

1

2 Evidence that Biosynthesis of the Second and Third Sugars of the Archaellin Tetrasaccharide in  
3 the Archaeon *Methanococcus maripaludis* Occurs by the Same Pathway Used by *Pseudomonas*  
4 *aeruginosa* to Make a Di-N-Acetylated Sugar

5

6 Sarah Siu<sup>1</sup>, Anna Robotham<sup>2</sup>, Susan M. Logan<sup>2</sup>, John F. Kelly<sup>2</sup>, Kaoru Uchida<sup>3</sup>, Shin-Ichi  
7 Aizawa<sup>3</sup> and Ken F. Jarrell<sup>1\*</sup>

8

9 <sup>1</sup>Department of Biomedical and Molecular Sciences, Queen's University, Kingston, Ontario,  
10 K7L 3N6, Canada

11 <sup>2</sup>Human Health Therapeutics Portfolio, National Research Council, Ottawa, Ontario,  
12 K1A 0R6, Canada

13 <sup>3</sup>Department of Life Sciences, Prefectural University of Hiroshima, 562 Nanatsuka, Shobara,  
14 Hiroshima 727-0023, Japan

15

16 \*corresponding author

17

18 Email: [jarrellk@queensu.ca](mailto:jarrellk@queensu.ca)

19 Phone: 613-533-2456

20 Fax: 613-533-6796

21

22 Running title: Biosynthesis of sugars of the archaellin N-glycan

23 Key words: archaellin, pilin, archaea, mass spectrometry, N-linked glycosylation

24

**25 ABSTRACT**

26 *Methanococcus maripaludis* has two surface appendages, archaella and type IV pili, which are composed  
27 of glycoprotein subunits. Archaellins are modified with a N-linked tetrasaccharide with the structure Sug-  
28 1,4- $\beta$ -ManNAc3NAmA6Thr-1,4- $\beta$ -GlcNAc3NAcA-1,3- $\beta$ -GalNAc, where Sug is (5S)-2-acetamido-2,4-  
29 dideoxy-5-O-methyl- $\alpha$ -L-erythro-hexos-5-ulose-1,5-pyranose. The pilin glycan has an additional hexose  
30 attached to GalNAc. In this study, genes located in two adjacent, divergently transcribed, operons  
31 (*mmp0350-mmp0354* and *mmp0359-mmp0355*) were targeted for study based on annotations suggesting  
32 their involvement in biosynthesis of N-glycan sugars. Mutants carrying deletions in *mmp0350*, *mmp0351*,  
33 *mmp0352* or *mmp0353* were nonarchaellated and synthesized archaellins modified with a 1-sugar glycan,  
34 as estimated from Western blots. Mass spectroscopy analysis of pili purified from the  $\Delta mmp0352$  strain  
35 confirmed a glycan with only GalNAc, suggesting *mmp0350-mmp0353* were all involved in biosynthesis  
36 of the second sugar (GlcNAc3NAcA). The  $\Delta mmp0357$  mutant was archaellated and had archaellins with  
37 a 2-sugar glycan, as confirmed by mass spectroscopy of purified archaella, indicating a role for  
38 MMP0357 in biosynthesis of the third sugar (ManNAc3NAmA6Thr). *M. maripaludis mmp0350*,  
39 *mmp0351*, *mmp0352*, *mmp0353* and *mmp0357* are proposed to be functionally equivalent to  
40 *Pseudomonas aeruginosa wbpABEDI* involved in converting UDP-N-acetylglucosamine to UDP-2,3-  
41 diacetamido-2,3-dideoxy-D-mannuronic acid, an O5-specific antigen sugar. Cross-domain  
42 complementation of the final step of the *P. aeruginosa* pathway with *mmp0357* supports this hypothesis.  
43

**44 IMPORTANCE.**

45 The work described identifies a series of genes in adjacent operons that are shown to encode the enzymes  
46 that complete the entire pathway for generation of the second and third sugars of the N-linked  
47 tetrasaccharide that modifies archaellins of *Methanococcus maripaludis*. This posttranslational  
48 modification of archaellins is important as it is necessary for archaella assembly. Pilins are modified with  
49 a different N-glycan consisting of the archaellin tetrasaccharide but with an additional hexose attached to  
50 the linking sugar. Mass spectrometry analysis of the pili of one mutant strain provided insight into how

51 this different glycan might ultimately be assembled. The studies include a rare example of an archaeal  
52 gene functionally replacing a bacterial gene in a complex sugar biosynthesis pathway.  
53

54 **INTRODUCTION**

55 While N-linked glycosylation was initially identified as an exclusively eukaryotic process, it is  
56 now well established that this pathway is present in the prokaryotic domains of Archaea and Bacteria as  
57 well. The post-translational modification of proteins with N-linked glycans is believed to be much more  
58 widespread in Archaea compared to Bacteria. In N-linked glycosylation systems, an  
59 oligosaccharyltransferase is required for the transfer of assembled oligosaccharides from the lipid carrier  
60 onto target proteins; genes encoding this key signature protein of the pathway have been identified in 166  
61 out of 168 sequenced archaeal genomes (1, 2). This is in contrast to Bacteria where the N-glycosylation  
62 pathway seems restricted to members of the  $\epsilon$ -subdivision of Proteobacteria (*Campylobacter*, *Wolinella*,  
63 *Helicobacter*) and a few Deltaproteobacteria, including *Desulfovibrio* species (3).

64 Though N-linked glycosylation is widespread in Archaea, study of N-linked glycosylation  
65 pathways by both genetic and structural methodologies is concentrated on a few model species (4),  
66 including the thermoacidophile *Sulfolobus acidocaldarius* (5-9), the halophile *Haloferax volcanii* (10-15)  
67 and the methanogens *Methanococcus voltae* (16, 17) and *Methanococcus maripaludis* (18-24). In general,  
68 the archaeal N-linked glycosylation pathway contains aspects of both the eukaryotic and bacterial  
69 systems, with additional domain-specific traits that include unique glycan structures, a variety of linking  
70 sugars and unusual dolichol carriers (4). Generally, archaeal N-linked glycans are initially assembled on  
71 the cytoplasmic face of the cytoplasmic membrane on dolichol lipid carriers through the activities of  
72 specific glycosyltransferases that catalyze the sequential addition of sugar monomers from nucleotide-  
73 charged sugar precursors. This lipid-linked oligosaccharide is then presumed to be flipped across the  
74 cytoplasmic membrane by a flippase prior to its *en bloc* transfer by the oligosaccharyltransferase AglB  
75 onto Asn residues within certain Asn-X-Ser/Thr sequons in the target protein, although additional sugars  
76 can still be added to the protein-bound glycan (4, 12).

77 N-linked glycans have been found to modify a variety of archaeal surface proteins including S-  
78 layer proteins (5, 13), as well as the subunits that comprise appendages including archaella (formerly  
79 archaeal flagella (25)) (14, 22) and pili (21). Typically, S-layer proteins and archaellins are used as

80 reporters of archaeal N-linked glycosylation (26). The importance of N-linked glycosylation for archaella  
81 assembly and function in *S. acidocaldarius* (7), *H. volcanii* (14), *M. voltae* (16) and *M. maripaludis* (22)  
82 is highlighted by archaellation and motility defects in mutants expressing archaellins that are either non-  
83 glycosylated or modified with truncated N-linked glycans.

84 The archaella of *M. maripaludis* S2 are composed of the major archaellins FlaB1 and FlaB2,  
85 located in the filament portion, and the minor archaellin FlaB3, found in the hook region (27). All three  
86 archaellins are modified by an N-linked tetrasaccharide at multiple sites (3 sites in FlaB1, 4 sites in  
87 FlaB2, and 2 sites in FlaB3) (18). The structure of the *M. maripaludis* S2 archaellin N-linked  
88 tetrasaccharide was determined to be Sug-1,4- $\beta$ -ManNAc3NAmA6Thr-1,4- $\beta$ -GlcNAc3NAcA-1,3- $\beta$ -  
89 GalNAc; Sug is [(5S)-2-acetamido-2,4,-dideoxy-5-O-methyl- $\alpha$ -L-erythro-hexos-5-ulose-1,5-pyranose]  
90 (18). An N-linked glycan also modifies the major pilin EpdE in *M. maripaludis* S2. However, the pilin  
91 glycan is a pentasaccharide that is identical to the archaellin tetrasaccharide except for an additional  
92 hexose that branches from the first sugar (N-acetylgalactosamine; GalNAc) (21).

93 Significant progress has been reported in the investigation of the N-linked glycosylation pathway  
94 in *M. maripaludis*. The oligosaccharyltransferase and the glycosyltransferases for the second, third, and  
95 fourth sugars of the archaellin tetrasaccharide have been identified as AglB (MMP1424), AglO  
96 (MMP1079), AglA (MMP1080), and AglL (MMP1088), respectively (22). Mutants carrying in-frame  
97 deletions of these genes exhibited archaellins bearing truncated N-linked glycans and defects in archaella  
98 assembly or function. Archaellins from *aglL* and *aglA* mutants are modified by 3- and 2-sugar glycans,  
99 respectively. Though  $\Delta aglL$  and  $\Delta aglA$  cells are archaellated, they are less motile than wild-type cells; the  
100 motility defects correlate with the degree of glycan truncation. Mutants with archaellins modified by a 1-  
101 sugar glycan, such as those found in an *aglO* mutant, or non-glycosylated archaellins, as present in an  
102 *aglB* mutant, are non-archaellated and thus, non-motile (22). Collectively, these results indicate that a  
103 minimum glycan of GlcNAc3NAcA-1,3- $\beta$ -GalNAc is required for archaella assembly in *M. maripaludis*  
104 S2; a similar requirement for archaellin modification by minimum N-linked glycans has been observed in

105 *M. voltae* (16) and *H. volcanii* (14). In *S. acidocaldarius*, mutants producing a truncated N-glycan were  
106 shown to have severely reduced motility (7) although recent studies indicate that strains engineered to  
107 lack all six glycosylation sites in the archaellin FlaB were still motile (9). Recent studies have also shown  
108 that for FlaB2 of *M. maripaludis*, at least one of the four sequons used for N-glycan attachment must be  
109 occupied for archaella assembly to occur (23) and for *H. volcanii* all sequons are necessary (14).

110 A number of gene products in *M. maripaludis* have been previously identified as involved in  
111 biosynthesis of the archaellin glycan. These include AglXYZ (MMP1081-MMP1083), responsible for the  
112 acetamidino group modification at C-3 of the third sugar (19), AglU (MMP1084), a threonine transferase  
113 that modifies the third sugar at C-6, and AglV (MMP1085), a methyltransferase for the methyl group at  
114 C-5 in the fourth sugar (20). MMP0350 was also identified as a putative acetyltransferase involved in  
115 biosynthesis of the second sugar (28). Additionally, *in vitro* assays using heterologously expressed  
116 proteins have identified other *M. maripaludis* gene products that are involved in acetamido sugar  
117 biosynthesis (29, 30), although genetic analysis of these is still lacking. Furthermore, the lipid carrier for  
118 the N-glycan in *M. maripaludis* was identified as a dolichol monophosphate (24).

119 In this work, we identify several *M. maripaludis* genes whose products are predicted to account  
120 for the complete biosynthesis of the second sugar (2,3-diacetamido-2,3-dideoxyglucuronic acid;  
121 GlcNAc3NAcA) of the archaellin glycan from UDP-GlcNAc as well as the biosynthesis of a direct  
122 precursor (2,3-diacetamido-2,3-dideoxymanuronic acid; ManNAc3NAcA) of the third sugar that is likely  
123 further modified by the activities of AglXYZ and AglU to generate the final form of the third sugar  
124 (ManNAc3NAmA6Thr). We also present evidence that biosynthesis of UDP-ManNAc3NAcA occurs in a  
125 manner similar to the WbpABEDI pathway that generates UDP-ManNAc3NAcA as a part of O-specific  
126 antigen (OSA) biosynthesis in *Pseudomonas aeruginosa* PAO1 lipopolysaccharide (31, 32).

## 127 MATERIALS AND METHODS

128       **Strains and growth conditions.** Strains used in this study are listed in Table A1. *Methanococcus*  
129 *maripaludis* S2  $\Delta$ hpt (Mm900) (33) and mutants derived from Mm900 were grown anaerobically under  
130 an atmosphere of CO<sub>2</sub>/H<sub>2</sub> (20:80) in Balch medium III (34) at 37°C with shaking. During the in-frame  
131 deletion mutagenesis process, Mm900-derived transformants were grown in McCas medium  
132 supplemented with neomycin (1 mg/mL) or 8-azahypoxanthine (250  $\mu$ g/mL, Acros Organics, NJ, USA)  
133 as required (33). In complementation experiments, *M. maripaludis* mutants harboring their respective  
134 complementation plasmids were grown in nitrogen-free medium containing puromycin (2.5  $\mu$ g/mL) for  
135 plasmid selection and supplemented with either L-alanine (10 mM) or NH<sub>4</sub>Cl (10 mM) as the lone  
136 nitrogen source (35). *Escherichia coli* TOP10 cells (Invitrogen, Burlington, ON, Canada) for plasmid  
137 cloning steps were grown in Luria-Bertani medium at 37°C supplemented with ampicillin (100  $\mu$ g/mL) as  
138 needed. For cross-domain complementation experiments, *Pseudomonas aeruginosa* strains were grown in  
139 Luria-Bertani medium supplemented with gentamicin (50  $\mu$ g/mL), carbenicillin (200  $\mu$ g/mL), and  
140 tetracycline (50  $\mu$ g/mL) when required. Antibiotics were obtained from Sigma-Aldrich (Oakville, ON,  
141 Canada).

142

143       **Construction of plasmids for in-frame deletions and complementations.** In-frame deletions  
144 were constructed for *mmp0351*, *mmp0352*, *mmp0353*, *mmp0355*, *mmp0357*, and *mmp0358* as previously  
145 described (33). In general, 1kb of the up- and downstream flanking regions were amplified by PCR using  
146 the corresponding primer pairs (P1 with P2, P3 with P4) listed in Table A2. Primers P2 and P3 introduced  
147 an AscI site so that ligation of the up- and downstream flanking regions yielded an in-frame deletion  
148 product for the target gene. BamHI or XbaI sites in P1 and P4 facilitated cloning of the in-frame deletion  
149 products into the vector pCRPrtNeo, resulting in pKJ1057 (*mmp0351*), pKJ1059 (*mmp0352*), pKJ980  
150 (*mmp0353*), pKJ1148 (*mmp0355*), pKJ978 (*mmp0357*), and pKJ1182 (*mmp0358*), respectively (Table  
151 A1).

152 For complementation of the *mmp0351*, *mmp0352*, *mmp0353*, and *mmp0357* deletion mutants,  
153 each gene was amplified by PCR using the complementation primers listed in Table A3. The resulting  
154 PCR products were digested with NsiI/MluI and cloned into NsiI/MluI-digested pHW40, a  
155 complementation vector in which transcription of the cloned gene is controlled by the inducible *nif*  
156 promoter (35). Internal NsiI sites were removed from *mmp0352* and *mmp0357* by a silent A-to-G and A-  
157 to-C change at nucleotide 891 and 663, respectively, using the site directed mutagenesis (SDM) primers  
158 listed in Table A3. The wild-type version of each respective gene was first cloned into the pCR2.1 TOPO  
159 vector and these TOPO vector constructs were then used as template for SDM.

160 The veracity of the in-frame deletion and complementation constructs were confirmed by DNA  
161 sequencing. Restriction enzymes were obtained from New England Biolabs, Fermentas, and Invitrogen. A  
162 PCR product purification kit from Qiagen (Toronto, Ontario Canada) was used according to the  
163 manufacturer's protocols.

164

165 **Generation of *M. maripaludis* mutants containing in-frame deletions of targeted genes.**

166 Previously, mutants containing markerless in frame deletions of *mmp0350* and *mmp0354* were reported  
167 (22, 28). Using the same methods established by Moore & Leigh (33), mutants containing markerless in-  
168 frame deletions of *mmp0351*, *mmp0352*, *mmp0353*, *mmp0355*, *mmp0357*, *mmp0358* and *mmp0359* were  
169 generated in *M. maripaludis* Mm900. PEG-mediated transformation (37) was used to introduce  
170 pCRPrtNeo in-frame deletion constructs into *M. maripaludis* Mm900 cells. After subsequent plating of  
171 transformants on McCas-Noble agar containing 8-azahypoxanthine, single colonies were picked, and then  
172 screened by PCR using primers designed to amplify across the target gene (listed in Table A4) in order to  
173 identify deletion mutants. Mutants identified by PCR screening were re-streaked, and screened by PCR  
174 again to check for purity.

175 The complementation vectors were transformed into their corresponding deletion mutants, again  
176 using the PEG precipitation methodology (37). These complementation strains were subsequently grown

177 under puromycin selection (2.5 µg/mL) in nitrogen-free medium supplemented with either L-alanine (*nif*  
178 promoter induced) or NH<sub>4</sub>Cl (*nif* promoter repressed) (35).

179

180         **RT-PCR.** In order to determine if the genes *mmp0350* to *mmp0354* are co-transcribed, and if the  
181 genes *mmp0359* to *mmp0355* are co-transcribed, reverse transcriptase polymerase chain reaction (RT-  
182 PCR) was performed using primers (listed in Table A5) designed to amplify across the intergenic regions  
183 between adjacent genes. RNA template was isolated from *M. maripaludis* S2 wild-type cells using a High  
184 Pure RNA Isolation Kit (Roche, Mississauga, Ontario, Canada) followed by an additional DNase  
185 treatment (TURBO DNA-free kit, Ambion, Burlington, Ontario, Canada). A One-step RT-PCR kit  
186 (Qiagen) was used to amplify cDNA, using the supplied protocol. Two additional standard PCR  
187 amplifications were conducted with each reaction using the same primer pairs but with different templates  
188 to verify amplicon sizes, using genomic DNA as template, and to rule out genomic DNA contamination  
189 of the RNA sample, using purified RNA that did not undergo reverse transcription.

190

191         **Western blot analysis of *M. maripaludis* archaellins and pilins.** Whole cell lysates of *M.*  
192 *maripaludis* were separated by SDS-PAGE and transferred onto Immobilon-P membrane (Millipore,  
193 Massachusetts, USA). For detection of archaellin FlaB2, 12.5% acrylamide gels were used while for  
194 detection of the pilin EpdE, 17.5% acrylamide gels were employed. FlaB2 was detected with anti-FlaB2  
195 specific polyclonal chicken antibodies (19); horseradish peroxidase-conjugated rabbit anti-chicken IgY  
196 (Jackson Laboratories Inc., PA, USA) was used as secondary antibody. Since antibody specific for the  
197 major pilin EpdE is unavailable, deletion mutants were transformed with pWLG40 carrying *epdE* with an  
198 additional C-terminal FLAG-tag to facilitate detection of EpdE-FLAG using monoclonal ANTI-FLAG®  
199 (Sigma-Aldrich); horseradish peroxidase-conjugated goat anti-mouse IgG (Jackson Laboratories) was  
200 used as secondary antibody. Blots were developed using a chemiluminescence kit (Roche) according to  
201 the manufacturer's protocols.

202

203       **Archaealla and pili isolation.** Archaealla from the *mmp0357* deletion mutant were isolated as  
204 previously described by Bardy *et al.* (38). Pili were isolated from the non-archaellated *mmp0352* deletion  
205 mutant using the same method (21).

206

207       **Mass spectrometry analysis.** Purified archaealla (50 µg) from *mmp0357* deletion mutant was  
208 digested overnight with trypsin (Promega, Madison WI) at a ratio of 30:1 (protein-enzyme [vol/vol]) in 50  
209 mM ammonium bicarbonate at 37°C. The pilin isolated from *mmp0352* deletion mutant was first resolved  
210 by SDS-PAGE and then stained with Coomassie blue. The pilin band was excised and subjected to an  
211 overnight in-gel trypic digestion. To improve peptide recovery, the digested pilin sample was then  
212 redigested with endoproteinase Asp-N (Roche, Indianapolis, IN). The digests were then analyzed by  
213 nano-liquid chromatography-tandem mass spectrometry (Nano-LC-MS/MS) using a NanoAquity UPLC  
214 system (Waters, Milford, MA) coupled to a QTOF Ultima hybrid quadrupole time-of-flight mass  
215 spectrometer (Waters). The digests were injected onto an Acclaim PepMax100 C<sub>18</sub> µ-precolumn (5mm by  
216 300 µm i.d.; Dionex/Thermo Scientific, Sunnyvale CA) and resolved on a 1.7- µm BEH130 C<sub>18</sub> column  
217 (100 µm by 100 mm i.d.; Waters, Milford, CA) using the following gradient conditions: 1 to 45%  
218 acetonitrile (ACN), 0.1% formic acid in 36 min and 45 to 95% ACN, 0.1% formic acid in 2 min. The  
219 flow rate was 400 nL/min. MS/MS spectra were acquired on doubly, triply and quadruply charged ions  
220 and searched against the NCBI nr database using the Mascot search engine (Matrix Science, Ltd., London,  
221 United Kingdom). The spectral datasets were searched for glycopeptide MS/MS spectra which were then  
222 interpreted by hand.

223

224       **Cross-domain complementation studies.** Given the sequence similarity between *mmp0353* and  
225 *mmp0357* to the *P. aeruginosa* genes *wbpA* and *wbpI* respectively, cross-domain complementation  
226 experiments were conducted to investigate these genes' potential functional equivalencies. *mmp0353* and  
227 *mmp0357* were each synthesized with a C-terminal His-tag using *P. aeruginosa* codon preferences  
228 (GenScript, NJ, USA); EcoRI/HindIII restriction sites were incorporated at the ends to facilitate cloning

229 into the *E. coli* – *P. aeruginosa* shuttle vectors pUCP18/pUCP19 and pUCP26/pUCP27. The vector pairs  
230 each have their multiple cloning sites in opposite orientations; thus, cloned genes are in the correct  
231 orientation when in pUCP18 and pUCP26. The Shine-Dalgarno and spacer sequence  
232 AGGAGGACAAGCT (39) was included at the start of each gene to facilitate expression.

233 To transform *P. aeruginosa* cells with the pUCP18/pUCP19 and pUCP26/pUCP27 constructs  
234 (listed in Table A1), electrocompetent cells were prepared according to Choi *et al.*(40), followed by  
235 electroporation using a MicroPulser (Bio-Rad). After a 2 hour recovery, cells were plated on LB  
236 containing gentamicin and carbenicillin (for pUCP18/pUCP19) or gentamicin and tetracycline (for  
237 pUCP26/pUCP27). Transformants were confirmed by a Gm<sup>R</sup>Cb<sup>R</sup> (those harboring pUCP18/pUCP19  
238 constructs) or Gm<sup>R</sup>Tc<sup>R</sup> (those harboring pUCP26/pUCP27 constructs) phenotype and recovery of the  
239 vector constructs using a QIAprep Spin Miniprep Kit (Qiagen).

240

241 **Bacteriophage D3 spot test and phage titration.** Restoration of wild-type O5-serotype OSA  
242 biosynthesis in complemented *P. aeruginosa* strains was assessed by spot test using bacteriophage D3  
243 (41), obtained from Dr. Andrew Kropinski. For these experiments, 10µl of a 2x10<sup>10</sup> pfu/ml stock of D3  
244 was used. Phage titration was conducted using the double overlay method described by Adams (42).

245

246 **LPS isolation and Western blot analysis.** LPS from overnight *P. aeruginosa* broth cultures was  
247 prepared as described by Hitchcock and Brown (43). LPS samples were separated on 15%-PAGE gels  
248 (44:0.8 acryl:bis, 0.0824 M NaCl) and detected by silver staining (43). The LPS samples were also  
249 analyzed by Western blotting using murine monoclonal antibody MF15-4 (O5-serotype OSA specific;  
250 obtained from Dr. Joe Lam) (31).

251

252 **Electron microscopy.** Electron microscopy of *M. maripaludis* cells was conducted using a  
253 Hitachi-7000 electron microscope operating at 75 kV. *M. maripaludis* cells were grown overnight; cells

254 were pelleted and briefly washed with phosphate-buffered saline. Cells were loaded onto Formvar-coated  
255 copper grids and negatively stained with 2% phosphotungstic acid (pH 7.0).

256

## 257 RESULTS

258 We have previously reported on a number of genes involved in the biosynthesis and assembly of  
259 the tetrasaccharide N-linked to archaellins of *M. maripaludis*. The majority of these genes lay in a large  
260 cluster (*mmp1079-mmp1088*) comprised of multiple operons (19, 20, 22). However, the deletion of one  
261 gene not located near this cluster, *mmp0350* which encodes a putative acetyltransferase, was also  
262 previously shown to affect glycan structure and archaella formation (28). In the immediate vicinity of  
263 *mmp0350* are a number of other genes annotated as potentially involved in the biosynthesis of the glycan  
264 sugars or assembly of the glycan. Of these, one (*mmp0354*, annotated as an oligosaccharide transporter)  
265 was deleted previously in an unsuccessful attempt to identify the flippase responsible for flipping the  
266 lipid-linked oligosaccharide prior to its transfer to target proteins by AglB (22). The MMP0352 protein  
267 was shown after expression and purification in *E. coli* to efficiently catalyze UDP-GlcNAc oxidation (29),  
268 although deletion analysis has yet to be reported. This report focuses on the role of the remaining genes in  
269 this locus in the N-glycosylation pathway of *M. maripaludis*.

270

### 271 Construction of in-frame deletions

272 In order to identify genes involved in sugar biosynthesis or glycan assembly in *M. maripaludis*,  
273 the genome was screened for clusters of genes whose annotations suggested roles in glycosylation. A  
274 cluster of ten genes, which included the previously studied putative acetyltransferase *mmp0350* (28), was  
275 identified and targeted for in-frame deletion to determine if any played a role in archaellin glycosylation.  
276 The annotations for these genes are listed in Table 1. In-frame deletion mutants were previously  
277 constructed for *mmp0350* and *mmp0354* (22, 28). Of the remaining eight genes, in-frame deletion mutants  
278 were successfully made in seven genes (*mmp0351*, *mmp0352*, *mmp0353*, *mmp0355*, *mmp0357*, *mmp0358*  
279 and *mmp0359*). Screening of over 100 transformants in multiple experiments for a deletion of *mmp0356*

280 yielded only wild-type colonies. Successful deletion of each gene was determined initially by a PCR  
281 screen (Figure A1), using the sequencing primers for each gene listed in Table A4. In all cases, the sizes  
282 of amplicons obtained by PCR in the mutants and wild-type strains were as predicted. The in-frame  
283 nature of each deletion was subsequently confirmed by DNA sequencing.

284

#### 285 **In-frame deletions result in a decrease in archaellin apparent molecular mass**

286 To determine if the in-frame deletion of *mmp0351*, *mmp0352*, *mmp0353*, *mmp0355*, *mmp0357*,  
287 *mmp0358* or *mmp0359* affected archaellin glycosylation, thereby resulting in archaellins of decreased  
288 apparent molecular mass, whole cell lysates were initially analyzed by Western blots probed with  
289 antibody specific for FlaB2, a major archaellin filament glycoprotein (19). Even minor truncations to the  
290 *M. maripaludis* archaellin tetrasaccharide result in decreases of archaellin apparent molecular mass that  
291 are detectable by Western blot as proteins of greater electrophoretic mobility (19, 20, 22). As previously  
292 reported, archaellins from the glycosyltransferase mutants  $\Delta aglL$ ,  $\Delta aglA$ , and  $\Delta aglO$  have glycan lengths  
293 of 3-, 2-, and 1-sugar, respectively, while archaellins from  $\Delta aglB$  are completely non-glycosylated (22).  
294 Whole cell lysates from  $\Delta aglL$ ,  $\Delta aglA$ ,  $\Delta aglO$ , and  $\Delta aglB$  strains collectively form a stepwise increase in  
295 FlaB2 electrophoretic mobility in Western blots which can be used effectively to initially estimate the  
296 degree of glycan truncation, if any, in novel deletion mutants. Archaellins from  $\Delta mmp0351$ ,  $\Delta mmp0352$ ,  
297 and  $\Delta mmp0353$  displayed the same increase in electrophoretic mobility as those from  $\Delta mmp0350$  and  
298  $\Delta aglO$  (i.e., with a predicted single sugar glycan) while archaellins from  $\Delta mmp0357$  displayed the same  
299 electrophoretic mobility as those from  $\Delta aglA$  (i.e., with a predicted two sugar glycan) (Figure 1A).

300 Due to their annotations suggesting roles in sugar biosynthesis and archaellin apparent molecular  
301 masses equal to those of  $\Delta aglO$  archaellins, it is believed that *mmp0351*, *mmp0352*, and *mmp0353* are all  
302 involved in biosynthesis of the second sugar (GlcNAc3NAcA) of the archaellin tetrasaccharide while the  
303 apparent molecular mass of archaellins from  $\Delta mmp0357$  suggested that *mmp0357* is involved in  
304 biosynthesis of the third sugar (ManNAc3NAmA6Thr).

305       Archaeellins from  $\Delta mmp0354$ ,  $\Delta mmp0355$ ,  $\Delta mmp0358$ , and  $\Delta mmp0359$  strains were all of  
306       apparent wild-type molecular mass as estimated from Western blots, suggesting that deletion of these  
307       genes did not affect archaellin glycosylation. The results for the *mmp0354* strain were consistent with the  
308       previous analysis of this deletion regarding effects on FlaB2 glycosylation (22). However, it is now  
309       known that EpdE, the major pilin of *M. maripaludis* is also a glycoprotein and it has an attached N-glycan  
310       that has the structure of the archaellin tetrasaccharide but with an extra hexose attached to the GalNAc  
311       linking sugar (21). Since *mmp0358* and *mmp0359* are annotated as glycosyltransferases, either could be  
312       the enzyme involved in the transfer of this hexose to the glycan. In addition, the annotation of *mmp0354*  
313       suggests it could act as a flippase (a wzx-like domain is present) that might be involved in the formation  
314       of the final pilin pentasaccharide. Thus, the possible involvement of all three of these genes was further  
315       examined for a role in pilin glycosylation. To investigate this, a FLAG-tagged version of EpdE was  
316       expressed in each of these mutants, as well as in wild-type cells, and the sizes of EpdE-FLAG determined  
317       by Western blots using anti-FLAG antibodies. If any of the deleted genes were involved in assembly of  
318       the pilin pentasaccharide, an increase in pilin electrophoretic mobility would be expected. However, in all  
319       cases the electrophoretic mobility of the EpdE-FLAG in the mutants was identical to that in the wild type  
320       (Figure 1B). We conclude that none of these three genes encode proteins that are involved in assembly of  
321       either the archaellin or pilin glycans.

322

### 323       **Gene complementation can restore archaellin molecular mass**

324       To confirm that the increase in electrophoretic mobility of archaellins in Western blots was due to  
325       the specific in-frame deletions, all deletion strains that showed an increase in the electrophoretic mobility  
326       of FlaB2 were complemented using a wild-type version of the deleted gene expressed *in trans*. *Mmp0350*,  
327       *mmp0351*, *mmp0352*, *mmp0353*, and *mmp0357* were each cloned into pHW40, a plasmid in which  
328       expression of the cloned gene is controlled by the inducible *nif* promoter (36). Complementation plasmids  
329       were transformed into their respective deletion mutants and maintained using puromycin selection.  
330       Expression of cloned genes from pHW40 was induced by growing cells in nitrogen-free media

331 supplemented with L-alanine. Cells were also grown in nitrogen-free media supplemented with NH<sub>4</sub>Cl.  
332 Though the nitrogen fixation pathway is not required in the presence of ammonia (44), a minor amount of  
333 transcription occurs from the *nif* promoter even in the presence of NH<sub>4</sub>Cl (22, 36).

334 When *mmp0351* was expressed in deletion mutants grown in N-free medium supplemented with  
335 alanine, archaellin apparent molecular mass was restored to that observed in wild-type cells. When the  
336 complemented strain was grown under NH<sub>4</sub>Cl-supplemented conditions, a ladder of FlaB2 molecular  
337 masses was detected, likely due to FlaB2 modified with partial glycans (Figure 2B). This is believed to be  
338 a result of limited *mmp0351* expression from pHW40 even when the *nif* promoter is repressed.

339 Complementation of the *mmp0352* and *mmp0357* deletion mutants also restored archaellin  
340 apparent molecular mass to that of wild-type cells when the complemented cells were grown in N-free  
341 medium with the addition of alanine (Figure 2C and Figure 2E). Unlike the example of the *mmp0351*  
342 complementation though, in these cases, archaellin apparent molecular mass was not restored to wildtype  
343 size, even partially, in cells grown in the presence of NH<sub>4</sub>Cl.

344 Complementation of *mmp0353* was also attempted multiple times. However, in each attempt, no  
345 change in the FlaB apparent molecular mass was observed upon complementation compared to the FlaB2  
346 size observed in the *mmp0353* mutant alone (Figure 2D). An analysis of the complementation of  
347 *mmp0350* by Western blot detecting FlaB2 was attempted in the initial study of this gene (28) but  
348 thwarted as the *mmp0350* deletion mutant stopped making FlaB2. In the current study, a new deletion of  
349 *mmp0350* was generated and again, complementation was attempted. Like before, the newly isolated  
350 mutant synthesized FlaB2 of increased electrophoretic mobility compared to the wild type when initially  
351 isolated. However, during the sub-culturing necessary in the course of the complementation experiment,  
352 the mutant again ceased making FlaB2, making the analysis of the complementation of archaellin  
353 glycosylation impossible (Figure 2A).

354

355 **In-frame deletions affect archaellation**

356       The presence or absence of archaella on the surface of *M. maripaludis* cells can be predicted from  
357 Western blot results; the minimum glycan length required for archaellation in *M. maripaludis* is the  
358 disaccharide GlcNAc3NAcA-1,3- $\beta$ -GalNAc (22). All deletion mutants were examined by electron  
359 microscopy to determine the effects of the deletions on archaellation (Figure 3).  $\Delta mmp0354$ ,  $\Delta mmp0355$ ,  
360  $\Delta mmp0358$ , and  $\Delta mmp0359$  were all archaellated, as predicted based on their archaellins' wild-type  
361 apparent molecular mass in Western blots.  $\Delta mmp0357$  archaellins were thought to be modified by a  
362 disaccharide based on Western blot results and these cells were predicted to be archaellated. This was  
363 confirmed by electron microscopy.  $\Delta mmp0351$ ,  $\Delta mmp0352$ , and  $\Delta mmp0353$  cells were all observed by  
364 electron microscopy to be non-archaellated, again consistent with their archaellins having a glycan of a  
365 single sugar as determined by Western blot analysis.

366        $\Delta mmp0351$  and  $\Delta mmp0352$  complemented cells were examined by electron microscopy to  
367 determine if the restoration of archaellin electrophoretic mobility to that of wild-type archaellin was  
368 accompanied by a restoration of the wild type archaellated phenotype. As expected, both  $\Delta mmp0351$  and  
369  $\Delta mmp0352$  complemented cells were archaellated (Figure 4). Complemented  $\Delta mmp0357$  cells were not  
370 examined by electron microscopy because the deletion mutant was already archaellated.

371

### 372 **Mass spectrometry analysis of archaellins from $\Delta mmp0357$ identifies a truncated glycan structure**

373       Of all the mutant strains generated in this study,  $\Delta mmp0357$  was the only strain that still produced  
374 archaella. To further characterize the glycan structure produced by this strain, archaella were purified,  
375 digested with trypsin and analyzed by mass spectroscopy (MS). MS analysis of tryptic glycopeptides  
376 confirmed that the glycan produced was indeed truncated compared to that found in the wildtype (Figure  
377 5a). Tryptic glycopeptides from  $\Delta mmp0357$  archaellin were shown to be modified with only a dimeric  
378 glycan species (258Da-203Da; GlcNAc3NAcA-1,3- $\beta$ -GalNAc). A representative spectrum of tryptic  
379 glycopeptide T<sub>53-81</sub> is presented in Figure 5b.

380

381   **Mass spectrometry analysis of pilins from *Δmmp0352* identifies a truncated glycan structure**

382

383           As the remaining mutants *Δmmp0350-Δmmp0353*, which all displayed an identical altered FlaB2  
384 migration pattern in Western blots, are unable to produce archaella, it was necessary to examine pili from  
385 these strains to determine what specific changes to the glycan structure had occurred. As a representative  
386 of these mutant strains, pili were isolated and analyzed from the *Δmmp0352* strain. Following  
387 trypsin/Asp-N double digestion of an excised gel band, a number of pilin glycopeptides were identified  
388 and characterized. In each case only a single monosaccharide moiety was present at each site of  
389 glycosylation (203Da; GalNAc). Representative spectra for pilin glycopeptides aa52-62 and aa34-47 are  
390 shown in Figure 6a and 6b, respectively. The glycopeptide aa34-47 contains two sequons that are adjacent  
391 to one another and both sites (with the modified asparagines separated by only 2 amino acids) could be  
392 found modified with a single GalNAc. The MS data for FlaB2 is consistent with the Western blot results  
393 and *in vitro* studies showing MMP0352 catalyzes UDP-GlcNAc oxidation at the 3" position (29), thus  
394 interfering with the synthesis of the second sugar of the N-linked glycan.

395

396   **RT-PCR analysis of *mmp0350-mmp0359* genes**

397           Based on the orientations and short intergenic regions, it was thought that *mmp0350-mmp0354*  
398 were likely co-transcribed, and that *mmp0359-mmp0355* also formed a single operon, transcribed from  
399 the complementary strand (Figure A2A). To test this hypothesis, primers were designed to amplify across  
400 the intergenic regions between adjacent genes within each possible operon. RT-PCR and standard PCR of  
401 RNA isolated from *M. maripaludis* Mm900 cells were performed as well as standard PCR with the same  
402 primers but using genomic DNA as template to confirm amplicon size and primer specificity (Figure  
403 A2B). In all cases, PCR products of the predicted size were amplified using DNA as template or RNA  
404 subjected to reverse transcription. No products were obtained when RNA was not subjected to the reverse  
405 transcriptase step prior to standard PCR, thereby indicating that the RNA samples were not contaminated

406 with DNA. The RT-PCR analysis shows that *mmp0350-mmp0354* are co-transcribed, and *mmp0359-*  
407 *mmp0355* are also co-transcribed.

408

409 **Cross-domain complementation restores O-specific antigen biosynthesis in *P. aeruginosa***

410 The gene annotations of *mmp0350*, *mmp0351*, *mmp0352*, *mmp0353*, and *mmp0357* suggested  
411 various roles in sugar biosynthesis (see Table 1). The amino acid sequences for each were used as queries  
412 for Basic Local Alignment Search Tool (BLAST; <http://blast.ncbi.nlm.nih.gov/Blast.cgi>). MMP0350,  
413 MMP0351, MMP0352, MMP0353, and MMP0357 were found to have high sequence similarity to the  
414 *Pseudomonas aeruginosa* proteins WbpD, WbpE, WbpB, WbpA, and WbpI, respectively. In *P.*  
415 *aeruginosa* PAO1, these five gene products are known to convert UDP-GlcNAc to UDP-  
416 ManNAc3NAcA in a stepwise fashion as part of O5-specific antigen biosynthesis (31, 32, 45).

417 Given the sequence similarities between the *P. aeruginosa* PAO1 gene products WbpD, WbpE,  
418 WbpB, WbpA, and WbpI and the *M. maripaludis* gene products MMP0350, MMP0351, MMP0352,  
419 MMP0353, and MMP0357, it was hypothesized that the enzymatic activities of the *M. maripaludis* gene  
420 products were responsible for the biosynthesis of UDP-ManNAc3NAcA from UDP-GlcNAc in a pathway  
421 that is functionally equivalent to the WbpABEDI pathway in *P. aeruginosa* PAO1. ManNAc3NAcA is a  
422 component of O-specific antigen (OSA) of O5 serotype *P. aeruginosa* PAO1; loss of UDP-  
423 ManNAc3NAcA biosynthesis results in loss of OSA expression (31, 32). ManNAc3NAcA is also likely a  
424 direct precursor of the third sugar of the *M. maripaludis* archaellin glycan, and could then undergo further  
425 modifications by AglXYZ and AglIU to generate ManNAc3NAmA6Thr.

426 Cross-domain complementation experiments were conducted for the first (WbpA) and last  
427 (WbpI) steps of the *P. aeruginosa* WbpABEDI pathway. Restoration of UDP-ManNAc3NAcA  
428 biosynthesis in *P. aeruginosa* was assessed by spot test using bacteriophage D3, a lysogenic  
429 bacteriophage that uses O5 OSA as its receptor (41). Thus, successfully complemented *P. aeruginosa*  
430 deletion mutants were predicted to become sensitive to bacteriophage D3, as expression of functionally

431 equivalent *M. maripaludis* gene products would restore UDP-ManNAc3NAcA biosynthesis and, in turn,  
432 the O5 OSA receptor for the phage.

433 MMP0353 and MMP0357 were expressed *in trans* in *P. aeruginosa*  $\Delta wbpA$  and *P. aeruginosa*  
434  $\Delta wbpI$ , respectively. For these experiments, *mmp0353* and *mmp0357* were synthesized using *P.*  
435 *aeruginosa* codon preferences, since the percent G+C content of the genome of each organism is  
436 dramatically different (*P. aeruginosa* PAO1 66.6% (46); *M. maripaludis* 33.1% (47)). In addition, Shine-  
437 Dalgarno sequences were added to the sequence for expression purposes (39).

438 Spot testing of *P. aeruginosa* strains using bacteriophage D3 was used to screen for O5 OSA  
439 expression. The control *P. aeruginosa* PAO1 exhibited a zone of lysis when spotted with bacteriophage  
440 D3. As expected, neither *P. aeruginosa*  $\Delta wbpA$  nor *P. aeruginosa*  $\Delta wbpI$  were sensitive to D3 in spot  
441 tests, confirming the absence of O5 OSA expression in these deletion strains (Figure 7). No lysis was  
442 observed when *P. aeruginosa*  $\Delta wbpA$  was complemented with MMP0353 expressed from pUCP18.  
443 Removal of the C-terminal His-tag from MMP0353 did not affect complementation results (data not  
444 shown); transcription of the non-His-tagged *mmp0353* from pUCP18 was confirmed by RT-PCR (see  
445 Figure A3). Additional attempts to complement *P. aeruginosa*  $\Delta wbpA$  using the non-His-tagged  
446 *mmp0353* expressed from the alternate shuttle vector pUCP26 (48) did not result in restoration of O5  
447 OSA expression when assessed by spot testing with bacteriophage D3 (data not shown).

448 In contrast to the complementation experiments with MMP0353, a zone of lysis was observed  
449 when *P. aeruginosa*  $\Delta wbpI$  was complemented using MMP0357 expressed from pUCP18 but not when  
450 the complementation vector was pUCP19 (Figure 7). Identical titers of D3 were obtained using *P.*  
451 *aeruginosa* PAO1 and the complemented *P. aeruginosa*  $\Delta wbpI$  as host. Further confirmation of the  
452 restoration of O5 OSA expression in *P. aeruginosa*  $\Delta wbpI$  when MMP0357 was expressed *in trans* was  
453 shown by Western blotting of isolated LPS using antibodies specific for the O5 OSA (Figure 8). As  
454 expected, no O5 OSA was detected in *P. aeruginosa*  $\Delta wbpI$  while the presence of O5 OSA was readily  
455 detected in the wild-type PAO1 strain. When *P. aeruginosa*  $\Delta wbpI$  was complemented with pUCP18

456 carrying *mmp0357*, restoration of O5 OSA synthesis was observed, in agreement with the sensitivity to  
457 D3. As expected, no O5 OSA was detected using *mmp0353* cloned into the pUCP19 complementation  
458 vector.

459 **DISCUSSION**

460 In this study, five genes involved in biosynthesis of the second and third sugars of the *M.*  
461 *maripaludis* archaellin N-linked tetrasaccharide glycan were identified; *mmp0350*, *mmp0351*, *mmp0352*,  
462 *mmp0353* and *mmp0357*. The in-frame deletion of these genes resulted in archaellins bearing truncated  
463 glycans, as shown by increased electrophoretic mobility of archaellins in western blots and by mass  
464 spectrometry analysis of archaellins and pilins. The genes collectively encode a five step pathway for the  
465 biosynthesis of UDP-ManNAc3NAcA from UDP-GlcNAc. The final step of the pathway is the 2-  
466 epimerization of UDP-GlcNAc3NAcA to UDP-ManNAc3NAcA; the function of the *M. maripaludis* gene  
467 product responsible for this step was demonstrated by cross-domain complementation of an equivalent *P.*  
468 *aeruginosa* deletion strain.

469 Deletions of *mmp0351*, *mmp0352*, and *mmp0353* all demonstrated a non-archaeallated phenotype,  
470 and these genes, along with  $\Delta mmp0350$  (28), contribute to biosynthesis of the second sugar  
471 (GlcNAc3NAcA). Western blotting of archaellins from  $\Delta mmp0351$ ,  $\Delta mmp0352$ , and  $\Delta mmp0353$  were all  
472 of decreased molecular mass indicative of modification by a 1-sugar glycan (GalNAc). In *M. maripaludis*,  
473 the presence of a 1-sugar glycan on archaellins is known to be insufficient for archaeallation although the  
474 cells remain pilated (22). The major pilin, EpdE, is known to be N-glycosylated with a variant of the  
475 archaellin N-glycan (21). The pilin glycan is a pentasaccharide that is identical to the archaellin glycan  
476 but with a hexose attached as a branch to the GalNAc linking sugar. This knowledge allowed for a  
477 determination of the effect of gene deletions on N-glycosylation even in non-archaeallated cells by the MS  
478 analysis of isolated pili. We isolated pili from the *mmp0352* deletion strain for analysis as a representative  
479 for the multiple deletion strains with archaellins bearing a 1-sugar glycan based on Western blots. MS  
480 analysis of EpdE pilins from  $\Delta mmp0352$  confirmed our prediction that this group of genes was involved  
481 in the synthesis of the second sugar as only tryptic glycopeptides bearing GalNAc were identified.  
482 Interestingly, the hexose branch found on GalNAc in pilin examined from wild-type cells was also absent.  
483 This result suggests that the attachment of the hexose cannot occur to a glycan consisting of GalNAc

484 alone and perhaps only occurs after the tetrasaccharide is completely assembled on the lipid carrier. This  
485 may even occur after the tetrasaccharide has been transferred to the pilins. It has been previously shown  
486 in *H. volcanii* that the transfer, by AlgS, of the terminal mannose residue in the pentasaccharide N-linked  
487 to S-layer occurs after the initial tetrasaccharide has been transferred to protein by AglB (49). In this  
488 scenario, the N-linked glycosylation system in *M. maripaludis* modifies both archaellins and pilins with  
489 the tetrasaccharide, however pilins are further modified by the addition of a hexose group by a specific  
490 glycosyltransferase after the tetrasaccharide is already attached to pilins. Three potential  
491 glycosyltransferases are present in the *mmp0350-mmp0359* gene cluster, namely *mmp0356*, *mmp0358* and  
492 *mmp0359*. However, when EpdE-FLAG was expressed in the *mmp0358* and *mmp0359* deletion strains,  
493 no reduced molecular mass compared to its expression in wild-type cells was detected in Western blots,  
494 appearing to eliminate these two genes as encoding the glycosyltransferase responsible for this pilin-  
495 specific step. All attempts to delete *mmp0356* were unsuccessful.

496 The gene *mmp0357* was also targeted for its potential role in sugar biosynthesis. Deletion of  
497 *mmp0357* resulted in archaellins of decreased apparent molecular mass but these cells were archaellated  
498 when examined by electron microscopy. Western blotting with anti-FlaB2 initially suggested the presence  
499 of a 2-sugar glycan on  $\Delta mmp0357$  archaellins and subsequent MS analysis of archaellins isolated from  
500  $\Delta mmp0357$  cells demonstrated a truncated glycan structure lacking ManNAc3NAmA6Thr and the  
501 terminal novel sugar [(5S)-2-acetamido-2,4,-dideoxy-5-O-methyl- $\alpha$ -L-erythro-hexos-5-ul-1,5-pyranose].  
502 Complementation of  $\Delta mmp0357$  using the wild-type gene restored archaellin molecular mass to that of  
503 the wild type. These data provide evidence that MMP0357 is involved in biosynthesis of the third sugar  
504 (ManNAc3NAmA) of the *M. maripaludis* archaellin glycan.

505 The five *M. maripaludis* genes (*mmp0350*, *mmp0351*, *mmp0352*, *mmp0353*, *mmp0357*) were  
506 assigned to specific steps of the UDP-ManNAc3NAcA biosynthesis pathway based on their high  
507 sequence similarity to gene products encoded by the *wbp* cluster in *P. aeruginosa* PAO1 (Figure 9). In *P.*  
508 *aeruginosa*, the conversion of UDP-GlcNAc to UDP-ManNAc3NAcA requires the sequential enzymatic

activities of WbpA, WbpB, WbpE, WbpD and WbpI. In-frame deletions and subsequent complementation studies have shown that all steps are essential for the synthesis of UDP-ManNAc3NAcA and that without this pathway the O5 OSA is not made (31, 32). Biosynthesis of UDP-ManNAc3NAcA in *Bordetella pertussis* proceeds by the identical scheme and successful cross-complementation of *P. aeruginosa* *wbp* deletion strains with the corresponding *B. pertussis* gene has been demonstrated (48). The evidence presented in this work demonstrates that this pathway also functions in Archaea. The evidence includes the extremely high sequence similarity of the five *M. maripaludis* genes to *wbpABEDI*, the MS analysis of N-linked glycans attached to either archaellins or pilins in the various *M. maripaludis* deletion strains which are fully consistent with the proposed pathway and the cross-domain complementation of the final step of the pathway in *P. aeruginosa* *ΔwbpI* by the corresponding *M. maripaludis* gene *mmp0357*, as evidenced by the restoration of O5 OSA synthesis. It is not clear why the cross complementation of the first step of the pathway carried out by WbpA using MMP0353 was unsuccessful. Such complementation was possible using the corresponding *B. pertussis* gene, but the *M. maripaludis* gene represent a much greater evolutionary distance. In all cases, C-terminal His-tagged versions of the complemented proteins were used and the same plasmid vectors were employed as well. Furthermore, it was not possible to show complementation even in the *M. maripaludis* *mmp0353* deletion mutant. We do not attribute this inability to be due to polar effects since the deletion of *mmp0353* was made in-frame and the only gene downstream and co-transcribed is *mmp0354* which when deleted on its own had no detectable effect on the N-linked glycan. It may be that MMP0353 forms a multi-enzyme complex and needs to be translationally coupled to the other members of the complex. Alternatively, the stoichiometry of the MMP0353 subunits in comparison with the other subunits of the proposed complex may be significantly different when *mmp0353* is expressed *in trans* from the *nif* promoter compared to its natural promoter. Unnatural variation in stoichiometry of subunits in a multi-subunit complex can lead to formation of a nonfunctional complex (50, 51).

Namboori and Graham (30) identified several enzymes involved in acetamido sugar biosynthesis in *M. maripaludis*. Particularly relevant to the current study are MMP0705 and MMP0706 which were

535 shown in *in vitro* assays, using proteins purified after expression in *E. coli*, to convert UDP-GlcNAc to  
536 UDP-ManNAc (via the 2-epimerase activity of MMP0705) and then to UDP-ManNAcA (via the UDP-  
537 ManNAc 6-dehydrogenase activity of MMP0706). *mmp0353* is a parologue of *mmp0706* and *mmp0357* is  
538 a parologue of *mmp0705* (30). Our deletion analysis indicates that *mmp0705* and *mmp0706* cannot  
539 compensate for the loss of *mmp0357* and *mmp0353*. Interestingly, Namboori and Graham (30) were able  
540 to demonstrate UDP-GlcNAc 2-epimerase activity for MMP0705 but not for MMP0357, despite the fact  
541 that MMP0357 is 53% identical to MMP0705. MMP0706 is homologous to not only UDP-ManNAc  
542 dehydrogenase but also to other enzymes, including UDP-Glc dehydrogenases, UDP-GalNAc  
543 dehydrogenase and UDP-glucose/GDP-mannose dehydrogenase. In some organisms, such as *B. pertussis*,  
544 it has been shown that some 6-dehydrogenases are promiscuous and can use UDP-GlcNAc, UDP-GalNAc  
545 and even UDP-Glc as substrates (48). While MMP0353 is 35% identical and 54% similar to MMP0706,  
546 its role in *M. maripaludis* is now believed to be as a UDP-GlcNAc dehydrogenase. Attempts to delete  
547 either of *mmp0705* or *mmp0706* have been unsuccessful to date (unpublished data), indicating they might  
548 play a role in an essential pathway, such as in the synthesis of coenzyme M, a critical coenzyme for  
549 methanogenesis (52) that is reported to be modified with ManNAcA (53). The *in vitro* work of Namboori  
550 and Graham on MMP0352 enzymatic activity demonstrated its ability to catalyze the NAD-dependent  
551 oxidation of the 3" position of UDP-GlcNAc (29). The genetic and MS data presented in the current work  
552 and the role proposed for MMP0352 in the pathway for generation of the second and ultimately the third  
553 sugars of the archaellin N-linked glycan are fully consistent with their findings.

554 The findings presented in this study significantly add to our knowledge of N-glycan biosynthesis  
555 in *M. maripaludis* as the five gene products investigated in this contribution complete the entire  
556 biosynthetic pathways for the second and third sugars of the archaellin tetrasaccharide. The combined  
557 activities of MMP0353, MMP0352, MMP0351 and MMP0350 would convert UDP-GlcNAc to UDP-  
558 GlcNAc3NAcA, which is the second sugar in the archaellin tetrasaccharide. This sugar would be  
559 transferred to the dolichol monophosphate lipid carrier by AglO. In addition, in continuation of the  
560 pathway, UDP-GlcNAc3NAcA would be converted through the 2-epimerase activity of MMP0357 to

561 UDP-ManNAc3NAcA. This sugar could be seen as an immediate precursor to the third sugar of the N-  
562 linked glycan, ManNAc3NAmA6Thr, which could be generated by the activities of gene products  
563 previously identified. AglXYZ has been shown to be responsible for the acetamidino group at position 3  
564 (19) and the glycosyltransferase AglA would transfer this intermediate to the lipid carrier (22). The final  
565 third sugar structure is obtained when AglU transfers a threonine to position 6. This threonine  
566 modification only occurs after the fourth sugar has been added to the lipid carrier (20). With the  
567 demonstrated involvement of *mmp0350*, *mmp0351*, *mmp0352*, *mmp0353* and *mmp0357* in the  
568 biosynthesis of the *M. maripaludis* N-linked glycan, these genes are now assigned *agl* gene designations  
569 (16) following the protocol proposed by Eichler *et al.* (54). These designations are: *agl17* (*mmp0350*),  
570 *agl18* (*mmp0351*), *agl19* (*mmp0352*), *agl20* (*mmp0353*) and *agl21* (*mmp0357*). A model for the N-linked  
571 glycosylation pathway of *M. maripaludis* incorporating the proposed biosynthetic roles of Agl17, Agl18,  
572 Agl19, Agl20, and Agl21 is depicted in Figure 10.

573

574     **ACKNOWLEDGEMENTS**

575  
576     This work was funded by a Discovery Grant from the Natural Sciences and Engineering Research  
577     Council of Canada (NSERC) (to KFJ). The authors thank Andrew Kropinski, Joe Lam and John Leigh for  
578     strains, bacteriophage and antibodies.

579

## 580 REFERENCES

- 581 1. **Magidovich, H., and J. Eichler.** 2009. Glycosyltransferases and oligosaccharyltransferases in  
582 Archaea: putative components of the N-glycosylation pathway in the third domain of life. FEMS  
583 Microbiol. Lett. **300**:122-130.
- 584 2. **Kaminski, L., M. N. Lurie-Weinberger, T. Allers, U. Gophna, and J. Eichler.** 2013. Phylogenetic-  
585 and genome-derived insight into the evolution of N-glycosylation in Archaea. Mol. Phylogenet. Evol.  
586 **68**:327-339.
- 587 3. **Nothaft, H., and C. M. Szymanski.** 2013. Bacterial protein N-glycosylation: new perspectives and  
588 applications. J. Biol. Chem. **288**:6912-6920.
- 589 4. **Jarrell, K. F., Y. Ding, B. H. Meyer, S. V. Albers, L. Kaminski, and J. Eichler.** 2014. N-Linked  
590 glycosylation in Archaea: a structural, functional, and genetic analysis. Microbiol. Mol. Biol. Rev.  
591 **78**:304-341.
- 592 5. **Peyfoon, E., B. Meyer, P. G. Hitchen, M. Panico, H. R. Morris, S. M. Haslam, S. V. Albers, and  
593 A. Dell.** 2010. The S-layer glycoprotein of the crenarchaeote *Sulfolobus acidocaldarius* is glycosylated at  
594 multiple sites with the chitobiose-linked N-glycans. Archaea doi: 10.1155/2010/754101.
- 595 6. **Meyer, B. H., B. Zolghadr, E. Peyfoon, M. Pabst, M. Panico, H. R. Morris, S. M. Haslam, P.  
596 Messner, C. Schäffer, A. Dell, and S. V. Albers.** 2011. Sulfoquinovose synthase - an important enzyme  
597 in the N-glycosylation pathway of *Sulfolobus acidocaldarius*. Mol. Microbiol. **82**:1150-1163.
- 598 7. **Meyer, B. H., E. Peyfoon, C. Dietrich, P. Hitchen, A. Dell, and S. V. Albers.** 2013. Agl16, a  
599 thermophilic glycosyltransferase, mediating the last step of the N-glycan biosynthesis in the  
600 thermoacidophilic crenarchaeon *Sulfolobus acidocaldarius*. J. Bacteriol. **195**:2177-2186.

- 601    8. **Meyer, B. H., and S. V. Albers.** 2013. Hot and sweet: protein glycosylation in Crenarchaeota.  
602    Biochem. Soc. Trans. **41**:384-392.
- 603    9. **Meyer, B. H., A. Birich, and S. V. Albers.** 2014. N-Glycosylation of the archaellum filament is not  
604    important for archaella assembly and motility, although N-Glycosylation is essential for motility in  
605    *Sulfolobus acidocaldarius*. Biochimie. . doi: 10.1016/j.biochi.2014.10.018.
- 606    10. **Guan, Z., S. Naparstek, D. Calo, and J. Eichler.** 2012. Protein glycosylation as an adaptive  
607    response in Archaea: growth at different salt concentrations leads to alterations in *Haloferax volcanii* S-  
608    layer glycoprotein N-glycosylation. Environ. Microbiol. **14**:743-753.
- 609    11. **Kaminski, L., Z. Guan, M. Abu-Qarn, Z. Konrad, and J. Eichler.** 2012. AglR is required for  
610    addition of the final mannose residue of the N-linked glycan decorating the *Haloferax volcanii* S-layer  
611    glycoprotein. Biochim. Biophys. Acta. **1820**:1664-1670.
- 612    12. **Eichler, J.** 2013. Extreme sweetness:protein glycosylation in Archaea. Nature Reviews Microbiology.  
613    **11**:151-156.
- 614    13. **Kaminski, L., Z. Guan, S. Yurist-Doutsch, and J. Eichler.** 2013. Two distinct N-glycosylation  
615    pathways process the *Haloferax volcanii* S-layer glycoprotein upon changes in environmental salinity.  
616    Mbio. **4**:e00716-13.
- 617    14. **Tripepi, M., J. You, S. Temel, Ö. Önder, D. Brisson, and M. Pohlschröder.** 2012. N-glycosylation  
618    of *Haloferax volcanii* flagellins requires known Agl proteins and is essential for biosynthesis of stable  
619    flagella. J. Bacteriol. **194**:4876-4887.
- 620    15. **Guan, Z., S. Naparstek, D. Calo, and J. Eichler.** 2012. Protein glycosylation as an adaptive  
621    response in Archaea: growth at different salt concentrations leads to alterations in *Haloferax volcanii* S-  
622    layer glycoprotein N-glycosylation. Environ. Microbiol. **14**:743-753.

- 623 16. **Chaban, B., S. Voisin, J. Kelly, S. M. Logan, and K. F. Jarrell.** 2006. Identification of genes  
624 involved in the biosynthesis and attachment of *Methanococcus voltae* N-linked glycans: insight into N-  
625 linked glycosylation pathways in Archaea. *Mol. Microbiol.* **61**:259-268.
- 626 17. **Voisin, S., R. S. Houliston, J. Kelly, J. R. Brisson, D. Watson, S. L. Bardy, K. F. Jarrell, and S.**  
627 **M. Logan.** 2005. Identification and characterization of the unique N-linked glycan common to the  
628 flagellins and S-layer glycoprotein of *Methanococcus voltae*. *J. Biol. Chem.* **280**:16586-16593.
- 629 18. **Kelly, J., S. M. Logan, K. F. Jarrell, D. J. Vandyke, and E. Vinogradov.** 2009. A novel N-linked  
630 flagellar glycan from *Methanococcus maripaludis*. *Carbohydr. Res.* **344**:648-653.
- 631 19. **Jones, G. M., J. Wu, Y. Ding, K. Uchida, S. Aizawa, A. Robotham, S. M. Logan, J. Kelly, and K.**  
632 **F. Jarrell.** 2012. Identification of genes involved in the acetamidino group modification of the flagellin  
633 N-linked glycan of *Methanococcus maripaludis*. *J. Bacteriol.* **194**:2693-2702.
- 634 20. **Ding, Y., G. M. Jones, K. Uchida, S. I. Aizawa, A. Robotham, S. M. Logan, J. Kelly, and K. F.**  
635 **Jarrell.** 2013. Identification of genes involved in the biosynthesis of the third and fourth sugars of the  
636 *Methanococcus maripaludis* archaellin N-linked tetrasaccharide. *J. Bacteriol.* **195**:4094-4104.
- 637 21. **Ng, S. Y. M., J. Wu, D. B. Nair, S. M. Logan, A. Robotham, L. Tessier, J. F. Kelly, K. Uchida, S.**  
638 **- Aizawa, and K. F. Jarrell.** 2011. Genetic and mass spectrometry analysis of the unusual type IV-like  
639 pili of the archaeon *Methanococcus maripaludis*. *J. Bacteriol.* **193**:804-814.
- 640 22. **Vandyke, D. J., J. Wu, S. M. Logan, J. F. Kelly, S. Mizuno, S. I. Aizawa, and K. F. Jarrell.** 2009.  
641 Identification of genes involved in the assembly and attachment of a novel flagellin N-linked  
642 tetrasaccharide important for motility in the archaeon *Methanococcus maripaludis*. *Mol. Microbiol.*  
643 **72**:633-644.

- 644 23. **Ding, Y., K. Uchida, S. I. Aizawa, K. Murphy, C. M. Khursigara, J. P. J. Chong, and K. F.**  
645 **Jarrell.** 2015. Effects of N-glycosylation site removal in archaellins on the assembly and function of  
646 archaella in *Methanococcus maripaludis*. PloS One. accepted manuscript.
- 647 24. **Larkin, A., M. M. Chang, G. E. Whitworth, and B. Imperiali.** 2013. Biochemical evidence for an  
648 alternate pathway in N-linked glycoprotein biosynthesis. *Nature Chem. Biol.* **9**:367-373.
- 649 25. **Jarrell, K. F., and S. V. Albers.** 2012. The archaellum: an old motility structure with a new name.  
650 *Trends Microbiol.* **20**:307-312. doi: 10.1016/j.tim.2012.04.007.
- 651 26. **Jarrell, K. F., G. M. Jones, L. Kandiba, D. B. Nair, and J. Eichler.** 2010. S-layer glycoproteins  
652 and flagellins: reporters of archaeal posttranslational modifications. *Archaea* doi: 10.1155/2010/612948.
- 653 27. **Chaban, B., S. Y. Ng, M. Kanbe, I. Saltzman, G. Nimmo, S. I. Aizawa, and K. F. Jarrell.** 2007.  
654 Systematic deletion analyses of the *fla* genes in the flagella operon identify several genes essential for  
655 proper assembly and function of flagella in the archaeon, *Methanococcus maripaludis*. *Mol. Microbiol.*  
656 **66**:596-609.
- 657 28. **VanDyke, D. J., J. Wu, S. Y. Ng, M. Kanbe, B. Chaban, S. I. Aizawa, and K. F. Jarrell.** 2008.  
658 Identification of putative acetyltransferase gene, MMP0350, which affects proper assembly of both  
659 flagella and pili in the archaeon *Methanococcus maripaludis*. *J. Bacteriol.* **190**:5300-5307.
- 660 29. **Namboori, S. C., and D. E. Graham.** 2008. Enzymatic analysis of uridine diphosphate N-acetyl-D-  
661 glucosamine. *Anal. Biochem.* **301**:94-100.
- 662 30. **Namboori, S. C., and D. E. Graham.** 2008. Acetamido sugar biosynthesis in the Euryarchaeia. *J.*  
663 *Bacteriol.* **190**:2987-2996.

- 664 31. **Burrows, L. L., K. E. Pigeon, and J. S. Lam.** 2000. *Pseudomonas aeruginosa* B-band  
665 lipopolysaccharide genes *wbpA* and *wbpI* and their *Escherichia coli* homologues *wecC* and *wecB* are not  
666 functionally interchangeable. FEMS Microbiol. Lett. **189**:135-141.
- 667 32. **Westman, E. L., D. J. McNally, A. Charchoglyan, D. Brewer, R. A. Field, and J. S. Lam.** 2009.  
668 Characterization of WbpB, WbpE, and WbpD and reconstitution of a pathway for the biosynthesis of  
669 UDP-2,3-diacetamido-2,3-dideoxy-D-mannuronic acid in *Pseudomonas aeruginosa*. J. Biol. Chem.  
670 **284**:11854-11862.
- 671 33. **Moore, B. C., and J. A. Leigh.** 2005. Markerless mutagenesis in *Methanococcus maripaludis*  
672 demonstrates roles for alanine dehydrogenase, alanine racemase, and alanine permease. J. Bacteriol.  
673 **187**:972-979.
- 674 34. **Balch, W. E., G. E. Fox, L. J. Magrum, C. R. Woese, and R. S. Wolfe.** 1979. Methanogens:  
675 reevaluation of a unique biological group. Microbiol. Rev. **43**:260-296.
- 676 35. **Lie, T. J., and J. A. Leigh.** 2003. A novel repressor of *nif* and *glnA* expression in the methanogenic  
677 archaeon *Methanococcus maripaludis*. Mol. Microbiol. **47**:235-246.
- 678 36. **Lie, T. J., G. E. Wood, and J. A. Leigh.** 2005. Regulation of *nif* expression in *Methanococcus*  
679 *maripaludis*: roles of the euryarchaeal repressor NrpR, 2-oxoglutarate, and two operators. J. Biol. Chem.  
680 **280**:5236-5241.
- 681 37. **Tumbula, D. L., R. A. Makula, and W. B. Whitman.** 1994. Transformation of *Methanococcus*  
682 *maripaludis* and identification of a *PstI*-like restriction system. FEMS Microbiol. Lett. **121**:309-314.
- 683 38. **Bardy, S. L., T. Mori, K. Komoriya, S. Aizawa, and K. F. Jarrell.** 2002. Identification and  
684 localization of flagellins FlaA and FlaB3 within flagella of *Methanococcus voltae*. J. Bacteriol. **184**:5223-  
685 5233.

- 686 39. **Chiang, P., M. Habash, and L. L. Burrows.** 2005. Disparate subcellular localization patterns of  
687 *Pseudomonas aeruginosa* Type IV pilus ATPases involved in twitching motility. *J. Bacteriol.* **187**:829-  
688 839.
- 689 40. **Choi, K. H., A. Kumar, and H. P. Schweizer.** 2006. A 10-min method for preparation of highly  
690 electrocompetent *Pseudomonas aeruginosa* cells: application for DNA fragment transfer between  
691 chromosomes and plasmid transformation. *J. Microbiol. Methods.* **64**:391-397.
- 692 41. **Taylor, V. L., M. L. Udaskin, S. T. Islam, and J. S. Lam.** 2013. The D3 bacteriophage  $\alpha$ -  
693 polymerase inhibitor (Iap) peptide disrupts O-Antigen biosynthesis through mimicry of the chain length  
694 regulator Wzz in *Pseudomonas aeruginosa*. *J. Bacteriol.* **195**:4735-4741.
- 695 42. **Adams, M. H.** 1959. Bacteriophages. Interscience Publishers Inc., New York.
- 696 43. **Hitchcock, P. J., and T. M. Brown.** 1983. Morphological heterogeneity among *Salmonella*  
697 lipopolysaccharide chemotypes in silver-stained polyacrylamide gels. *J. Bacteriol.* **154**:269-277.
- 698 44. **Kessler, P. S., and J. A. Leigh.** 1999. Genetics of nitrogen regulation in *Methanococcus maripaludis*.  
699 *Genetics.* **152**:1343-1351.
- 700 45. **Wenzel, C. Q., C. Daniels, R. A. B. Keates, D. Brewer, and J. S. Lam.** 2005. Evidence that WbpD  
701 is an N-acetyltransferase belonging to the hexapeptide acyltransferase superfamily and an important  
702 protein for O-antigen biosynthesis in *Pseudomonas aeruginosa* PAO1. *Mol. Microbiol.* **57**:1288-1303.
- 703 46. **Stover, C. K., X. Q. Pham, A. L. Erwin, S. D. Mizoguchi, P. Warrener, M. J. Hickey, F. S. L.  
704 Brinkman, W. O. Hufnagle, D. J. Kowalik, M. Lagrou, R. L. Garber, L. Goltby, E. Tolentino, S.  
705 Westbrock-Wadman, Y. Yuan, L. L. Brody, S. N. Coulter, K. R. Folger, A. Kas, K. Larbig, R. Lim,  
706 K. Smith, D. Spencer, G. K. S. Wong, Z. Wu, I. T. Paulsen, J. Reizer, M. H. Saier, R. E. W.**

- 707   **Hancock, S. Lory, and M. V. Olson.** 200. Complete genome sequence of *Pseudomonas aeruginosa*  
708   PAO1, an opportunistic pathogen. *Nature*. **406**:959-964.
- 709   47. **Hendrickson, E. L., R. Kaul, Y. Zhou, D. Bovee, P. Chapman, J. Chung, E. Conway de Macario,**  
710   **J. A. Dodsworth, W. Gillett, D. E. Graham, M. Hackett, A. K. Haydock, A. Kang, M. L. Land, R.**  
711   **Levy, T. J. Lie, T. A. Major, B. C. Moore, I. Porat, I. Palmeiri, G. Rouse, C. Saenphimmachak, D.**  
712   **Soll, S. Van Dien, T. Wang, W. B. Whitman, Q. Xia, Y. Zhang, F. W. Larimer, M. V. Olson, and J.**  
713   **A. Leigh.** 2004. Complete genome sequence of the genetically tractable hydrogenotrophic methanogen  
714   *Methanococcus maripaludis*. *J. Bacteriol.* **186**:6956-6969.
- 715   48. **Westman, E. L., A. Preston, R. A. Field, and J. S. Lam.** 2008. Biosynthesis of a rare di-N-  
716   acetylated sugar in the lipopolysaccharides of both *Pseudomonas aeruginosa* and *Bordetella pertussis*  
717   occurs via an identical scheme despite different gene clusters. *J. Bacteriol.* **190**:6060-6069.
- 718   49. **Cohen-Rosenzweig, C., S. Yurist-Doutsch, and J. Eichler.** 2012. AglS, a novel component of the  
719   *Haloflexax volcanii* N-glycosylation pathway, is a dolichol phosphate-mannose mannosyltransferase. *J.*  
720   *Bacteriol.* **194**:6909-6916.
- 721   50. **Cieslewicz, M., and E. Vimr.** 1997. Reduced polysialic acid capsule expression in *Escherichia coli*  
722   K1 mutants with chromosomal defects in *kpsF*. *Mol. Microbiol.* **26**:237-249.
- 723   51. **Løvdok, L., K. Bentele, N. Vladimirov, A. Müller, F. S. Pop, D. Lebiedz, M. Kollmann, and V.**  
724   **Sourjik.** 2009. Role of translational coupling in robustness of bacterial chemotaxis pathway. *PloS Biol.*  
725   **7**: e1000171.
- 726   52. **DiMarco, A. A., T. A. Bobik, and R. S. Wolfe.** 1990. Unusual coenzymes of methanogenesis. *Annu.*  
727   *Rev. Biochem.* **59**:355-394.

- 728 53. **Sauer, F. D., B. A. Blackwell, J. K. Kramer, and B. J. Marsden.** 1990. Structure of a novel  
729 cofactor containing N-(7-mercaptopentanoyl)-O-3-phosphothreonine. *Biochemistry*. **29**:7593-7600.
- 730 54. **Eichler, J., K. Jarrell, and S. Albers.** 2013. A proposal for the naming of N-glycosylation pathway  
731 components in Archaea. *Glycobiology*. **23**:620-621.
- 732
- 733

734 **Figure Legends**

735

736

737

738 **Figure 1. Western blot detection of archaellin FlaB2 and pilin EpdE-FLAG from in-frame deletion**

739 **mutants.** A. Blots were developed with *M. maripaludis* anti-FlaB2. FlaB2 modified by known glycan

740 lengths are as indicated for wild type (Mm900), glycosyltransferase mutants (*aglO*, *aglA*, *aglL*) and the

741 oligosaccharyltransferase mutant (*aglB*). B. Wildtype cells (Mm900) and the indicated deletion mutants

742 were transformed with pKJ1107 and EpdE-FLAG detected with anti-FLAG antibodies.

743

744 **Figure 2. Western blot analysis of complementation experiments.** Whole cell lysates from (A)

745 *mmp0350*, (B) *mmp0351*, (C) *mmp0352*, (D) *mmp0353* and (E) *mmp0357* deletion mutants as well as the

746 same deletion mutants harboring their respective complementation vectors were separated by SDS-PAGE,

747 transferred to Immobilon membrane and probed with anti-FlaB2. Complemented strains were grown in

748 N-free medium with addition of either alanine (transcription on) or NH<sub>4</sub>Cl (transcription off). Wildtype

749 cell lysates (Mm900) were run as control.

750

751 **Figure 3. Electron micrographs of wild type and in-frame deletion mutants showing archaellation**

752 **state.** Whole cells are shown in the insert with a portion of the whole cell enlarged to more readily

753 observe archaella. All samples were negatively stained with 2% phosphotungstic acid (pH 7.0). Arrows

754 indicate archaella. Scale bars represent 0.5 μm.

755

756 **Figure 4. Electron micrographs of (A) *mmp0351* and (B) *mmp0352* deletion mutants complemented**

757 **with the wild-type version of the respective gene.** The deletions alone are shown in the left panel

758 followed by the complement-on cells grown in N-free medium with addition of alanine and complement-

759 off cells grown in N-free medium with the addition of NH<sub>4</sub>Cl. All samples were negatively stained with  
760 2% phosphotungstic acid (pH 7.0). Arrows indicate archaella. Scale bars represent 0.5 μm.

761

762 **Figure 5. NanoLC-MS/MS analysis of the FlaB2 tryptic glycopeptide, T<sup>53-81</sup>, from the (a) WT and**  
763 **the (b) Δmmp 0357 mutant strain.** The FlaB2 tryptic peptide T<sup>53-81</sup> contains one site of N-glycosylation  
764 (DSTEQVASGLQIMGISGYQAGTANANITK). The major carbohydrate oxonium ions are identified in  
765 the MS/MS spectra using symbols to indicate the sugar residues present (■: GalNAc, ●:  
766 GlcNAc3NAcA, ▲: ManNAc3NAmA6Thr, ♠: ManNAc3NAmA, ♦: (5S)-2-acetamido-2,4-dideoxy-5-  
767 O-methyl-α-L-erythro-hexos- 5-ulose-1,5-pyranose). The b and y ions arising from fragmentation of the  
768 peptide bonds are also shown. The structure of tetrameric WT glycan has been described previously  
769 (Kelly *et al*, 2009). The glycopeptide from the Δmmp0357 mutant strain (panel b) is modified with a  
770 disaccharide lacking the 3<sup>rd</sup> and 4<sup>th</sup> sugar residue of the WT glycan.

771

772 **Figure 6. NanoLC-MS/MS analysis of Δmmp0352 pilin tryptic/Asp-N glycopeptides.** (a) MS/MS  
773 spectrum of the doubly protonated glycopeptide ion at m/z 693.3 corresponding to <sup>52</sup>DATSQMSNITD<sup>62</sup>  
774 with a single GalNAc (■). The same peptide was observed with no GalNAc residues attached (data not  
775 shown). (b) MS/MS spectrum of the doubly protonated glycopeptide ion at m/z 949.9 corresponding to  
776 <sup>34</sup>QITNSTQTTQALA<sup>47</sup> modified with two GalNAc residues. The same peptide was also observed with  
777 0 and 1 GalNAc modifications (data not shown). Both MS/MS spectra are complicated by the strong loss  
778 of ammonia (-17 Da) and water (-18 Da) from the major fragment ions. Nevertheless, there is sufficient  
779 evidence in these spectra to confirm that these peptides are modified with one and two GalNAc residues,  
780 respectively.

781

782

783 **Figure 7. Spot test plates of *P. aeruginosa* strains using bacteriophage D3.** Quadrants marked “D3”  
784 were spotted with 10 µL bacteriophage D3 ( $2 \times 10^{10}$  PFU/mL stock). Quadrants marked “LB” were spotted  
785 with 10 µL of LB media as a negative control. Complementation of *P. aeruginosa*  $\Delta wbpI$  using  
786 MMP0357 expressed *in trans* restored D3 sensitivity.

787

788 **Figure 8. Restoration of O5 OSA synthesis in the complemented *P. aeruginosa*  $\Delta wbpI$  strain. A.**  
789 Silver staining of LPS isolated from *P. aeruginosa* strains. B. Western blot of LPS isolated from *P.*  
790 *aeruginosa* strains; the blot was developed using MF15-4, antibodies specific for O5 OSA.  
791 Complementation of *P. aeruginosa*  $\Delta wbpI$  using MMP0357 expressed *in trans* restored O5 OSA  
792 biosynthesis.

793

794 **Figure 9. Proposed biosynthetic pathway for UDP-D-ManNAc3NAcA in *M. maripaludis*, based on**  
795 **the WbpABEDI pathway determined for *P. aeruginosa*.** *P. aeruginosa* proteins are shown in brackets  
796 for each step. Additional steps required for generation of the final *M. maripaludis* sugars are indicated, as  
797 well.

798

799 **Figure 10. Working model of the N-glycosylation pathway in *M. maripaludis*, highlighting the gene**  
800 **products analyzed in the current paper.** The MMP nomenclature has been replaced with Agl  
801 designations: Agl17 (MMP0350), Agl18 (MMP0351), Agl19 (MMP0352), Agl20 (MMP0353) and Agl21  
802 (MMP0357).

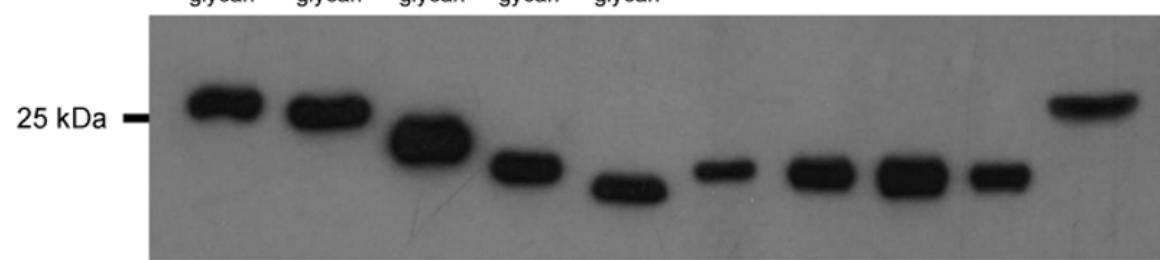
803      Table 1: *M. maripaludis* genes targeted for involvement in N-linked glycosylation in this study  
 804  
 805

Gene	Annotation	BLAST results for <i>P. aeruginosa</i> Wbp proteins (% coverage E-value)
<i>mmp0350</i> ( <i>agl17</i> )	Putative acetyltransferase	<i>P. aeruginosa</i> WbpD (82% 1e-59)
<i>mmp0351</i> ( <i>agl18</i> )	Predicted pyridoxal phosphate-dependent aminotransferase	<i>P. aeruginosa</i> WbpE (94% 8e-66)
<i>mmp0352</i> ( <i>agl19</i> )	Putative oxidoreductase	<i>P. aeruginosa</i> WbpB (37% 4e-6)
<i>mmp0353</i> ( <i>agl20</i> )	UDP-glucose/GDP-mannose dehydrogenase-related protein	<i>P. aeruginosa</i> WbpA (93% 6e-52)
<i>mmp0354</i>	Putative oligosaccharide transporter (flippase)	-
<i>mmp0355</i>	Conserved hypothetical protein	-
<i>mmp0356</i>	Glycosyltransferase, group 1	-
<i>mmp0357</i> ( <i>agl21</i> )	UDP-N-acetylglucosamine 2-epimerase	<i>P. aeruginosa</i> WbpI (100% 5e-118)
<i>mmp0358</i>	Glycosyltransferase, family 1	-
<i>mmp0359</i>	Glycosyltransferase, family 2	-

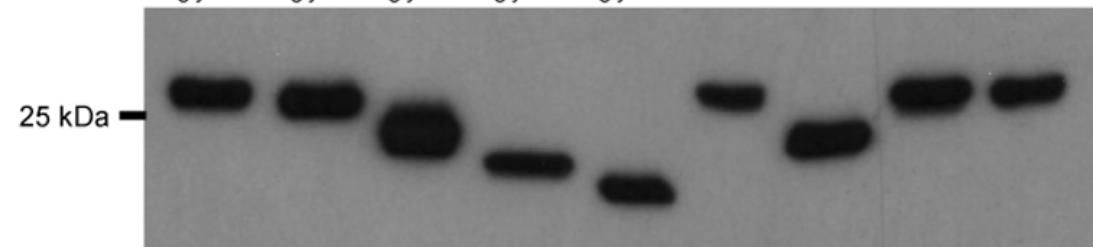
806

**A**

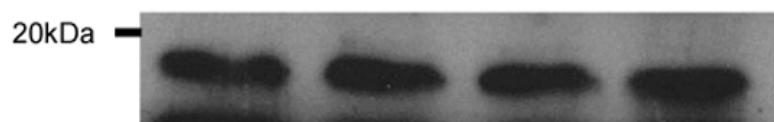
Mm900	<i>aglL</i>	<i>aglA</i>	<i>aglO</i>	<i>aglB</i>	0350	0351	0352	0353	0354
4-sugar glycan	3-sugar glycan	2-sugar glycan	1-sugar glycan	No glycan					

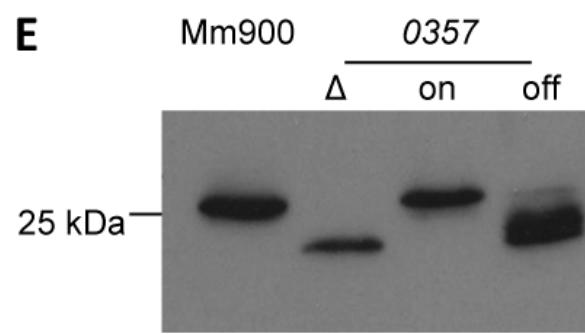
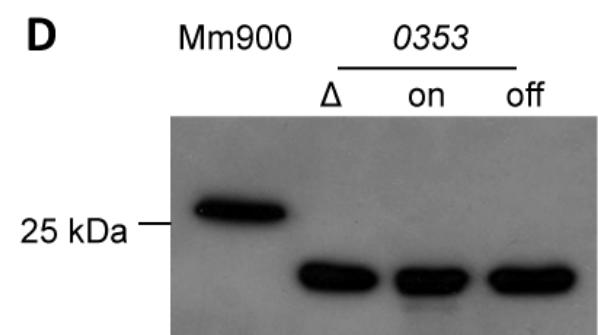
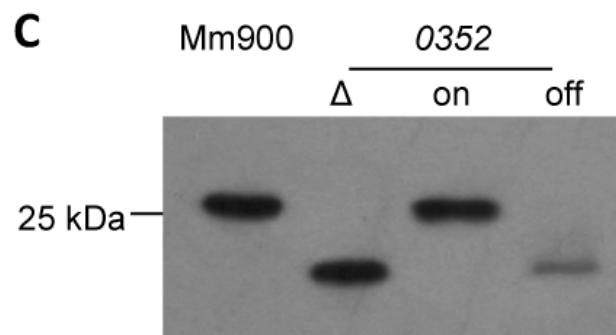
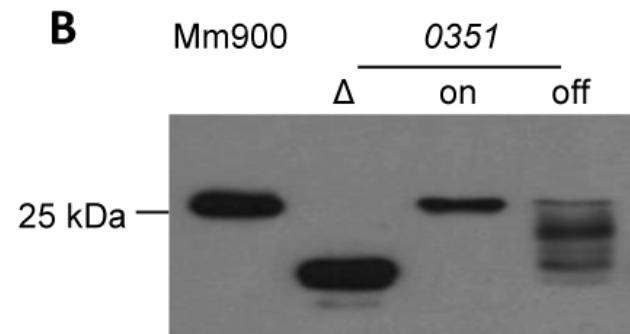
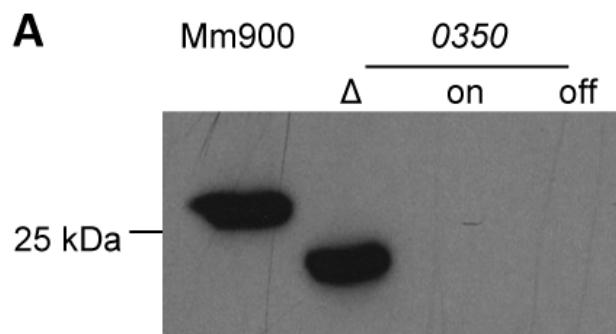


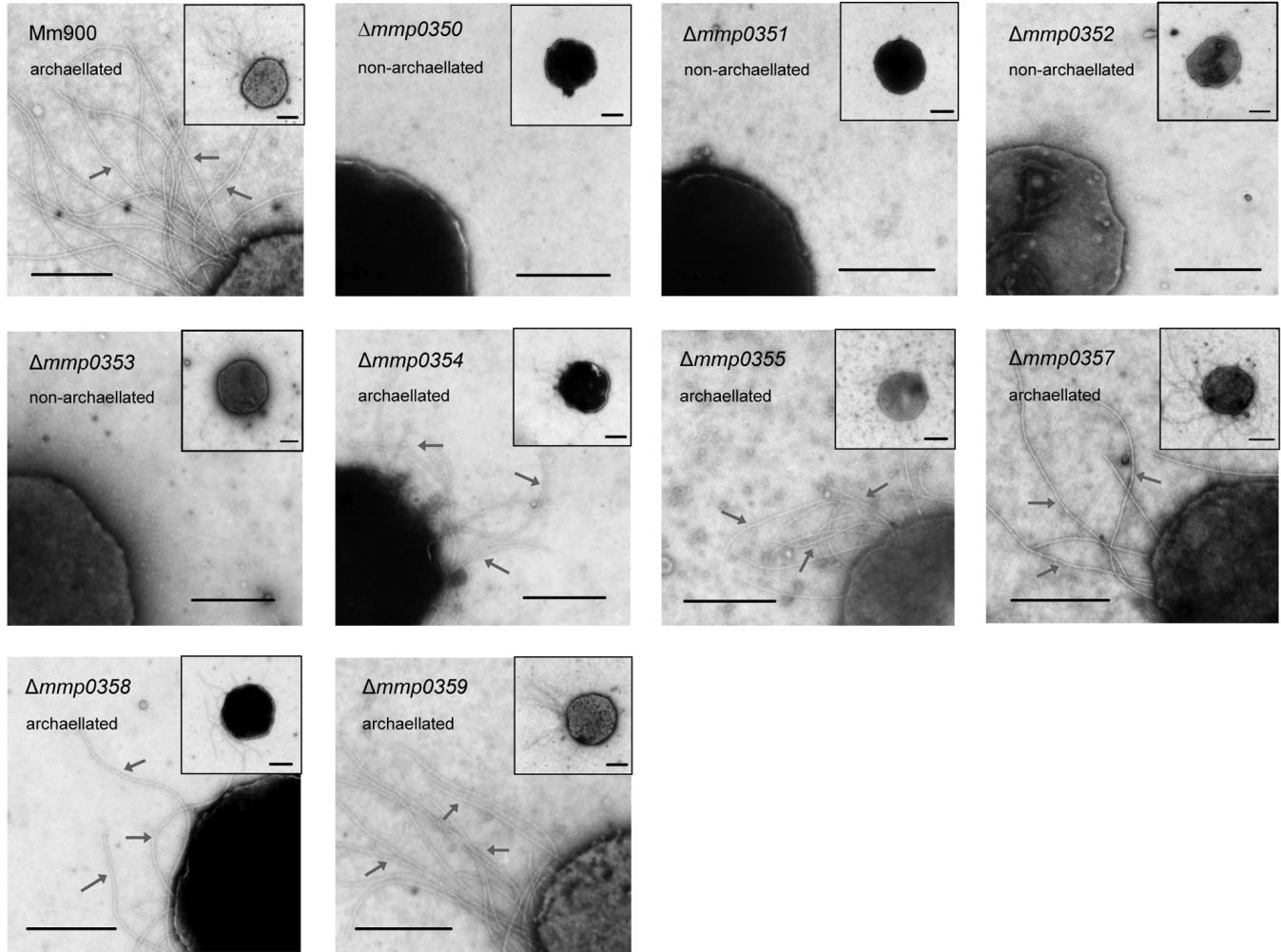
Mm900	<i>aglL</i>	<i>aglA</i>	<i>aglO</i>	<i>aglB</i>	0355	0357	0358	0359
4-sugar glycan	3-sugar glycan	2-sugar glycan	1-sugar glycan	No glycan				

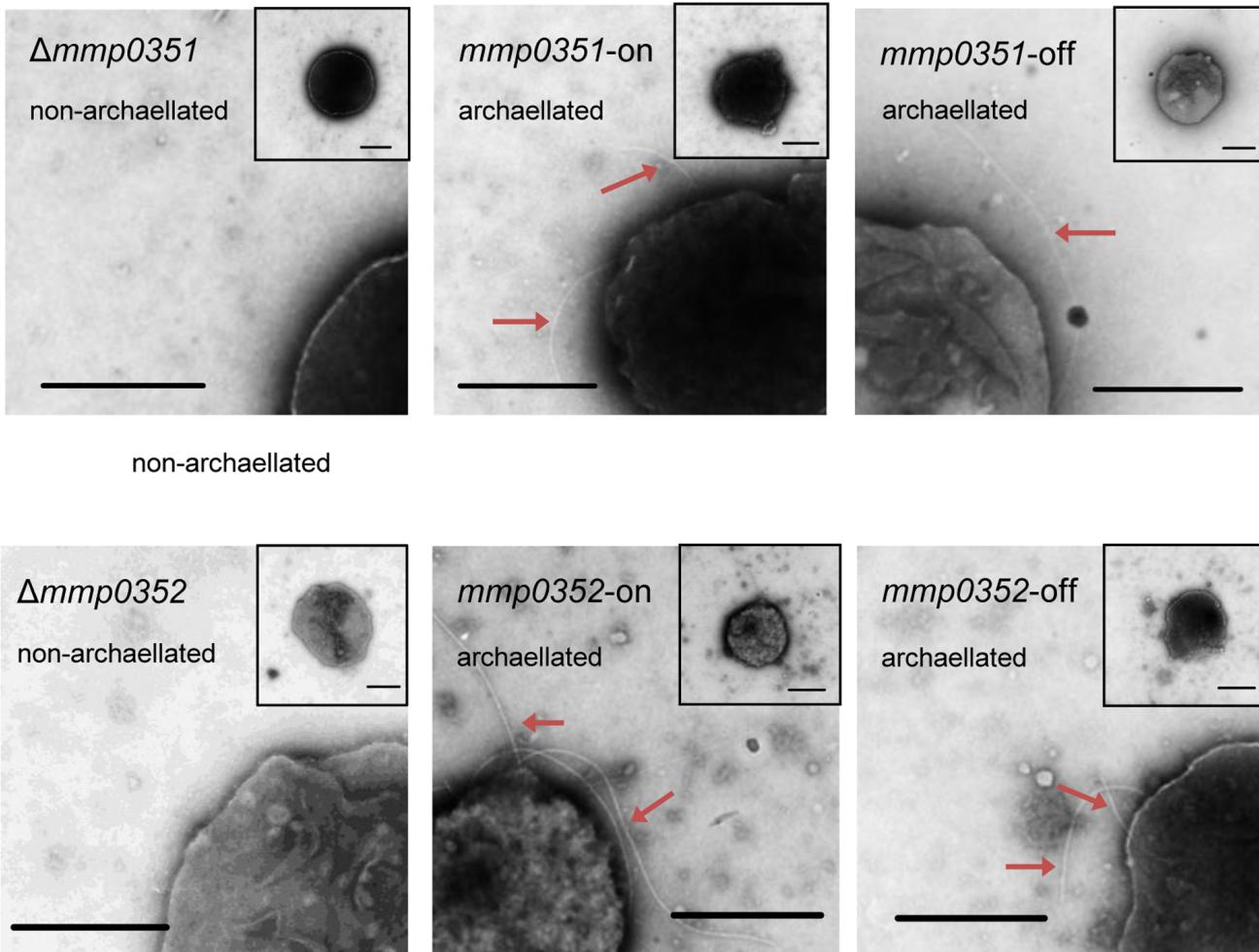
**B**

Mm900	0354	0358	0359
-------	------	------	------

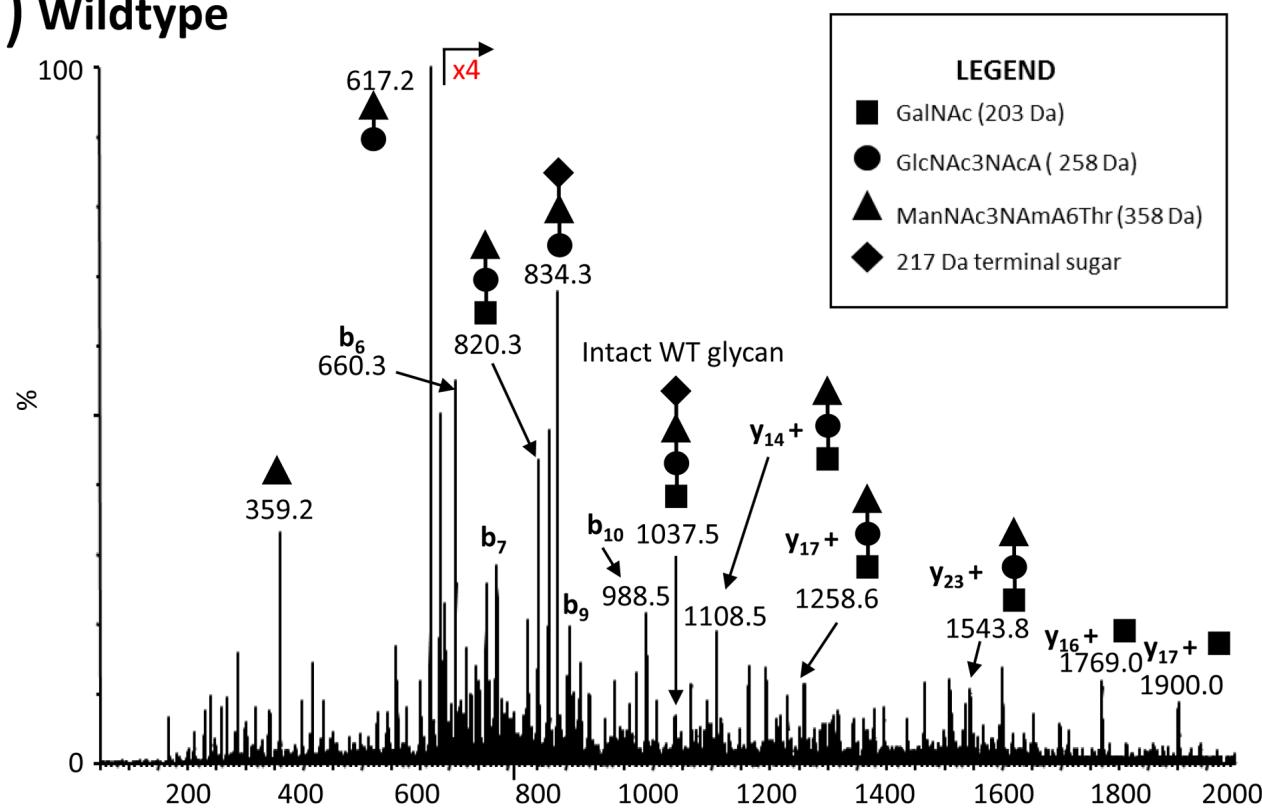




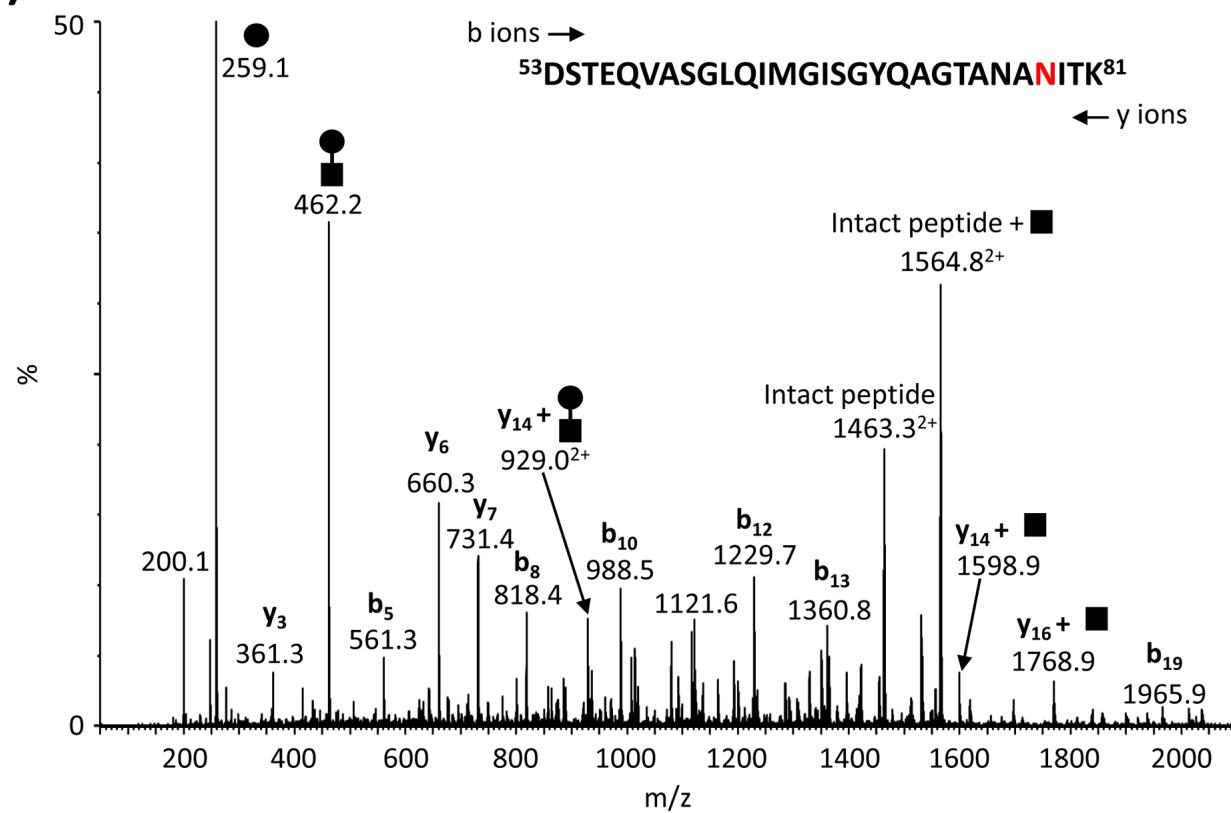




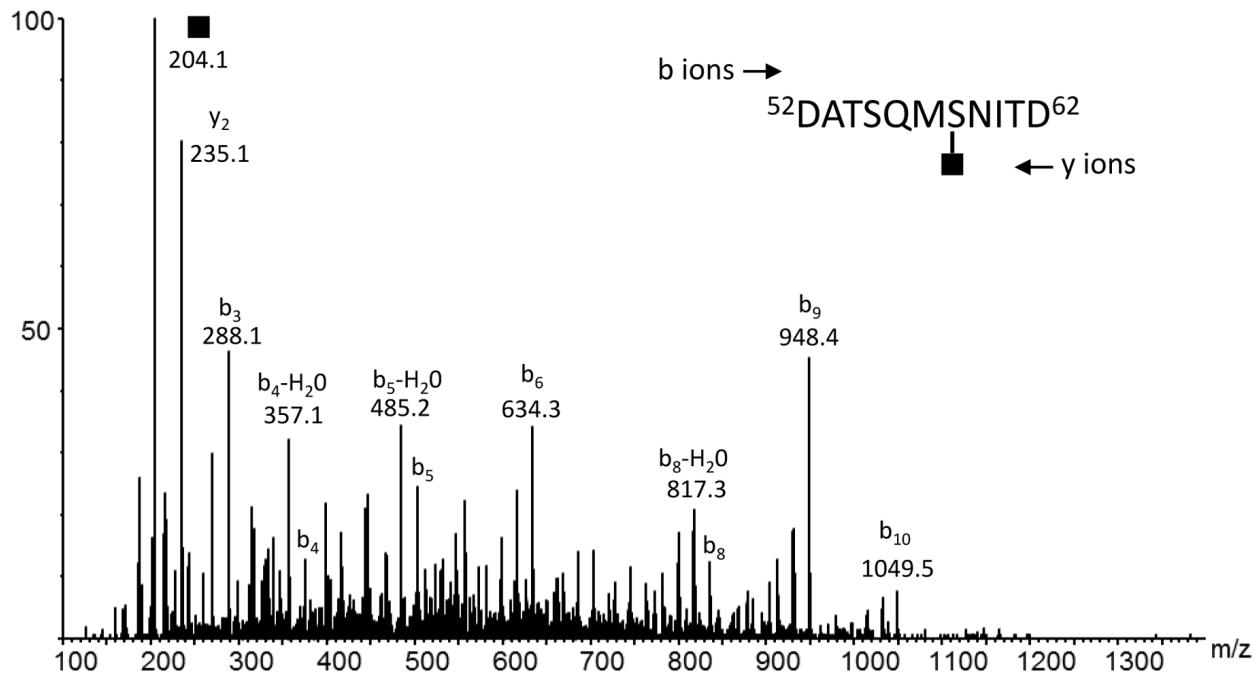
## a) Wildtype



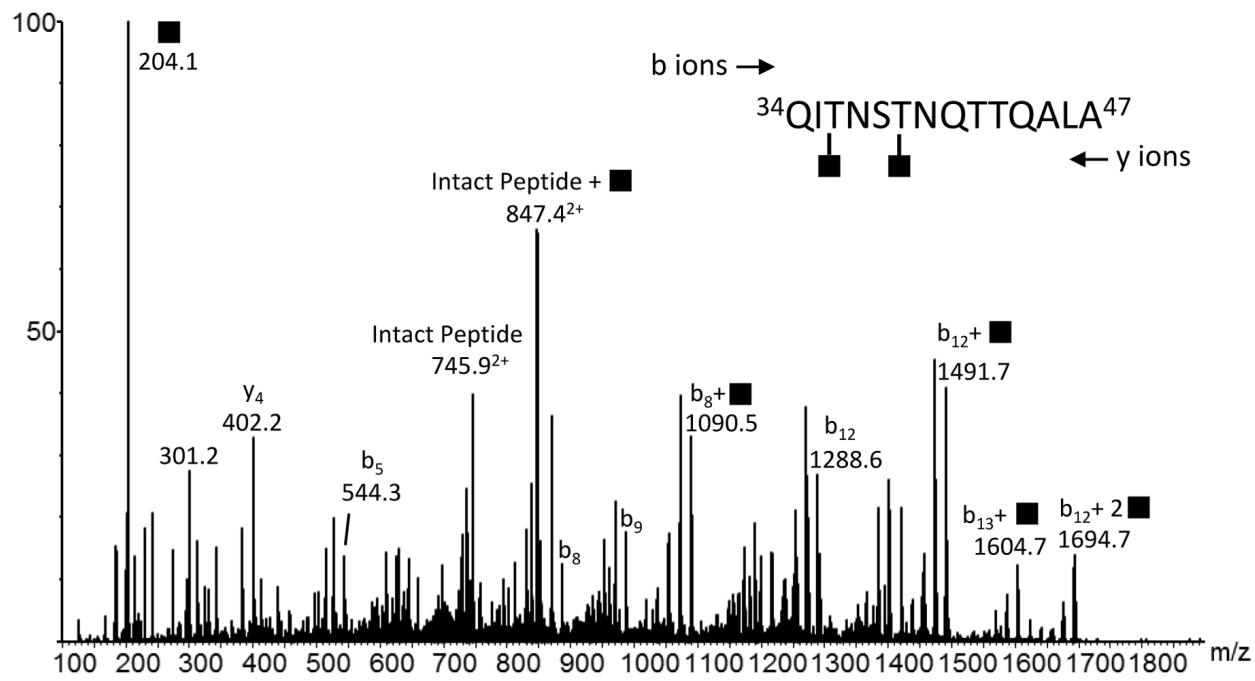
## b) Mutant 0357

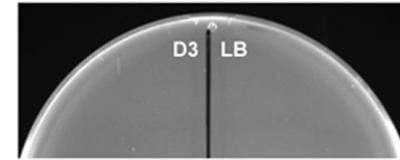
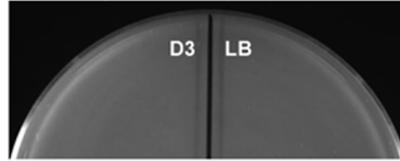
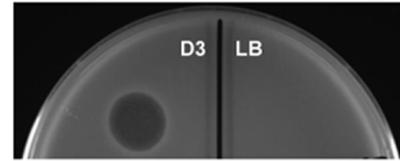
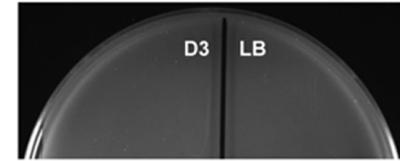


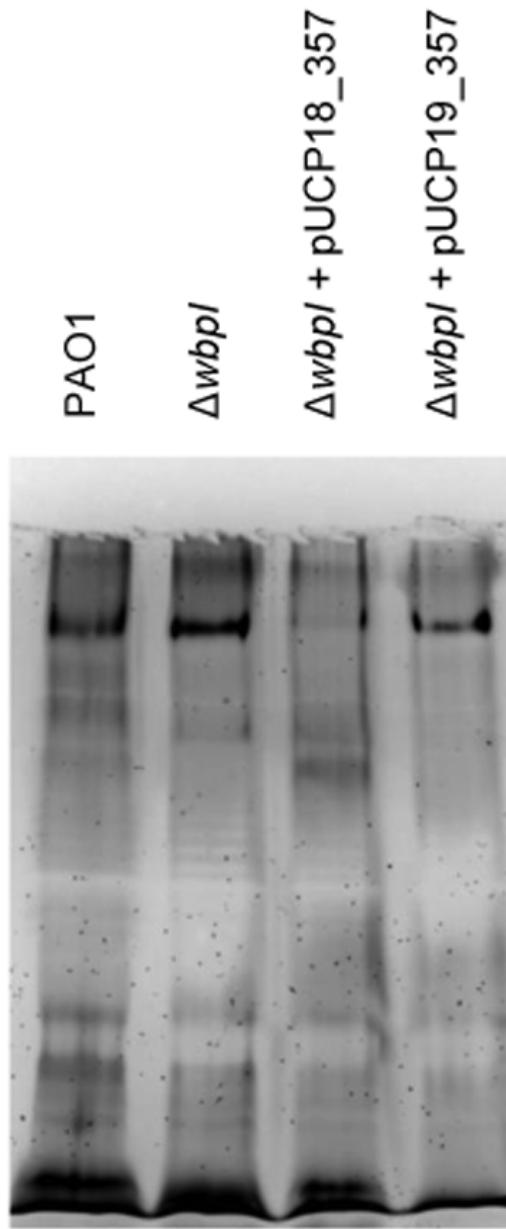
a)



b)



*P. aeruginosa* PAO1*P. aeruginosa*  $\Delta wbpA$ *P. aeruginosa*  $\Delta wbpA + pUCP18\_353$ *P. aeruginosa*  $\Delta wbpA + pUCP19\_353$ *P. aeruginosa*  $\Delta wbpI$ *P. aeruginosa*  $\Delta wbpI + pUCP18\_357$ *P. aeruginosa*  $\Delta wbpI + pUCP19\_357$ 

**A****B**