Molecular Microbiology (2009) 72(3), 633–644 ■
doi:10.1111/j.1365-2958.2009.06671.x
First published online 13 April 2009

Identification of genes involved in the assembly and attachment of a novel flagellin *N*-linked tetrasaccharide important for motility in the archaeon *Methanococcus maripaludis*

David J. VanDyke,[1] John Wu,[1] Susan M. Logan,[2] John F. Kelly,[2] Shino Mizuno,[3] Shin-Ichi Aizawa[3] and Ken F. Jarrell[1]*

[1]*Department of Microbiology and Immunology, Queen's University, Kingston, Ontario, K7L 3 N6, Canada.*
[2]*Institute for Biological Sciences, National Research Council, Ottawa, Ontario, K1A 0R6, Canada.*
[3]*Department of Life Sciences, Prefectural University of Hiroshima, 562 Nanatsuka, Shobara, Hiroshima 727–0023, Japan.*

Summary

Recently, the flagellin proteins of *Methanococcus maripaludis* were found to harbour an *N*-linked tetrasaccharide composed of *N*-acetylgalactosamine, di-acetylated glucuronic acid, an acetylated and acetamidino-modified mannuronic acid linked to threonine, and a novel terminal sugar [(*5S*)-2-acetamido-2,4-dideoxy-5-O-methyl-α-L-*erythro*-hexos-5-ulo-1,5-pyranose]. To identify genes involved in the assembly and attachment of this glycan, in-frame deletions were constructed in putative glycan assembly genes. Successful deletion of genes encoding three glycosyltransferases and an oligosaccharyltransferase (Stt3p homologue) resulted in flagellins of decreased molecular masses as evidenced by immunoblotting, indicating partial or completely absent glycan structures. Deletion of the oligosaccharyltransferase or the glycosyltransferase responsible for the transfer of the second sugar in the chain resulted in flagellins that were not assembled into flagella filaments, as evidenced by electron microscopy. Deletions of the glycosyltransferases responsible for the addition of the third and terminal sugars in the glycan were confirmed by mass spectrometry analysis of purified flagellins from these mutants. Although flagellated, these mutants had decreased motility as evidenced by semi-swarm plate analysis with the presence of each additional sugar improving movement capabilities.

Accepted 13 March, 2009. *For correspondence. E-mail jarrellk@queensu.ca; Tel. (+1) 613 533 2456; Fax (+1) 613 533 6796.

Introduction

Glycosylation, a post-translational modification involving the covalent attachment of sugar groups to target proteins, has been shown to be utilized across all three domains of life. *N*-linked glycosylation, specifically, refers to the linking of glycan moieties to asparagine residues residing within an Asn-X-Ser/Thr consensus sequence (where X represents any amino acid except proline) (Gavel and von Heijne, 1990; Kowarik *et al*., 2006; Abu-Qarn and Eichler, 2007). The processes involved in the assembly and attachment of *N*-linked glycans appear to be quite similar across the three domains, as all cases require the assembly of the glycan onto a lipid carrier embedded in a cellular membrane, its translocation across the membrane and, finally, its transfer to the target protein (Szymanski and Wren, 2005; Weerapana and Imperiali, 2006; Yurist-Doutsch *et al*., 2008a).

Although the bacterial and eukaryotic *N*-glycosylation systems have been relatively well characterized, only recently has the archaeal equivalent been investigated in studies conducted on archaeal flagellin and surface (S)-layer proteins (Voisin *et al*., 2005; Chaban *et al*., 2006; Abu-Qarn *et al*., 2007; 2008; Plavner and Eichler, 2008; Shams-Eldin *et al*., 2008; VanDyke *et al*., 2008; Yurist-Doutsch *et al*., 2008b; Chaban *et al*., 2009; Kelly *et al*., 2009). This has led to the identification of numerous archaeal glycosylation (*agl*) genes (Yurist-Doutsch *et al*., 2008a). Putative glycosylation has been reported in the flagellins from various archaeal species based mainly on glycan-detecting stains or inhibitors that affect glycosylation (Eichler and Adams, 2005), but the structures of attached glycans have been elucidated only in *Halobacterium salinarum* and *Methanococcus* spp. Of the archaeal flagellin glycans that have been characterized, all have been reported to be *N*-linked. *H. salinarum* was found to have flagellins possessing a linear oligosaccharide

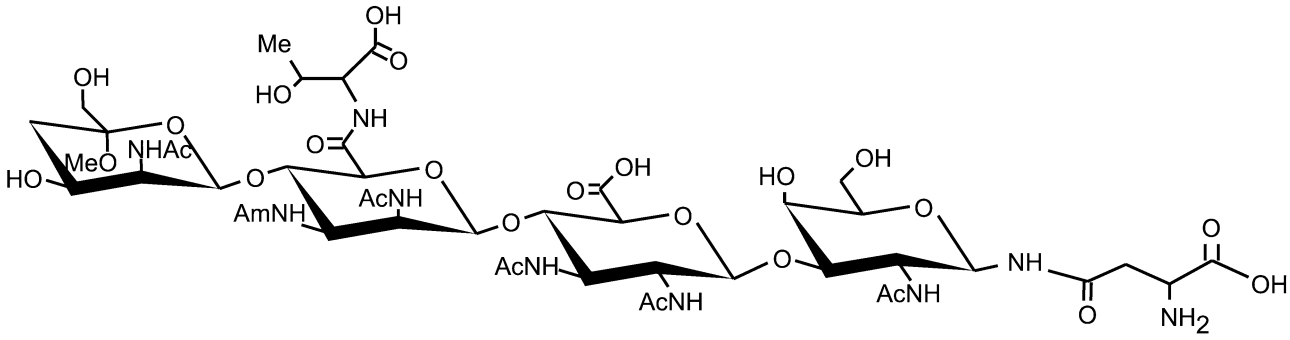

Fig. 1. Structure of the glycan *N*-linked to the flagellins of *M. maripaludis*.

composed of glucose and sulphated glucuronic and iduronic acids (Sumper, 1987; Wieland, 1988). The glycan identified on *Methanococcus voltae* flagellin proteins was shown to be a trisaccharide composed of *N*-acetylglucosamine, di-acetylated glucuronic acid and a modified mannuronic acid linked to a threonine residue (Voisin *et al*., 2005), but recent evidence shows that flagellins from other strains possess a fourth sugar (Chaban *et al*., 2009). The glycans observed on *M. voltae* and *H. salinarum* flagellins were also found to be attached to the S-layer proteins suggesting a common mechanism of assembly and attachment.

Although bacterial and archaeal flagella provide motility in the same fashion, they are actually quite different in terms of their structure and presumed mode of assembly. In fact, archaeal flagella have been noted to more closely resemble type IV pili systems than bacterial flagella by several groups (Peabody *et al*., 2003; Ng *et al*., 2006; Trachtenberg and Cohen-Krausz, 2006; Jarrell *et al*., 2007). In the cases where bacterial flagellins have been found to be glycosylated, the glycans are limited to *O*-glycosidic linkages (Logan, 2006). Bacterial flagella contain a central channel and assembly occurs as flagellin proteins travel through the structure from the cytoplasm to be incorporated at the growing distal end under the capping protein, HAP2 (Macnab, 2003). Archaeal flagella, however, do not possess a central channel large enough for flagellin proteins to pass through (Trachtenberg and Cohen-Krausz, 2006; Cohen-Krausz and Trachtenberg, 2008). Thus flagellin proteins, which possess solely *N*-linked as opposed to *O*-linked glycosylation, are apparently translocated across the cytoplasmic membrane to the outer cell surface, where the *N*-glycan attachment is believed to occur, prior to their incorporation at the base of the growing structure (Jarrell and McBride, 2008).

The assembly of the *M. voltae* flagellin glycan is carried out by the actions of the glycosyltransferases AglH, AglC/AglK and AglA. *M. voltae aglH*, believed to be responsible for the addition of GlcNAc to the dolichol carrier, was identified by its ability to complement a conditional lethal *alg7* mutation in yeast, as eukaryotes also employ GlcNAc as the linking sugar and Alg7 is known to carry out this same reaction (Shams-Eldin *et al*., 2008). Insertional inactivation studies conducted in *M. voltae* revealed that AglC/AglK and AglA catalyse the addition of the second and third sugars to the glycan respectively, with disruption of either adversely affecting flagellar assembly and function (Chaban *et al*., 2006; 2009). The oligosaccharyltransferase (AglB), responsible for the transfer of the complete sugar from the lipid carrier to target asparagine residues, was also identified (Chaban *et al*., 2006).

Recently, flagellin proteins isolated from *M. maripaludis* were found to possess an *N*-linked glycan differing in size, structure and linking sugar from that of the related *M. voltae*. *M. maripaludis* has three flagellins, FlaB1, FlaB2 and FlaB3 (Chaban *et al*., 2007), that comprise the flagellar structure and analysis of the amino acid sequences of these proteins indicates that they possess four, five and two asparagine residues residing within potential *N*-linked glycosylation sequons respectively. Of these 11 potential sites, nine were found to have an attached glycan. Only N^{26} of both FlaB1 and FlaB2 was found not to be glycosylated (Kelly *et al*., 2009). The *M. maripaludis* glycan is a tetrasaccharide (Fig. 1) consisting of Sug-4-β-ManNAc3NAmA6Thr-4-β-GlcNAc3NAcA-3-β-GalNAc-Asn where Sug is a novel monosaccharide – methyl (*5S*)-2-acetamido-2,4-dideoxy-α-L-*erythro*-hexos-5-ulo-1,5-pyranose (2NAc-2,4 dideoxy-hex-5-ulose) (Kelly *et al*., 2009). The identification of this unusual glycan, peculiar for its novel terminal sugar as well as its use of GalNAc as the linking sugar exemplifies unique archaeal traits and, thus, prompted the search for novel components involved in its assembly. In addition, the function that the glycan is serving is unknown, although the deletion of an acetyltransferase believed to modify the *M. maripaludis* flagellar glycan was demonstrated to impair flagellin *N*-glycosylation as well as flagellar assembly and pili attachment (VanDyke *et al*., 2008). In this study, we take advantage of genetic tools recently established by Moore

Table 1. Putative *M. maripaludis* genes involved in *N*-glycosylation targeted in this study.

Gene	Annotation	Assigned gene name
MMP1080	Glycosyltransferase, group 1	*AglA*
MMP1424	Oligosaccharyl transferase STT3 subunit	*AglB*
MMP1079	Glycosyltransferase, family 2	*AglO*
MMP0354	Oligosaccharide transporter (flippase)	
MMP1423	Glycosyltransferase, family 4	*AglH*
MMP1170	Glycosyltransferase, family 2	
MMP1089	Polysaccharide biosynthesis protein (flippase)	
MMP1088	Glycosyltransferase, group 1	*AglL*

and Leigh (2005) for the creation and complementation of in-frame deletion mutations in *M. maripaludis* to study the assembly of the *N*-linked flagellin glycan. We identify several glycosyltransferases and the oligosaccharyltransferase involved in *N*-glycan assembly and correlate the deletion of the genes with their effects on the glycan structure as well as flagella assembly and function.

Results

Construction of in-frame deletions

In an effort to elucidate the genetics of flagellin glycosylation in *M. maripaludis,* the genome was screened for coding regions bearing sequence similarity to known components of archaeal, bacterial and eukaryotic *N*-linked glycosylation systems. Eight genes were identified and targeted for in-frame deletion to determine if their removal had effects on flagellin glycosylation. These genes and their annotations are listed in Table 1. In-frame deletion plasmids for all eight genes targeted were constructed. However, in-frame deletions were successfully made in only five genes (*MMP0354*, *MMP1079*, *MMP1080*, *MMP1088* and *MMP1424*) while the screening of transformants of the three others resulted in only wild-type colonies. Deletion of the gene *MMP0354*, targeted for a potential role as an *N*-glycan membrane transporter (flippase), was subsequently shown to not affect flagellin glycosylation as evidenced by immunoblotting with anti-flagellin antisera (data not shown) and hence, it was not characterized further. The genes *MMP1079*, *MMP1080*, *MMP1088* and *MMP1424* were successfully deleted and shown to affect flagellin glycosylation. Deletion of each of the four genes was shown initially through a PCR screen and further confirmed by Southern blot analysis (Fig. 2).

In-frame deletions decrease flagellin molecular mass

To determine if the loss of *MMP1079*, *MMP1080, MMP1088* or *MMP1424* had an effect on flagellin glycosylation that altered flagellin molecular mass, whole-cell lysates were subjected to immunoblotting using anti-flagellin antisera (Fig. 3). Compared with wild-type flagellins, there was a detectable decrease in the flagellin molecular mass of all four deletion mutants, with the deletions in *MMP1088*, followed by *MMP1080*, *MMP1079* and *MMP1424* resulting in progressively lower-molecular-

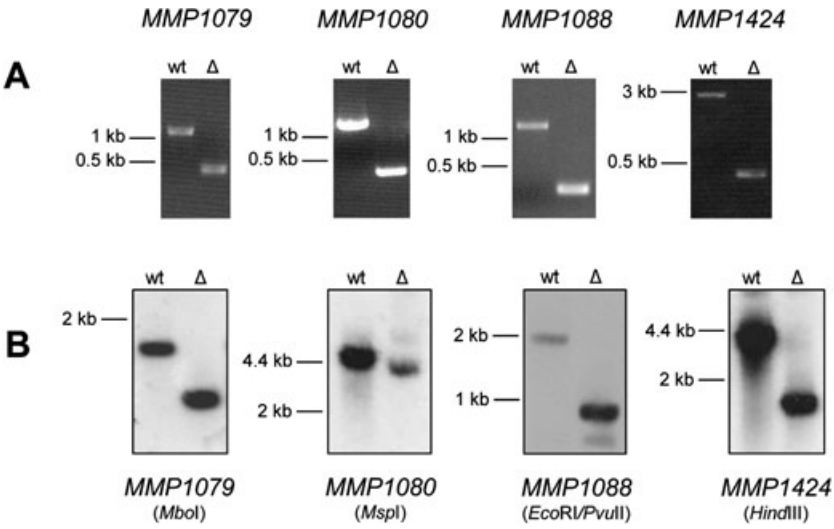

Fig. 2. Confirmation of in-frame deletions of targeted genes.
A. PCR screens of *M. maripaludis* wild-type (wt) and in-frame deletion mutants (Δ). Wild-type and mutant cells amplified bands that matched the predicted sizes of 1104 and 363 bp for Δ*MMP1079* screens, 1384 and 337 bp for Δ*MMP1080,* 1388 and 155 bp for Δ*MMP1088* and 2882 and 362 bp for Δ*MMP1424*.
B. Southern blot analysis on wild-type and mutant genomic DNA. DNA was digested with an appropriate restriction enzyme and hybridized with a probe specific for the target gene surrounding region. Wild-type and mutant DNA yielded bands that matched the predicted sizes of 1421 and 680 bp for Δ*MMP1079* analysis, 5098 and 4051 bp for Δ*MMP1080,* 2088 and 856 bp for Δ*MMP1088* and 4130 and 1610 bp for Δ*MMP1424*.

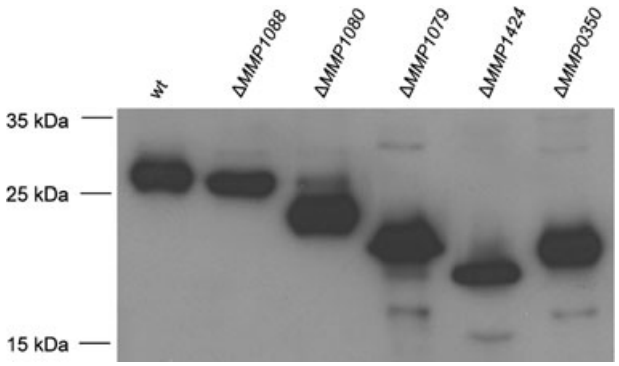

Fig. 3. Immunoblot of flagellin proteins from *M. maripaludis* in-frame deletion mutants. Whole-cell lysates were electrophoresed through a 12.5% acrylamide gel and the blot was developed with *M. voltae* anti-FlaB2.

weight flagellins. The entire range of molecular weights varied from 27 kDa for the wild-type to approximately 20 kDa for the *MMP1424* deletion. This suggested that the mutants contained flagellins with a lesser degree of glycosylation. The flagellins in the Δ*MMP1424* cells were also found to migrate with lower molecular mass than flagellins from an acetyltransferase mutant (Δ*MMP350*) that had been suspected of being minimally glycosylated in a previous study (VanDyke *et al*., 2008). Deletion of *MMP1079* resulted in the flagellin proteins migrating as though slightly larger than those of Δ*MMP1424* and equivalent to Δ*MMP350*.

Due to its annotation as the Stt3p oligosaccharyltransferase, it is believed that MMP1424 is responsible for attaching the glycan to the Asn residue of the protein and that its deletion results in completely unmodified flagellins. As the glycan is a tetrasaccharide, four glycosyltransferases would be expected to be involved in its assembly. The stepwise increases in apparent molecular mass of flagellins in mutants deleted for *MMP1079*, *MMP1080* and *MMP1088* (all annotated as glycosyltransferases) suggested that MMP1079, MMP1080 and MMP1088 are likely responsible for addition of the second, third and fourth sugars respectively. A deletion in the glycosyltransferase responsible for attachment of the first sugar should leave flagellins totally unglycosylated and so migrating at the same molecular weight as the flagellins from the oligosaccharyltransferase mutant (Δ*MMP1424*). None of the glycosyltransferase mutants isolated had this flagellin phenotype. Perhaps this is a role for one of the annotated glycosyltransferases for which deletions were unattainable (i.e. *MMP1170* or *MMP1423*). The similar flagellin molecular masses observed between flagellins of Δ*MMP1079* cells and those from the Δ*MMP350* acetyltransferase mutant suggests that these two enzymes may be both acting on the second sugar (diacetylated glucuronic acid).

Trans gene complementation restores flagellin molecular mass

To confirm the role of the deleted genes in flagellin glycosylation, mutant cells were complemented with a plasmid expressing the deleted gene to try to restore function. C-terminal histidine-tagged versions of *MMP1079*, *MMP1080*, *MMP1088* and *MMP1424* were cloned into pHW40, a plasmid that allows expression of the cloned genes under control of the inducible *nif* promoter (Kessler and Leigh, 1999). Complementation plasmids were transformed into the corresponding mutant cells and maintained by selection with puromycin. The cells were then transferred to nitrogen-free media supplemented with either L-alanine as a nitrogen source to induce expression from the *nif* promoter or NH_4Cl as a nitrogen source to limit expression (Kessler and Leigh, 1999). In nitrogen-free medium supplemented with L-alanine, it was found that complemented Δ*MMP1080* cells had flagellins of increased molecular mass, but not exactly to wild-type size (Fig. 4B). Flagellins of increased size were also found in cells that were grown in nitrogen-free medium supplemented with NH_4Cl, but only in very minor amounts. This suggested that there was some minor expression of the cloned gene when the promoter is believed to be uninduced. This is consistent with previous studies (Lie *et al*., 2005).

When *MMP1079* was induced in mutant cells, flagellin molecular mass was fully restored to wild-type size. Interestingly, when NH_4Cl was used as a nitrogen source, the flagellin proteins occurred in a ladder, likely representing the different partial glycosylation states including glycans consisting of one, two, three or four sugars (Fig. 4A). Again, this is believed to be a result of minor amounts of *MMP1079* being expressed even when the *nif* promoter is expected to be uninduced.

Complementation of the Δ*MMP1088* mutant with the cloned *MMP1088* gene resulted in a restoration of some of the flagellins to the molecular weight of wild-type flagellins, more so in the *nif* on conditions than the uninduced conditions (Fig. 4C).

Complementation of Δ*MMP1424* was attempted as the complementation vector was created and successfully transformed. However, further examination indicated that, over the course of the complementation experiment, Δ*MMP1424* cells stopped making sufficient amounts of flagellin proteins to be detectable by immunoblotting. Thus, the ability of the complemented *MMP1424* gene to restore the flagellins to wild-type molecular mass could not be assessed. This is similar to a previous attempt to complement the Δ*MMP350* acetyltransferase mutant (VanDyke *et al*., 2008), which could also not be achieved due to lack of flagellin production, an observation that was shown to be associated with the loss of another protein

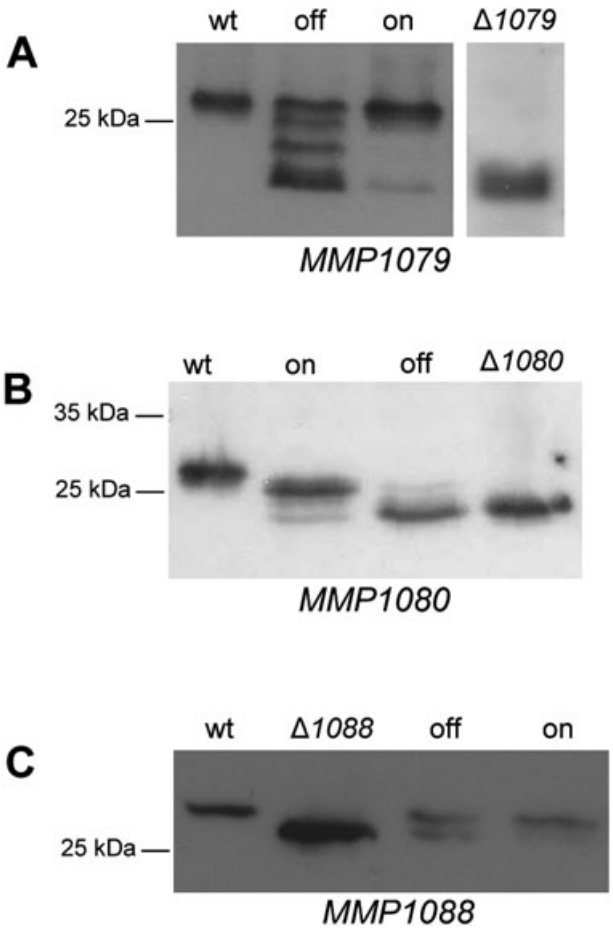

Fig. 4. Immunoblots of flagellins from (A) *MMP1079*, (B) *MMP1080* and (C) *MMP1088 M. maripaludis* deletion mutants complemented with a his-tagged version of the respective gene under the control of a nitrogen-inducible promoter. Cells were grown in nitrogen-free media supplemented with L-alanine (*nif* promoter on) or NH_4Cl (*nif* promoter off). The blot was developed with *M. voltae* anti-FlaB2 antibodies.

encoded within the *fla* operon. A similar cessation of flagellin production was also noted in Δ*MMP1079* cells after a series of subcultures, but as these cells were able to maintain flagellin production over the course of the complementation experiment, it did not seem to be as immediate of an effect as in the Δ*MMP1424* mutant.

In-frame deletions affect flagellar filament assembly

To observe the effects that the deletions had on flagellar filament structure, wild-type *M. maripaludis* and all mutant cells were examined by electron microscopy (EM) (Fig. 5). Wild-type cells possessed numerous flagella on their cell surface (Fig. 5A), as did Δ*MMP1088* (Fig. 5B) and Δ*MMP1080* (Fig. 5C) mutant cells. Upon examination of many cells, there was no obvious or consistent difference in the appearance or number of flagella found on either of these mutant cells compared with wild-type. The deletion of *MMP1079* (Fig. 5D) and *MMP1424* (Fig. 5E), however, resulted in cells that were completely non-flagellated, leaving only pili as surface appendages. As Δ*MMP1079* and Δ*MMP1424* cells are believed to have flagellin proteins lacking most or all of the glycan, this suggested that a minimal size of glycan is required for the assembly/stability of the flagellar structures.

To observe the effects that complementation had on flagellar filament structure, complemented Δ*MMP1080*, Δ*MMP1088* and Δ*MMP1079* cells, grown in nitrogen-free media supplemented with either L-alanine or NH_4Cl, were examined by EM. As Δ*MMP1080* and Δ*MMP1088* mutant cells still had the ability to assemble flagellar filaments, it was not surprising that complemented Δ*MMP1080* and Δ*MMP1088* cells grown in either NH_4Cl or L-alanine were all flagellated (data not shown). There was no notable difference in the number, appearance or location of flagella between Δ*MMP1080* and Δ*MMP1088* mutant cells and complemented cells supplemented with either L-alanine or NH_4Cl. In the case of *MMP1079* complementation, cells grown in NH_4Cl did not have any flagellar filaments (Fig. S1A), which is comparable to uncomplemented Δ*MMP1079* cells. However, when complemented cells were induced to make the MMP1079 glycosyltransferase by growing in L-alanine as a sole nitrogen source, flagellar filaments were observed on the cell surface (Fig. S1B).

Mass spectrometry analysis of flagellins from flagellated mutants identifies partial glycan structures

To more accurately assign the role of respective glycosyltransferase enzymes encoded by *MMP1080* and *MMP1088*, flagella filaments from each mutant strain were purified and the flagellin structural proteins were examined by mass spectrometry (MS). Tryptic glycopeptides were analysed by nano-LC-MS/MS confirming that unique glycan structures were made in each case. Analysis of flagellar tryptic glycopeptides from Δ*MMP1080* flagellin indicated that the *N*-linked glycan made by this strain was a disaccharide composed of the linking GalNAc and GlcNAc3NAcA residues (Fig. 6B). In contrast, flagellar tryptic glycopeptides from Δ*MMP1088* appeared to be glycosylated with an *N*-linked trimeric glycan composed of GalNAc, GlcNAc3NAcA and ManNAc3NAmA (Fig. 6C). In the parent flagellar glycan (Fig. 6A), this latter monosaccharide is decorated with a threonine modification that is clearly absent in the glycan made by the Δ*MMP1088* strain. Analysis of tryptic glycopeptides of flagellins from complemented Δ*MMP1080* and Δ*MMP1088* indicated the presence of both truncated and full-length glycans, demonstrating that these constructs partially restored the wild-type profile of glycosylation (data not shown).

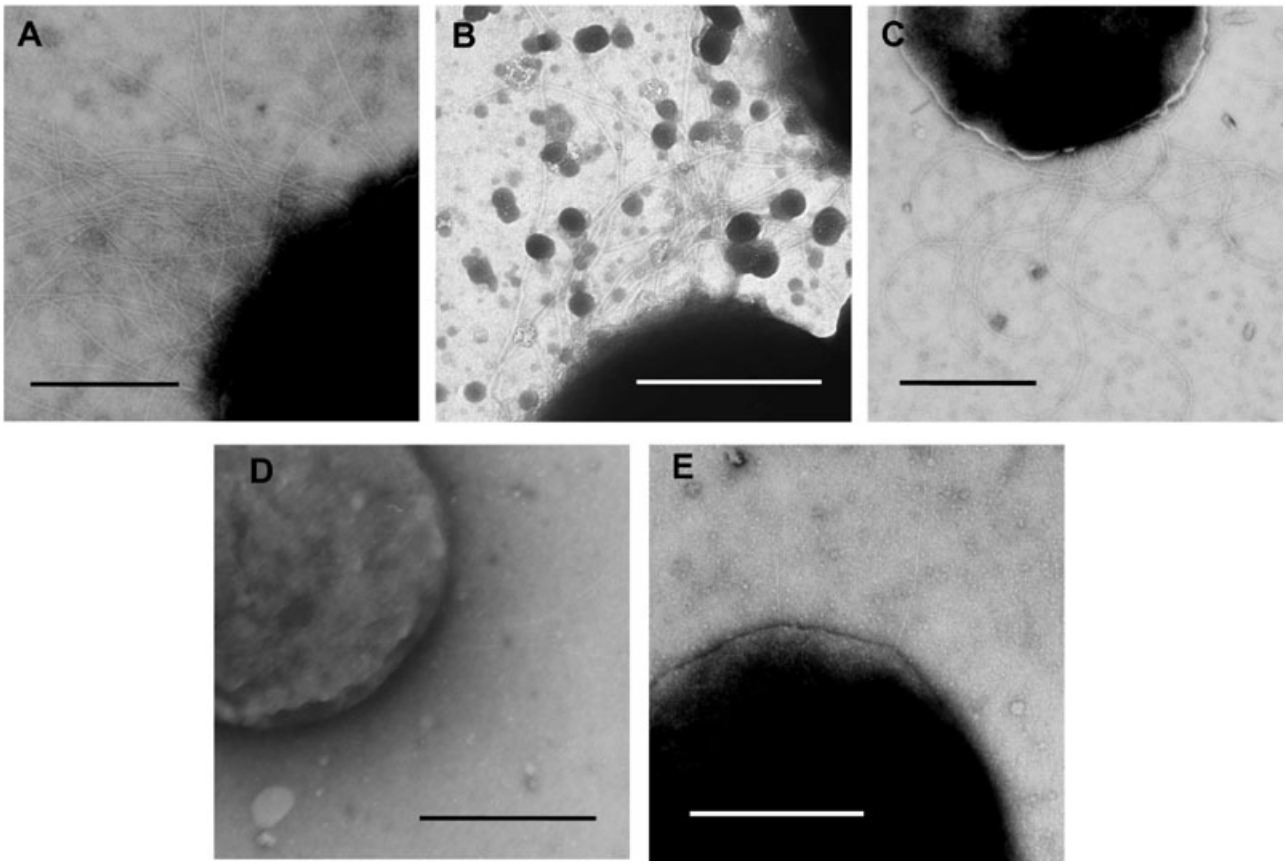

Fig. 5. Electron micrographs of (A) *M. maripaludis* strain Mm900, (B) Δ*MMP1088*, (C) Δ*MMP1080*, (D) Δ*MMP1079* and (E) Δ*MMP1424*. The cells were negatively stained with 2% phosphotungstic acid. The scale bars represent 0.5 μm.

Complete flagellin glycan is required for efficient motility

To determine if the deletion of genes involved in flagellin glycosylation affected cell motility, wild-type and mutant cells were assessed by semi-swarm plate analysis. 0.25% agar plates were inoculated in the centre with wild-type, Δ*MMP1079*, Δ*MMP1080*, Δ*MMP1088* and Δ*MMP1424* cultures and analysed after 7 days of growth. Wild-type cells had the largest zone of motility (Fig. 7). The mutants determined to be non-flagellated by EM (Δ*MMP1079*, Δ*MMP1424*) did not swim out from the inoculation stab, as expected. Although the mutants Δ*MMP1080* and Δ*MMP1088* have normal looking flagella as observed by EM, the presence of only partial glycans on these cells impeded motility compared with wild-type. Δ*MMP1080* mutant cells, having flagellins with only two of the four sugars in the attached glycan, formed a motility zone with a diameter that was approximately 54% (±9.4) of the wild-type zone (range of 45–64% in six duplicates) while Δ*MMP1088*, lacking only the terminal sugar and amino acid modification of the third sugar had a zone of motility that reached approximately 74% (±2.9) of the wild-type zone (range of 71–79% in six duplicates).

Discussion

Four genes involved in the assembly of the *M. maripaludis* *N*-linked glycan and its attachment to flagellin proteins were identified. The in-frame deletion of these genes was shown to result in either no glycan attachment or the attachment of only partial glycan structures to the flagellins, as exemplified by lower apparent masses in immunoblots compared with wild-type flagellins and MS analysis of purified flagellin proteins. The presence of a disaccharide glycan structure was shown to be sufficient for flagellar assembly, although the entire glycan was shown to be necessary for optimal swimming as judged by semi-swarm plate analysis.

The gene *MMP1424* was targeted for its role as the possible oligosaccharyltransferase involved in flagellin glycan attachment due to its annotation and significant homology to the *M. voltae* oligosaccharyltransferase AglB (Chaban *et al.*, 2006). Deletion of *MMP1424* resulted in a large decrease in flagellin molecular mass as evidenced by immunoblotting with anti-flagellin antisera and the mutant cells were non-flagellated. Due to its sequence similarity to the Stt3p oligosaccharyltransferase, it is

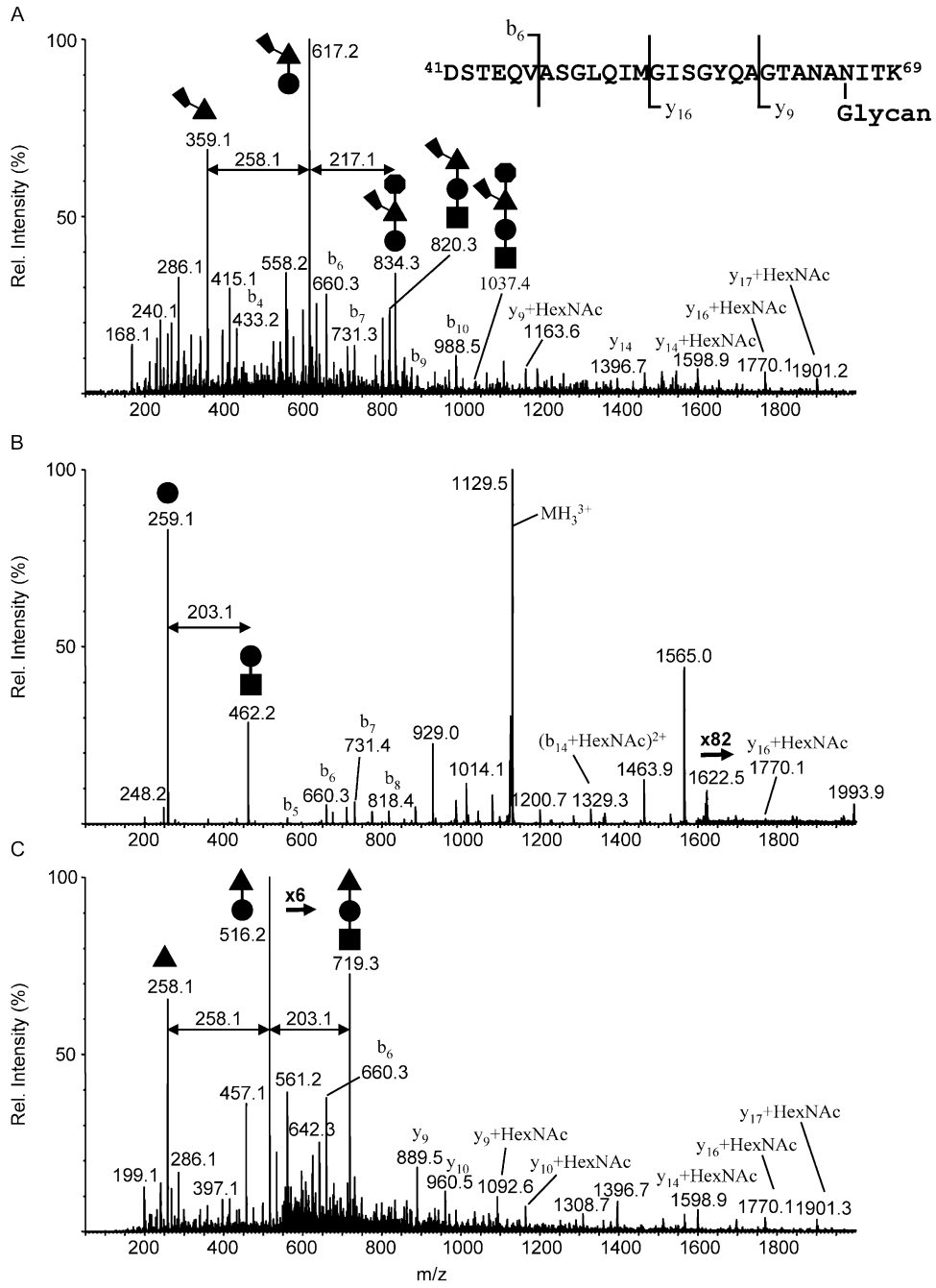

Fig. 6. MS/MS spectra of the triply protonated ion corresponding to the $T^{41\text{-}469}$ glycopeptide of the mature FlaB2 protein at (A) m/z 1321.3 in the WT, (B) m/z 1129.5 in the Δ*MMP1080* mutant and (C) m/z 1215.3 in the Δ*MMP1088* mutant. The amino acid sequence of this peptide is presented in the inset in (A) and the site of *N*-glycosylation is indicated. Some of the b and y fragment ions arising from fragmentation of the peptide bonds are indicated in the spectrum. The major carbohydrate oxonium ions are identified in the spectra using symbols to indicate the sugar residues present (GalNAc: ■ 203 Da; (5S)-2-acetamido-2,4-dideoxy-α-L-*erythro*-hexos-5-ulo-1,5-pyranose 217 Da; ManNAc3NamA6: ▲ 257 Da; GlcNAc3NAcA: ● 258 Da and ManNAc3NAmA6Thr: ◆ 358 Da). The MS/MS spectrum for the wild-type glycopeptide (panel a) is dominated by the glycan oxonium ions at m/z 617.2 (● + ◆) and m/z 359.1 (◆), although a weak ion signal for the intact WT glycan is observed at m/z 1037.4. MS/MS analysis of the equivalent FlaB2 glycopeptide from the 1080 mutant revealed that the *N*-linked glycan is composed only of the linking GalNAc (■) and GlcNAc3NAcA (●). On the other hand, MS/MS analysis of the equivalent glycopeptide from the 1088 mutant indicated the presence of a trimeric glycan composed of the GlcNAc (■), GlcNAc3NAcA (●) and ManNAc3NAmA6 (▲) without the threonine modification. Regions of the MS/MS spectra in (B) and (C) have been expanded (B, ×82; C, ×6) to aid identification of key fragment ions.

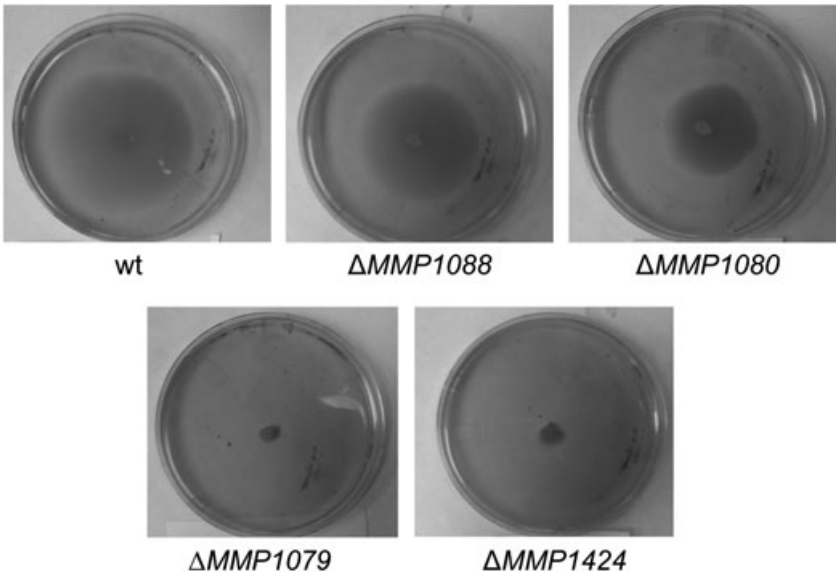

Fig. 7. Semi-swarm plates of wild-type, Δ*MMP1088*, Δ*MMP1080*, Δ*MMP1079*, and Δ*MMP1424* mutants. 0.25% agar plates were inoculated in the centre and analysed after 7 days of growth at 30°C.

believed that *MMP1424* is responsible for transferring the completed glycan to the protein and that the flagellin proteins from the Δ*MMP1424* mutant therefore represent unmodified flagellins with no attached glycan. This observation of unmodified flagellins and non-flagellation is similar to that found by disruption of the Stt3p homologue in *M. voltae* (Chaban *et al*., 2006) and, thus, it is proposed that MMP1424 should be classified as AglB in the *M. maripaludis* system accordingly, although the specificity of each enzyme may be different due to GalNAc serving as the linking sugar in *M. maripaludis* as opposed to GlcNAc in *M. voltae*.

The genes *MMP1079*, *MMP1080* and *MMP1088* were targeted for their role as potential glycosyltransferases involved in *M. maripaludis* flagellin glycosylation as their annotations were as glycosyltransferases and they all group within glycosyltransferase clusters of orthologous genes. Each deletion resulted in flagellins of decreased molecular mass compared with wild-type, as evidenced by immunoblotting, suggestive of different degrees of partial glycosylation. MS analysis carried out on flagellins isolated from a Δ*MMP1088* mutant demonstrated that the flagellins from Δ*MMP1088* were missing the terminal sugar of the glycan and, unexpectedly, the threonine residue attached to the third sugar. Complementation restored the wild-type glycan structure. These data provide evidence that MMP1088, which has conserved glycosyltransferase domains, is likely responsible for the addition of the fourth monosaccharide to the flagellar glycan. This structure is then recognized as the substrate to which threonine is added to the ManNAc3NAmA. Alternatively, MMP1088 may be responsible for the addition of threonine to ManNAc3NAmA and, without this addition, the fourth monosaccharide cannot be attached to the flagellar glycan. As the protein encoded by *MMP1088* is critical for glycan assembly we propose the designation AglL for MMP1088.

MS analysis carried out on flagellins isolated from a Δ*MMP1080* mutant demonstrated that Δ*MMP1080* flagellins were lacking the modified mannuronic acid and 2NAc-2,4 dideoxy-hex-5-ulose. MS analysis of flagellins from complemented Δ*MMP1080* demonstrated that the complemented gene could restore the wild-type glycan profile (data not shown). These data provide evidence that MMP1080 is responsible for adding ManNAc3NAmA to the flagellar glycan, in an identical fashion to AglA from *M. voltae* that has been shown to add the related monosaccharide, ManNAcA (Chaban *et al*., 2006). In both organisms, the modified mannuronic acid is added to the identical diacetylated glucuronic acid. Because of this and the strong sequence similarity of MMP1080 to AglA of *M. voltae* (35% identical, 55% similar), it is proposed that MMP1080 be designated as the *M. maripaludis* AglA.

Unlike mutants in *MMP1088* and *MMP1080*, Δ*MMP1079* cells did not form flagellar filaments and, thus, MS analysis on these flagellins could not be performed. In immunoblots, Δ*MMP1079* flagellin proteins migrated as though slightly larger than those seen for the oligosaccharyltransferase Δ*MMP1424* mutant, indicating that one of the earlier glycosyltransferases in the pathway may be affected. Complementation of the Δ*MMP1079* mutant fully restored flagellins to wild-type molecular mass and to flagellated cells. When NH_4Cl was used as the nitrogen source in the complementation experiment, minor levels of expression occurred. Interestingly, this resulted in a ladder of flagellins, likely representing the different partial glycosylation states. Δ*MMP1079* mutant flagellins are slightly larger than the flagellins from the oligosaccharyl-

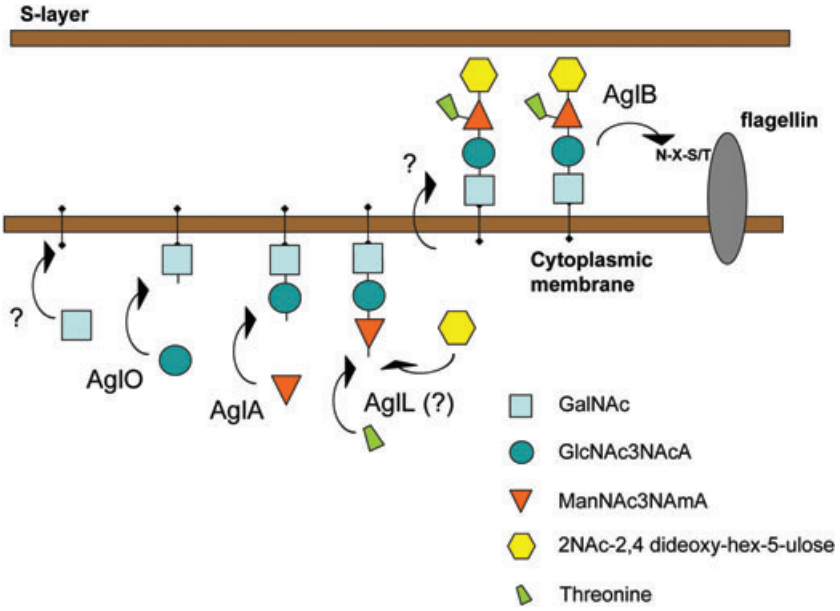

Fig. 8. Proposed model for flagellin *N*-linked glycosylation in *M. maripaludis*. The tetrasaccharide is assembled onto the lipid carrier using a number of glycosyltransferases. The glycosyltransferases encoded by *MMP1079* (AglO) and *MMP1080* (AglA) add the second and third sugars respectively. That encoded by *MMP1088* (AglL) is either responsible for the addition of the terminal sugar or the threonine residue to ManNAc3NAmA. The completed glycan is flipped across the cytoplasmic membrane and transferred to an asparagine residue on the flagellin protein by the oligosaccharyltransferase encoded by *MMP1424* (AglB).

transferase deletion; therefore, it is believed that these mutants may still have the first sugar (GalNAc) in the chain attached and that MMP1079 carries out the transfer of GlcNAc3NAcA. The increased molecular mass of flagellins from this strain suggests that the oligosaccharyltransferase (MMP1424) is still able to transfer a single sugar from the lipid-linked carrier to the flagellin monomers. The four bands observed in the complementation experiment, where minor expression occurred, may represent the Δ*MMP1079* mutant flagellins (with one sugar attached) and flagellins containing a glycan of two, three and finally four sugars (wild-type). In *M. voltae,* AglC and AglK have been implicated in the attachment of GlcNAc3NAcA to the GlcNAc linking sugar (Chaban *et al*., 2009). In *M. maripaludis*, MMP1079 transfers the same GlcNAc3NAcA but to the different linking sugar, GalNAc. Because of this difference and the low sequence similarity of MMP1079 to either AglC or AglK of *M. voltae*, we propose that MMP1079 is not a homologue of either of these *M. voltae* genes and should therefore be designated as AglO. Presumably, it is the transfer to a different linking sugar that requires a unique glycosyltransferase in the two methanococcal species. Flagellins from the previously characterized Δ*MMP350* acetyltransferase mutant (VanDyke *et al*., 2008) migrated at a molecular weight comparable to flagellins from Δ*MMP1079* cells, indicating that these proteins also likely had one sugar attached. Therefore, the MMP350 acetyltransferase appears to be critical in the biosynthetic pathway of GlcNAc3NAcA and, upon deletion, results in cells that cannot synthesize the substrate of MMP1079 resulting in no incorporation into the flagellin glycan. A model for the *N*-linked glycosylation pathway of *M. maripaludis* flagellin proteins, incorporating the proposed roles of MMP1424 (AglB), MMP1079 (AglO), MMP1080 (AglA) and MMP1088 (AglL), is shown in Fig. 8.

The requirement of the flagellin *N*-linked glycan for proper flagellar assembly and function was investigated by assessing all four mutants for motility by semi-swarm plate analysis and for the presence of flagellar filaments by EM. Deletion of either *MMP1424* or *MMP1079* resulted in cells that were non-motile and did not have flagella demonstrating the requirement of at least some partial glycan for flagellin subunit assembly. *M. maripaludis* mutants possessing flagellins with at least two sugars of the tetrasaccharide can be assembled into filaments with the presence of each additional sugar resulting in increased motility, as measured on semi-swarm plates. Δ*MMP1080* cells, with flagellins possessing a two-sugar glycan migrated less far in the semi-swarm assay than did the Δ*MMP1088* mutant (lacking the terminal sugar and threonine modification), which in turn migrated less than wild-type cells containing the complete four-sugar glycan. If only the first two sugars are required to assemble flagella, the contribution of each of the additional two sugars appears to be relevant for motility. In bacteria, the glycans found on flagellins thus far are limited to *O*-linked structures (Logan, 2006) and the requirement for glycosylation in flagellar assembly is variable. Glycosylation mutants in *Campylobacter* spp. (Goon *et al*., 2003), *Helicobacter pylori* (Schirm *et al*., 2003) and *Caulobacter crescentus* (Leclerc *et al*., 1998) failed to assemble flagellar filaments. In contrast, lack of glycosylation had no effect on flagellar assembly or subsequent motility in *Pseudomonas* spp. and *Listeria monocytogenes* (Takeuchi *et al*. 2003; Schirm *et al*., 2004a,b; Shen *et al*., 2006).

The glycosyltransferase responsible for the transfer of the first sugar (GalNAc) to the lipid carrier still remains to be determined in *M. maripaludis*. In *M. voltae,* where the first sugar in the glycan is GlcNAc, it was demonstrated by complementation of a yeast conditional lethal mutant in Alg7 (the GlcNAc-phosphate transferase responsible for adding GlcNAc, to the dolichol-P carrier in eukaryotic *N*-linked glycosylation) that the enzyme responsible for the GlcNAc addition was AglH (Shams-Eldin *et al*., 2008). A homologue of AglH is found in *M. maripaludis* (*MMP1423*). Like *aglH* in *M. voltae*, *MMP1423* could not be deleted, indicating that it may function in similar fashion and be responsible for the transfer of either GalNAc or GlcNAc. This could depend on the relative amounts of these sugars in intracellular pools in the two methanogens, which are data that are currently lacking. The essential role played by MMP1423 is likely in the biosynthesis of the methanogenic cofactor, cofactor B (Namboori and Graham, 2008). Also yet to be discovered in the archaeal *N*-linked glycosylation pathway is the flippase responsible for transporting the glycan across the cytoplasmic membrane. Although disruption of several gene candidates has been carried out, both in this study (*MMP354* and *MMP1089*) and previously (Chaban *et al*., 2006), this component remains elusive. The struggle to identify this component of the archaeal *N*-glycosylation pathway could be reflective of a unique nature of the element responsible for this task.

The data presented in this study identify several components involved in the assembly and attachment of this unique flagellar *N*-linked glycan. The importance of this glycan was demonstrated with respect to proper flagellin assembly and function as a minimal portion of the glycan was required to facilitate flagellar assembly and the presence of each additional sugar improved motility. This information contributes to the limited data available on archaeal *N*-glycosylation and enzymes responsible for biosynthesis and assembly of the novel archaeal flagellar glycans.

Experimental procedures

Strains and growth conditions

The *M. maripaludis* strain Mm900 was used to create in-frame deletions of target genes. This strain, derived from wild-type strain S2, contains a deletion in the *hpt* gene (encoding hypoxanthine phosphoribosyltransferase), making it resistant to the toxic base analogue 8-azahypoxanthine (Moore and Leigh, 2005). Mm900 was grown in McCas medium (Moore and Leigh, 2005) at 30°C under an atmosphere of CO_2/H_2 (20:80). Neomycin (1 mg ml^{-1}), puromycin (2.5 μg ml^{-1}) and 8-azahypoxanthine (240 μg ml^{-1}) were added for selection when required. For complementation studies, cells were cultured in nitrogen-free minimal medium (Blank *et al*., 1995) supplemented with NH_4Cl (10 mM) or L-alanine (10 mM) as a nitrogen source. *Escherichia coli* DH5α, grown at 37°C in Luria–Bertani medium, was used for all plasmid cloning, and ampicillin (100 μg ml^{-1}) was added for selection when necessary.

Plasmid construction

Deletion plasmids of all targeted genes were constructed as previously described (Moore and Leigh, 2005; VanDyke *et al*., 2008). These plasmids are listed in Table S1.

To generate plasmids for complementation under an inducible promoter (see Table S1) the genes *MMP1079*, *MMP1080*, *MMP1088* and *MMP1424* were amplified by PCR, incorporating C-terminal histidine tags, and cloned into pHW40 (obtained from J. Leigh). When cloned into pHW40, expression of each gene was under control of the *nif* promoter (Kessler and Leigh, 1999). Expression was either induced by growing cells in nitrogen-free media supplemented with L-alanine or prevented by supplementing the media with NH_4Cl. The veracity of all plasmid constructs was confirmed by DNA sequencing.

Generation of M. maripaludis *mutants*

Markerless in-frame deletions were constructed using a double-recombination event as demonstrated previously (Moore and Leigh, 2005). Potential deletion mutants were identified by PCR across the target gene using washed cells as template and the deletion was confirmed by Southern blot analysis.

Southern blot analysis

Chromosomal DNA was digested with MboI, MspI, EcoRI/PvuII or HindIII, for analysis of Δ*MMP1079,* Δ*MMP1080*, Δ*MMP1088* and Δ*MMP1424* mutants, respectively. Restriction enzymes were chosen following examination of target gene regions for restriction sites located on either side of, but not within, the target gene. Southern blotting was carried out as previously described (Chaban *et al*., 2007).

Immunoblotting

Immunoblots of whole-cell lysates were run and developed as previously described (VanDyke *et al*., 2008).

Flagella isolation

Flagella were isolated from *M. maripaludis* strain Mm900 and the deletion strains Δ*MMP1080* and Δ*MMP1088* as well as from the complemented Δ*MMP1080* and ΔMMP1088 strains, as previously described (Bardy *et al*., 2002).

Semi-swarm plate analysis

Overnight cultures (5 ml) of *M. maripaludis* Mm900 and mutants Δ*MMP1079*, Δ*MMP1080*, Δ*MMP1088* and Δ*MMP1424* were pelleted anaerobically in sealed centrifuge

tubes. The pellets were resuspended gently in 200 µl of Balch medium III (Balch *et al*., 1979) and 10 µl was used to inoculate the centre of Balch III medium plates containing 0.25% agar. Plates were incubated for 7 days at 30°C, after which, motility was assessed by measuring the diameter of the swarm zone.

Electron microscopy

The *M. maripaludis* cells were washed with 50 mM $MgSO_4$ and negatively stained with 2% phosphotungstic acid (pH 7.0). Preparations were transferred to formvar-coated gold grids and viewed on a Hitachi H-700 transmission electron microscope operating at 75 kV.

Mass spectrometry

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

Acknowledgements

We thank J. Leigh for supplying *M. maripaludis* Mm900, as well as the plasmids used in this study. This work was supported by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (NSERC, to K.F.J.) and by the National Research Council of Canada (to S.M.L. and J.F.K.).

References

Abu-Qarn, M., and Eichler, J. (2007) An analysis of amino acid sequences surrounding archaeal glycoprotein sequons. *Archaea* **2:** 73–81.

Abu-Qarn, M., Yurist-Doutsch, S., Giordano, A., Trauner, A., Morris, H.R., Hitchen, P., *et al.* (2007) *Haloferax volcanii* AglB and AglD are involved in N-glycosylation of the S-layer glycoprotein and proper assembly of the surface layer. *J Mol Biol* **374:** 1224–1236.

Abu-Qarn, M., Giordano, A., Battaglia, F., Trauner, A., Hitchen, P.G., Morris, H.R., *et al.* (2008) Identification of AglE, a second glycosyltransferase involved in N-glycosylation of the *Haloferax volcanii* S-layer glycoprotein. *J Bacteriol* **190:** 3140–3146.

Balch, W.E., Fox, G.E., Magrum, L.J., Woese, C.R., and Wolfe, R.S. (1979) Methanogens: reevaluation of a unique biological group. *Microbiol Rev* **43:** 260–296.

Bardy, S.L., Mori, T., Komoriya, K., Aizawa, S.-I., and Jarrell, K.F. (2002) Identification and localization of flagellins FlaA and FlaB3 within flagella of *Methanococcus voltae*. *J Bacteriol* **184:** 5223–5233.

Blank, C.E., Kessler, P.S., and Leigh, J.A. (1995) Genetics in methanogens: transposon insertion mutagenesis of a *Methanococcus maripaludis nif*H gene. *J Bacteriol* **177:** 5773–5777.

Chaban, B., Voisin, S., Kelly, J., Logan, S.M., and Jarrell, K.F. (2006) Identification of genes involved in the biosynthesis and attachment of *Methanococcus voltae* N-linked glycans: insight into N-linked glycosylation pathways in archaea. *Mol Microbiol* **61:** 259–268.

Chaban, B., Ng, S.Y., Kanbe, M., Saltzman, I., Nimmo, G., Aizawa, S.I., and Jarrell, K.F. (2007) Systematic deletion analyses of the *fla* genes in the flagella operon identify several genes essential for proper assembly and function of flagella in the archaeon, *Methanococcus maripaludis*. *Mol Microbiol* **66:** 596–609.

Chaban, B., Logan, S.M., Kelly, J.F., and Jarrell, K.F. (2009) AglC and AglK are involved in biosynthesis and attachment of diacetylated glucuronic acid to the N-glycan in *Methanococcus voltae*. *J Bacteriol* **191:** 187–195.

Cohen-Krausz, S., and Trachtenberg, S. (2008) The flagellar filament structure of the extreme acidothermophile *Sulfolobus shibatae* B12 suggests that archaeabacterial flagella have a unique and common symmetry and design. *J Mol Biol* **375:** 1113–1124.

Eichler, J., and Adams, M.W.W. (2005) Posttranslational protein modification in archaea. *Microbiol Mol Biol Rev* **69:** 393–425.

Gavel, Y., and von Heijne, G. (1990) Sequence differences between glycosylated and non-glycosylated Asn-X-Thr/Ser acceptor sites: implications for protein engineering. *Protein Eng* **3:** 433–442.

Goon, S., Kelly, J.F., Logan, S.M., Ewing, C.P., and Guerry, P. (2003) Pseudaminic acid, the major modification on *Campylobacter* flagellin, is synthesized via the Cj1293 gene. *Mol Microbiol* **50:** 659–671.

Jarrell, K.F., and McBride, M.J. (2008) The surprisingly diverse ways that prokaryotes move. *Nat Rev Microbiol* **6:** 466–476.

Jarrell, K.F., Ng, S.Y., and Chaban, B. (2007) Flagellation and chemotaxis. In *Archaea: Molecular and Cellular Biology*. Cavicchioli, R. (ed.). Washington, DC: American Society for Microbiology Press, pp. 385–410.

Kelly, J., Logan, S.M., Jarrell, K.F., Vandyke, D.J., and Vinogradov, E. (2009) A novel N-linked flagellar glycan from *Methanococcus maripaludis*. *Carbohydr Res* **344:** 648–653.

Kessler, P.S., and Leigh, J.A. (1999) Genetics of nitrogen regulation in *Methanococcus maripaludis*. *Genetics* **152:** 1343–1351.

Kowarik, M., Young, N.M., Numao, S., Schulz, B.L., Hug, I., Callewaert, N., *et al.* (2006) Definition of the bacterial N-glycosylation site consensus sequence. *EMBO J* **25:** 1957–1966.

Leclerc, G., Wang, S.P., and Ely, B. (1998) A new class of *Caulobacter crescentus* flagellar genes. *J Bacteriol* **180:** 5010–5019.

Lie, T.J., Wood, G.E., and Leigh, J.A. (2005) Regulation of *nif* expression in *Methanococcus maripaludis*: roles of the euryarchaeal repressor NrpR, 2-oxoglutarate, and two operators. *J Biol Chem* **280:** 5236–5241.

Logan, S.M. (2006) Flagellar glycosylation – a new component of the motility repertoire? *Microbiology* **152:** 1249–1262.

Macnab, R.M. (2003) How bacteria assemble flagella. *Annu Rev Microbiol* **57:** 77–100.
Moore, B.C., and Leigh, J.A. (2005) Markerless mutagenesis in *Methanococcus maripaludis* demonstrates roles for alanine dehydrogenase, alanine racemase, and alanine permease. *J Bacteriol* **187:** 972–979.
Namboori, S.C., and Graham, D.E. (2008) Acetamido sugar biosynthesis in the euryarchaea. *J Bacteriol* **190:** 2987–2996.
Ng, S.Y., Chaban, B., and Jarrell, K.F. (2006) Archaeal flagella, bacterial flagella and type IV pili: a comparison of genes and posttranslational modifications. *J Mol Microbiol Biotechnol* **11:** 167–191.
Peabody, C.R., Chung, Y.J., Yen, M.R., Vidal-Ingigliardi, D., Pugsley, A.P., and Saier, M.H., Jr, (2003) Type II protein secretion and its relationship to bacterial type IV pili and archaeal flagella. *Microbiology* **149:** 3051–3072.
Plavner, N., and Eichler, J. (2008) Defining the topology of the N-glycosylation pathway in the halophilic archaeon *Haloferax volcanii*. *J Bacteriol* **190:** 8045–8052.
Schirm, M., Soo, E.C., Aubry, A.J., Austin, J., Thibault, P., and Logan, S.M. (2003) Structural, genetic and functional characterization of the flagellin glycosylation process in *Helicobacter pylori*. *Mol Microbiol* **48:** 1579–1592.
Schirm, M., Arora, S.K., Verma, A., Vinogradov, E., Thibault, P., Ramphal, R., and Logan, S.M. (2004a) Structural and genetic characterization of glycosylation of type a flagellin in *Pseudomonas aeruginosa*. *J Bacteriol* **186:** 2523–2531.
Schirm, M., Kalmokoff, M., Aubry, A., Thibault, P., Sandoz, M., and Logan, S.M. (2004b) Flagellin from *Listeria monocytogenes* is glycosylated with beta-O-linked N-acetylglucosamine. *J Bacteriol* **186:** 6721–6727.
Shams-Eldin, H., Chaban, B., Niehus, S., Schwarz, R.T., and Jarrell, K.F. (2008) Identification of the archaeal *alg7* gene homolog (encoding N-acetylglucosamine-1-phosphate transferase) of the N-linked glycosylation system by cross-domain complementation in *Saccharomyces cerevisiae*. *J Bacteriol* **190:** 2217–2220.
Shen, A., Kamp, H.D., Grundling, A., and Higgins, D.E. (2006) A bifunctional O-GlcNAc transferase governs flagellar motility through anti-repression. *Genes Dev* **20:** 3283–3295.
Sumper, M. (1987) Halobacterial glycoprotein biosynthesis. *Biochim Biophys Acta* **906:** 69–79.
Szymanski, C.M., and Wren, B.W. (2005) Protein glycosylation in bacterial mucosal pathogens. *Nat Rev Microbiol* **3:** 225–237.
Takeuchi, K., Taguchi, F., Inagaki, Y., Toyoda, K., Shiraishi, T., and Ichinose, Y. (2003) Flagellin glycosylation island in *Pseudomonas syringae* pv. glycinea and its role in host specificity. *J Bacteriol* **185:** 6658–6665.
Trachtenberg, S., and Cohen-Krausz, S. (2006) The archaeabacterial flagellar filament: a bacterial propeller with a pilus-like structure. *J Mol Microbiol Biotechnol* **11:** 208–220.
Twine, S.M., Paul, C.J., Vinogradov, E., McNally, D.J., Brisson, J.R., Mullen, J.A., *et al.* (2008) Flagellar glycosylation in *Clostridium botulinum*. *FEBS J* **275:** 4428–4444.
VanDyke, D.J., Wu, J., Ng, S.Y., Kanbe, M., Chaban, B., Aizawa, S.I., and Jarrell, K.F. (2008) Identification of putative acetyltransferase gene, MMP0350, which affects proper assembly of both flagella and pili in the archaeon *Methanococcus maripaludis*. *J Bacteriol* **190:** 5300–5307.
Voisin, S., Houliston, R.S., Kelly, J., Brisson, J.R., Watson, D., Bardy, S.L., *et al.* (2005) Identification and characterization of the unique N-linked glycan common to the flagellins and S-layer glycoprotein of *Methanococcus voltae*. *J Biol Chem* **280:** 16586–16593.
Weerapana, E., and Imperiali, B. (2006) Asparagine-linked protein glycosylation: from eukaryotic to prokaryotic systems. *Glycobiology* **16:** 91R–101R.
Wieland, F. (1988) Structure and biosynthesis of prokaryotic glycoproteins. *Biochimie* **70:** 1493–1504.
Yurist-Doutsch, S., Chaban, B., VanDyke, D.J., Jarrell, K.F., and Eichler, J. (2008a) Sweet to the extreme: protein glycosylation in archaea. *Mol Microbiol* **68:** 1079–1084.
Yurist-Doutsch, S., Abu-Qarn, M., Battaglia, F., Morris, H.R., Hitchen, P.G., Dell, A., and Eichler, J. (2008b) AglF, aglG and aglI, novel members of a gene island involved in the N-glycosylation of the *Haloferax volcanii* S-layer glycoprotein. *Mol Microbiol* **69:** 1234–1245.

Supporting Information

Additional supporting information may be found in the online version of this article.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the Correspondence for the article.

