>5'-CAATATGAGGCTGATGGA-3'</td>
</tr>
<tr>
<td></td>
<td>Seq-rev</td>
<td>5'-CAATATGAGGCTGATGGA-3'</td>
</tr>
<tr>
<td></td>
<td>Sdm-for</td>
<td>5'-CAATATGAGGCTGATGGA-3'</td>
</tr>
<tr>
<td></td>
<td>Sdm-rev</td>
<td>5'-CAATATGAGGCTGATGGA-3'</td>
</tr>
<tr>
<td>MMP0281</td>
<td>P1</td>
<td>5'-GGGATCCCAAGTAGCTTGAATGATTCTG-3'</td>
</tr>
<tr>
<td></td>
<td>P2</td>
<td>5'-TTGGGCGGCGCCTACACTTCTTGCACC-3'</td>
</tr>
<tr>
<td></td>
<td>P3</td>
<td>5'-GGGATCCGACGAGATGATGATTCTTGCACC-3'</td>
</tr>
<tr>
<td></td>
<td>P4</td>
<td>5'-GGGATCCGACGAGATGATGATTCTTGCACC-3'</td>
</tr>
<tr>
<td></td>
<td>Seq-for</td>
<td>5'-CAATATGAGGCTGATGGA-3'</td>
</tr>
<tr>
<td></td>
<td>Seq-rev</td>
<td>5'-CAATATGAGGCTGATGGA-3'</td>
</tr>
<tr>
<td></td>
<td>Comp-for</td>
<td>5'-CAATATGAGGCTGATGGA-3'</td>
</tr>
<tr>
<td></td>
<td>Comp-rev</td>
<td>5'-CAATATGAGGCTGATGGA-3'</td>
</tr>
<tr>
<td>MMP0038</td>
<td>P1</td>
<td>5'-GGGATCCCAAGTAGCTTGAATGATTCTG-3'</td>
</tr>
<tr>
<td></td>
<td>P2</td>
<td>5'-TTGGGCGGCGCCTACACTTCTTGCACC-3'</td>
</tr>
<tr>
<td></td>
<td>P3</td>
<td>5'-GGGATCCGACGAGATGATGATTCTTGCACC-3'</td>
</tr>
<tr>
<td></td>
<td>P4</td>
<td>5'-GGGATCCGACGAGATGATGATTCTTGCACC-3'</td>
</tr>
<tr>
<td></td>
<td>Seq-for</td>
<td>5'-CAATATGAGGCTGATGGA-3'</td>
</tr>
<tr>
<td></td>
<td>Seq-rev</td>
<td>5'-CAATATGAGGCTGATGGA-3'</td>
</tr>
<tr>
<td></td>
<td>Comp-for</td>
<td>5'-CAATATGAGGCTGATGGA-3'</td>
</tr>
<tr>
<td></td>
<td>Comp-rev</td>
<td>5'-CAATATGAGGCTGATGGA-3'</td>
</tr>
<tr>
<td>MMP0039</td>
<td>P1</td>
<td>5'-GGGATCCCAAGTAGCTTGAATGATTCTG-3'</td>
</tr>
<tr>
<td></td>
<td>P2</td>
<td>5'-TTGGGCGGCGCCTACACTTCTTGCACC-3'</td>
</tr>
<tr>
<td></td>
<td>P3</td>
<td>5'-GGGATCCGACGAGATGATGATTCTTGCACC-3'</td>
</tr>
<tr>
<td></td>
<td>P4</td>
<td>5'-GGGATCCGACGAGATGATGATTCTTGCACC-3'</td>
</tr>
<tr>
<td></td>
<td>Seq-for</td>
<td>5'-CAATATGAGGCTGATGGA-3'</td>
</tr>
<tr>
<td></td>
<td>Seq-rev</td>
<td>5'-CAATATGAGGCTGATGGA-3'</td>
</tr>
<tr>
<td></td>
<td>Comp-for</td>
<td>5'-CAATATGAGGCTGATGGA-3'</td>
</tr>
<tr>
<td></td>
<td>Comp-rev</td>
<td>5'-CAATATGAGGCTGATGGA-3'</td>
</tr>
</tbody>
</table>
Supplemental Table 3.2: Plasmids used in this study.

<table>
<thead>
<tr>
<th>PLASMIDS</th>
<th>DESCRIPTION/GENOTYPE</th>
<th>SOURCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>pCRPrtNeo</td>
<td><em>hmv</em> promoter-<em>hpt</em> fusion plus 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>pKJ833</td>
<td>pCRPrtNeo with in-frame deletion of <em>mmp0231</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKJ697</td>
<td>pCRPrtNeo with in-frame deletion of <em>mmp0232</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKJ774</td>
<td>pCRPrtNeo with in-frame deletion of <em>mmp0234</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKJ753</td>
<td>pCRPrtNeo with in-frame deletion of <em>mmp0235</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKJ829</td>
<td>pCRPrtNeo with in-frame deletion of <em>mmp0238</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKJ830</td>
<td>pCRPrtNeo with in-frame deletion of <em>mmp0239</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKJ832</td>
<td>pCRPrtNeo with in-frame deletion of <em>mmp0240</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKJ834</td>
<td>pCRPrtNeo with in-frame deletion of <em>mmp0241</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKJ714</td>
<td>pCRPrtNeo with in-frame deletion of <em>mmp0040</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKJ787</td>
<td>pCRPrtNeo with in-frame deletion of <em>mmp0281</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKJ1011</td>
<td>pCRPrtNeo with in-frame deletion of <em>mmp0038</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKJ1010</td>
<td>pCRPrtNeo with in-frame deletion of <em>mmp0039</em></td>
<td>This study</td>
</tr>
<tr>
<td>pHW40</td>
<td><em>nif</em> promoter-<em>lacZ</em> fusion plus Pur cassette; Amp&lt;sup&gt;r&lt;/sup&gt;</td>
<td>John Leigh</td>
</tr>
<tr>
<td>pKJ711</td>
<td>pHW40 with <em>mmp0232</em> complement</td>
<td>This study</td>
</tr>
<tr>
<td>pKJ954</td>
<td>pHW40 with <em>mmp0234</em> complement</td>
<td>This study</td>
</tr>
<tr>
<td>pKJ969</td>
<td>pHW40 with <em>mmp0239</em> complement</td>
<td>This study</td>
</tr>
<tr>
<td>pKJ966</td>
<td>pHW40 with <em>mmp0240</em> complement</td>
<td>This study</td>
</tr>
<tr>
<td>pKJ968</td>
<td>pHW40 with <em>mmp0241</em> complement</td>
<td>This study</td>
</tr>
<tr>
<td>pKJ780</td>
<td>pHW40 with <em>mmp0040</em> complement</td>
<td>This study</td>
</tr>
<tr>
<td>pKJ1027</td>
<td>pHW40 with <em>mmp0038</em> complement</td>
<td>This study</td>
</tr>
<tr>
<td>pKJ1043</td>
<td>pHW40 with <em>mmp0039</em> complement</td>
<td>This study</td>
</tr>
</tbody>
</table>
wild type sequence or result in a deletion of the targeted gene. The culture was plated onto McCas agar containing 8-azahypoxanthine (240 µg/ml), which would be lethal to any cells that retained the vector-borne hpt gene. Following incubation at 37°C in an anaerobic canister for 1 week, individual colonies were picked and inoculated into Balch medium III for analysis. The individual transformants were screened by using washed whole cells resuspended in 2% NaCl as template for PCR along with sequencing primers (Supplemental Table 3.1) designed to amplify across the target gene. The PCR products were examined by agarose gel electrophoresis and the size compared to that predicted for the wildtype and deletion versions of the gene in order to identify mutants for specific gene deletions. Transformants showing the deletion size PCR products were restreaked onto Balch medium III plates and single colonies picked and again screened by PCR to confirm their purity.

Complementation of the gene deletion strains

Plasmids used for complementation of the gene deletions were constructed as previously described (Chaban et al. 2007, Lie et al. 2005) and listed in Supplemental Table 3.2. The vector used for complementation was the self-replicating plasmid, pHW40, a derivative of pWLG40 (Gardner and Whitman 1999), in which the transcription of the cloned gene is under the control of an inducible nif promoter (Kessler and Leigh 1999). pHW40 plasmids with the wildtype version of the complementing gene were transformed into the corresponding deletion strain using the PEG procedure (Tumbula et al. 1994). Complemented strains were grown in nitrogen free medium supplemented with either 10 mM NH₄Cl (where transcription from the nif promoter is repressed) or 10mM alanine (where transcription from the nif promoter is on). The PCR primers used to amplify the wildtype versions of each gene for creating the complementation vectors for each gene are listed in Supplemental Table 3.1 (gene specific primers labelled Comp-for and Comp-rev). NsII and MluI restriction sites were added to the
forward and reverse primers, respectively, for cloning purposes. In the case of *mmp0232* complementation, site directed mutagenesis was used to remove an internal Nsil site, using primers listed in Supplemental Table 3.1, prior to cloning into the complementation vector.

Reverse transcriptase-PCR

Reverse transcriptase PCR (RT-PCR) experiments were performed to determine if all 11 genes in the proposed type IV pilus-like operon were co-transcribed. Primers (Supplemental Table 3.3) were designed such that a sequence linking two neighboring genes across the intergenic region would be amplified if the genes were co-transcribed. RNA template was extracted from wildtype cells using an RNeasy Mini Kit (Qiagen Inc. Canada Mississauga, ON) with optional DNase digestion (Qiagen Inc.) as per the manufacturer’s protocol. cDNA was amplified using a One-Step RT-PCR kit (Qiagen Inc.) in accordance with the supplied protocol. In addition to using cDNA as template, PCR reactions were done using the same primer combinations with the purified RNA without the RT step as template as a control for possible DNA contamination of the RNA samples. PCR reactions were also run with genomic DNA as template to ensure the primers amplified the predicted size fragments.

The same strategy was employed to examine the possible co-transcription of the ATPase and conserved pilus membrane component genes.

Electron microscopy

Cells were grown overnight and washed with 50 mM MgSO$_4$ prior to treatment with 2% phosphotungstic acid to negatively stain the samples. Cells were examined on Formvar-coated gold grids and imaged under a Hitachi 7000 electron microscope operating at an accelerating voltage of 75 kV.
Supplemental Table 3.3: Primers used in this study for RT-PCR experiments.

<table>
<thead>
<tr>
<th>GENES LINKED</th>
<th>PRIMER NAME</th>
<th>PRIMER SEQUENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMP0231-MMP0232</td>
<td>RT-P1</td>
<td>TTGCGAATATTATCCATGC</td>
</tr>
<tr>
<td></td>
<td>RT-P2</td>
<td>TATGTTGCAGTCAGTATCAG</td>
</tr>
<tr>
<td>MMP0232-MMP0233</td>
<td>RT-P3</td>
<td>CGCTCGTGATATGCATTTA</td>
</tr>
<tr>
<td></td>
<td>RT-P4</td>
<td>AAAGATATCTGACCCCGTT</td>
</tr>
<tr>
<td>MMP0233-MMP0234</td>
<td>RT-P5</td>
<td>GCGGTTCCCTGTGAATTC</td>
</tr>
<tr>
<td></td>
<td>RT-P6</td>
<td>CCACTATACTCTTCAGCAGG</td>
</tr>
<tr>
<td>MMP0234-MMP0235</td>
<td>RT-P7</td>
<td>GATATAGTGAAAGAGATCC</td>
</tr>
<tr>
<td></td>
<td>RT-P8</td>
<td>AGTTCAGAATGTCATTTG</td>
</tr>
<tr>
<td>MMP0235-MMP0236</td>
<td>RT-P9</td>
<td>GATACCGATGTGCTGAGTAT</td>
</tr>
<tr>
<td></td>
<td>RT-P10</td>
<td>AATTCTACTGAACCTGTCC</td>
</tr>
<tr>
<td>MMP0236-MMP0237</td>
<td>RT-P11</td>
<td>TGTATCTACAATGACTCCAG</td>
</tr>
<tr>
<td></td>
<td>RT-P12</td>
<td>AATCGAAATGAAACCTGTCT</td>
</tr>
<tr>
<td>MMP0237-MMP0238</td>
<td>RT-P13</td>
<td>GAACATGCTCATGCAATTTCG</td>
</tr>
<tr>
<td></td>
<td>RT-P14</td>
<td>TTCCAAAGCATTGGATGATG</td>
</tr>
<tr>
<td>MMP0238-MMP0239</td>
<td>RT-P15</td>
<td>GAGTTATAATCGTTATGGA</td>
</tr>
<tr>
<td></td>
<td>RT-P16</td>
<td>AACTGCAACAGCCTCATAAG</td>
</tr>
<tr>
<td>MMP0239-MMP0240</td>
<td>RT-P17</td>
<td>GAAGCAGATAATATCTCTCA</td>
</tr>
<tr>
<td></td>
<td>RT-P18</td>
<td>ATCGCAAATTACATGTCGCT</td>
</tr>
<tr>
<td>MMP0240-MMP0241</td>
<td>RT-P19</td>
<td>CTGGTTTCTACAGTGGTGC</td>
</tr>
<tr>
<td></td>
<td>RT-P20</td>
<td>GGAATCGATGTTTCCACAAAC</td>
</tr>
</tbody>
</table>
Results

A type IV pili-like locus encoding 11 potential genes (\textit{mmp0231-mmp0241}) was initially identified by Szabo et al (Szabo et al. 2007) (Figure 3.1A). We recently demonstrated the essential involvement of the three pilin-like genes (\textit{epdA}, \textit{epdB} and \textit{epdC}) of the locus in the normal assembly of surface pili (Ng et al. 2011) while \textit{eppA} was previously shown to encode the prepilin peptidase (Szabo et al. 2007). However, the possible involvement of the remaining genes in the operon in pili biosynthesis remained to be determined. These predicted proteins do not show homology to any bacterial type IV pilus genes or to pili genes in \textit{Sulfolobus solfataricus}. The proteins encoded by these genes were analyzed by various bioinformatics programs (PRED-SIGNAL, PSORT, TmPred, InterProScan), which at times gave conflicting predictions (Table 3.1). Most of the proteins are annotated as conserved hypothetical proteins and most have significant BLAST matches to only archaeal proteins and sometimes only to \textit{Methanococcus} proteins. Most lack the presence of motifs that would be useful in deciphering a possible function (Table 3.1). Two are predicted to have signal peptides (MMP0235 and MMP0238) while MMP0241 contains a domain of unknown function (DUF2341) that is found in various bacterial proteins that form proton channels (MotA, TolQ, ExbB) or act as transport proteins. Other proteins are predicted to be cytoplasmic or membrane located (refer to Table 3.1).

RT-PCR experiments

All 11 genes in the type IV pilus locus are oriented in the same direction and the intergenic region between each adjoining gene is very small (ranging from overlapping to 52 nucleotides), with the exception of the distance between \textit{mmp0233-mmp0234}. Agarose gel electrophoresis (Figure 3.1B) shows amplified PCR products of the correct predicted length in all the lanes representing the 10 intergenic region between the 11 genes when
Figure 3.1: Analysis of the major type IV pilus locus. A) The major type IV pilus locus showing all 11 genes with the three pilin-like genes (epdA, epdB and epdC, formerly mmp0233, mmp0236 and mmp0237 respectively) and the prepilin peptidase EppA (formerly mmp0232) indicated. b) RT-PCR experiment indicating co-transcription of all 11 genes. For each pair of adjacent genes the triplet of lanes represent; 1) Standard PCRs using Mm900 genomic DNA as template and the respective RT primers which amplify across the intergenic regions to indicate the expected amplicon size and primer specificity; 2) RT-PCR run using total RNA extracted from Mm900 cells as template with the same RT primers. The RT lanes which have bands at the same size as the DNA lanes indicate the co-transcription of the indicated genes; 3) Standard PCRs performed using total RNA that did not undergo reverse transcription as template to rule out possible DNA contamination of the RNA sample.
Table 3.1: Properties of the genes/proteins in the pili locus of Methanococcus maripaludis.

<table>
<thead>
<tr>
<th>GENE</th>
<th>ANNOTATION</th>
<th>GENE SIZE</th>
<th>PREDICTION OF SIGNAL PEPTIDE (Pred-Signal)</th>
<th>PREDICTED TMDS (TmPred)</th>
<th>LOCATION IN THE CELL (PSortb)</th>
<th>MOTIFS/COMMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMP0231</td>
<td>Zn finger containing protein</td>
<td>369 bp</td>
<td>NO</td>
<td>0</td>
<td>Cytoplasm</td>
<td>Zinc finger, probable metal binding domain</td>
</tr>
<tr>
<td>MMP0232</td>
<td>Conserved hypothetical protein</td>
<td>1041 bp</td>
<td>NO</td>
<td>8</td>
<td>Cytoplasmic membrane</td>
<td>pfam01478, Type IV prepilin peptidase family (Peptidase activity demonstrated)</td>
</tr>
<tr>
<td>MMP0234</td>
<td>Conserved hypothetical protein</td>
<td>1149 bp</td>
<td>NO</td>
<td>1</td>
<td>Cytoplasm</td>
<td>DUF515, methanogen specific protein</td>
</tr>
<tr>
<td>MMP0235</td>
<td>Conserved hypothetical protein</td>
<td>894 bp</td>
<td>YES</td>
<td>1</td>
<td>Unknown</td>
<td>Methanococcus specific protein</td>
</tr>
<tr>
<td>MMP0238</td>
<td>Conserved hypothetical protein</td>
<td>486 bp</td>
<td>YES</td>
<td>2</td>
<td>Unknown</td>
<td>Methanogen specific protein</td>
</tr>
<tr>
<td>MMP0239</td>
<td>Conserved hypothetical protein</td>
<td>858 bp</td>
<td>NO</td>
<td>1</td>
<td>Cytoplasm</td>
<td>Methanogen specific protein</td>
</tr>
<tr>
<td>MMP0240</td>
<td>Conserved hypothetical protein</td>
<td>534 bp</td>
<td>NO</td>
<td>1</td>
<td>unknown</td>
<td>Methanogen specific protein</td>
</tr>
<tr>
<td>MMP0241</td>
<td>Conserved hypothetical protein</td>
<td>1764 bp</td>
<td>NO</td>
<td>2</td>
<td>Extracellular</td>
<td>DUF2341, found in MotA/TolQ/Exb B proton channels and</td>
</tr>
<tr>
<td>ID</td>
<td>Description</td>
<td>Length</td>
<td>Membrane Location</td>
<td>Function</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>------------------------------------------------------------------------------</td>
<td>--------</td>
<td>-------------------</td>
<td>-----------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMP0038</td>
<td>Bacterial type II secretion system protein F domain protein</td>
<td>897 bp</td>
<td>Cytoplasmic membrane</td>
<td>Type II secretion system protein F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMP0039</td>
<td>Bacterial type II secretion system protein F domain protein</td>
<td>1059 bp</td>
<td>Cytoplasmic membrane</td>
<td>Pfam 00482, type II secretion system protein F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMP0040</td>
<td>Type II secretion system protein putative; subunit 1</td>
<td>1650 bp</td>
<td>Cytoplasmic membrane</td>
<td>VirB11-like ATPase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMP0281</td>
<td>P-loop ATPase of the PilT family</td>
<td>1893 bp</td>
<td>Unknown</td>
<td>VirB11-like ATPase, PilT N-terminus domain</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
the RNA was first reverse transcribed, suggesting that all 11 genes form a single operon. These PCR products were the same size as those obtained using genomic DNA as template. No PCR products were obtained when the RT step was omitted (Figure 3.1B), indicating the RNA samples were free of contaminating genomic DNA.

**Targeted internal gene deletions of pili operon genes**

To determine if each of the genes found in the pili locus was essential for piliation, an internal deletion of each gene was attempted (Moore and Leigh 2005, Ng et al. 2011). These deletions were all created in a *flaK* deletion strain of *M. maripaludis*. This strain lacks the signal peptidase necessary to process archaellins (Bardy and Jarrell 2002, Bardy and Jarrell 2003) making the cells nonarchaellated and leaving only the less numerous and thinner pili as the sole surface appendages (Ng et al. 2011). In spite of repeated attempts, an internal deletion of *mmp0231* was not successful; however, deletions were created in all the remaining genes. Deletion strains were identified among the transformants by whole cell PCR using gene specific primers that amplified across the deleted region, resulting in a smaller PCR product in a mutant strain when compared to the products obtained with the same primer pair using the wildtype cells as template (Figure 3.2). Electron microscopic examination of each mutant revealed that deletion of any of the genes, with the exception of *mmp0238* and *mmp0235*, led to nonpiliated cells (Figure 3.3). For all the deletions that led to nonpiliated cells, at least 50 cells of each mutant were examined. EM examination of the *mmp0235* and *mmp0238* deletion strains, on the other hand, revealed a wildtype number and appearance of pili demonstrating that neither of these genes was essential for piliation (Figure 3.3). For these two mutants as well as the parent *flaK* deletion strain, the number of pili per cell was small (*flaK* 3.1+/− 2.5 pili per cell, N=16; *mmp0235* 2.5+/− 1.8 pili per cell, N=27; *mmp0238* 4.0+/− 2.8 pili per cell, N=19). Even in the *flaK*
Figure 3.2: Confirmation of the deletions of the targeted pili locus genes by PCR. Washed whole cells of each deletion strain as well as wildtype cells were used as template for the PCR confirmation of each deletion, using corresponding sequencing primers (Supplemental Table 3-1). Results for each indicated deletion are shown in pairs with the first lane showing the PCR amplicon using wildtype cells as template and the second lane representing the amplicon from the gene deletion strain. The amplicon size in each lane is the predicted size. M indicates 1 kb DNA ladder from New England Biolabs.
Figure 3.3: Electron micrographs of strains carrying the indicated pili locus gene deletion showing its effect on piliation. All deletion strains are non-piliated except for Δmmp0235 and Δmmp0238. The parent strain, ΔflaK, is shown for comparison. Enlargements of a section of a cell are presented to more easily visualize the thin pili, when present. Samples were negatively stained with 2% phosphotungstic acid (pH 7.0). Arrows indicate pili. Bar, 200nm.
cells not all cells were piliated (35/39 or 83%). All of the gene deletions that resulted in loss of piliation were complemented with plasmid-borne wildtype versions of the appropriate gene. In all cases, complementation of the deletion mutants resulted in a return to piliation (Figure 3.4), although in the case of ∆mmp0232, ∆mmp0234 and ∆mmp0239 only a small percentage (4%, N=75) of complemented cells were piliated.

**Identification of type IV pilin ATPase and conserved membrane protein genes**

Type IV pili systems in both Bacteria and Archaea require at least a single ATPase to provide energy for the assembly process. Most have two ATPases, one for extension and one for retraction of the pili, enabling a type of surface motility called twitching (Burrows 2012, Mattick 2002). In addition, type IV pili systems have a conserved membrane component that acts as a platform for assembly and can interact with the ATPases (Burrows 2012, Takhar et al. 2013). The type IV pili-like locus in *M. maripaludis* does not contain homologues of a pilus ATPase or conserved membrane component gene. However, two potential type IV pilus ATPase genes, *mmp0040* and *mmp0281*, were found outside the locus upon examination of the annotated sequenced *M. maripaludis* genome and a search for homologues to pilus ATPases from *Pseudomonas aeruginosa* using the BLAST algorithm. Of the two genes, *mmp0040* had a higher sequence similarity to the *P. aeruginosa* pilus ATPases. It was also found to be in a gene cluster containing two genes encoding type IV pili conserved membrane components (type 2 secretion system protein F, PilC in *P. aeruginosa*) (Figure 3.5A). These three genes were shown by RT-PCR to be co-transcribed as part of a five gene operon (Figure 3.5B), along with genes for transcription initiation factor B (*mmp0041*) and H/ACA RNA-protein complex
Figure 3.4: Electron micrographs of deletion strains complemented with a plasmid-borne wildtype copy of the deleted gene. All strains were returned to a piliated state following complementation, although only a small percentage of cells were piliated in the complementations of mmp0232, mmp0234 and mmp0239. Enlargements of a section of a cell are presented to more easily visualize the thin pili. Samples were negatively stained with 2% phosphotungstic acid (pH 7.0). Arrows indicate pili. Bar, 200nm.
Figure 3.5: The five gene locus containing genes for the pili assembly ATPase and two copies of the platform (pilC-like) protein. a) Organization of the genetic locus. *mmp0040* encodes the ATPase and *mmp0038* and *mmp0039* encode the conserved membrane proteins. b) RT-PCR experiment indicating co-transcription of all 5 genes. For each pair of adjacent genes the triplet of lanes represent; 1) Standard PCRs using Mm900 genomic DNA as template and the respective RT primers which amplify across the intergenic regions to indicate the expected amplicon size and primer specificity; 2) RT-PCR run using total RNA extracted from Mm900 cells as template with the same RT primers. The RT lanes which have bands at the same size as the DNA lanes indicate the co-transcription of the indicated genes; 3) Standard PCRs performed using total RNA that did not undergo reverse transcription as template to rule out possible DNA contamination of the RNA sample. c) Confirmation of in-frame deletions of the targeted pili locus genes by PCR. Washed whole cells of each deletion strain as well as wildtype cells were used as template for the PCR confirmation of each in-frame deletion, using corresponding sequencing primers (Supplemental Table 3-1). Results for each indicated deletion are shown in pairs with the first lane showing the PCR amplicon using wildtype cells as template and the second lane representing the amplicon from the gene deletion strain. In all cases, the obtained PCR products were the predicted size. M indicates 1 kb DNA ladder from New England Biolabs.
component Gar1 \((mmp0042)\). When the MMP0040 protein sequence was used as a query to search \textit{Pseudomonas aeruginosa} genomes in a BLAST search \((\text{http://blast.ncbi.nlm.nih.gov/Blast.cgi})\) it retrieved matches to TadA ATPases \((4e-60; 61\% \text{ coverage})\) with less significant alignments to PilB \((1e-08)\) and PilT \((3e-08)\). Using the protein sequence of MMP0281 in a similar search retrieved significant alignments to trb conjugation ATPase \((2e-04)\) and GspE \((0.006)\) with coverage of less than 30\%. When the protein sequence of \(P.\ aeruginosa\) TadA, was used as query to search the \textit{M. maripaludis} genomes, it resulted in hits to MMP0040 \((3e-62, 79\% \text{ coverage})\) as well as Flal \((3e-38, 78\% \text{ coverage})\) and MMP0281 \((2e-07, 31\% \text{ coverage})\). Flal is the ATPase for archaella assembly and rotation \((\text{Banerjee et al. 2012, Thomas et al. 2002})\). When the protein sequences of \(P.\ aeruginosa\) PilT or PilB were used as queries in a similar search, it returned significant alignments to Flal \((1e-10, 31\% \text{ coverage for PilT and 1e-10, 23\% coverage for PilB})\). The PilB search also identified significant alignment to MMP0040 but with low coverage \((1e-07, 12\% \text{ coverage})\). Deletions were created in each of the two potential ATPase genes and each was examined for the presence of pili \((\text{Figure 3.5C})\). The \textit{mmp0040} deletion strain was shown to be nonpiliated \((\text{Figure 3.6})\) while the \textit{mmp0281} deletion strain remained piliated \((\text{Figure 3.6})\) to a similar extent as the \textit{flaK} cells \((\text{for the mmp0281 deletion strain 3.1+/- 1.9 pili per cell, N=22})\). Internal deletions were created in each of the two pilus platform genes as well \((\text{Figure 3.5C})\) and each of these deletion strains was examined for the presence of pili. Interestingly, both \textit{mmp0038} and \textit{mmp0039} were required for piliation \((\text{Figure 3.6})\). While both MMP0038 and MMP0039 are type II secretion system protein F domain proteins, the amino acid similarity between the two proteins is very low. Complementation of the \textit{mmp0038, mmp0039} and \textit{mmp0040} deletion strains all resulted in a return to the piliated state, confirming their involvement in piliation \((\text{Figure 3.6})\). A summary of
Figure 3.6: Electron micrographs of strains deleted for the type IV pili ATPase and platform protein genes and their complemented strains. All deletion strains were nonpiliated, except Δmmp0281, but were returned to a piliated state following complementation. In the case of mmp0040, the complementation returned cells to a wildtype level of piliation while for the complementations of both mmp0038 and mmp0039 only a small percentage of cells were piliated. Samples were negatively stained with 2% phosphotungstic acid (pH 7.0). Bar, 200nm.
the effects of all the studied gene deletions and complementations on piliation, including our previous results on the pilus structural genes (Ng et al. 2011), are presented in Table 3-2.

**Discussion**

The 11 gene type IV pili-like locus (*mmp00231-mmp0241*) includes a gene for a prepilin peptidase EppA (*mmp0232*) and three type IV-pilin like genes (*mmp0233* [epdA], *mmp0236* [epdB] and *mmp0237* [epdC]) (Szabo et al. 2007). However, the possible involvement of the remaining genes in pili formation has not been previously addressed and was a major focus of this work. The results of RT-PCR experiments demonstrated that all 11 genes formed a single transcriptional unit. With the exception of the prepilin peptidase and the pilins, the seven other genes in the locus represent novel genes necessary for assembly of type IV pili-like structures; they bear no homology to known type IV pili genes in bacteria or to ones so far identified in other Archaea like the Aap and Ups pili of *S. solfataricus* and *S. acidocaldarius* (Frols et al. 2008, Henche et al. 2012a). Some of the genes appear to be unique to *Methanococcus* and encode proteins predicted to be located in the cytoplasm, cytoplasmic membrane or to be secreted. Elucidation of the functions of these gene products will be a future challenge. Even in well studied bacterial type IV pili systems, the functions of many conserved gene products remain a mystery with the type IVa and type IVb (along with the tad/flp subgrouping) all having unique components with unknown functions (Burrows 2012). It could be speculated that all the proteins encoded within this operon may be needed in relatively small numbers and so found in an operon separate from the gene encoding the major pilin subunit, MMP1685. The latter would be needed in significantly higher numbers than the minor pilins EpdA, EpdB and EpdC and may be transcribed from a much stronger promoter. This may be how *M. maripaludis* obtains the optimal amounts of the different structural proteins present in significantly different stoichiometries in the final assembled pili.
Table 3.2: Summary of the targeted genes effects on piliation.

<table>
<thead>
<tr>
<th>GENE</th>
<th>FUNCTION</th>
<th>RESULT OF GENE DELETION</th>
<th>RESULT OF GENE COMPLEMENTATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMP0231</td>
<td>Unknown</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>MMP0232 (eppA)</td>
<td>Prepilin peptidase</td>
<td>Non-piliated</td>
<td>Restores piliation, but poorly</td>
</tr>
<tr>
<td>MMP0233 (epdA)</td>
<td>Minor pilin</td>
<td>Reduced number of pili</td>
<td>Restores normal piliation</td>
</tr>
<tr>
<td>MMP0234 (epdF)</td>
<td>Unknown</td>
<td>Non-piliated</td>
<td>Restores piliation, but poorly</td>
</tr>
<tr>
<td>MMP0235</td>
<td>Unknown</td>
<td>Piliated</td>
<td>ND</td>
</tr>
<tr>
<td>MMP0236 (epdB)</td>
<td>Minor pilin</td>
<td>Non-piliated</td>
<td>Restores piliation</td>
</tr>
<tr>
<td>MMP0237 (epdC)</td>
<td>Minor pilin</td>
<td>Non-piliated</td>
<td>Restores piliation</td>
</tr>
<tr>
<td>MMP0238</td>
<td>Unknown</td>
<td>Piliated</td>
<td>ND</td>
</tr>
<tr>
<td>MMP0239 (epdG)</td>
<td>Unknown</td>
<td>Non-piliated</td>
<td>Restores piliation, but poorly</td>
</tr>
<tr>
<td>MMP0240 (epdH)</td>
<td>Unknown</td>
<td>Non-piliated</td>
<td>Restores piliation</td>
</tr>
<tr>
<td>MMP0241 (epdI)</td>
<td>Unknown</td>
<td>Non-piliated</td>
<td>Restores piliation</td>
</tr>
<tr>
<td>MMP1685 (epdE)</td>
<td>Major pilin</td>
<td>Non-piliated</td>
<td>Restores piliation</td>
</tr>
<tr>
<td>MMP1283 (epdD)</td>
<td>Minor pilin</td>
<td>Non-piliated</td>
<td>Restores piliation</td>
</tr>
<tr>
<td>MMP0038 (epdJ)</td>
<td>PilC-like membrane component</td>
<td>Non-piliated</td>
<td>Restores piliation, but poorly</td>
</tr>
<tr>
<td>MMP0039 (epdK)</td>
<td>PilC-like membrane component</td>
<td>Non-piliated</td>
<td>Restores piliation, but poorly</td>
</tr>
<tr>
<td>MMP0040 (epdL)</td>
<td>Type IV pilin ATPase</td>
<td>Non-piliated</td>
<td>Restores piliation</td>
</tr>
</tbody>
</table>
Electron microscopic analysis of cells containing deletions in each of seven other genes of this major pilus operon indicated that all of the pilus locus genes, except mmp0235 and mmp0238, were essential for piliation. Successful complementation of all gene deletions indicated that the effects on piliation observed were due to the specific deleted gene and not from polar effects generated in the creation of each gene deletion. Not all complementations, however, returned the mutants to wildtype levels of piliation. In some examples, the complementation only restored piliation to a small percentage of cells. In the P. aeruginosa type IV system (Giltner et al. 2010), overexpression of minor pilins can be detrimental to pilus biosynthesis and it has been suggested that the overabundant minor pilins may titrate out other key chromosomally encoded pilus proteins. In the cases of complementations that did not return mutant cells to a wildtype state of piliation in our experiments, it is possible that gene products were produced at above normal levels and this may have interfered with normal pilus assembly.

The two genes in the locus not found to be necessary for pili formation, namely mmp0235 and mmp0238, might still have a role to play in pili function, such as attachment. Adhesins have been identified in certain bacterial type IV pili systems, including Neisseria species. Here PilC1 has been identified as an adhesin located at the tip of the pili (Rudel et al. 1995). PilC1 is made initially with a signal peptide (Morand et al. 2001) and pili can be formed without the adhesin present (Rudel et al. 1995). Both MMP0235 and MMP0238 are also made with a predicted signal peptide and if either of these two proteins is the adhesin then their absence may affect pili function but not its assembly.

Of the nonpilin genes studied in the major pilus locus, only mmp0232 (eppA) has an assigned function as encoding the prepilin peptidase, required for processing of all the pilin-like proteins prior to their incorporation into the pilus structure (Szabo et al. 2007). As expected, deletion of this gene led to nonpiliated cells. Among the studied Archaea, Methanococcus
appears unique in possessing two prepilin peptidases involved in assembly of two different surface appendages. FlaK was the first identified prepilin peptidase in Archaea and its essential role in the processing of pre-archaellins for archaella biosynthesis in *Methanococcus sp.* has been well established (Bardy and Jarrell 2002, Bardy and Jarrell 2003). Clearly, FlaK cannot compensate for the loss of EppA and each enzyme is restricted in its substrates to either pilins or archaellins despite the similarities in the signal peptide and N-termini of the two substrates (Szabo et al. 2007). Other archaea, such as *S. solfataricus* and *H. volcanii*, possess a variety of proteins with class 3 signal peptides (including archaellins and pilins) but these all appear to be processed by a single enzyme, designated PibD, that possesses broad substrate specificity, (Albers and Pohlschroder 2009, Lassak et al. 2012a, Tripepi et al. 2010). Interestingly, *S. solfataricus* PibD was shown to be able to process *M. voltae* archaellin in *invitro* assays (Ng et al. 2009).

Examination of the 11 gene pili cluster indicated that genes encoding conserved type IV pili ATPases and the membrane component were missing. In most type IV pili systems, an ATPase (PilB) is needed for incorporation of new subunits into the structure for pilus extension while a separate ATPase (PilT) removes subunits from the base of the structure leading to retraction of the pilus (Burrows 2012, Burrows 2005). This extension and retraction of the pilus leads to the movement of cells across a solid surface in a process called twitching (Burrows 2012). The conserved membrane protein appears to interact with the ATPases to form a platform for assembly/disassembly of the pilus (Burrows 2012, Crowther et al. 2004). Archaeal species have not been shown to twitch and in Archaea, only homologues to PilB have been identified (Peabody et al. 2003). In *S. acidocaldarius*, at least two different type IV pili, produced under different growth conditions, are already known. A single ATPase and a single homologue of the conserved membrane protein have been identified in both Ups and Aap pili systems in a locus that also contains genes for pilins. Deletion of either the ATPase or the conserved
membrane protein gene in either pilus system led to the inability of those mutants to assemble that particular pilus type (Hence et al. 2012a, van Wolferen et al. 2013). Deletion analysis identified a small locus in the M. maripaludis genome containing a single pilus ATPase homologue adjacent to two homologues of the conserved pilus membrane protein that were all essential for piliation. Based on this evidence, M. maripaludis pili are predicted to be unable to retract unless MMP0040 is able to perform both polymerization and removal of subunits from the structure, an idea already considered for the single ATPase Sulfolobus pili systems (Albers and Pohlschroder 2009). The presence of two divergent copies of the conserved membrane component gene is relatively rare in bacterial type IV pili systems, but found in the tad pili systems of P. aeruginosa and Aggregatibacter (previously Actinobacillus) actinomycetemcomitans (Burrows 2012). In the latter case, both membrane component genes (tadB and tadC) are required for piliation (Kachlany et al. 2000). Why two versions of this protein are required for in the Tad pili system and here in M. maripaludis is not known but it has been suggested that perhaps the single ATPase of the Tad system may interact with one version of the conserved membrane proteins in pilin addition and with the other in pilin removal (Burrows 2012). In P. aeruginosa, recent evidence led to speculation that the platform protein PilC is likely a dimer which might reside within the lumen of the hexameric ATPase (Takhar et al. 2013). The polymerization and retraction ATPases might then interact with the two cytoplasmic domains of PilC with the PilB ATPase interacting with the N-terminal domain and PilT ATPase with the C-terminal domain (Takhar et al. 2013). In systems with two platform proteins as in M. maripaludis, it is possible that a heterodimer is formed and the single ATPase interacts with one member of the dimer for extension but with the other for retraction (Burrows 2012).

The results presented in this report, coupled to our earlier investigations (Nair et al. 2013, Ng et al. 2011), indicate that genes responsible for the type IV-like pili in M. maripaludis
are spread around the genome, in sharp contrast to pili loci in *Sulfolobus* (Lassak et al. 2012a). The 11 gene operon investigated here contains three minor pilins, the prepilin peptidase necessary for prepilin processing as well as numerous genes that have no obvious counterpart in either bacterial or other archaeal pili systems. The gene for the major structural protein MMP1685 (Ng et al. 2011) is located at an entirely different genetic locus as is the gene for an additional minor pilin MMP1283 (Nair et al. 2013) and as shown here, the genes encoding the highly conserved assembly ATPase and membrane proteins are located at a fourth distinct locus. With the involvement of five structural proteins (EpdA, EpdB, EpdC, MMP1283 and MMP1685) already determined to be necessary for normal piliation, the structure of pili of *M. maripaludis* seems to be more complex than *Sulfolobus* systems where only two pilins have been reported (Frols et al. 2008, Henche et al. 2012a, van Wolferen et al. 2013). The pili of *M. maripaludis* are usually observed in small numbers under routine growth conditions and, unlike in *Sulfolobus* (Orell et al. 2013b, Reimann et al. 2012, Vassart et al. 2012), no studies on their possible regulation have been reported. The initial publication on the 11 gene pilus locus designated *mmp0233* as *eppA* to denote a novel subclass of a euryarchaeal type IV prepilin peptidase (Szabo et al. 2007). The pilin genes were designated *epdA* (*mmp0233*), *epdB* (*mmp0236*) and *epdC* (*mmp0237*) since they were dependent on EppA for signal peptide removal and hence EppA-dependent proteins. Since these are the only known pili in *M. maripaludis* and the structures are currently lacking a specific designation, we feel they could rightly be called Epd pili since the pili themselves are EppA-dependent. Since we have identified a number of genes essential for Epd pili formation in this and prior publications (Nair et al. 2013, Ng et al. 2011), we propose all of these genes now be given *epd* designations. For the pilins MMP1283 and MMP1685, they are given the designations EpdD and EpdE, respectively. The other genes whose deletions led to nonpiliated cells are designated as follows: *mmp0234: epdF; mmp0239: epdG; mmp0240: epdH; mmp0241: epdI; mmp0038: epdJ; mmp0039: epdK and
Delineating the roles of the novel gene products involved in *M. maripaludis* pili biosynthesis represent immediate challenges for the field.

**Acknowledgments.** This research was funded by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) to KFJ.
Literature cited


Jarrell KF, Stark M, Nair DB, Chong JPJ (2011) Flagella and pili are both necessary for efficient attachment of *Methanococcus maripaludis* to surfaces. FEMS Microbiol Lett 319:44-50


Morand PC, Tattevin P, Eugene E, Beretti JL, Nassif X (2001) The adhesive property of the type IV pilus-associated component PilC1 of pathogenic Neisseria is supported by the conformational structure of the N-terminal part of the molecule. Mol Microbiol 40:846-856


Vassart A, Van Wolferen M, Orell A, Hong Y, Peeters E, Albers SV, Charlier D (2012) Sa-Lrp from *Sulfolobus acidocaldarius* is a versatile, glutamine-responsive, and architectural transcriptional regulator. Microbiology Open 2:75-93


Chapter 4

Identification of an Additional Minor Pilin Essential for Piliation in the Archaeon

* Methanococcus maripaludis *

Abstract

*Methanococcus maripaludis* is an archaeon with two studied surface appendages, archaella and type IV-like pili. Previously, the major structural pilin was identified as MMP1685 and three additional proteins were designated as minor pilins (EpdA, EpdB and EpdC). All of the proteins are likely processed by the pilin-specific prepilin peptidase EppA. Six other genes were identified earlier as likely encoding pilin proteins processed also by EppA. In this study, each of the six genes (*mmp0528, mmp0600, mmp0601, mmp0709, mmp0903* and *mmp1283*) was deleted and the mutants examined by electron microscopy to determine their essentiality for pili formation. While mRNA transcripts of all genes were detected by RT-PCR, only the deletion of *mmp1283* led to nonpiliated cells. This strain could be complemented back to a piliated state by supplying a wildtype copy of the *mmp1283* gene in trans. This study adds to the complexity of the type IV pili system in *M. maripaludis* and raises questions about the functions of the remaining five pilin-like genes and whether *M. maripaludis* under other growth conditions may be able to assemble additional pili-like structures.
Introduction

Type IV pili are a very common type of surface appendage found in a variety of Gram-negative and Gram positive bacteria, as well as certain members of the Domain Archaea (Burrows 2012a, Giltner et al. 2012, Imam et al. 2011, Ng et al. 2008, Pelicic 2008, Pohlschroder et al. 2011). They are involved in a wide variety of processes including adherence, aggregation, DNA transfer in transformation and conjugation, biofilm formation, electron transfer and a type of surface motility termed twitching (Boesen and Nielsen 2013, Burrows 2012a, Giltner et al. 2012, Lovley et al. 2011). The core components of a type IV pili system includes structural proteins with class 3 signal peptides, a prepilin signal peptidase, one or more ATPases and a conserved membrane (platform) protein (Giltner et al. 2012, Pohlschroder et al. 2005, Takhar et al. 2013). One ATPase powers the incorporation of new subunits into the growing filament while in many cases, the presence of a second, depolymerizing ATPase acts to remove subunits from the structure. The combined activities of the two ATPases results in extension and retraction of the pili, leading to the twitching motility associated with type IV pili in many bacteria (Burrows 2005). The conserved inner membrane or platform protein is considered to interact with the ATPase(s) and form an export complex for the structural proteins and to be involved in both pilus assembly and disassembly (Takhar et al. 2013). In addition to these conserved components, type IV pilus systems in different organisms often have other components whose role in pilus assembly and function remain unknown (Burrows 2012a).

The structural subunits of the type IV pilus consist of a major pilin and typically several other pilins, termed minor pilins due to their much lower abundance, all synthesized initially with class 3 signal peptides that are specifically processed by the prepilin peptidase (Strom and Lory 1993, Strom et al. 1993). Minor pilins have been shown to be necessary for pili formation in several different systems (Alm and Mattick 1995, Alm et al. 1996, Carbonnelle et al. 2006,
Winther-Larsen et al. 2005) and they have been detected in sheared pili samples of *N. gonorrhoeae* and *P. aeruginosa* (Giltner et al. 2010, Winther-Larsen et al. 2005), although other roles for minor pilins as activators of pilus assembly without incorporation into the structure have also been proposed in other systems (Burrows 2012b, Cisneros et al. 2012). In some type IV pili systems, evidence for a minor pilus constituent acting as a specific adhesin has been presented (Heiniger et al. 2010, Ishiwa and Komano 2004, Rudel et al. 1995, Winther-Larsen et al. 2001).

Archaea are known to use the type IV pili-like model to assemble numerous surface structures (Jarrell et al. 2013, Lassak et al. 2012a, Ng et al. 2008, Pohlschroder et al. 2011) including type IV-like pili (Frols et al. 2008, Henche et al. 2012a, Jarrell et al. 2013, Lassak et al. 2012a, Ng et al. 2011, Szabo et al. 2007), the bindosome for substrate uptake in *Sulfolobus solfataricus* (Lassak et al. 2012a, Zolghadr et al. 2011, Zolghadr et al. 2007), likely the unusual Iho670 fibres of *Ignicoccus hospitalis* (Muller et al. 2009, Yu et al. 2012) and the best studied example, namely the archaellum (Ghosh and Albers 2011, Jarrell et al. 2009, Jarrell et al. 2010, Jarrell et al. 1996, Lassak et al. 2012a, Lassak et al. 2012b). The name “archaellum” has been proposed to replace the term “archaeal flagellum” (Jarrell and Albers 2012) since the archaeal structure, while involved in swimming (as well as other functions), is not homologous to the bacterial flagellum and is related instead to type IV pili in structure and likely assembly (Ghosh and Albers 2011, Ng et al. 2006, Peabody et al. 2003). This proposal is still under discussion in the scientific community and further arguments, both pro and con, have been presented (Eichler 2012, Wirth 2012).

Recent studies in *Methanococcus*, *Haloferax* and especially *Sulfolobus* have been devoted specifically to the study of the type IV-like pili (Ajon et al. 2011, Esquivel et al. 2013, Frols et al. 2008, Henche et al. 2012a, Ng et al. 2011, Orell et al. 2013, Wang et al. 2008). *Sulfolobus* has been shown to produce two different type IV pili structures. One, called UV-inducible type IV pili (Ups pili; (Ajon et al. 2011, Frols et al. 2008)), is widespread throughout the
Sulfolobales while the second called archaeal adhesive pili (Aap pili) is limited so far to S. acidocaldarius (Henche et al. 2012a, Henche et al. 2012b, Lassak et al. 2012a). Ups pili are upregulated under conditions that lead to DNA double stranded breaks such as UV light and their formation leads to cell aggregation that promotes DNA exchange that might help in repairing the DNA damage (Ajon et al. 2011, Frols et al. 2008). On the other hand, Aap pili are the most abundant surface appendage observed on S. acidocaldarius under normal growth conditions in nutrient rich medium (Lassak et al. 2012a). Aap pili are adhesion structures primarily but they also influence biofilms, promoting the formation of tower-like structures (Henche et al. 2012b). The loci identified as encoding the biosynthesis of both Aap pili and Ups pili were shown to consist of only five genes (Frols et al. 2008, Henche et al. 2012a, Pohlschroder et al. 2011). In each case, there are genes for two prepilins, a single pilin assembly ATPase, the conserved pilus membrane/platform protein and one additional gene in each operon that has an unknown function. In the Aap system, AapB appears to be the major pilin and AapA the minor pilin (Henche et al. 2012a). Mutational studies demonstrated that all five aap genes were necessary for pili formation (Henche et al. 2012a). In the Ups system, deletion of the gene encoding the ATPase (upsE) eliminated UV-induced aggregation (Frols et al. 2008). The assignment of UpsA or UpsB proteins as the major or minor pilins has not been reported.

In Haloferax volcanii, six novel type IV pilins termed PilA (1-6) involved in adhesion have been recently studied (Esquivel et al. 2013). Each protein contains a highly conserved domain of unknown function (Duf1628) at their N-terminus and all are processed by PibD. Strains carrying deletions of up to five of the six pilin genes are still able to adhere while a mutant with all six genes deleted cannot. Complementation of the strain deleted for all six pilin genes with any single pilin gene results in cells that have functional pili.
In *Methanococcus maripaludis*, an *in-silico* study (Szabo et al. 2007) identified the existence of a type IV pili-like locus, consisting of 11 potential genes, including three encoding prepilin-like proteins and one encoding a prepilin peptidase. Subsequent genetic studies demonstrated that two of the pilin genes (*epdB* and *epdC*) were essential for pili formation while deletion of the third (*epdA*) resulted in cells with a reduced number of pili (Ng et al. 2011). Unexpectedly, none of these genes encoded the major structural protein which was revealed by mass spectrometry of purified pili samples to be MMP1685, located at a distant locus (Ng et al. 2011). Several other genes in the 11 gene operon are also essential for piliation (Nair et al. 2014). Unlike the case with *Sulfolobus* pili loci, the assembly ATPase and conserved pilus membrane protein genes are not found within this operon (Nair et al. 2014). The novel additional essential genes found in the pili locus in *M. maripaludis*, the presence of at least four pilin structural proteins, coupled with the separate locations of both the major pilin and the ATPase and membrane component genes and the presence of two essential copies of the membrane component gene (Nair et al. 2014) all suggest that the pili of this methanogen may be more complex than those found in *Sulfolobus* species.

In addition to the four identified pilin genes (*epdA*, *epdB*, *epdC* and *mmp1685*), the initial *in-silico* study also identified six other genes in the *M. maripaludis* genome that were predicted to be type IV pilin-like genes processed by EppA (*mmp0528, mmp0600, mmp0601, mmp0709, mmp0903 and mmp1283*) (Szabo et al. 2007). Whether these genes encode additional essential structural components of the type IV pili assembled mainly from MMP1685 subunits has not yet been addressed. These genes were the focus of the current study where it was shown that only *mmp1283* was essential for pili formation.
Materials and Methods

Strains and growth conditions

*Methanococcus maripaludis* MM900 (Moore and Leigh 2005) and a non-archaellated ΔflaK mutant strain derived from MM900 (Ng et al. 2009) were grown in Balch medium III (Balch et al. 1979) at 35°C under a headspace gas of CO₂/H₂ (20:80). For transformations, cells were grown in McCas medium (Moore and Leigh 2005) supplemented at various steps with neomycin (1 mg/ml) or 8-azohypoxanthine (240 μg/ml) for selection. For complementation experiments, transformants were grown in Balch medium III with added puromycin (2.5μg/ml) to select for uptake of the vectors. *E. coli* TOP10 cells (Invitrogen), used for various cloning steps, were grown in Luria-Bertani medium supplemented with ampicillin (100 μg/ml) as needed for transformations.

Reverse transcriptase-PCR

RT-PCR was done to determine if all the six genes used in this study were transcribed. Primer pairs were designed to amplify an internal fragment of each particular gene. The template RNA was extracted from the wildtype cells using an RNeasy Mini Kit (Qiagen Inc. Canada Mississauga, ON) with optional DNase digestion as per the manufacturer’s instructions. A One-Step RT-PCR kit (Qiagen Inc.) was used to amplify the cDNA. Using the same primer pairs, additional templates were also used in PCR reactions, including purified RNA not subjected to the reverse transcriptase step (as a control for DNA contamination of the RNA preparation) and genomic DNA to verify amplicon size and specificity of the primers.

Construction of gene deletion plasmids.

Plasmids were generated as described earlier (Moore and Leigh 2005, VanDyke et al. 2008) to make inframe deletions of each of the targeted potential minor pilin genes. The P1/P2 primers
(Table 4.1) were used to amplify an approximately 1 kb fragment upstream of the target gene and the P3/P4 primers to amplify an approximately 1 kb fragment downstream of the gene. P2 and P3 primers were designed with added AscI restriction sites so that the upstream and downstream fragments could be digested with AscI and then ligated together, resulting in an inframe deletion of most of the targeted gene. The ligated piece was used as template for another PCR using the P1 and P4 primers, which were designed with added BamHI sites. The approximately 2 kb PCR product was purified, digested with BamHI and cloned into pCRPrtNeo to create the plasmids (Table 4.2) used in generating the deletion strains (Moore and Leigh 2005).

*M. maripaludis* mutant generation

The pCRPrtNeo plasmid derivatives carrying the in-frame deletions of possible minor pilins were transformed into *M. maripaludis* ΔflaK using the PEG precipitation method (Tumbula et al. 1994). Following growth overnight without selection, the transformation mixture was subcultured into McCas medium with neomycin to select for plasmid integration. Subsequently, the culture was used to inoculate McCas medium without antibiotic selection. After growth overnight in non-selective McCas medium, aliquots were plated onto 8-azahypoxanthine (240 µg/ml)-containing McCas-Noble agar plates and incubated at 37°C for a week in an anaerobic steel canister. Individual transformant colonies were subsequently picked and grown in Balch III medium. Cells from the various individual colonies were washed, resuspended in 2% NaCl and used as template for PCR using the respective sequencing primers (Table 4.1) which were designed to amplify across the targeted genes. Amplicons were analyzed by agarose gel electrophoresis to identify the smaller fragments expected of the deletion mutants. Potential mutants identified this way were restreaked for purity, single colonies again picked,
Table 4.1: Primers used in this study. The restriction sites used are underlined (GGATCC – BamH1; GGCACGGCC – Asc1; ATGCAT – Nsi1; ACGCGT – Mu1).

<table>
<thead>
<tr>
<th>TECHNIQUE AND PRIMER NAME</th>
<th>SEQUENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>RT-PCR</td>
<td></td>
</tr>
<tr>
<td>mmp0528_ RT_ for</td>
<td>5’-AATGGAATGGGAATCTTTGTTG</td>
</tr>
<tr>
<td>mmp0528_ RT_ rev</td>
<td>5’-AACGTTTATAAGGGCATTACTG</td>
</tr>
<tr>
<td>mmp0600_ RT_ for</td>
<td>5’-GTGACATATTATCGACGTTAGG</td>
</tr>
<tr>
<td>mmp0600_ RT_ rev</td>
<td>5’-GTGATTTCTCAGATATGGTACAG</td>
</tr>
<tr>
<td>mmp0601_ RT_ for</td>
<td>5’-GGAACGATATTCGCGCAGACATG</td>
</tr>
<tr>
<td>mmp0601_ RT_ rev</td>
<td>5’-CTGGGTGAACCACATCAGGTAAGG</td>
</tr>
<tr>
<td>mmp0709_ RT_ for</td>
<td>5’-GAGACTACTGCTGTTAATGATGTCAGG</td>
</tr>
<tr>
<td>mmp0709_ RT_ rev</td>
<td>5’-CGTAAGTAGTACTGTCGCACACCAGTTC</td>
</tr>
<tr>
<td>mmp0903_ RT_ for</td>
<td>5’-CCAGATATCTCTTTGACCTTGG</td>
</tr>
<tr>
<td>mmp0903_ RT_ rev</td>
<td>5’-CTGCACCTGTTATACATGGCAGG</td>
</tr>
<tr>
<td>mmp1283_ RT_ for</td>
<td>5’-GCTGTTCTTGCAGTCCATACAG</td>
</tr>
<tr>
<td>mmp1283_ RT_ rev</td>
<td>5’-AAGCATCTATTGTCTGATTGAG</td>
</tr>
</tbody>
</table>

Screening of deletion mutants

<table>
<thead>
<tr>
<th>TECHNIQUE AND PRIMER NAME</th>
<th>SEQUENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>mmp0528_ seq_ for</td>
<td>5’-GGATCATGGGAGATGACC</td>
</tr>
<tr>
<td>mmp0528_ seq_ rev</td>
<td>5’-GTGCACGGACATACATCAGG</td>
</tr>
<tr>
<td>mmp0600_ seq_ rev</td>
<td>5’-TACTGCGACTATTATACATTGG</td>
</tr>
<tr>
<td>mmp0600_ seq_ rev</td>
<td>5’-GAAATAGTATTTTGCAGTCTAGGTTCCG</td>
</tr>
<tr>
<td>mmp0601_ seq_ for</td>
<td>5’-TGACTACGTTATATTGGAAGAG</td>
</tr>
<tr>
<td>mmp0601_ seq_ rev</td>
<td>5’-GTATTAATCTCATGAAAATCTGG</td>
</tr>
<tr>
<td>mmp0709_ seq_ for</td>
<td>5’-GAGTATGCCCTTCATGAGGAG</td>
</tr>
<tr>
<td>mmp0709_ seq_ rev</td>
<td>5’-GTTCCGTCACGGTAATTACC</td>
</tr>
<tr>
<td>mmp0903_ seq_ for</td>
<td>5’-CCGACCTAAATTTAGGAG</td>
</tr>
<tr>
<td>mmp0903_ seq_ rev</td>
<td>5’-CCGAGAATTTGCACCTTGGTACG</td>
</tr>
<tr>
<td>mmp1283_ seq_ for</td>
<td>5’-GAAGTTTGTATTGGTATAC</td>
</tr>
<tr>
<td>mmp1283_ seq_ rev</td>
<td>5’-GAGACATGGTACTGCAATCG</td>
</tr>
<tr>
<td>In-frame deletions/plasmids</td>
<td>mmp0528_P1</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>------------</td>
</tr>
<tr>
<td></td>
<td>mmp0528_P2</td>
</tr>
<tr>
<td></td>
<td>mmp0528_P3</td>
</tr>
<tr>
<td></td>
<td>mmp0528_P4</td>
</tr>
<tr>
<td></td>
<td>mmp0600_P1</td>
</tr>
<tr>
<td></td>
<td>mmp0600_P2</td>
</tr>
<tr>
<td></td>
<td>mmp0600_P3</td>
</tr>
<tr>
<td></td>
<td>mmp0600-P4</td>
</tr>
<tr>
<td></td>
<td>mmp0601_P1</td>
</tr>
<tr>
<td></td>
<td>mmp0601_P2</td>
</tr>
<tr>
<td></td>
<td>mmp0601_P3</td>
</tr>
<tr>
<td></td>
<td>mmp0601_P4</td>
</tr>
<tr>
<td></td>
<td>mmp0709_P1</td>
</tr>
<tr>
<td></td>
<td>mmp0709_P2</td>
</tr>
<tr>
<td></td>
<td>mmp0709_P3</td>
</tr>
<tr>
<td></td>
<td>mmp0709_P4</td>
</tr>
<tr>
<td></td>
<td>mmp0903_P1</td>
</tr>
<tr>
<td></td>
<td>mmp0903_P2</td>
</tr>
<tr>
<td></td>
<td>mmp0903_P3</td>
</tr>
<tr>
<td></td>
<td>mmp0903_P4</td>
</tr>
<tr>
<td></td>
<td>mmp1283_P1</td>
</tr>
<tr>
<td></td>
<td>mmp1283_P2</td>
</tr>
<tr>
<td></td>
<td>mmp1283_P3</td>
</tr>
<tr>
<td></td>
<td>mmp1283_P4</td>
</tr>
</tbody>
</table>

**Complementation**

<table>
<thead>
<tr>
<th>mmp1283_comp_for</th>
<th>5'-CGGATCCGGATATTACCGAAAGGAAAGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>mmp1283_comp_rev</td>
<td>5'-AGGCGCGTTTTAAACTTACCTATATGAGGAGC</td>
</tr>
</tbody>
</table>
Table 4.2: Plasmids used in this study

<table>
<thead>
<tr>
<th>PLASMID</th>
<th>DESCRIPTION/GENOTYPE</th>
<th>SOURCE OR REFERENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>pCRPrtNeo</td>
<td><em>hmv</em> promoter-<em>hpt</em> fusion plus Neo&lt;sup&gt;+&lt;/sup&gt; cassette in pCR2.1Topo; Amp&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Moore and Leigh, 2005</td>
</tr>
<tr>
<td>pKJ976</td>
<td>pCRPrtNeo with in-frame deletion of <em>mmp1283</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKJ1016</td>
<td>pCRPrtNeo with in-frame deletion of <em>mmp0600</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKJ1053</td>
<td>pCRPrtNeo with in-frame deletion of <em>mmp528</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKJ1056</td>
<td>pCRPrtNeo with in-frame deletion of <em>mmp0903</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKJ1068</td>
<td>pCRPrtNeo with in-frame deletion of <em>mmp0709</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKJ1104</td>
<td>pCRPrtNeo with in-frame deletion of <em>mmp0601</em></td>
<td>This study</td>
</tr>
<tr>
<td>pWLG40</td>
<td><em>hmv</em> promoter-<em>lacZ</em> fusion plus Pur cassette; Amp&lt;sup&gt;+&lt;/sup&gt;</td>
<td>John Leigh</td>
</tr>
<tr>
<td>pKJ1007</td>
<td>pWLG40 with <em>mmp1283</em> complement</td>
<td>This study</td>
</tr>
</tbody>
</table>
grown in Balch medium III and screened by PCR with the sequencing primers.

Complementation of the ∆mmp1283 deletion strain

Complementation of the ∆mmp1283 gene deletion were done in pWLG40 (Lie and Leigh 2003) in which the complementing gene is under the control of the constitutive, strong hmv promoter (Chaban et al. 2007, Lie et al. 2005, Ng et al. 2011). For this, the mmp1283 gene was amplified using the forward and reverse complementation primers (Table 4-1) containing added NsiI and MluI restriction sites, respectively. The PCR product was digested with NsiI and MluI and cloned into pWLG40, to generate pKJ1007. This plasmid was transformed into M. maripaludis ∆mmp1283 using puromycin for selection.

Electron microscopy.

Overnight cultures were washed with 50 mM MgSO₄, and negatively stained with 2% phosphotungstic acid. Cells were examined on Formvar-coated gold grids and imaged under a Hitachi 7000 electron microscope operating at an accelerating voltage of 75 kV.

Results

Using the FlaFind program, Szabo et al. 2007 (Szabo et al. 2007) identified 14 proteins in M. maripaludis that had class 3 signal peptides characteristic of archaellins and bacterial type IV pilins. Of these 14 genes, three were previously identified as archaellins (Chaban et al. 2007) and several were already shown to be involved in pili formation (Ng et al. 2011). The latter included epdA, epdB and epdC as well as the major pilin gene mmp1685. One of the 14 genes is a NAD+ synthase-related protein while the remaining six genes encode type IV pilin-like proteins with a DUF361 Pfam domain. The six pilin-like proteins were predicted to be processed by EppA, a pre pilin peptidase that was already shown to process the DUF361 domain-
containing pilins EpdA and EpdC (Szabo et al. 2007). The six genes under study here as 
encoding potential pili structural proteins are mmp0528, mmp0600, mmp0601, mmp0709, 
mmp0903 and mmp1283. All of these proteins have a class 3 signal peptide of 5-13 amino 
acids ending with a glycine, a conserved +5 glutamic acid, and a conserved +1 glutamine 
(Figure 4.1), which may be required for EppA processing (Szabo et al. 2007). Of the six genes, 
mmp0600 and mmp0601 appear to be in an operon while the four remaining genes are not. 
MMP0528, MMP0903 and MMP1283 are very small proteins of 67-76 amino acids in length (54- 
63 amino acids after signal peptide removal), almost identical to the size of the major structural 
pilin, MMP1685 which is 74 amino acids in length (62 amino acids after signal peptide removal). 
The minor pilins already identified (EpdA, EpdB and EpdC) are about twice as long (130-156 
amino acids). MMP0600, MMP0601 and MMP0709 are much larger proteins (200-299 amino 
acids). MMP1685 is known to be a glycoprotein with an attached N-linked pentasaccharide 
identical in structure to the tetrasaccharide identified attached to archaellins (Kelly et al. 2009) 
but with an additional hexose attached as a branch to the linking sugar N-acetyl-galactosamine 
(Ng et al. 2011). Analysis of the sequence of the 6 putative minor pilins identified 1-6 N- 
glycosylation sequons (N-X-S/T, where X is not proline) which indicates that all of these proteins 
could also be N-glycosylated (Figure 4.1).

As a first step to determining that the 6 putative minor pilin genes corresponded to true 
genes, evidence for an mRNA transcript of each gene was sought using RT-PCR since 
detection of transcript provides support that an ORF indeed encodes a true protein (Abu-Qarn 
and Eichler 2006). RT-PCR was performed on isolated total RNA using primers that would 
amplify an internal fragment of each gene. In all cases, a PCR product was obtained of the 
predicted size only when the RNA was subjected to a reverse transcription step and not from 
the RNA sample itself (Figure 4.2), indicating that the product arose from cDNA and not
Figure 4.1: Sequences of the six pilin-like proteins as well as the major pilin (MMP1685) of *M. maripaludis* studied in this report. The demonstrated or predicted signal peptides are shown in bold, conserved +1Q and +5E are shown in red and possible N-linked glycosylation sequons are highlighted in green.
Figure 4.2: RT-PCR analysis reveals that all six pilin-like genes are transcribed under standard laboratory conditions. Primers were designed that would amplify an internal fragment for each of the six genes. For each gene, the triplet of lanes indicates PCR products obtained using as template either genomic DNA, purified RNA subjected first to reverse transcription or purified RNA not subjected to a reverse transcription step. M indicates 1 kb DNA ladder from New England Biolabs.
contaminating genomic DNA present in the RNA preparation. The product of the RT-PCR was, in each case, identical in size to that obtained using genomic DNA as template. Sequencing of each PCR product confirmed their identities. Thus, all 6 putative pilin genes were transcribed under standard growth conditions in Balch medium III at 35°C.

Each of the 6 genes was then targeted for deletion. The parent strain for these deletions was *M. maripaludis ΔflaK* (Ng et al. 2011). This strain is deleted for *flaK* which encodes the prepilin peptidase essential for signal peptide removal from archaellins (Bardy and Jarrell 2002, Bardy and Jarrell 2003). Without this processing, the cells cannot assemble archaella so that the only surface structures remaining are pili. This makes analysis of effects on piliation by specific gene deletions easier to visualize. Transformants were screened using a PCR method with whole cells as template and primers that would amplify across the targeted genes. Successful deletion of each gene would result in a smaller PCR amplification product, whose size can be predicted from the site of the primers used in the PCR. Mutants carrying a deletion of each of the 6 putative minor pilin genes were obtained (Figure 4.3).

To investigate whether any of the targeted genes played an essential role in piliation, all mutants were examined by electron microscopy for the presence and abundance of pili. In bacterial type IV pili systems usually there is one major pilin and a number of minor pilins (Burrows 2012b, Giltner et al. 2012). The major pilin, as well as three minor pilins, were already identified in *M. maripaludis* so it seemed unlikely that all six putative minor pilin genes studied here would be involved in assembly of the MMP1685 pili as this would result in a total of ten different structural proteins. The electron microscopic examination of the various mutants supported this contention as only the strain carrying the deletion of *mmp1283* was nonpiliated (Figure 4.4). Strains with deletions in *mmp0528, mmp0600, mmp0601, mmp0709* and *mmp0903* were all piliated to the extent of the parent *M. maripaludis ΔflaK* cells (compare piliation here to that seen in *M. maripaludis ΔflaK* cells in Figure 4.5). In the case of the
Figure 4.3: Confirmation of the deletion of each of the six pilin-like genes. PCR reactions used whole cells of the wildtype or the deletion strains as template with gene specific primers. In the case of each gene, the first lane is the PCR product obtained with DNA from wildtype cells as template and the second lane is the PCR product obtained with DNA from the deletion strain as a template. In all cases a smaller PCR product is obtained for the deletion strain and the predicted sizes of the amplicons were obtained. M indicates 1 kb DNA ladder from New England Biolabs.
Figure 4.4: Electron micrographs of strains carrying deletions of each of the six pilin-like genes.

An enlargement of a portion of each mutant cell is presented below the intact cell to enhance visualization of pili. Arrows indicate pili on the cell surface. Only the *M. maripaludis Δmmp1283* strain is nonpiliated.

Bar, 0.5µm.
Figure 4.5: Electron micrographs showing that complementation restores piliation to the *M. maripaludis* Δmmp1283 strain. The *M. maripaludis* ΔflaK strain (non-archaeellated) used as the parent for the pilin gene deletion studies is shown for comparison. An enlargement of a portion of each mutant cell is presented below the intact cell to enhance visualization of pili. The *M. maripaludis* Δmmp1283 strain was complemented with a plasmid-borne wildtype version of the *mmp1283* gene under the control of the constitutive *hmv* promoter. Arrows indicate pili on the cell surface. Bar, 0.5µm.
mmp1283 deletion strain, complementation with a wildtype version of mmp1283 supplied in trans restored the cells to a piliation state comparable to \( M. \ maripaludis \Delta flaK \) cells (Figure 4.5)

**Discussion**

*Methanococcus maripaludis* is known to have at least two surface appendages that are assembled in a bacterial type IV pili mode, namely archaella and type IV-like pili (Jarrell et al. 2011, Jarrell et al. 2013, Ng et al. 2008). Unusual for Archaea is that the processing of the structural subunits, i.e. archaellins and pilins, in *M. maripaludis* has been demonstrated to occur through the actions of two different prepilin peptidase-like enzymes whose substrate specificities do not overlap. FlaK processes only archaellins and EppA only pilins (Bardy and Jarrell 2002, Ng et al. 2007, Szabo et al. 2007). In other studied Archaea, a single enzyme designated PibD is thought to remove the signal peptide from all pilin-like substrates, including archaellins (Albers et al. 2003, Lassak et al. 2012a, Tripepi et al. 2010). As more Archaea are studied, it seems likely that the division of labor in processing prepilin-like substrates by two separate prepilin peptidase-like enzymes reported so far only in *M. maripaludis* may be found in other members of the domain. It has been reported that several members of the Euryarchaeota, mainly Methanococcales as well as *Pyrococcus* and *Thermococcus* species harbor both *flaK* and *eppA* homologs in their genomes (Szabo et al. 2007). There are also a limited number of Euryarchaeota that have been reported to possess more than one copy of *flaK* (Desmond et al. 2007) but the roles and substrates of these potential prepilin peptidases, some found in species reported to be non-archaellated cells, has not been studied.

Previous studies have demonstrated roles for four pilins in the biosynthesis of the *M. maripaludis* pili. Genes encoding three minor pilins, EpdA, EpdB and EpdC are found in a single large gene cluster (Szabo et al. 2007) that also includes EppA and several other genes shown
to be essential for pili formation (Nair et al. 2014). Deletions of the genes for the three pilins result in either completely nonpiliated cells or cells in which the number of pili is significantly reduced (Ng et al. 2011). The major structural protein was identified as MMP1685 and deletion of mmp1685 led to nonpiliated cells. Interestingly, MMP1685 had been previously identified by bioinformatics analysis as a predicted substrate for EppA (Szabo et al. 2007). In addition, that study also predicted that six other genes encoded pilin-like proteins likely to be EppA substrates. In this report, deletions were created in all six genes to investigate the potential role of the encoded pilin-like proteins in the biosynthesis of the surface pili of M. maripaludis.

Three of the pilin-like proteins MMP0528, MMP0903 and MMP1283 are of similar size to the previously identified major pilin structural protein MMP1685 (Ng et al. 2011) while the other three (MMP0600, MMP0601 and MMP0709) are much larger. The smaller pilin sizes are typical lengths for type IV pilins of the Flp (Tad) class and the presence of +1 glutamine is also common in Gram positive Flp pilins (Burrows 2012a, Giltner et al. 2012, Imam et al. 2011). M. maripaludis pilins, however, lack the +6 tyrosine and the so called Flp motif of Flp pilins. All six pilin-like proteins possess a +5 glutamic acid, conserved in most bacterial type IV pilins (Giltner et al. 2012) but absent in the pilins of both Sulfolobus (Henche et al. 2012a, Pohlschroder et al. 2011) and Haloferax (Esquivel et al. 2013). Examination of mutant strains carrying deletions of each of the targeted genes by electron microscopy indicated that only mmp1283 was essential for piliation as all other deletion strains had similar numbers of pili per cell as wildtype cells. A piliated state could be restored to the ∆mmp1283 strain by supplying a wildtype copy of the gene in trans under the control of a constitutive hmv promoter. Like the major pilin MMP1685 and all the previously identified minor pilins (EpdA, EpdB, and EpdC), MMP1283 carries an amino acid sequon necessary for N-linked glycan attachment, suggesting that MMP1283 may be modified by the pentasaccharide found attached to MMP1685 (Ng et al. 2011). While the M. maripaludis major pilins are modified with the N-linked glycan, this posttranslational modification
is not needed for pilus formation as pili are formed even in an aglB mutant that is missing the oligosaccharyltransferase necessary to transfer the glycan from its lipid carrier to the protein target (Vandyke et al. 2009). It is not yet known if these assembled pili, however, are functional in surface attachment, the only known function attributed to the pili (Jarrell et al. 2011).

The function of the other pilin-like genes shown not to be essential for the MMP1685 pili is unknown but several possibilities exist. They could still be involved in the MMP1685 pili structure but dispensable proteins. In Neisseria meningitidis type IV pili, five different pilin proteins are necessary for piliation while three others that are normally incorporated into the pilus are dispensable for piliation but play important roles in function. Interestingly, these three minor pilins (PilX, ComP and PilV) all have different functions (Brown et al. 2010). Alternatively, the pilin-like proteins could be structural proteins of an entirely separate pilus-like structure. It is known in the thermoacidophilic archaean Sulfolobus acidocaldarius that two different types of type IV pili are made (Ajon et al. 2011, Henche et al. 2012a, Lassak et al. 2012a) with one type (Ups pili) only made after UV induction or other DNA damaging treatment (Ajon et al. 2011, Frols et al. 2008). Perhaps, under still undefined growth conditions or stress, M. maripaludis has the capacity to assemble a novel pilus type composed of one or more of the remaining pilin-like proteins that currently have no function. Although mRNA transcripts were detected for all five of the other pilin-like genes by RT-PCR, it is possible that posttranscriptional regulation mechanisms prevent pilin protein synthesis. The regulation of archaella assembly, for example, in S. acidocaldarius seems to involve both transcriptional and posttranscriptional control (Lassak et al. 2012a, Reimann et al. 2012). Alternatively, in the case of the five pilin-like proteins for which deletion did not affect formation of MMP1685 pili, other key components essential for pili formation from these pilin-like proteins may not be made under tested growth conditions, even though the pilins themselves are made. In the case of S. acidocaldarius, archaella synthesis is induced under tryptone starvation conditions even though archaella core
proteins are constitutively produced. This is because the major structural protein (the archaellin FlaB) is only made under starvation conditions (Lassak et al. 2012a). A third possibility is that M. maripaludis makes a very short pilus-like structure from one or more of these proteins that has gone undetected by electron microscopy. In Sulfolobus solfataricus, it is known that sugar binding proteins are pilin-like glycoproteins that form a macromolecular cell-surface-associated structure (Lassak et al. 2012a, Zolghadr et al. 2011) that may form a short pilus-like structure that extends only from the cytoplasmic membrane to the S-layer (Ng et al. 2008). The putative bindosome pilus-like structure has never been observed in electron microscopic studies. Type two secretion systems in bacteria use type IV pilin-like proteins to produce a very short pilus-like piston (pseudopilus) proposed to push exoproteins through an outer membrane channel (Korotkov et al. 2012). This pseudopilus would extend only from the cytoplasmic membrane to the outer membrane and likely be dynamic in nature. While M. maripaludis does not utilize sugars as substrates or possess a type II secretion system, there may be other functions in the cells that may require such a short type IV pilus-like structure. The recent identification of putative diverse type IV pili in a variety of Gram positive bacteria using a program called PilFind (Imam et al. 2011) suggest that all possible functions for these structures have not likely been identified yet.

In bacteria, it is not unusual for type IV pili to be composed of a major pilin and multiple minor pilins. For example, in P. aeruginosa, the pilus is comprised of the major pilin PilA and five minor pilins (FimU, PilV, PilW, PIIX and PilE) which were all shown to be incorporated into the pilus by immunogold labelling experiments (Giltner et al. 2010). Among the studied Archaea, however, the type IV pilus locus of M. maripaludis appears to be considerably more complex than the two gene clusters encoding Aap and Ups pili in Sulfolobus. In both the Aap and Ups pili systems, there appear to be only two pilin genes and they are encoded along with the conserved ATPase and membrane component genes. Interestingly, in the case of Aap pilins of
S. acidocaldarius and Ups pili of S. solfataricus, the lengths (138-168 amino acids with signal peptides) are much larger than seen with MMP1685. In the case of the six recently described Hfx. volcanii pilins, none are co-transcribed with the pilus ATPase and conserved membrane protein genes (Esquivel et al. 2013). The M. maripaludis pili genes are known to lie now in at least four separate locales around the chromosome. One major operon encoding EppA, EpdA, EpdB and EpdC along with other essential genes has already been analyzed to some extent (Ng et al. 2011, Szabo et al. 2007). Furthermore, there is a separate locus encoding the ATPase and two membrane protein components (Nair et al. 2014) as well as the major pilin subunit MMP1685 (Ng et al. 2011) and now, in this report, the minor pilin MMP1283. It is not yet known if type IV pili formation is constitutive in M. maripaludis or whether it can be induced or repressed under specific environmental circumstances such as attachment or planktonic conditions, as is observed in bacterial type IV pili systems (Giltner et al. 2012). If pili formation is not constitutive then the cells must regulate transcription of several essential gene clusters located at some distance from each other. Regulation of minor pilin expression has not been well studied even in bacteria. In type IVb systems where minor pilins are clustered with other components of the pilus system, they are likely co-regulated with them. However, in type IVa systems, minor pilins are often unlinked to other pilus component genes, as found in M. maripaludis, and they can be differentially regulated, sometimes by two-component systems (Belete et al. 2008, Burrows 2012a, Giltner et al. 2012).

In Sulfolobus species, studies on the regulation of pili have already been initiated, with intriguing findings reported. In S. acidocaldarius, a two component regulatory system (ArnA and ArnB) was found to repress archaella expression. Interestingly, overproduction of ArnA also resulted in a strong enhancement of Aap pili production, suggesting there is a regulation of the two surface organelles that involves cross-talk between the two systems (Reimann et al. 2012). Recently, the product of an Lrs14 regulator gene saci0446, was shown to bind to promoters of
both archaellum \((fla)\) genes and \(aap\) pili genes and result in an upregulation of \(aap\) genes and downregulation of \(fla\) genes, again showing that regulation of different surface structures in this archaeon is connected (Orell et al. 2013).

This report adds to our knowledge about the complexity of the type IV-like pili in \(M. maripaludis\) by identifying a fourth minor pilin that is essential for piliation. Unlike other archaeal systems, piliation in \(M. maripaludis\) requires a separate pilin-specific signal peptidase, two conserved membrane (platform) proteins, four minor pilins and additional novel essential proteins (Ng et al. 2011, Szabo et al. 2007). This report also eliminates five other pilin-like proteins as playing an essential role in MMP1685 pili formation and hints that they may form additional type IV pili-like surface structures under appropriate growth conditions.
Literature cited


Burrows LL (2012b) Prime time for minor subunits of the type II secretion and type IV pilus systems. Mol Microbiol 86:765-769


Jarrell KF, Stark M, Nair DB, Chong JP (2011) Flagella and pili are both necessary for efficient attachment of Methanococcus maripaludis to surfaces. FEMS Microbiol Lett 319:44-50


Addendum to Chapter 4

Identification of MMP1685 as the Major Pilin in *Methanococcus maripaludis*


While the lead author was Sandy Ng who conducted portions of this work as part of her Ph.D. thesis, my contribution to the paper was in determining that the gene *mmp1685* encoded the major pilin. This portion of the work is presented as an addendum to Chapter 4 since Chapter 4 identifies a new essential minor pilin while the following describes the identification of MMP1685 (subsequently renamed EdpE; Nair et al. 2014) as the major pilin.
Mass spectrometry analysis of purified pili isolated from a ΔflaK mutant of *M. maripaludis* identified the major component to be a small molecular mass protein with extensive posttranslational modification. The identified sequence matched that of MMP1685, a type IV pilin-like protein with a class III signal peptide, later renamed EpdE (Nair et al. 2014). EpdE possessed a +1Q (relative to the signal peptide cleavage site) as found in the known type IV pilins like MMP0233 (EpdA), MMP0236 (EpdB) and MMP0237 (EpdC) and a +5E as found in the EpdA and EpdB proteins (Ng et al. 2011; Szabo et al. 2007). Surprisingly, analysis of the N-linked glycan attached to EpdE revealed that it was not identical to the tetrasaccharide attached to the archaellins but instead had an extra hexose as a branch linked to the N-acetylgalactosamine (Ng et al. 2011). It appeared that EpdE was glycosylated at three of the four potential N-linked glycosylation sites by this pentasaccharide of mass 1,196 Da (Figure 1).

If EpdE was the major pilin, then deletion of *epdE* would be expected to lead to nonpiliated cells. Therefore, a complementation and deletion analysis was done to find out if *epdE* was essential for piliation. To make an inframe deletion of the gene, approximately 1 kb upstream of the gene was amplified using the primers P1 (5'-GGGGATCCCAATA
TCAACAGCTCAGCATCATCG; added BamHI restriction site is underlined) and P2 (5'-TTGGCGCGCTAAAGAATACCGAGTTCTGCTATTTGACC; added Ascl restriction site is underlined) while an approximately 1 kb fragment downstream of the gene was amplified using the primers P3 (5'-TTGGCGCGCTCAAAATATTACCGGATTATATATC; added Ascl restriction site is underlined) and P4 (5'-CCGGATCCCTATTACTTCATTTGCACTTTGTTGG; added BamHI restriction site is underlined). The upstream and the downstream fragments were purified and ligated together following Ascl digestion. This product was used in PCR with primers P1 and P4 to generate a 2kb fragment in which a large internal portion of the *epdE* gene
A. Cleavage site of the prepilin peptidase homolog in *M. maripaludis* pilins.

MMP0233 (EpdA)  MFKNFNRG  QISFE

MMP0236 (EpdB)  MSKQ  QVSVE

MMP0237 (EpdC)  [MIK]MLQLPFNKKQ  QVSFD

MMP1685 (EpdE)  MKFLEKLTSDKG  QIAM

B. Presence of potential sites for N-linked glycosylation in *M. maripaludis* pilin. Sites are bolded.

> MMP1685 (EpdE)
MKFLEKLTSDKGQIAMELGILVMAAVAAIAAYFYANVSTGQITNSTDQTTQALADAISSDNTATQSMSNITD

Addendum Figure 1: Posttranslational modifications of *M. maripaludis* pilins.

A. Alignment of the N-terminus of the *M. maripaludis* type IV pilins showing the cleavage sites for processing by the prepilin peptidase homolog, EppA (indicated by the arrow head). B. Presence of potential N-linked glycosylation sites. Sites in EpdE are shown in bold type. (From Ng et al. 2011)
was deleted. This fragment was cloned into the BamHI site of plasmid pCRPrtNeo creating plasmid pKJ836.

This plasmid was used to transform *M. maripaludis ΔflaK* and mutants carrying a deletion in *epdE* were found by screening individual transformant colonies by PCR using the primers mmp1685_seq_for (5’-AATAGGGAGTATGGTGTAGTCTGG) and mmp1685_seq_rev (5’-AATGAATCGTTCTGATTGCCTG) (Figure 2). Mutants were restreaked for purity and the *epdE* deletion reconfirmed by PCR, using various PCR primer pairs (Figure 2). The ΔflaKΔepdE deletion strain was shown by electron microscopy to no longer produce pili at the cell surface (Figure 3A, 3B). Pili production was restored to the ΔflaKΔepdE mutant strain upon complementation with a plasmid borne wildtype version of *epdE* gene (pKJ880) that was amplified by PCR with the primers, mmp1685_comp_for (5’-CCAATGCATGAAATTTTAGAAAAAATACACATC) and mmp1685_comp_rev (5’-AGCACGCGTTAATCCGTAATTTGACATTGTGAGG) using wildtype (MM900) *M. maripaludis* cells as template (Figure 3C, 3D). In this plasmid construct, expression of the *epdE* gene is under control of the strong, constitutive *hmv* promoter present in plasmid pWLG40 (Lie et al. 2005).

The deletion and complementation analysis of *epdE*, coupled with the identification of EpdE in the purified pili samples by mass spectrometry and electron microscopy of the nonpiliated ΔepdE mutant, all indicate that EpdE is the major structural protein of the Epd pili of *M. maripaludis*. 

124
Addendum Figure 2: PCR analysis of wildtype and the epdE deletion strain. Figure shows location of various primer pairs for amplification. Primer pairs 2 and 3 have forward (Pfor: 5’-GACAACACAAGCGCTAGCTGACG) and reverse (Prev: 5’-CGATTGCTGCAACTGCCACTG) primers, respectively, located within the deleted portion of epdE and so give predicted size products with wildtype cells as template but not with the epdE deletion strain cells as template. Primer pair 1 has both primers located outside of the epdE deletion and so amplify products with either wildtype cells or the epdE deletion strain as template, with the product reduced in size in the deletion strain. M indicates 1 kb DNA ladder from New England Biolabs.
Addendum Figure 3: Electron microscopy examination of *M. maripaludis ΔflaKΔepdE* mutant cells carrying the vector control (A, B) and carrying the complementation construct (C, D). The images shown in B and D are enlargements of the boxed areas depicted in A and C, respectively. Arrows indicate pili. Samples were negatively stained with 2% phosphotungstic acid (pH 7.0). Bar: 200 nm. (From Ng et al. 2011)
Literature cited


Chapter 5

Pilin Processing Follows a Different Temporal Route Than That of Archaellins in the Archaeon *Methanococcus maripaludis*
Abstract

Unlike other studied archaea, *Methanococcus maripaludis* has two prepilin peptidase-like enzymes FlaK and EppA that cleave signal peptides from its two type IV pilin-like proteins. FlaK specifically processes the signal peptides of the archaellins and EppA processes the pilin signal peptides. While pilins and archaellin share significant similarities in their signal peptide and N-termini, the prepilin peptidases are not interchangeable. We investigated the possible key role that the +3 position (relative to the signal peptide cleavage site) played in making a type IV pilin-like protein a FlaK substrate. Epitope-tagged mutant pilin EpdE, carrying the +3 amino acid of archaellins was expressed in various mutant backgrounds (strains deleted for *eppA*, *flaK* or both genes) and examined for processing. While this single change did not convert the pilin to a substrate for FlaK, the results indicated that signal peptide removal was necessary before the pilin could be glycosylated. This was further investigated by expressing epitope-tagged pilins in mutant backgrounds where signal peptide removal (*ΔeppA*), N-linked glycosylation (*ΔaglB*) or both posttranslational modifications (*ΔeppAΔaglB*) were prevented. The data indicated that while signal peptide removal was required before the pilins could be glycosylated, glycosylation was not necessary for signal peptide removal. Unexpectedly, this is unlike the situation for archaellins, where the two posttranslational modifications were shown to occur independently of each other.
Introduction

*M. maripaludis* possesses two surface structures that have been studied, namely the archaella (formerly archaeal flagella (Jarrell and Albers 2012) and type IV-like pili (Jarrell et al. 2011, Jarrell et al. 2013, Nair et al. 2013, Nair et al. 2014, Ng et al. 2011). Both of these structures are believed to be assembled via a bacterial type IV pilus-like model (Burrows 2005, Jarrell et al. 2013, Lassak et al. 2012a). Indeed, the many similarities of archaella to bacterial type IV pili and their lack of homology to bacterial flagella, other than their shared involvement in swimming motility, initially led to the proposal of a distinct name for the archaeal structure (Jarrell and Albers 2012). Both archaella and archaeal type IV pili have structural proteins synthesized initially as preproteins with a class III (type IV pilin-like) signal peptide which is subsequently cleaved by a dedicated prepilin peptidase-like enzyme (Albers et al. 2003, Bardy and Jarrell 2002, Bardy and Jarrell 2003, Esquivel et al. 2013, Ng et al. 2007, Szabo et al. 2007, Tripepi et al. 2010). Also shared with type IV pili systems of bacteria is the presence of an essential ATPase involved in polymerization of the subunits into a filament and a conserved membrane platform protein thought to interact with the ATPase (Albers and Pohlschroder 2009, Banerjee et al. 2012, Chaban et al. 2007, Ghosh and Albers 2011, Ghosh et al. 2011, Lassak et al. 2012b, Patenge et al. 2001, Thomas et al. 2002).

In *M. maripaludis*, the archaella are composed of three structural glycoproteins (the archaellins, FlaB1, FlaB2 and FlaB3) all modified at multiple positions with an N-linked
tetrasaccharide (Chaban et al. 2007, Kelly et al. 2009, Vandyke et al. 2009). The prepilin peptidase-like enzyme responsible for signal peptide removal from archaellins was identified as FlaK and its activity shown to be necessary for archaella assembly (Bardy and Jarrell 2003, Ng et al. 2009). Gene deletion studies, coupled with mass spectrometry of purified archaella filaments, have identified a number of steps in both the assembly of the glycan (glycosyltransferases, oligosaccharyltransferase; (Vandyke et al. 2009)) and in the pathways to synthesize the individual sugar components (Ding et al. 2013, Jones et al. 2012, VanDyke et al. 2008). Electron microscopy studies have further demonstrated that interference with the N-linked glycosylation pathway has severe effects on archaellation. In mutants of *M. maripaludis* where the glycan structure has been truncated to less than two sugars, cells cannot assemble archaella (Vandyke et al. 2009).

Less numerous and thinner than archaella are the Epd pili of *M. maripaludis* (Nair et al. 2014). At least five different type IV pilin-like genes (*epdA, epdB, epdC, epdD* and *epdE*) have been shown by deletion analysis to be necessary for normal piliation (Nair et al. 2013, Ng et al. 2011), with *epdE* (*mmp1685*) shown by mass spectrometry analysis of purified pili to encode the major structural subunit (Ng et al. 2011). Mass spectrometry analysis of purified pili also revealed that EpdE was a glycoprotein with multiple sites occupied by an N-linked glycan (Ng et al. 2011). Unexpectedly, however, the attached glycan was not identical to the tetrasaccharide of the archaellins (Kelly et al. 2009) but instead was a pentasaccharide that had an extra unidentified hexose attached to the linking sugar of the archaellin tetrasaccharide (Ng et al. 2011).

Unusually, the pilins have their own dedicated prepilin peptidase, EppA, which cleaves the signal peptide from pilins but not archaellins (Szabo et al. 2007). This is in contrast to other studied Archaea which have a single prepilin peptidase-like enzyme, PibD which cleaves all type IV pilin-like proteins, including archaellins, pilins and sugar binding proteins (Albers et al.
2003, Esquivel et al. 2013, Tripepi et al. 2010). For *M. maripaludis*, it was shown by clever swapping experiments of the -2 to +2 amino acids around the conserved cleavage site that mutant archaellins that contained the -2 to +2 amino acids of pilins could then be processed by EppA (Szabo et al. 2007). Similar switching of sites of archaellins to pilins however did not result in pilins that could be processed by FlaK. However, previous work in *Methanococcus voltae* established that the +3 glycine of the archaellin mature protein was essential for signal peptide removal in that methanogen (Thomas et al. 2001b). The failure of the -2 to +2 archaellin to pilin amino acid swap to lead to a hybrid pilin that could be processed by FlaK may be thus explained, since *M. maripaludis* pilins do not possess glycine at the +3 position (Nair et al. 2013, Ng et al. 2011, Szabo et al. 2007). Here, we investigated whether a simple change of the +3 amino acid of the pilin EpdE to a glycine might make it susceptible to processing by FlaK.

During the course of these studies, it became evident that pilins were not glycosylated unless they were first processed by EppA. This is in contrast to archaellins where it was shown that signal peptide-bearing archaellins of a flaK deletion strain were still glycosylated. Hence, it seems that pilins follow a different order of posttranslational modification that is unlike that of the other type IV pilin-like proteins in *M. maripaludis*, namely the archaellins.

**Materials and Methods**

Strains and growth conditions

*Methanococcus maripaludis MM900* (Moore and Leigh 2005) and various single and double deletion mutants derived from it were used in this study. These mutants include; ΔflaK, a non-archaellated strain deleted for the pre-archaellin peptidase (Ng et al. 2009), ΔeppA, a non-piliated strain deleted for the prepilin peptidase (Jarrell et al. 2011); ΔaglB, a strain lacking the oligosaccharyltransferase necessary for N-glycosylation (Vandyke et al. 2009) and the double mutants, ΔflaKΔeppA (Jarrell et al. 2011), ΔflaKΔaglB and ΔeppAΔaglB, the latter two created in
this study. All strains were routinely grown in Balch medium III (Balch et al. 1979) at 35°C under a headspace gas of CO₂/H₂ (20/80). For steps of the inframe mutant generation procedure as described below, cells were grown in McCas medium (Moore and Leigh 2005). In complementation experiments, transformants were selected by the addition of puromycin (2.5µg/ml) to Balch medium III. *Escherichia coli* TOP10 cells (Invitrogen), used for various cloning steps, were grown in Luria-Bertani medium supplemented with ampicillin (100 µg/ml) or ampicillin (100 µg/ml) and chloramphenicol (30 µg/ml) as needed. *E. coli* strain BL21 (DE3)/pLysS was used as host for the overexpression of a C-terminal histagged version of EpdE.

*M. maripaludis* mutant generation

Plasmid pKJ574 (Vandyke et al. 2009) was used for the generation of an inframe deletion of *aglB* in the ∆flaK strain (Ng et al. 2009) and plasmid pKJ697 for the generation of an eppA deletion in the pre-existing ∆aglB strain (Vandyke et al. 2009) to create markerless double mutants, using procedures as previously described (Moore and Leigh 2005, VanDyke et al. 2008). pKJ574 and pKJ697 were transformed into *M. maripaludis* ∆flaK and *M. maripaludis* ∆aglB, respectively, using the PEG precipitation method (Tumbula et al. 1994). Individual transformant colonies from the McCas plates containing hypoxanthine were inoculated into Balch medium III. Deletion mutants were identified by using washed whole cells resuspended in 2% NaCl as template for PCR along with sequencing primers (Table 5.1) designed to amplify across the targeted gene deletion. The PCR products were examined by agarose gel electrophoresis and the size compared to that predicted for the wildtype and the deletion version of *aglB*. Transformants that generated the predicted deletion size PCR product were restreaked onto Balch medium III plates and single colonies picked and again screened by PCR to confirm
Table 5.1: Primers used in this study for PCR.

<table>
<thead>
<tr>
<th>PRIMERS</th>
<th>PRIMER SEQUENCE 5’ TO 3’</th>
<th>RESTRICTION SITE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1685+3 sdm-for Nsi1</td>
<td>CCAATGCATGAAATTTTTAGAAAAACTAACATCAA AAAAGGTCAAATAGGAATGGGAACTCGG</td>
<td>Nsi1</td>
</tr>
<tr>
<td>1685_Mlu_rev</td>
<td>AGCACGCGTTTAATCCGTAATATTTGACATTGTG AGG</td>
<td>Mlu1</td>
</tr>
<tr>
<td>1685_Nsi1 for</td>
<td>CCAATGCATGAAATTTTTAGAAAAACTAACATCA</td>
<td>Nsi1</td>
</tr>
<tr>
<td>1685_histag_Mlu_rev</td>
<td>CGCGACGCCTTTAGGTGATGGTGATGGTGATGGTGATGC CGTAATATTTGACATTGTGAG</td>
<td>Mlu1</td>
</tr>
<tr>
<td>1685_FLAG_Mlu_rev</td>
<td>CGACGCCTTTTATTGTCATCGTCATTTGTAATCA TCCGTAATATTTGACATTGTGAG</td>
<td>Mlu1</td>
</tr>
<tr>
<td>1283_Nsi1_for</td>
<td>CCAATGCATGTCTGGCTTTAAAGAGGTTTTTTTGAAAACG</td>
<td>Nsi1</td>
</tr>
<tr>
<td>1283_FLAG_Mlu_rev</td>
<td>GCTACGCCTTTTATTGTCATCGTCATTTGTAATCA TCCGTAATATTTGACATTGTGAG</td>
<td>Mlu1</td>
</tr>
<tr>
<td>FlaK_seq_for</td>
<td>AATATCTGGCGGATACAGG</td>
<td></td>
</tr>
<tr>
<td>FlaK_seq_rev</td>
<td>TTCAAGCCATAGATACTGC</td>
<td></td>
</tr>
<tr>
<td>EppA_Seq_for</td>
<td>CTGGAGCTGTATGAAATGCAAC</td>
<td></td>
</tr>
<tr>
<td>EppA_Seq_rev</td>
<td>GGATGACCTGGGATAATGCGG</td>
<td></td>
</tr>
<tr>
<td>AglB_Seq_for</td>
<td>CATAAACCATATTTGTAATTAAC</td>
<td></td>
</tr>
<tr>
<td>AglB_Seq_rev</td>
<td>CTCAATAGCCATAAAATCACC</td>
<td></td>
</tr>
</tbody>
</table>
their purity. The pre-existing deletions of flaK and aglB in these double mutants were also confirmed by PCR using the sequencing primers listed in Table 5.1.

Complementation experiments

Plasmid pKJ711, containing eppA in pHW40 under the control of the nif promoter (Gardner and Whitman 1999), was used to complement the ΔeppA mutant, as reported previously (Nair et al. 2014).

Expression of epitope-tagged versions of EpdD and EpdE in various M. maripaludis strains

PCR using primers 1685+3_sdm_For_Nsi1 and either 1685_histag_mlu_rev or 1685_FLAG_Mlu_Rev was used to change the +3 position (relative to the prepilin peptidase cleavage site) of EpdE from alanine to glycine. The long forward primer (Table 5.1) was designed to include the desired change while the reverse primer was used to create either a Histag or FLAG tag at the C-terminus. NsiI and MluI restriction sites were added to the forward and reverse primers, respectively, to facilitate cloning steps. The PCR product was digested with NsiI and MluI and cloned into pWLG40, where its expression is driven by the strong constitutive hmv promoter (Lie et al. 2005). This created pKJ1079 for the His-tagged version of EpdE with the +3 glycine change and pKJ1108 for the FLAG-tagged version (Table 5.2). Successful complementation of a ΔepdE strain with a wildtype copy of the epdE gene (mmp1685) cloned into this vector was previously demonstrated (Ng et al. 2011).

A C-terminal histagged version of wildtype EpdE was generated by PCR using primers 1685_Nsil_For and 1685_histag_Mlu_Rev. The PCR product was digested with NsiI and MluI and cloned into the vector pWLG40. This plasmid was designated pKJ1072. Similarly, a C-terminal FLAG tagged version of EpdE was amplified by PCR using primers1685_Nsil_For and 1685_FLAG_Mlu_Rev and cloned into pWLG40 creating pKJ1107. A C-terminal FLAG tagged
### Table 5.2: Strains and plasmids used in this study

<table>
<thead>
<tr>
<th>STRAINS</th>
<th>DESCRIPTION AND / OR GENOTYPE</th>
<th>SOURCE OR REFERENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. coli</em> K113</td>
<td>BL21(DE3)/pLysS; expression host, Cm&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Novagen</td>
</tr>
<tr>
<td><em>Methanococcus maripaludis</em></td>
<td>∆\text{hpt}</td>
<td>Moore and Leigh 2005</td>
</tr>
<tr>
<td>Mm900</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>M. maripaludis</em> ∆\text{flaK}</td>
<td>Mm900.∆\text{flaK}</td>
<td>Ng et al. 2009</td>
</tr>
<tr>
<td><em>M. maripaludis</em> ∆\text{eppA}</td>
<td>Mm900.∆\text{eppA}</td>
<td>Jarrell et al. 2011</td>
</tr>
<tr>
<td><em>M. maripaludis</em> ∆\text{flaK} \text{eppA}</td>
<td>Mm900.∆\text{flaK} ∆\text{eppA}</td>
<td>Jarrell et al. 2011</td>
</tr>
<tr>
<td><em>M. maripaludis</em> ∆\text{aglB}</td>
<td>Mm900.∆\text{aglB}</td>
<td>Vandyke et al. 2009</td>
</tr>
<tr>
<td><em>M. maripaludis</em> ∆\text{flaK} \text{aglB}</td>
<td>Mm900.∆\text{flaK} ∆\text{aglB}</td>
<td>This study</td>
</tr>
<tr>
<td><em>M. maripaludis</em> ∆\text{eppA} \text{aglB}</td>
<td>Mm900.∆\text{eppA} ∆\text{aglB}</td>
<td>This study</td>
</tr>
</tbody>
</table>

**Plasmids**

<table>
<thead>
<tr>
<th>Plasmids</th>
<th>DESCRIPTION</th>
<th>SOURCE OR REFERENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>pET23a+</td>
<td>T7 promoter-based expression vector, Amp&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Novagen</td>
</tr>
<tr>
<td>pKJ900</td>
<td>pET23a+ with epdE Nde1-Xho1 fragment C-terminal histagged</td>
<td>This study</td>
</tr>
<tr>
<td>pWLG40</td>
<td>hmv promoter- lacZ fusion plus Pur&lt;sup&gt;+&lt;/sup&gt; cassette; Amp&lt;sup&gt;+&lt;/sup&gt;</td>
<td>William B. Whitman</td>
</tr>
<tr>
<td>pKJ880</td>
<td>pWLG40 with epdE complement</td>
<td>Ng et al. 2011</td>
</tr>
<tr>
<td>pKJ1072</td>
<td>pWLG40 with epdE C-terminal histagged complement</td>
<td>This study</td>
</tr>
<tr>
<td>PKJ1079</td>
<td>pWLG40 with epdE (+3 Gly) C-terminal histagged complement</td>
<td>This study</td>
</tr>
<tr>
<td>pKJ1107</td>
<td>pWLG40 with epdE FLAG complement</td>
<td>This study</td>
</tr>
<tr>
<td>pKJ1108</td>
<td>pWLG40 with epdE (+3 Gly) FLAG complement</td>
<td>This study</td>
</tr>
<tr>
<td>pHW40</td>
<td>\text{nif} promoter-lacZ fusion plus Pur&lt;sup&gt;+&lt;/sup&gt; cassette; Amp&lt;sup&gt;+&lt;/sup&gt;</td>
<td>John Leigh</td>
</tr>
<tr>
<td>pKJ1169</td>
<td>pHW40 with epdD FLAG complement</td>
<td>This study</td>
</tr>
</tbody>
</table>
version of the minor pilin EpdD (Nair et al. 2013) was generated by PCR using primers
1283_NsiI_For and 1283_FLAG_Mlu_Rev and the product cloned into pHW40 creating
pKJ1169. The various plasmids carrying epitope-tagged versions of the pilin genes (Table 5.2)
were transferred into the wildtype and mutant strains via the PEG procedure (Tumbula et al.
1994).

Overexpression and purification of His-tagged EpdE

The epdE gene was amplified by PCR from wildtype cells using primers 1685-for-exp
and 1685-rev-exp. The PCR product was digested with NdeI and XhoI and cloned into the
expression vector pET23a+, creating pKJ900. This plasmid was transformed into E. coli
strain BL21 (DE3)/pLysS and colonies that had taken up the plasmid selected by plating on LB agar
containing ampicillin (100 µg/ml) and chloramphenicol (30 µg/ml). Subsequently, these cells
were grown in LB broth with appropriate antibiotics at 37°C to an OD600 of 0.6–1 and induced
with 0.4 mM IPTG (Life Technologies, Burlington, ON, Canada) for 2 hrs. The induced cells
were collected by centrifugation and the resulting pellet frozen and subsequently thawed after
the addition of 5 ml 10 mM Tris–EDTA buffer. Lysis and shearing of DNA was aided by brief
sonication. After a low speed centrifugation to remove unbroken cells, a crude membrane
fraction was obtained by centrifugation at 18,000xg/30 minutes. Purification of EpdE from this
fraction was carried out using nickel affinity purification under denaturing conditions by using a
His-bind kit (Novagen), as directed by the manufacturer. Purification was followed by SDS-
PAGE (Laemmli 1970). Gels were stained with a Coomassie brilliant blue G250-perchloric acid
solution and destained in water (Faguy et al. 1996). Purified protein was concentrated and used
to raise antibodies in chickens (Chaban et al. 2007) (RCH Antibodies, Sydenham, Ontario
Canada). The anti-EpdE antibodies, isolated from yolks, were further subjected to affinity
purification. Purified Histagged EpdE was electrophoresed through a 1.5 mm thick gel (15%)
and blotted (Towbin et al. 1979) to supported nitrocellulose (0.22 micron, GE Water and
Process Technologies). Transferred protein was located on the membrane by Ponceau S staining (Harper and Speicher 2001). This region was cut into small pieces and blocked overnight at 4°C, washed and incubated with 1 ml crude antibody again at 4°C overnight. After extensive washing, the antibodies were removed by a 5 min incubation with 500 µl 0.2M glycine, pH 2.8, followed by neutralization using 125 µl 1M Tris-HCl, pH 8. These antibodies were used at 1:100 dilution in western blots.

Western blot analysis

*M. maripaludis* whole-cell lysates were electrophoresed through 15% SDS-PAGE and then transferred to an Immunobilon-P membrane (Millipore, Bedford, MA) (Towbin et al. 1979). Major archaellin FlaB2 was detected with chicken anti-FlaB2 specific antibodies (Jones et al. 2012). Major pilin EpdE was detected with chicken anti-EpdE specific antibodies, affinity purified as described above. Horseradish peroxidase-conjugated rabbit anti-chicken immunoglobulin Y (Jackson Immuno Research Laboratories, West Grove, PA) was used as secondary antibody. HisTagged proteins were detected using a rabbit polyclonal IgG against the hexahistidine tag (Santa Cruz, Biotechnology, Dallas Tx) as the primary antibody and horseradish peroxidase-conjugated goat anti-rabbit IgG (BioRad Laboratories, Mississauga, ON, Canada) as the secondary antibody. To detect FLAG-tagged proteins, a rabbit polyclonal Anti-FLAG primary antibody (SIGMA-Aldrich, St. Louis, MO, USA) and a horseradish peroxidase-conjugated goat anti-rabbit IgG secondary antibody (BioRad Laboratories) were used. All blots were developed with a chemiluminescent kit according to the manufacturer’s instructions (Roche Molecular Biochemicals, Laval, QC, Canada).
Results and Discussion

*M. maripaludis* has two prepilin peptidase-like enzymes. In addition to the one that processes archaellins, identified previously by Bardy and Jarrell (2002) and designated FlaK, Szabo et al (2007) discovered EppA which processes pilins. Both pilins and archaellins share significant amino acid sequence similarities in their signal peptides, including key conserved residues near the cleavage site itself (positions -1 to -3; Figure 5.1), and mature N-terminus (Jarrell 2011, Jarrell et al. 2013, Thomas et al. 2001b) but FlaK can only process archaellins and EppA only pilins (Szabo et al. 2007). Swapping of only four amino acids, the -2 to +2 positions around the signal peptide cleavage site, from a pilin to an archaellin was enough to transform the archaellin into a substrate for EppA (Szabo et al. 2007). A similar swap of the -2 to +2 amino acids from an archaellin to a pilin did not allow FlaK to cleave the modified pilin, although archaellin containing the -2 to +2 amino acids of a pilin was still cleaved by FlaK. However, it had been previously shown that in *M. voltae* that changing the +3 glycine to valine in an archaellin prevented cleavage of the signal peptide in *in vitro* experiments (Thomas et al. 2001b). Since glycine is universally found in the +3 position in archaellins (Bardy et al. 2004), we reasoned that the failure of the initial swapping experiment might be because the hybrid pilin carrying the -2 to +2 archaellin amino acids still retained the pilin +3 position, which is not glycine. Thus, we used a PCR-based protocol to change the +3 position of EpdE from alanine to glycine to investigate whether this single amino change would make the mutant EpdE pilin a substrate for FlaK. Both the wildtype version and the +3 glycine version of EpdE were expressed as C-terminal histagged versions from a complementation vector in either wildtype cells or in mutants deleted for either *eppA*, *flaK* or both *eppA* and *flaK*. Western blots were developed with anti-His antibodies. In the two mutant strains carrying a deletion of *eppA* (Δ*eppA* and Δ*flaKΔeppA*), a faster migrating band was observed in western blots that was not seen in strains that still possessed *eppA* (Figure 5.2). In all four strains with the complementation vector,
Figure 5.1: Alignment of the N-terminus region including the signal peptide of type IV pilin-like proteins of *M. maripaludis*.
Figure 5.2: Western blot analysis of various *M. maripaludis* mutant strains expressing a plasmid borne C-terminal histagged version of EpdE. Whole cell lysates were separated by SDS-PAGE (17.5% gel), transferred to Immobilon membrane and developed with antibodies to the Histag. Note the 16 kDa band that is present in wildtype cells without the vector.
a band was observed at approximately 16 kDa, the size expected for EpdE (Ng et al. 2011). However, a band of similar mobility was detected even in control cells that had not received the complementation plasmid (Figure 5.2). This is likely the 16.1 kDa sirohydrochlorin colbatochelatase which has 12 histidine residues in a 15 amino acid stretch (Chaban et al. 2007). As this band complicated interpretation of the pilin processing, the EpdE constructs were remade with a C-terminal FLAG-tag and transformed into the wildtype and mutant strains. Western blots developed to detect this epitope tag indicated that the electrophoretic mobility of the FLAG-tagged wildtype EpdE was the same in both wildtype cells and the flaK mutant and at a position expected from the molecular weight of EpdE found in isolated pili samples (Ng et al. 2011). However, in mutants where eppA was deleted (∆eppA and ∆eppA∆flaK strains), no EpdE was detected at this position in western blots but instead a faster migrating band was observed (Figure 5.3). This was a striking and unexpected observation since in the strains deleted for eppA the signal peptide would still be attached to EpdE and the protein would be 12 amino acids larger. The same pattern was observed when the +3 position of EpdE was changed to glycine (Figure 5.3). In the case of archaellins expressed in a ∆flaK background, the unprocessed archaellins run as slightly larger than in a wildtype background where the signal peptide is removed, as expected and as previously demonstrated in M. voltae (Bardy and Jarrell 2003). Since EpdE is known to be a protein modified with an N-linked pentasaccharide at multiple sites (Ng et al. 2011), a logical explanation for the apparent sizes of EpdE in the various mutants is that the EpdE protein is not N-glycosylated unless the signal peptide is first removed by EppA. The nonglycosylated EpdE would exhibit increased electrophoretic mobility compared to the glycosylated version. Significant amounts of this faster migrating anti-FLAG reactive band were also present in the wildtype and flaK deletion mutants expressing either the wildtype EpdE or the EpdE +3 version. We believe that this is due to an inability of the cell machinery to cope with the unnatural overproduction of EpdE driven by the powerful constitutive hmv promoter on
the complementation plasmid. The very large bands seen in the ΔeppA and ΔeppAΔflaK double deletion lanes are not reflective of any excess amounts of protein at that molecular weight when gels were stained with Coomassie Blue and may reflect a greater sensitivity of the FLAG antibody for the lower molecular weight, unglycosylated form. This is clearly not the case for archaellins. These results indicated that the single amino change at the +3 position to glycine was not sufficient to turn the EpdE pilin into a FlaK substrate since if FlaK was now able to process this mutant version of EpdE, then in the ΔeppA mutant which still possesses flaK, EpdE would have had its signal peptide removed and then be glycosylated and appear on the western blot at about 16 kDa. From the cleavage site swapping experiments it is clear that the amino acids in the immediate vicinity of the cleavage site are enough to allow EppA processing (Szabo et al. 2007), considering the other amino acid positions that are highly conserved between the archaellins and the pilins, i.e. the identical KKG sequence immediately prior to the cleavage site found in EpdC, EpdE and the three archaellins and the conserved hydrophobic stretch of amino acids at the N-terminus of all of the mature proteins. Szabo et al. (Szabo et al. 2007) concluded from their swapping experiments that EppA may have a critical requirement for glutamine at +1. In archaellins, the +1 position is a highly conserved alanine (Thomas et al. 2001a) although glutamine has been found in certain species, notably, for example in Hfx. volcanii where the prepilin peptidase enzyme is PibD (Tripepi et al. 2010). PibD is much more lenient in its acceptance of amino acid substitutions in its substrates at key positions than FlaK and this is reflected in its much greater range of substrates to include all proteins with class III signal peptides (Albers et al. 2003). Previous investigations of essential amino acid positions in and around the archaellin signal peptide in M. voltae did not include any changes to the +1 alanine (Thomas et al. 2001b) so the acceptable variation in this position is unknown. The +2 position is one of the most variable in the N-terminus of archaellins (Thomas et al. 2001a) and in both M. maripaludis and M. voltae there are examples of archaellins bearing serine, valine or
Figure 5.3: Western blot analysis of various *M. maripaludis* mutant strains expressing a plasmid borne C-terminal FLAG-tagged version of EpdE or EpdE with the +3 amino acid position (relative to the signal peptide cleavage site) changed to glycine. Whole cell lysates were separated by SDS-PAGE (17.5% gel), transferred to Immobilon membrane and developed with antibodies to the FLAG-tag.
threonine at this position. Since one of the pilins (EpdC) of *M. maripaludis* also possesses valine at +2, this position does not seem to be important in distinguishing substrate specificity of the two prepilin peptidases. The fact that FlaK can cleave archaellins with the -2 to +2 amino acids of pilins (Szabo et al. 2007) indicates that having the +1 glutamine is not the reason why the +3 glycine containing pilin was not processed by FlaK.

To examine the unusual finding regarding the two posttranslational modifications of the pilins more thoroughly, we raised antibodies to EpdE and looked at the size of native EpdE in wildtype cells and the prepilin peptidase mutants. EpdE was expressed as an N-terminally His-tagged protein using a pET vector system and subsequently purified from induced *E. coli* cells using Ni-affinity chromatography. The purified protein (Figure 5.4) was used to raise antibodies in chickens. Western blots using the purified IgY were not specific, however, requiring a subsequent affinity purification step. Antibodies eluted from this step were used in subsequent western blots. The anti-EpdE antibodies, affinity purified from EpdE protein expressed in *E. coli* where it would not be glycosylated, reacted much more strongly to non-processed pilin compared to the wildtype pilin glycoprotein. In addition, these antibodies detected a band that migrated just below the non processed pilin band (indicated by the arrow in Figure 5.5). This band is unrelated to EpdE as it is also observed in the *epdE* deletion strain (Figure 5.5B). As seen initially with the EpdE FLAG-tagged protein expressed in trans, native EpdE in *eppA* deletion strains also migrated as a much smaller protein than in wildtype cells or the *flaK* deletion strain (Figure 5.5A). For a size comparison we also included the purified C-terminal Histagged version of EpdE expressed in *E. coli* and used to generate antibodies. This version of EpdE would be nonglycosylated and still have its signal peptide attached, as *E. coli* lacks the genes necessary for the two posttranslational modifications. It would be slightly larger than the native EpdE protein by virtue of the extra six C-terminal histidine residues. The size of EpdE as detected by western blots using anti-EpdE antibodies is smaller in the Δ*eppA* and Δ*eppA ΔflaK*
Figure 5.4: Coomassie stained SDS-PAGE gel following the purification of C-terminal Histagged EpdE by Ni-affinity chromatography Log phase *E. coli* strain BL21 (DE3)/pLysS cells containing pKJ900 were induced by the addition of IPTG for 2 hours. Histagged EpdE was purified from induced cells using nickel affinity columns under denaturing conditions. Purified EpdE, indicated by the arrow, was further concentrated for use as antigen for antibody generation in chickens.
Figure 5.5: Western blot detection of native EpdE in various mutant backgrounds. A). Western blot analysis to detect EpdE in various *M. maripaludis* mutant strains using anti-EpdE antibody. C-terminal histagged EpdE was expressed in *E.coli* and purified by a Ni-affinity column is used for size comparison. B). Detection of EpdE in Western blots of the ∆eppA mutant and in the ∆eppA mutant following complementation with a plasmid borne copy of *eppA*. Shown are samples after the complemented cells were grown for 3 transfers in N-free medium supplemented with alanine (promoter on conditions) and the same cells grown in N-free medium supplemented with NH$_4$Cl (promoter off conditions). In both A and B, whole cell lysates were separated by SDS-PAGE (17.5% gel), transferred to Immobilon membrane and developed with antibodies to EpdE. The red arrows indicate the cross-reacting band, also seen in ∆epdE.
strains than the *E. coli*-produced histagged version. This would be consistent with the *ΔeppA* and *ΔeppA ΔflaK* strain version being nonglycosylated and with its signal peptide attached, as in the *E. coli* produced version but without the hexahistidine C-terminal extension. When the *ΔeppA* strain was complemented with a plasmid borne copy of *eppA*, the EpdE detected in western blots returned to the size observed in the wildtype cells (Figure 5.5B), proving that EpdE is, in fact, an EppA substrate, as initially predicted by Szabo et al (2007) and that signal peptide removal was necessary for the dramatic shift of the pilin to the higher apparent molecular mass.

To explore further the order of the signal peptide removal and N-glycosylation modifications of archaellins and pilins, we looked at western blots of the two types of proteins in cells in which the signal peptide removal or N-glycosylation or both was prevented. Prevention of glycosylation was obtained by using a *ΔaglB* strain since AglB is the ligosaccharyltransferase responsible for the terminal step in the N-linked glycosylation pathway (Chaban et al. 2006, Jarrell et al. 2014, Vandyke et al. 2009). For analysis of the archaellin FlaB2, a *ΔflaKΔaglB* double deletion strain was created, using the *ΔflaK* mutant as the starting strain. PCR analysis confirmed that the mutant was deleted for both *flaK* and *aglB* (Figure 5.6A). FlaB2 was detected in western blots of whole cell lysates of wildtype, *ΔflaK*, *ΔaglB* and the *ΔflaKΔaglB* double mutant strains (Figure 5.7A). FlaB2 runs as a slightly larger protein in the *ΔflaK* mutant compared to wildtype, as a result of it retaining the 12 amino acid signal peptide (Bardy and Jarrell 2002, Bardy and Jarrell 2003). In a *ΔaglB* mutant, where N-glycosylation at 4 sequons is now prevented (Vandyke et al. 2009), the protein migrates much faster than in the wildtype cells as expected from the lack of attached glycan. In the *ΔflaKΔaglB* double mutant, the majority of the FlaB2 detected is slightly larger in molecular mass than that observed in the *ΔaglB* mutant. This band most logically represents the unglycosylated FlaB2 with its signal peptide intact, meaning that the smaller molecular mass band observed in the *ΔaglB* mutant is most likely the
Figure 5.6. Confirmation of the double gene deletion strains by PCR. A). PCR screening of \( aglB \) and \( flaK \) in the double mutant \( \Delta flaK \Delta aglB \) using whole cells of the wildtype and the deletion strain as template with gene specific primers. In each case, the first lane shows the PCR product obtained when wildtype cells were used as template and the second lane shows the result of using the deletion strain as the template. B). PCR screening of \( aglB \) and \( eppA \) in the double mutant \( \Delta eppA \Delta aglB \) using whole cells of wildtype and the deletion strain as template. In each case, the first lane shows the PCR product obtained when the wildtype cells were used as template and the second lane shows the PCR product obtained when the deletion strain was used as template. 1Kb and 100 bp DNA ladders were obtained from New England Biolabs.
Figure 5.7. Examination of post-translational modifications of archaellins and pilins in various mutant backgrounds. A). Western blot detection of archaellin FlaB2 in various *M. maripaludis* mutant strains. Whole cell lysates were separated by SDS-PAGE (15% gel), transferred to Immobilon membrane and the blot developed with antibodies to FlaB2. B). Western blot detection of EpdD in various *M. maripaludis* mutant strains expressing a plasmid borne C-terminal FLAG-tagged version of EpdD. Whole cell lysates were separated by SDS-PAGE (17.5% gel), transferred to Immobilon membrane and developed with antibodies to the FLAG-tag.
nonglycosylated protein but with its signal peptide removed. Thus for archaellins, the two posttranslational modifications can occur independently of the other; N-linked glycosylation of the protein can occur whether the signal peptide is removed or not and it is also possible for the protein to have its signal peptide removed whether the protein is first glycosylated or not.

We then turned our attention to signal peptide processing and N-glycosylation of the pilins in appropriate mutant strains. For this, we attempted to examine both EpdE as well as EpdD (MMP1283), a recently identified minor pilin essential for pili formation in *M. maripaludis* (Nair et al. 2013). In addition to the \( \Delta eppA \) and \( \Delta aglB \) strains, a \( \Delta eppA \ \Delta aglB \) double mutant was constructed for this work using the \( \Delta aglB \) strain as the starting strain for a deletion of *eppA*. PCR screening of transformants identified the double mutant (Figure 5.6B). While it proved difficult to determine small differences in the electrophoretic mobility of EpdE-FLAG tag expressed in the three mutant strains, we were more successful examining the migration of EpdD in western blots of various mutant strains. For these experiments, we created a FLAG-tagged version of EpdD expressed from the regulatable *nif* promoter in a second available complementation vector which we transformed into the various strains. It is possible that the expected lower level of expression of EpdD-FLAG from the *nif* promoter compared to the expression of EpdE-FLAG from the strong constitutive *hmv* promoter may have be responsible for the improved resolution of bands obtained in western blots for EpdD. In the \( \Delta eppA \) mutant where the signal peptide would not be removed from EpdD, the protein migrated faster than the same protein in wildtype cells (Figure 5.7B), despite the fact the protein would be larger by the additional signal peptide in the \( \Delta eppA \) strain, as observed for EpdE earlier. In the \( \Delta aglB \) mutant, EpdD migrated as an even faster migrating protein than in the \( \Delta eppA \) strain indicating that the one possible N-linked glycosylation site is EpdD is, in fact, occupied in wildtype cells and that AglB is necessary for this N-glycan attachment. In the \( \Delta aglB \Delta eppA \) double mutant, the electrophoretic mobility of EpdD is the same as in the \( \Delta eppA \) strain. The EpdD band observed in
western blots of the ∆eppA mutant most likely represents a nonglycosylated pilin with its signal peptide still attached since the electrophoretic mobility of EpdD-FLAG in the ∆eppA ∆aglB double mutant is the same, and glycosylation would be prevented in the double mutant. In the ∆aglB mutant alone, the EpdD-FLAG migrates as a smaller protein consistent with this pilin having its signal peptide removed in addition to being nonglycosylated. This indicates that EpdD is an EppA substrate, as predicted previously (Szabo et al. 2007). Thus, in the case of pilins, the timing of the addition of the N-glycan to pilins is dependent on the signal peptide first being cleaved. This seems to be an unusual requirement as this is not the case with archaellins and, in *P. aeruginosa*, O-glycosylation of type IV pilins was shown to occur in a pilD mutant indicating that signal peptide removal was not a prerequisite for glycosylation in this system either (Horzempa et al. 2006). On the other hand, signal peptide removal can occur even if the pilins are nonglycosylated, as observed in archaellins.

The presence in one organism of two different proteins having different N-linked glycans in Archaea is rare but is found for the pilins and archaellins of *M. maripaludis* (Kelly et al. 2009; Ng et al. 2011). In *M. voltae*, the same trisaccharide glycan has been shown to be attached to both archaellins and the S-layer protein (Voisin et al. 2005). In *Hfx. volcanii*, the same *agl* genes involved in formation of the glycan initially reported to be attached to the S-layer glycoprotein (Kaminski et al. 2013b) were also found to be involved in the N-glycosylation of archaellins (Tripepi et al. 2012), although it is unknown whether the glycan is identical. In different extreme halophiles, the same protein has been shown to be modified with two different N-linked glycans. This was first reported in the S-layer protein of *Halobacterium salinarum* where there is a single site which has an attached repeating unit N-linked glycan and 10 sites which contain a sulfated oligosaccharide (Lechner and Sumper 1987, Wieland 1988). Interestingly, archaellins in *Hb. salinarum* are modified with the sulfated oligosaccharide but not with the repeating unit glycan (Lechner and Sumper 1989, Sumper 1987, Wieland et al. 1985). Recently, in *Hfx. volcanii*, two
different N-linked glycans have been found on the S-layer protein when cells are cultured in medium containing reduced salt concentrations (Kaminski et al. 2013a). How the cell delivers the different glycan to the different sites on the same protein is unknown but critically in the *Hfx volcanii* case, the transfer of the low salt glycan has been shown to occur independently of the lone identified oligosaccharyltransferase (Kaminski et al. 2013a). This is not the case in *M. maripaludis* where the glycan attached to both archaellins and pilins depends upon the presence of AgIB since, in a ΔagIB mutant, both archaellins and pilins migrate as much smaller proteins indicative of being nonglycosylated. How the pilins get decorated with a pentasaccharide glycan while the archaellins are modified with a tetrasaccharide is unknown.

In some bacteria, appendages are localized to defined regions, such as the cell pole (Chiang et al. 2005, Keilberg and Søgaard-Andersen 2014, Kim and McCarter 2004, Vignon et al. 2003). Type IV pili are polar organelles in *Pseudomonas aeruginosa* (Chiang et al. 2005), for example, and in *Caulobacter crescentus* flagella, pili, the stalk and holdfast all assemble at the so-called polar region of the cell, thought to be a distinct functional and biochemical entity (Hinz et al. 2003, Smit 1987). In such situations, it is possible for the cell to differentially locate specific proteins to the cell pole (Bardy and Maddock 2007, Chiang et al. 2005, Dworkin 2007, Hinz et al. 2003, Keilberg and Søgaard-Andersen 2014, Smit 1987). For *M. maripaludis*, it could then be possible for a localized pilus-specific assembly apparatus to exist that might contain, among other proteins, EppA and a unique N-glycosylation pathway complex. However, in *M. maripaludis*, both archaella and pili are located peritrichously on the cell surface, leaving open the question of how these archaea are able to differentiate their type IV pilin-like proteins in terms of signal peptide removal and N-glycosylation.
Acknowledgments. This research was funded by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) to KFJ. The authors thank Richard Casselman for advice on antibody purification.
**Literature cited**


Jarrell KF, Stark M, Nair DB, Chong JPJ (2011) Flagella and pili are both necessary for efficient attachment of Methanococcus maripaludis to surfaces. FEMS Microbiol Lett 319:44-50


Kaminski L, Guan Z, Yurist-Doutsch S, Eichler J (2013a) Two distinct N-glycosylation pathways process the Haloferax volcanii S-layer glycoprotein upon changes in environmental salinity. Mbio 4:e00716-13


Thomas NA, Chao ED, Jarrell KF (2001b) Identification of amino acids in the leader peptide of *Methanococcus voltae* preflagellin that are important in posttranslational processing. Arch Microbiol 175:263-269


Chapter 6

General Discussion

Various kinds of surface structures are found in Archaea, some that are similar to bacterial ones and others that are unique (Jarrell et al. 2013; Lassak et al. 2012). One common appendage found in both Euryarcheota and Crenarcheota is the type IV-like pilus. Both archaella and type IV-like pili have components that are homologous to ones found in bacterial type IV pili assembly systems and hence both are considered to be type IV pili-like appendages. Little was known about archaeal type IV-like pili until recently, when studies were published on the genetics, structure and function of the organelle in a limited number of archaea, mainly in *Sulfolobus* species and *M. maripaludis* (Jarrell et al. 2013; Lassak et al. 2012; Pohlschroder et al. 2011). In *M. maripaludis*, a putative 11 gene type IV pilus locus was first identified by bioinformatics (Szabo et al. 2007). Later, genetic and electron microscopic investigations by the Jarrell group determined that the three pilin-like genes in the locus were all essential for normal piliation (Ng et al. 2011). The structure of the pilus, the first determined for any pili in archaea, was also determined to unique (Wang et al. 2008).

The work reported in this thesis significantly advances our understanding of the complexity and novelty of type IV-like pili in *M. maripaludis*. In Chapter 2, we analysed the involvement of other, previously unstudied, genes in the type IV pilus-like locus of *M. maripaludis*, identifying, through inframe deletion analysis, complementation studies and electron microscopy, several novel genes essential for piliation (Nair et al. 2014). In addition, we identified the polymerizing ATPase gene, as well as two tandem copies of genes encoding the conserved membrane (platform) protein involved in pili formation, at a location in the
genome distant from the previously identified type IV pilus locus. All three of these genes were demonstrated to be essential for piliation. In Chapter 3, we screened all potential pilin-like genes in the *M. maripaludis* genome by deletion analysis and identified a further minor pilin MMP1283 (EpdD) essential for pili formation, also at a locus distinct from other genes already known to be involved in pili formation (Nair et al 2013). The major pilin encoded by *mmp1685*, and later designated *epdE* (Nair et al. 2014), was identified by a combination of mass spectroscopy of purified pili, gene deletion analysis and electron microscopy (Ng et al. 2011). In Chapter 4, we focused on posttranslational modifications of the pilin proteins in *M. maripaludis*. By expressing an epitope-tagged version of EpdD in different mutant strains, we provided evidence that EpdD was both processed by EppA and also a glycoprotein with an attached N-linked glycan. Further studies with FLAG-tagged versions of EpdE and EpdD expressed in various mutant strains suggested strongly that the pilins in *M. maripaludis* could only be glycosylated if the pilins first had their signal peptide removed by EppA. This is unlike the case for archaellins, where N-glycosylation of the protein can occur whether the signal peptide has been first removed or not. However, like archaellins, the pilins do not need to be glycosylated to have the signal peptide removed. This suggests that the order of posttranslational modification is important for pilins and that the order must be signal peptide removal by EppA and then glycosylation through the activity of the oligosaccharyltransferase AglB.

**Genes Involved in *M. maripaludis* Piliation are Scattered Around the Genome**

Our results indicated that genes found in at least four distinct locations in the genome are necessary for piliation. These loci are the initial type IV pilus-like operon first identified by Szabo et al. (2007), the operon encoding the polymerizing ATPase and two copies of the conserved type IV pilus membrane protein, as well as the separate gene locations for the minor pilin *epdD* and the major pilin *epdE*. These findings, coupled with the fact that four genes in the
11 gene locus (mmp0234 [epdF], mmp0239 [epdG], mmp0240 [epdH], mmp0241 [epdl]) shown to be essential for piliation do not show any homology to known type IV pili genes presently identified in either Bacteria or Archaea, suggest that the type IV-like pilus system in M. *maripaludis* is much more complex than that of other archaea and perhaps more on par in complexity to type IV pili systems in many bacterial species. In *Sulfolobus*, two distinct type IV pili structures have been identified, Ups pili and Aap pili. However in both cases, only a 5 gene cluster has been identified. Both loci contain genes for two pilins, a pilus-like ATPase and a pilus-like conserved membrane protein (van Wolferen et al. 2013; Henche et al. 2012a). In the Ups system, the fifth gene (*upsX*) encodes a predicted cytoplasmic protein with no specific functional domains detected by BLAST (van Wolferen et al. 2013). Deletion of each of the pilin genes alone still left cells piliated, suggesting each pilin alone can form a pilus. As expected, deletion of the ATPase or membrane protein gene resulted in nonpiliated cells. Mutants carrying a deletion of *upsX* still made pili and aggregated following UV treatment but DNA transfer was impaired (van Wolferen et al. 2013). In the Aap pili system, the fifth gene (*aapX*) encodes a putative iron-sulfur oxidoreductase (Hence et al. 2012a). Here, deletion analysis revealed all five genes were essential for piliation. In the case of Hfx. *volcanii*, where type IV-like pili are thought to be involved in adhesion (Tripepi et al. 2010), a more complex story is emerging but these studies are in an early stage and have focused solely on pilins (Esquivel et al. 2013). In this extreme halophile, six pilins were characterized (PilA1-6). Strains that were deleted for up to five of the pilin genes had no significant adhesion defects while a strain carrying deletions in all six pilin genes could not adhere to tested surfaces. Complementation experiments showed that expression of any of the six pilin genes in trans in the strain deleted for all six pilin genes led to the production of functional pili. As in *M. maripaludis*, co-regulation of the pilin genes with the ATPase and conserved membrane protein does not occur. In bacteria, a larger number of genes is required for the assembly and function of the type IV pili system. In
*P. aeruginosa*, for example, about 40 genes have been identified that are involved in the structure and function of type IV pili (Burrows 2012; Mattick 2002). As is typical of many bacterial type IV systems, especially type IVa ones (Pellicic 2008), in *P. aeruginosa* these genes are found in various locations around the genome. The polymerizing ATPase, conserved membrane protein and the prepilin peptidase genes are found in a single operon in *P. aeruginosa* that is not co-transcribed with pilin genes, although the major pilin gene, *pilA*, is located adjacent to *pilB* on the complementary strand (Burrows 2012). Five minor and one major pilin are found in *P. aeruginosa* and all have been shown to be incorporated into pili by immunolocalization experiments (Giltner et al. 2010). Likewise, about two dozen genes have been shown to be involved in type IV pili biosynthesis and function in *Neisseria meningitidis* (Brown et al. 2010). However, a much smaller number of genes may be necessary for just the biosynthesis of the pilus; only 12 for example for bundle-forming pili synthesis in enteropathogenic *E. coli* [EPEC] (Stone et al 1996; Milgotina and Donnenberg 2009).

### Deciphering the roles of *M. maripaludis* epd genes of unknown function

The *M. maripaludis* 11 gene operon encoding three minor pilins and the prepilin peptidase also contains genes with no homology to any type IV pilin genes found in other archaea or bacteria. Some of these are methanogen-specific genes and yet they are essential for formation of pili. Determining the function of the proteins encoded by these genes will be a major challenge since it is not unusual for various type IV pili systems, even well studied ones, to have essential genes for which functions have still not be assigned (Burrows 2012; Carbonelle et al. 2006; Koomey 2009; Pellicic 2008). A further complication is evidence demonstrating that a given protein may have multiple functions in type IV assembly and function, as shown recently for PilW in *N. meningitides* (Szeto et al. 2011). A first step in
assigning a role to these proteins in *M. maripaludis* is to determine their cellular location by cell fractionation. Lack of antibodies to these proteins means that an epitope-tagged version of each would need to be generated, expressed in cells and tested for its ability to restore piliation when the corresponding gene has been deleted. Such an experiment would indicate that the added tag did not interfere with the normal location of the protein within the cell. While cellular location can be predicted by programs such as PSORT, experimental confirmation must be done. There are many instances in type IV pili systems where proteins predicted to be cytoplasmic were found associated with the cytoplasmic membrane due to strong interaction with integral membrane proteins. A primary example of this is the interaction of pilus system ATPases (predicted to be cytoplasmic) with the integral membrane platform protein in *Thermus thermophilus* (Karuppiah et al. 2013), *P. aeruginosa* (Takhar et al. 2013), *Neisseria gonorrhoeae* (Koomey 2009) and *E. coli* (Milgotina and Donnenberg 2009).

In type IV pili systems, most pilus related-proteins are found in the cytoplasmic membrane. However, several different subcomplexes have been identified in various cellular compartments, especially in well-studied Gram negative systems. These include a cytoplasmic membrane complex, an outer membrane complex, a transmembrane complex and the pilin (major and minor) complex that makes up the appendage itself (Burrows 2012). Since *M. maripaludis* lacks an outer membrane, one would anticipate that most of the Epd proteins with unknown function might be localized to the cytoplasmic membrane. PSORT predictions for EpdF and EpdG are that they are cytoplasmic proteins although each contains a single predicted transmembrane domain. EpdH has no predicted location by PSORT although it also has a single predicted transmembrane domain while EpdI is predicted to be extracellular although it has two predicted transmembrane domains. Specific interactions between Epd proteins with unknown functions could be investigated by a variety of techniques used successfully in bacterial type IV pili systems (Pellicic 2008). These methodologies include two hybrid methodologies, in vivo
cross-linking and co-affinity purification studies (Crowther et al. 2004; Takhar et al. 2013; Tripathi and Taylor 2007; Milgotina and Donnenberg 2009). In addition, the stability and localization of various Epd proteins in mutants deleted for each of the other epd genes could be examined as often proteins fail to localize properly or are unstable if their natural interaction partner is absent (Georgiadou et al. 2012; Ramer et al. 2002; Tripathi and Taylor 2007). A drawback to such studies is the current lack of antibodies against any of the Epd proteins.

Type IV pili can be involved in numerous and varied activities but most commonly the functions are adhesion and motility (Conrad et al. 2011; Gibiansky et al. 2010; Giltner et al. 2012; Hahn 1997). In regards to type IV pili and adhesion, it seems that at least two possible mechanisms are known. The major pilin of the type IV pili can act as the adhesive element as is the case for P. aeruginosa (Lee et al. 1994) and for bundle forming pili of E. coli (Hyland et al. 2008). However, specific tip adhesins (which are non pilins) have also been reported, as in the case of PilC in N. gonorrhoeae (Rudel et al. 1995). PilC was reported to be synthesized initially with a signal peptide (Rudel et al. 1995) and pili were still assembled by the cell in the absence of the adhesin (Rudel et al. 1992). In the 11 gene pilus cluster of M. maripaludis, there are two genes that encode proteins predicted to have signal peptides (mmp0235 and mmp0238). In both cases, deletion of the gene did not have any detectable effect on piliation. The predicted presence of signal peptides suggests that if the proteins were involved in piliation, as seems likely since they are encoded by genes located within an operon dedicated to type IV pili production, one, at least, may be involved as a tip adhesin and require the signal peptide for export from the cell. Investigation of this possibility could involve at least two avenues of research. Antibodies could be raised against each of the two proteins after expression and purification of an epitope-tagged version in E. coli. These antibodies could be used in immunogold labelling experiments to see if either protein localizes to the tip of pili. Such experiments would be best performed on cells already lacking archaella, such as in a flaK
deletion strain (Ng et al. 2011), so that pili are the sole surface appendages. Such an
immunolabeling approach was used to show the tip location of PilC in *N. gonorrhoeae* (Rudel et
al. 1995). A second approach would complement the first. It is known that the type IV pili of *M.
maripaludis* are necessary for adhesion of cells to surfaces, but this adhesion requires both
archaella and pili (Jarrell et al. 2011). If either *mmp0235* or *mmp0238* encode the adhesin for
the type IV pili then their deletion may not affect the appearance of pili on the cell surface but
should have a drastic effect on their function in adhesion. For clarity in electron microscopic
examination of strains deleted for genes potentially involved in piliation, all the pilin gene
deletions were created in a strain that was already nonarchaellated due to the deletion of *flaK.*

To study the possible effects of loss of *mmp0235* or *mmp0238* on adhesion, these deletions
would need to be generated in a wildtype (i.e. archaellated) background. The resulting strains
would have archaella and pili but the pili may be lacking the putative adhesin. Such mutants can
be tested for attachment to various surfaces by scanning electron microscopy, as previously
described (Jarrell et al. 2011).

*M. maripaludis* Type IV pilus ATPase and Conserved Membrane Protein

Unlike in other studied archaea, a gene encoding the type IV pilus polymerizing ATPase
was missing from the type IV pili locus found in the genome of *M. maripaludis.* However, a
BLAST search of the genome of *M. maripaludis* showed two potential candidates that might be
type IV pilus ATPase, MMP0040 and MMP0281. Deletion and complementation analysis
showed that MMP0040 alone was involved in piliation, since deletion of this gene left cells
nonpiliated while deletion of *mmp0281* had no discernible effect on cell piliation. If *mmp0281*
had encoded the depolymerizing ATPase, deletion of this gene should have resulted in
hyperpiliated cells. Most bacterial type IV pili systems have two associated ATPases. One is
involved in the incorporation of subunits into the base of the structure resulting in extension of
the pilus while the other is responsible for retraction of the pilus by removal of subunits from the base into the cytoplasmic membrane (Burrows 2012; Burrows 2005; Craig and Li 2008; Morand et al. 2004). Rounds of extension and retraction of the type IV pili are responsible for the surface dependent motility known as twitching (Bradley 1980; Burrows 2012). Not all bacterial type IV pili are, however, known to retract: subgroups such as type IVb (e.g. V. cholera TCP pili) and Tad pili are not usually associated with twitching and lack the retraction or depolymerizing ATPase (Burrows 2012). Bundle forming pili of E. coli are an exception in being a type IVb pilus system that possesses a retraction ATPase (Milgotina and Donnenberg 2009) and where retraction of type IV pili was shown to play a role in pathogenesis (Aroeti et al. 2012). To date, homologues of retraction ATPases have not been reported in any archaeon, suggesting that archaeal type IV pili do not retract and cells do not twitch and indeed, no reports of twitching in archaea have been reported.

*M. maripaludis* is unique so far amongst the studied archaeal pilus systems in having two genes encoding the conserved membrane protein that is considered a platform for assembly of the pilus. While rare amongst bacterial type IV pili systems, this situation is found in Tad pili systems of *P. aeruginosa* (Bernard et al. 2009) and *Aggregatibacter actinomycetemcomitans* (Kachlany et al. 2000) and in the latter case both versions of the gene (*tad*B and *tad*C) are essential for piliation, as found for *M. maripaludis*. Why this is so is unclear. The pilins of *M. maripaludis* bear some similarities to the Tad subclass of type IV pilins (Ng et al. 2011). These similarities include short signal peptides and very small mature pilin proteins (typically between 50-80 amino acids). The presence of the two platform genes raises an interesting question regarding retraction of pili and twitching. It has been suggested by different groups that it may be possible for a single ATPase to perform both extension and retraction of the pili by interacting with the two different platform proteins, with the ATPase interacting with one version resulting in extension and the same ATPase interacting with the other platform protein to
facilitate retraction of the pilus (Albers and Pohlschroder 2009; Burrows 2012). It would seem unlikely that this is occurring in *M. maripaludis*, however, since deletion of the gene encoding the platform protein responsible for the depolymerisation process should lead to hyperpiliated cells rather than to the nonpiliated cells observed in strains carrying a deletion in either *epdJ* or *epdK*. Nonetheless, it would be of interest to determine if the ATPase does interact with both platform proteins. It is already established that in *P. aeruginosa* the polymerizing (PilB) and depolymerizing (PilT) ATPases interact with different cytoplasmic domains of the platform protein PilC and that PilC is essential for the type IV pili system, controlling both pilus assembly and disassembly (Takhar et al. 2013). To show potential interaction between the one type IV pilin ATPase of the *M. maripaludis* system (EpdL) and the two platform protein (EpdJ and EpdK), in vitro co-affinity purification experiments could be performed. This would involve expressing an epitope tagged (e.g: histagged) version of the ATPase in *M. maripaludis* and seeing if either or both of the platform proteins co-purified following Ni-affinity column purification. A current drawback to this type of experiment is the lack of antibodies available to any of the pili system proteins in order to readily identify interaction partners by Western blotting. It may be possible to identify co-purified proteins by mass spectrometry obviating the need for specific antibodies.

**Minor pilins in *M. maripaludis***

Four minor pilins (EpdA, EpdB, EpdC and EpdD) have been identified in *M. maripaludis*. Three of them are found in the main locus along with the prepilin peptidase essential for the processing of these pilins and the other one (EpdD) is found in a different locus. Mass spectrometry and western blotting analysis of the purified pili samples as well as immunolabelling experiments on whole cells can be done to determine if and where the minor pilins are incorporated into the pilus. During our effort to find the major pilin protein using mass
spectrometry, a single peptide corresponding to the minor pilin EpdB was identified in the purified pili samples, which suggests that this pilin at least is being incorporated into the structural pilus. Western blotting of the purified pili samples using antibodies specific to each of the minor pilins may have identified further minor pilins in the samples but such antibody is unavailable. Similar techniques have been used to identify minor pilins in bacterial type IV pili (Giltner et al. 2010) although sometimes interpretation of these experiments is complicated by the fact researchers have used a sheared/vortexed fraction in the western blotting experiments without reporting control experiments to ensure that these pili samples are not contaminated with cytoplasmic membrane fragments that may contain the pili inner membrane assembly complexes (Winther-Larsen et al. 2005; Helaine et al. 2005; Giltner et al. 2010). Early studies on minor pilin distribution in *P. aeruginosa* suggested that the minor pilins stayed in the cytoplasmic membrane (Alm et al. 1996) but later immunolabelling experiments revealed their incorporation into the fiber (Giltner et al. 2010). In many bacterial examples, deletion of minor pilin genes resulted in loss of piliation (Winther-Larsen et al. 2005; Carbonnelle et al. 2006; Giltner et al. 2010). However, when the minor pilin gene deletion is created in a retraction ATPase deficient background, piliation is restored indicating that the minor pilins are not essential for pili formation but pili assembly is suboptimal in their absence.

So what is the role of the minor pilins? Many functions have been proposed in the literature including initiation of pilus assembly, antagonism of retraction and contact-dependent opening of the secretin (Burrows. 2012). In *M. maripaludis*, there is no outer membrane so there is no secretin which must be opened. In addition, there is no evidence of pilus retraction since a retraction ATPase has not been found in any archaeon to date. Those two possible roles for the minor pilins can be dismissed. Since the minor pilins are essential for piliation they may be playing a key role in either initiation of assembly or stability of the structure. A role for minor pilins in the initiation of the assembly has been proposed that could take on different formats. A
minor pilin complex has been thought to prime pseudopilus assembly in the type II secretion system either by forming the nucleation site for assembly or by activating the assembly ATPase and platform protein (Burrows. 2012; Cisneros. et al 2012). The idea of a priming complex could mean that the minor pilins are not actually incorporated into the complex although there is now clear immunolabelling that shows that in P. aeruginosa they are incorporated (Giltner et al. 2010). The priming complex could still occur but as a result the complex of minor pilins also gets incorporated. In this view, it is thought that the addition of major pilin subunits occurs under the initiation complex, pushing that complex to the tip in the final structure (Giltner et al. 2012). The location of minor pilins at positions other than the tips in immunolabelling experiments, however, indicate that minor pilins may be incorporated throughout the structure.

Five additional type IV pilin-like genes (mmp0528, mmp0600, mmp0601, mmp0709 and mmp0903) were deleted but electron microscopic examination of these mutant cells showed all had piliation indistinguishable from wildtype cells. While obviously not essential for Epd pil formation, these proteins could still be necessary for efficient assembly or functioning of the pilus. These effects would not reveal themselves in examination of the cells by electron microscopy but may be discovered in functional analysis (i.e., attachment studies) of cells deleted for the corresponding genes. The other possibility is that these genes may be involved in the formation of a completely different type IV pili-like structure on the M. maripaludis cells, which either may not assemble under the particular growth condition provided or be of such a length to be missed in electron microscopic examination of cells. This could be particularly true for mmp0600 and mmp0601, since they are located adjacent to each other and are considerably larger in size compared to the other pilin proteins. A precedent for multiple type IV pili-like appendages in a single archaeal cell has already be shown to occur in S.acidocaldarius where two types of pili, archaella and the bindosome may all co-exist, although not all are expressed under all growth conditions (Lassak et al. 2012). The bindosome in Sulfolobus has
never been observed in electron micrographs, likely because it may extend only from the
cytoplasmic membrane to the S-layer (Zolghadr et al. 2011; Lassak et al. 2012). Since
methanogens do not utilize sugars and homologues of genes encoding sugar-binding proteins
are not found in their sequenced genomes, the bindosome will not be present. However, it is
possible the pilin-like proteins of *M. maripaludis* currently without function may incorporate into
such a short structure that carries out an unknown function and that this short length has
hindered its detection so far. Alternatively, these pilin proteins may simply form a pilus structure
that, like the Ups pili in *Sulfolobus* species (Frols et al. 2008), is only assembled under certain
growth conditions or if cells are stressed.

**Pilin posttranslational Modifications**

*M. maripaludis* pilins are known to undergo two post-translational modifications: signal
peptide removal and N-glycosylation. The results presented in this thesis show that, unlike the
case seen in archaellins, the pilins require signal peptide removal prior to N-glycosylation of the
protein. Two different prepilin peptidases are found in the genome of *M. maripaludis*, which is an
unusual occurrence in archaea. Other archaea, like *Sulfolobus* and *Halofex*, are known to
have a single prepilin peptidase (PibD) responsible for the processing of both archaellins and
pilins as well as all other proteins bearing class III signal peptides, such as various sugar-
binding proteins (Albers et al. 2003; Lassak et al. 2012; Tripepi et al. 2010). In *M. maripaludis*,
FlaK specifically cleaves the archaellins and the pilins are processed by yet another dedicated
peptidase, EppA (Bardy and Jarrell 2002; Bardy and Jarrell 2003; Szabo et al. 2007). Though it
was known earlier that the +3 glycine is essential for the processing of archaellins by FlaK
(Thomas et al. 2001), a mere change in the +3 position of pilins to glycine was insufficient to
allow FlaK to process the pilins, even though both archaellins and pilins share a number of
conserved amino acids in the cleavage site region shown to be important for FlaK activity. For
instance, in both EpdC and EpdE pilins, the signal peptide ends in KKG and all three of these amino acids are important for FlaK recognition of the archaellin substrate (Thomas et al. 2001). Site-directed mutagenesis of EpdE to change its +3 position to glycine means that the mutant EpdE has the -1,-2,-3 and +3 key residues as well as the conserved N-terminal hydrophobic stretch of amino acids that is a hallmark of type IV pilin-like proteins but it still was not recognized by FlaK. Therefore, some other feature or features of the sequence must be necessary for substrate recognition by FlaK. For EppA, a key position for substrate recognition appears to be the +1 Q (Szabo et al. 2007). Alanine is the most common amino acid found at at the +1 position in archaellins (Thomas et al. 2001) and all three archaellins of *M. maripaludis* have alanine at this position. However, since FlaK can process archaellins with the -2 to +2 amino acids of pilins (Szabo et al. 2007), the presence of glutamine at +1 does not inhibit FlaK activity. These results indicate that the specific features of a type IV pilin-like protein that make it a FlaK substrate extend further than the immediate cleavage site, unlike the case for EppA.

Unlike archaella, N-glycosylation of *M. maripaludis* pili is not essential for the assembly of the fiber, since an aglB deletion mutant is known to be piliated (VanDyke et al. 2009). So far, it is not known if the loss of N-glycosylation can affect the function of the assembled pili. In bacteria, the effects of O-glycosylation on type IV pilus assembly and function have also reported to vary in different organisms. Marceau et al (1998) ruled out a major role for pilin glycosylation in piliation or adherence in *Neisseria meningitidis*. In *N. gonorrhoeae*, no significant effects on either piliation or associated phenotypes were found in glycosylation mutants (Vik et al. 2012). In *P. aeruginosa*, two different O-glycosylation systems have been reported. One is mediated by an oligosaccharyltransferase TfpO and this can be dispensed with without effects on piliation, although it seems that this modification is a significant virulence factor (Smedley et al. 2005). Other *P. aeruginosa* strains lack *tfpO* but have O-glycosylation of pilins mediated through a different pathway. Mutants in which this pathway was disrupted
produced unmodified pilins and only a few pili indicating a key role for the glycosylation in assembly or stability (Kus et al. 2008).

The timing of the addition of the N-glycan to the *M. maripaludis* pilins is dependent on the signal peptide first being cleaved. This seems to be an unusual requirement as this is not the case with archaellins and in *P. aeruginosa*, O-glycosylation of type IV pilins was shown to occur in a *pilD* mutant indicating that signal peptide removal was not a prerequisite for glycosylation in this system either (Horzempa et al. 2006). An intriguing question raised by the findings in this thesis is the differential modification of the two classes of type IV pilin proteins with glycans. How does the cell distinguish the two classes and attach a pentasaccharide to pilins and a tetrasaccharide to the archaellins. The assembly mechanism reported for the pentasaccharide of the S-layer of *Hfx. volcanii* may provide clues (Jarrell et al. 2014; Kaminski et al. 2013). In this halophile, the first four sugars are assembled on a dolichol phosphate carrier by distinct glycosyltransferases. This is then flipped to the external face of the cytoplasmic membrane and transferred to specific sites on the S-layer protein by AglB. The terminal mannose of the pentasaccharide meanwhile is added by the glycosyltransferase AglD to a separate dolichol phosphate carrier which is flipped by AglR across the membrane. This final mannose is delivered from its carrier to the tetrasaccharide already present on the S-layer protein by AglS (Cohen-Rosenzweig et al. 2012). In *M. maripaludis*, it is conceivable that all proteins to be N-glycosylated receive the tetrasaccharide present on archaellins but that a separate enzyme then adds the final hexose of the pilin pentasaccharide after the tetrasaccharide is already present on the pilins. The identity of the glycosyltransferase responsible for the hexose addition is unknown. It is also not currently known whether the hexose is assembled onto a dolichol carrier separate from the dolichol carrying the tetrasaccharide. This could be investigated by mass spectrometry analysis of isolated dolichol phosphate pools for the identification of the glycan charged species (Guan et al. 2010; Larkin et
The presence of a dolichol phosphate charged with the pentasaccharide would rule out his hypothesis but if the pool showed only dolichol species charged with the tetrasaccharide and its precursors and dolichol phosphate charged with a hexose only this proposal would be supported. The hypothesis would need to include a mechanism by which this hexose glycosyltransferase specifically recognizes the pilins or at least is specifically localized to where pilins are assembled. Since glycosylation is not needed for pilus assembly, it raises another intriguing possibility, namely that the glycosylation of the pilins occurs after assembly of the subunit into the pilus structure. This would then be similar to the transfer of the last sugar of the N-linked pentasaccharide glycan of *Hfx. volcanii* which is added after the tetrasaccharide is already attached to the S-layer protein (Cohen-Rosenzweig et al. 2012).

The interpretation of events of the posttranslational modifications of the pilins is based on migration in western blots of pilins in mutants deleted for *eppA* and *aglB*. Confirmation of these interpretations, i.e. the presence or absence of the signal peptide and the N-linked pentasaccharide by direct methods such as N-terminal sequencing and mass spectrometric analysis of pilins is problematic. Generally such studies have been done on purified appendages (Ding et al. 2013; Jones et al. 2012; Ng et al. 2011; Vandyke et al. 2009) and mutants in *eppA* do not make pili to isolate (Nair et al. 2014). Furthermore, previous attempts at N-terminal sequencing of pilins from pili preparations were unsuccessful (Ng et al. 2011). It may be possible to purify enough of the epitope-tagged pilin samples for analysis but Coomassie-stained cell lysates from strains expressing tagged versions of EpdE do not indicate any noticeable excess of protein bands at the molecular weight detected in western blots. It should be emphasized that the functions of both EppA and AglB are both well established and the explanations of the western blotting results in mutants deleted for these genes are consistent with these functions.
Model of assembly of *M. maripaludis* Pili

A hypothetical model of pilus assembly in *M. maripaludis* that includes the major findings of this thesis is shown in Figure 6.1. The major prepilin (EpdE) and other prepilins (EpdA, EpdB, EpdC and EpdD) are targeted to the cytoplasmic membrane where their signal peptide is removed by the prepilin peptidase (EppA. The ATPase (EpdL), interacting with the two conserved platform proteins, EpdJ and EpdK provides the energy for the assembly of the various processed pilins into the appendage. The glycosylation of the pilin is known to occur only after the removal of the signal peptide. While many possibilities exist for the assembly of the pentasaccharide, I have shown the involvement of two dolichol carriers, one carrying the tetrasaccharide also found on archaellins and a separate dolichol carrier charged with only the additional branched hexose. The two non-essential proteins predicted to be secreted proteins, MMP0235 and MMP0238 are shown in this model to be an adhesion and an adapter protein needed to connect the adhesion to the remainder of the pilus fibre formed primarily by EpdE. There is no experimental evidence for this at present and the functions of the two proteins shown here could also be reversed. EpdI, predicted to be an extracellular protein, nonetheless has two transmembrane domains and a domain found in transport proteins. It is shown at the base of the structure where it may help to stabilize the pilus.
Figure 6.1: A hypothetical model of pili assembly in *M. maripaludis*. The assembly of the pilin proteins is shown and is assumed to occur at the base of the structure due to the energy provided by the polymerizing ATPase EpdL, known to function as an octamer, and thought to interact with the platform proteins EpdJ and EpdK. The glycosylation of pilins is depicted to occur only after the signal peptide removal of the pilins, as shown in this thesis. The assembly on a dolichol lipid carrier of the branched pentasaccharide found N-linked to pilins is shown on the left. The Epd proteins are labelled with the fourth letter of their designations. For clarity S-layer is not shown. See text for more details.
Conclusions and Future Directions

This thesis work presents the successful identification of all the core components essential for the formation of the type IV-like Epd pili found in *M. maripaludis*, including the major pilin, minor pilins, the prepilin peptidase, ATPase and two platform proteins. Many additional experiments have been suggested throughout the General Discussion that could shed extra light on aspects of piliation in *M. maripaludis*. These proposed experiments represent possible fruitful avenues of research on this important structure. These include; determining roles of the unique proteins found in the *M. maripaludis* pili system; the mechanism for the order of the posttranslational modifications; the roles of the many minor pilins and the purpose and pathway of the differential glycosylation found on pilins compared to archaellins. Not addressed in this thesis, but of considerable interest based on studies in *Sulfolobus* species, is the regulation of the type IV pili system in *M. maripaludis.*

The study of regulation of type IV pili systems in Archaea is in its infancy and limited solely to *Sulfolobus* species. Here, an intriguing interplay between archaella and Aap pili production has been observed (Reimann et al. 2012). For example, in an *aapF* deletion mutant unable to make Aap pili, the expression of archaella was highly induced (Hench et al. 2012a; Lassak et al. 2012). In addition, overexpression of the archaella repressor ArnA led to hyperpiliation. Recent evidence points to the involvement of a member of the leucine-responsive regulatory protein (Sa-Lrp) family in the transcription of genes of the *ups* operon, perhaps in concert with other transcriptional factors (Vassart et al. 2013). Analysis of the Epd pili genes has not reached that stage.

Unlike the case, for example, in the Ups pili system where all five genes, including pilins, ATPase and membrane component are all co-transcribed, the genes for the Epd pili are scattered in at least four different locations in the genome and this suggests a regulatory control that may be more complex in order to maintain the proper stoichiometry of various subunits for
pili assembly. Obviously, the major pilin will need to be expressed in larger amounts than the other, minor pilins and accessory proteins like the ATPase and platform proteins. It is known in some type IV pili systems that the stoichiometric ratio of the various minor pilins is critical for function (Giltner et al. 2010). Another possibility is that pilus core components are always made and whenever necessary the major pilin is expressed and the pilus can then be quickly assembled. It has been previously shown in S. acidocaldarius that a constitutive promoter regulates the basal level of all the genes encoding the archaella accessory and core elements while a separate, inducible promoter controls the expression of the major filament protein FlaB1 (Lassak et al. 2012). The hypothesis is that the core structure is already in place so that archaella can rapidly be assembled whenever the environmental conditions result in transcription of the flaB1 gene.

Nothing is currently known about the regulation of pili production in M. maripaludis or about the environmental factors that might influence their production. One could suppose that pili production might be altered in planktonic versus attached states, since surface contact leads to an increase in pilin expression in many bacterial species (Bertrand et al. 2010). In S. acidocaldarius, the two pili systems have different effects on biofilm formation with deletion of aap genes leading to a very dense biofilm and deletion of ups genes leading to an open biofilm structure (Hence et al. 2012b; Orell et al. 2013). Both systems appear to be upregulated in biofilm conditions (Hence et al. 2012b; Koerdt et al. 2011). It is known that type IV pili are necessary, in conjunction with archaella, for attachment to various surfaces in M. maripaludis (Jarrell et al. 2011). Whether transcription of epd genes is upregulated in attached cells could be analyzed by qRT-PCR following isolation of RNA from both attached and planktonic cells, as reported for Sulfolobus species (Koerdt et al. 2011). In addition, the appearance of Aap pili is strongly influenced by growth phase with much reduced piliation observed in stationary phase compared to logarithmic phase (Hence et al. 2012a). Ups pili are only observed following UV
treatment or other stress conditions that lead to DNA double strand breaks (Frols et al. 2008). It is known in bacteria that environmental conditions and nutrient availability can affect type IV pili production (Tomich et al. 2006; Bertrand et al. 2010; Coil and Anne 2009) and examining conditions that could affect piliation in *M. maripaludis* is a possible useful avenue of research to pursue.

Awareness of vital contributions of archaea to several environmental processes, such as methanogenesis, anaerobic methane oxidation and ammonia oxidation has been improved by our understanding of their diversity and distribution and as more species become amenable to genetic manipulation, and characterization. Several cell surfaces structures especially S-layers and various appendages including archaella and pili are key interfaces between the archaeal cell and its environment. Since the pili of archaea are often type IV-like, their study could help elucidate roadblocks in the study of their bacterial counterparts, important since type IV pili are confirmed virulence factors for numerous pathogens (Burrows. 2012; Nudleman and Kaiser 2004). There are already instances where studies of archaeal proteins have helped to elucidate abacterial or eukaryotic mechanism. For instance, it was the crystal structure of an archaeal oligosaccharyltransferase that led to the identification of a new catalytic motif which was then shown to be essential in the yeast enzyme (Igura et al 2008). Furthermore, it was examination of the crystal structure of the *Archaeoglobus fulgidus* secretion ATPase GspE that led to a unified mechanism for the secretion ATPases found in type IV pili and related systems that include type II secretion and archaellation (Yamagata and Tainer 2007). It is expected that continued studies on archaeal structures like their type IV pili will continue to aid in our understanding of not only this archaeal appendage but also archaella, bacterial type IV pili as well as the type II secretion systems (Craig and Li 2008).
Literature cited


Jarrell KF, Stark M, Nair DB, Chong JP (2011) Flagella and pili are both necessary for efficient attachment of Methanococcus maripaludis to surfaces. FEMS Microbiol Lett 319:44-50


Kaminski L, Guan Z, Yurist-Doutsch S, Eichler J (2013) Two distinct N-glycosylation pathways process the Haloferax volcanii S-layer glycoprotein upon changes in environmental salinity. MBio 4:e00716-13


184


Vassart A, Van Wolferen M, Orell A, Hong Y, Peeters E, Albers SV, Charlier D (2013) Sa-Lrp from *Sulfolobus acidocaldarius* is a versatile, glutamine-responsive, and architectural transcriptional regulator. Microbiologyopen 2:75-93


