JB Journals.ASM.org

Identification of Genes Involved in the Biosynthesis of the Third and Fourth Sugars of the *Methanococcus maripaludis* Archaellin N-Linked Tetrasaccharide

Yan Ding,[a] Gareth M. Jones,[a] Kaoru Uchida,[c] Shin-Ichi Aizawa,[c] Anna Robotham,[b] Susan M. Logan,[b] John Kelly,[b] Ken F. Jarrell[a]

Department of Biomedical and Molecular Sciences, Queen's University, Kingston, Ontario, Canada[a]; Human Health Therapeutics Portfolio, National Research Council, Ottawa, Ontario, Canada[b]; Department of Life Sciences, Prefectural University of Hiroshima, Shobara, Hiroshima, Japan[c]

N-glycosylation is a protein posttranslational modification found in all three domains of life. Many surface proteins in *Archaea*, including S-layer proteins, pilins, and archaellins (archaeal flagellins) are known to contain N-linked glycans. In *Methanococcus maripaludis*, the archaellins are modified at multiple sites with an N-linked tetrasaccharide with the structure Sug-1,4-β-ManNAc3NAmA6Thr-1,4-β-GlcNAc3NAcA-1,3-β-GalNAc, where Sug is the unique sugar (5*S*)-2-acetamido-2,4-dideoxy-5-*O*-methyl-α-L-erythro-hexos-5-ulo-1,5-pyranose. In this study, four genes—*mmp1084*, *mmp1085*, *mmp1086*, and *mmp1087*—were targeted to determine their potential involvement of the biosynthesis of the sugar components in the N-glycan, based on bioinformatics analysis and proximity to a number of genes which have been previously demonstrated to be involved in the N-glycosylation pathway. The genes *mmp1084* to *mmp1087* were shown to be cotranscribed, and in-frame deletions of each gene as well as a Δ*mmp1086*Δ*mmp1087* double mutant were successfully generated. All mutants were archaellated and motile. Mass spectrometry examination of purified archaella revealed that in Δ*mmp1084* mutant cells, the threonine linked to the third sugar of the glycan was missing, indicating a putative threonine transferase function of MMP1084. Similar analysis of the archaella of the Δ*mmp1085* mutant cells demonstrated that the glycan lacked the methyl group at the C-5 position of the terminal sugar, indicating that MMP1085 is a methyltransferase involved in the biosynthesis of this unique sugar. Deletion of the remaining two genes, *mmp1086* and *mmp1087*, either singularly or together, had no effect on the structure of the archaellin N-glycan. Because of their demonstrated involvement in the N-glycosylation pathway, we designated *mmp1084* as *aglU* and *mmp1085* as *aglV*.

N-glycosylation is a prevalent protein posttranslational modification that is not limited to eukaryotes as originally thought but is also found in both prokaryotic domains, the *Bacteria* and *Archaea* (1–6). Unlike the bacterial N-glycosylation system, which is thought to be restricted to only a few bacterial subgroupings, such as *Deltaproteobacteria* and *Epsilonproteobacteria* (6), N-glycosylation is much more prevalent in the *Archaea*. A gene encoding an oligosaccharyltransferase, required for the transfer of the assembled glycan from the lipid carrier to the target protein, has been detected in all but two of sequenced archaeal genomes, representing all five archaeal phyla (7, 8). In addition, the N-glycosylation pathway in *Archaea* differs from the other domains in the diversity of N-glycan structures reported, variations in the lipid carriers and, at least in the halophiles, the versatility in the pathways used to assemble similar glycans in different species (2, 4, 9–11).

Although N-glycosylation is widely distributed in *Archaea*, studies combining both genetic and structural information about N-glycosylation pathways are focused mainly on a few model species (2, 9, 12), such as the mesophilic methanogens *Methanococcus voltae* and *Methanococcus maripaludis* (13–17), the thermoacidophile *Sulfolobus acidocaldarius* (9, 18–20), and the halophiles *Haloferax volcanii* and *Haloarcula marismortui* (2, 11, 21–26). As seen for many archaeal traits, the archaeal N-glycosylation pathway appears as a partial amalgam of the pathway found in eukaryotes and bacteria (1, 2, 27). In general, the archaeal N-glycan precursor is first synthesized onto a lipid carrier (mono- or pyrophosphorylated dolichol) by sequential addition of sugar monomers by glycosyltransferases using sugar-nucleotide donors on the cytoplasmic side of the cytoplasmic membrane. The lipid-linked oligosaccharide is then thought to be flipped across the cytoplasmic membrane by a flippase and transferred *en bloc* by the oligosaccharyltransferase AglB onto Asn residues within selected Asn-Xaa-Ser/Thr motifs in the target protein. To date, the identification of the putative flippases has remained elusive, except that in *H. volcanii* an involvement of AglR has been demonstrated in the flipping of one of the charged dolichol carriers (26).

Reporter proteins used for the genetic and structural analysis of archaeal N-glycosylation pathways are usually the S-layer protein (13, 18, 23) and archaellin (formerly archaeal flagellin [28]) (13, 16, 29), although other proteins have also been discovered to be modified with N-glycans, including pilins (15), cytochrome $b_{558/556}$ of *Sulfolobus acidocaldarius* (30), and even archaeal viral proteins (31).

In *M. maripaludis* S2, archaella are composed of three archaellin proteins: the major archaellins FlaB1 and FlaB2 form the filament, while the minor archaellin FlaB3 is responsible for the hook region (32). All three archaellins are modified by an N-linked tetrasaccharide attached at multiple sites: three sites in FlaB1, four sites in FlaB2 and two sites in FlaB3 (16). The structure of this N-glycan was determined to be Sug-1,4-β-ManNAc3NAmA6Thr-1,4-β-GlcNAc3NAcA-1,3-β-GalNAc, where Sug is [(5*S*)-2-acetamido-2,4-dideoxy-5-*O*-meth-

Received 4 June 2013 Accepted 3 July 2013
Published ahead of print 8 July 2013
Address correspondence to Ken F. Jarrell, jarrellk@queensu.ca, or John Kelly, john.kelly@nrc-cnrc.gc.ca.

doi:10.1128/JB.00668-13

yl-α-L-erythro-hexos-5-ulo-1,5-pyranose), found exclusively in this species (16). Interestingly, the N-glycan attached to the major pilin in *M. maripaludis* is a slightly modified pentasaccharide containing an extra hexose attached as a branch to the first sugar, *N*-acetyl-galactosamine (15).

In *M. maripaludis* S2, the oligosaccharyltransferase and glycosyltransferases responsible for the addition of the second, third, and fourth sugar residues of the N-glycan have been identified as AglB (MMP1424), AglO (MMP1079), AglA (MMP1080), and AlgL (MMP1088), respectively (17). The assembly and function of archaella were significantly affected in mutant cells carrying in-frame deletions of these genes, which all resulted in a truncated N-glycan. Archaellins modified with N-glycans lacking the last one (*algL* mutant) or two (*aglA* mutant) sugar residues are still assembled into archaella on the cell surface, although the motility of mutant cells is impaired compared to wild-type cells, with motility directly correlated to the length of the N-glycans. Further reduction in glycan size to a single sugar, as in an *aglO* mutant, or blocking the transfer of any N-glycan to archaellins, as in an *aglB* mutant, led to nonmotile, completely nonarchaellated cells. These results indicated that a minimum disaccharide N-glycan attachment to archaellins is essential for archaella assembly (17). Similarly, the addition of N-glycan onto archaellins has also been shown to be necessary for stable archaella formation and motility function in *S. acidocaldarius* (19) and *H. volcanii* (29). Interestingly, the addition of N-glycan is not required for pilins to assemble into pili in *M. maripaludis*, since *aglB* mutant cells still assemble a wild-type number of pili, even though the major structural pilin is normally a glycoprotein (17).

Besides the genes involved in the assembly of the glycan onto the lipid carrier, a number of genes involved in the biosynthesis pathways for the individual sugars of the glycan have been identified or characterized in either *in vivo* or *in vitro* studies (17, 33). Most recently, the products of the *mmp1081-mmp1082-mmp1083* (*aglXYZ*) gene cluster were shown to be responsible for the acetamidino group modification of the third sugar (34). Several of these genes are clustered in the genome, including ones encoding enzymes involved in sugar biosynthesis (*mmp1076*, *mmp1077*, *mmp1090*, *aglX*, *aglY*, and *aglZ*) (33, 34; Y. Ding, I. Brockhausen, and K. Jarrell, unpublished data) and glycosyltransferases (*aglO*, *aglA*, and *aglL* [17]) (see Fig. 1). Here, we identify a four-gene operon, *mmp1084* to *mmp1087*, located between *aglZ* and *aglL* and demonstrate an involvement of two of these genes in the modification of sugars in the N-linked glycan. We present evidence that MMP1084, which belongs to the glutamine amidotransferase class II (Gn_AT_II) superfamily, is involved in the transfer of the threonine residue linked to the third sugar of the N-glycan, whereas MMP1085 is responsible for the addition of the methyl group linked to C-5 on the fourth sugar, consistent with its predicted function as an *S*-adenosylmethionine (SAM)-dependent methyltransferase. Deletion of the remaining two genes, *mmp1086* and *mmp1087*, either singularly or together, had no effect on the structure of the final attached N-glycan.

MATERIALS AND METHODS

Strains and growth conditions. *M. maripaludis* S2 Δ*hpt* (Mm900) (35) and subsequent mutants derived from Mm900 were grown anaerobically in Balch medium III (36) under an atmosphere of CO_2-H_2 (20:80) at 37°C with shaking. In the in-frame deletion mutagenesis experiments, Mm900 derived transformants were grown in McCas medium with 1 mg of neomycin/ml or 240 μg of 8-aza-hypoxanthine/ml as required at the various steps (35). In complementation experiments, *M. maripaludis* mutant strains harboring the appropriate complementation plasmid were grown in nitrogen-free medium supplemented with either 10 mM L-alanine or 10 mM NH_4Cl as the sole nitrogen source (37) in the presence of 2.5 μg of puromycin/ml for plasmid selection. *Escherichia coli* TOP10 cells used for cloning steps were grown in Luria-Bertani medium at 37°C with 100 μg of ampicillin/ml for selection when required.

RT-PCR. To determine whether the genes *mmp1084* to *mmp1087* formed a cotranscribed unit, reverse transcriptase PCR (RT-PCR) was performed using primers (listed in Table 1) designed to amplify across the intergenic regions of neighboring genes. RNA template was extracted from wild-type cells using an RNeasy Minikit (Qiagen, Inc., Mississauga, Ontario, Canada) with optional DNase digestion (Qiagen) according to the manufacturer's protocol. cDNA was amplified using a One-Step RT-PCR kit (Qiagen) in accordance with the supplied protocol. In addition, PCR amplifications were performed with the same primer combinations using genomic DNA to verify amplicon size and specificity of primer pairs or the purified RNA not subjected to reverse transcription as a control to rule out possible contamination of the RNA sample with genomic DNA.

In-frame gene deletion plasmid construction. Plasmids used for in-frame deletions of *mmp1084*, *mmp1085*, *mmp1086*, and *mmp1087* were constructed as previously described (35), using the primers listed in Table 1. Confirmation of the in-frame nature of each deletion was obtained by DNA sequencing. Table 2 lists the plasmids used in the present study.

Generation of *M. maripaludis* mutants carrying in-frame deletions in targeted genes. Mutants carrying markerless in-frame deletions of *mmp1084*, *mmp1085*, *mmp1086*, and *mmp1087* were generated as described previously (35). Single transformant colonies picked from McCas-Noble agar containing 8-azahypoxanthine were screened by PCR using primers (listed in Table 1) that would amplify across the target gene to identify deletion mutants. Mutants identified in this screen were restreaked, and single colonies were again screened by PCR to ensure purity. To create the mutant deleted for both *mmp1086* and *mmp1087*, *mmp1086* was deleted from the *mmp1087* deletion mutant. Table 2 lists the strains generated in this study.

Complementation studies. *mmp1084* and *mmp1085* were amplified using the complementation primers listed in Table 1, digested with NsiI and MluI, and ligated into the complementation vector pHW40, resulting in the inserted gene being under the control of the inducible *nif* promoter (37). To facilitate cloning, an internal NsiI site in *mmp1085* was removed by a silent T-to-C change at nucleotide position 72 using the long forward primer 1085comp_F (Table 1). The fidelity of the cloned genes in all of the complementation plasmids (listed in Table 2) was confirmed by DNA sequencing. The complementation vectors were transformed into the corresponding deletion mutants under the selection of 2.5 μg of puromycin/ml, as described previously (37, 38). Complemented mutants were grown in the presence of puromycin in nitrogen-free medium supplemented with either L-alanine (*nif* promoter induced) or NH_4Cl (*nif* promoter repressed) (32, 37).

Western blot analysis. *M. maripaludis* whole-cell lysates were subjected to SDS–15% PAGE (39) and then transferred to an Immobilon-P membrane (Millipore, Bedford, MA) (40). Major archaellin FlaB2 was detected with chicken anti-FlaB2 specific antibodies (34). Horseradish peroxidase-conjugated rabbit anti-chicken immunoglobulin Y (Jackson ImmunoResearch Laboratories, West Grove, PA) was used as secondary antibody. Blots were developed with an enhanced chemiluminescence kit (Millipore) according to the manufacturer's instructions.

Archaella isolation. Archaella from the various single gene deletion strains, as well as the Δ*mmp1086* Δ*mmp1087* double mutant, were isolated as previously described (41).

Semisolid swarm motility plate assay. Overnight wild-type cells and various mutants were pelleted anaerobically and resuspended in Balch III medium to a final optical density at 600 nm of 3.0. Then, 5 μl of each resuspension was inoculated into semisolid swarm plates (Balch III me-

TABLE 1 Primers used in this study

Technique and primer	Sequence (5′–3′)[a]	Restriction site
RT-PCR		
1083–1084_F	GTTTTGGGTGCTGGAGAACTCG	
1083–1084_R	AATTTGGCGAATACTACGAAGC	
1084–1085_F	TGTATCTTCGGGATTAAATCGC	
1084–1085_R	TATAACTGCACCAAATGTCTGG	
1085–1086_F	ATAAGTGCTCAATTACATCAAGG	
1085–1086_R	AATCTTGGAAATTGTAATGAGG	
1086–1087_F	TTGGATCCAACAAAGATTCTCC	BamHI*
1086–1087_R	ATCTCACACTTTCATACCTACG	
1087–1088_F	TTTCACGAGTAACTTTCCTTCG	
1087–1088_R	AATCTTGGAAAACATTACTTGC	
In-frame deletions/plasmid construction		
1084del-up_F	ACCAAGGATCCGTAAATAACAATGTAAATGGATG	BamHI
1084del-up_R	AAAGGCGCGCCCCTATATTTCCGGCATAAAATC	AscI
1084del-down_F	AAAGGCGCGCCATGGATACCTATTATGACGCTAC	AscI
1084del-down_R	TTAAGGATCCCAAGTGCTGTTTCGAAGGATAAC	BamHI
1085del-up_F	AAAGGATCCAACACTATGAATTTTATAGAACG	BamHI
1085del-up_R	AAAGGCGCGCCATGCATGCAATAGATAATCATG	AscI
1085del-down_F	AAAGGCGCGCCCCACATATGCGTAATGTTTTA	AscI
1085del-down_R	AAAGGATCCGTGCTATAATATATCTTGAATCC	BamHI
1086del-up_F	ATTCTAGATCTAGAGTAGAACTAACTAATGATTGTAACTATCAGTG	XbaI
1086del-up_R	GCATGGCGCGCCGGTGTGTACAATATTACATTGTG	AscI
1086del-down_F	GCATGGCGCGCCCCATATTATGCTTTACTTCAATGTAGGC	AscI
1086del-down_R	ATTCTAGATCTAGATCGTGTATTTTATCGAGCCCATGAGC	XbaI
1087del-up_F	ATTCTAGATCTAGAACCACAACCTAAAGCTTCGAACATCACTGTG	XbaI
1087del-up_R	GCATGGCGCGCCCAATCATTAGTTAGTTCTACTATAGC	AscI
1087del-down_F	GCATGGCGCGCCGTGAAGCTCCATCGCCATTGTTC	AscI
1087del-down_R	ATTCTAGATCTAGAAATTCACCGCTAACATCCATGTGTG	XbaI
In-frame deletions/screening		
1084seq_F	ATTTTCAGGATATTCTGTTAAAGG	
1084seq_R	ACGATGCAATTAATGTGGGAGC	
1085seq_F	AGTAAATAACAATGTAAATGG	
1085seq_R	ATATTTTCTGTATCTTCGGG	
1086seq_F	TACCTACGGCAAAATGTGAA	
1086seq_R	AATTCACCGCTAACATCCAT	
1087seq_F	CTCTTGTCGGGATAACATGG	
1087seq_R	GTGTGGACAATACTTACATTGTG	
Complementation		
1084comp_F	GGCCATGCATGCCGGAAATATACCTTGAA	NsiI
1084comp_R	CCGGACGCGTTTATTTATATTTTATATCATAATATTTTTCAATGTATTC	MluI
1085comp_F	GGCCATGCATGGATGTTAGCGGTGAATTTATGAAAACTAATACCTGGTTAAACT ATAAACATGATTATCTATTGC**ACGCAT**	NsiI
1085comp_R	CCGGACGCGTTCACCGATTTGAATTTATCTCC	MluI

[a] Restriction sites added to primers are underlined. *, Natural BamHI site in this primer. The abolished NsiI site after a silent T-to-C mutation is indicated in boldface.

dium containing 0.25% agar), followed by incubation at 37°C for 5 days (17).

Mass spectrometry analysis. Purified archaella (50 μg) from the single mutants were digested overnight with trypsin (Promega, Madison, WI) at a ratio of 30:1 (protein-enzyme [vol/vol]) in 50 mM ammonium bicarbonate at 37°C. The archaella isolated from the Δ*mmp1086* Δ*mmp1087* double mutant was first resolved by SDS–15% PAGE and then stained with collioidal Coomassie blue. The archaellin band was excised and subjected to an overnight in-gel tryptic digestion. All of the archaellin digests were analyzed by nano-liquid chromatography-tandem mass spectrometry (Nano-LC-MS/MS) using a NanoAquity UPLC system (Waters, Milford, MA) coupled to a QTOF Ultima hybrid quadrupole time-of-flight mass spectrometer (Waters). The digests were injected onto an Acclaim PepMap100 C_{18} μ-precolumn (5 mm by 300 μm [inner diameter]; Dionex/Thermo Scientific, Sunnyvale, CA) and resolved on a 1.7-μm BEH130 C_{18} column (100 μm by 100 mm [inner diameter]; Waters) using the following gradient conditions: 5 to 45% acetonitrile (ACN) in 0.1% formic acid for 35 min and 45 to 95% ACN for 5 min. The flow rate was 450 nl/min. MS/MS spectra were acquired on doubly, triply, and quadruply charged ions and searched against the NCBInr database using the Mascot search engine (Matrix Science, Ltd., London, United Kingdom). The spectral data sets were searched for glycopeptide MS/MS spectra, which were then interpreted by hand.

Electron microscopy. *M. maripaludis* S2 wild-type and deletion mutants were grown overnight. A 1-ml portion of the cells was briefly washed and resuspended in phosphate-buffered saline. Samples were negatively

TABLE 2 Strains and plasmids used in this study

Strain or plasmid	Description and/or genotype[a]	Source or reference
Strains		
E. coli TOP10	F$^-$ *mcrA* Δ(*mrr-hsdRMS-mcrBC*) ϕ80*lacZ*ΔM15 Δ *lacX74 nupG recA1 araD139* Δ(*ara-leu*)*7697 galE15 galK16 rpsL*(Str^r) *endA1* λ$^-$	Invitrogen
M. maripaludis S2		
Mm900	Δ*hpt*	35
Δ*mmp1084*	Mm900Δ*mmp1084*	This study
Δ*mmp1085*	Mm900Δ*mmp1085*	This study
Δ*mmp1086*	Mm900Δ*mmp1086*	This study
Δ*mmp1087*	Mm900Δ*mmp1087*	This study
Δ*mmp1086*Δ*mmp1087*	Mm900Δ*mmp1086* Δ*mmp1087*	This study
Plasmids		
pCRPrtNeo	*hmv* promoter-*hpt* fusion plus Neo^r cassette in pCR2.1Topo; Amp^r	35
pKJ795	pCRPrtNeo with in-frame deletion of *mmp1084*	This study
pKJ715	pCRPrtNeo with in-frame deletion of *mmp1085*	This study
pKJ882	pCRPrtNeo with in-frame deletion of *mmp1086*	This study
pKJ783	pCRPrtNeo with in-frame deletion of *mmp1087*	This study
pHW40	*nif* promoter-*lacZ* fusion plus Pur^r cassette; Amp^r	John Leigh
pKJ959	pHW40 with *mmp1084* complement	This study
pKJ955	pCRII-TOPO with *mmp1085*	This study
pKJ958	pHW40 with *mmp1085* complement	This study

[a] Neo^r, neomycin resistance; Pur^r, puromycin resistance; Amp^r, ampicillin resistance; Str^r, streptomycin resistance.

stained with 2% phosphotungstic acid and supported on carbon-Formvar-coated copper grids. The samples were then examined in a Hitachi 7000 electron microscope operating at an accelerating voltage of 75 kV.

RESULTS

The genomic location of *mmp1084* to *mmp1087* between genes already shown to be involved in the assembly or biosynthesis of the archaellin N-linked glycan (Fig. 1) led us to examine these genes for a possible role in the N-linked glycosylation pathway.

Bioinformatic analysis of *mmp1084* to *mmp1087*. As a first step in obtaining information about the putative functions of MMP1084 to MMP1087, bioinformatics-based analyses were performed with these protein sequences. When MMP1084 was used as a query in a BLAST search (i.e., the Basic Local Alignment Search Tool [http://blast.ncbi.nlm.nih.gov/]) (42), the results showed significant similarities to numerous sequences, most of which are asparagine synthases from a variety of different species. BLAST analysis also indicated that MMP1084 possessed a Gn_AT_II superfamily domain that is found at the N terminus of various enzymes, including asparagine synthase. MMP1085 analysis revealed a conserved domain with specific hits to SAM-dependent methyltransferases. MMP1086 and MMP1087 share 56% identity, and both have an N-terminal radical-SAM domain with a FeS/SAM binding site, including the conserved CxxxCxxC motif, which coordinates the iron-sulfur cluster. Radical SAM enzymes are widespread and are known to be involved in such diverse processes as metabolism, DNA repair, and the biosynthesis of vitamins, coenzymes, and antibiotics. The two proteins also contain a C-terminal DUF4008 domain. DUF4008, domain of unknown function 4008, is a functionally uncharacterized domain found in bacteria and archaea. Similar conserved domains were found using another protein functional analysis server, InterProScan (http://www.ebi.ac.uk/Tools/pfa/iprscan/ [43]). The potential cellular location of the four proteins was examined using PRED-SIGNAL (http://bioinformatics.biol.uoa.gr/PRED-SIGNAL/) (44), TMHMM (http://www.cbs.dtu.dk/services/TMHMM/), and PSortb (http://www.psort.org/) (45). No signal peptides or transmembrane domains were found, and the location of the proteins was predicted by PSortb (set to archaeal proteins) to be cytoplasmic, except for MMP1086 for which a location could not be predicted. Interestingly, homologues of *mmp1084* to *mmp1087* are not found in the four other sequenced *M. maripaludis* strains (C5, C6, C7, and X1).

FIG 1 Genomic area surrounding the *mmp1084*-to-*mmp1087* region, where a number of genes involved in the N-glycosylation pathway are located. *mmp1077* encodes a phosphomutase converting D-glucosamine-6-phosphate to D-glucosamine-1-phosphate. MMP1076 is a GlcN-1-P uridylyltransferase/acetyltransferase that catalyzes the acetylation of the 2-amino group of GlcN-1-P and the transfer of GlcNAc-1-P to UTP to form UDP-GlcNAc (33). *aglO*, *aglA*, and *aglL* encode the glycosyltransferases for the second, third and fourth sugars, respectively (17) The *aglXYZ* gene cluster is responsible for the acetamidino group of the third sugar (34). *mmp1090* encodes UDP-Gal/GalNAc 4-epimerase (Ding et al., unpublished).

RT-PCR analysis of the *mmp1084*-to-*mmp1087* gene cluster. *mmp1084* to *mmp1087* are all oriented in the same direction and are separated by short intergenic regions (5, 37, and 45 bp, respectively), suggesting that they likely form a single transcriptional unit. To test this, primers were designed to amplify across the intergenic regions between adjacent two genes from *mmp1083* to *mmp1088*, as shown in Fig. 2A. In the RNA samples extracted from Mm900 cells subjected to reverse transcription, amplification products of the predicted size linking *mmp1084* and *mmp1085* (563 bp), *mmp1085* and *mmp1086* (550 bp), and *mmp1086* and *mmp1087* (484 bp) were obtained, whereas, as expected, no products were amplified between *mmp1083* and *mmp1084* (expected size, 584 bp) or between *mmp1087* and *mmp1088* (expected size, 510 bp), which are genes transcribed in

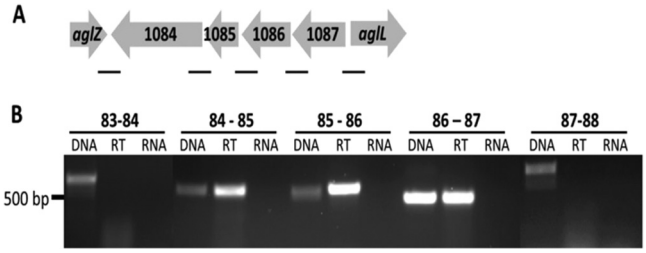

FIG 2 (A) *mmp1084-mmp1087* operon. The gene region surrounding *mmp1084* to *mmp1087*, which was targeted for RT-PCR to determine potential cotranscription. Black lines below the genes represent the predicted amplicons obtained from RT-PCR. (B) RT-PCR experiment indicating cotranscription of *mmp1084*, *mmp1085*, *mmp1086*, and *mmp1087*. Standard PCRs using Mm900 genomic DNA (DNA lanes) and the respective RT primers that amplify the intergenic regions were performed for amplicon size confirmation. RT-PCR (RT lanes) was run using total RNA extracted from Mm900 cells with the same RT primers. The RT lanes that have bands at the same size as the DNA lanes indicate the cotranscription of the according genes. Standard PCRs were also performed using total RNA that did not undergo reverse transcription as a template (RNA lanes) to rule out possible DNA contamination of the RNA sample.

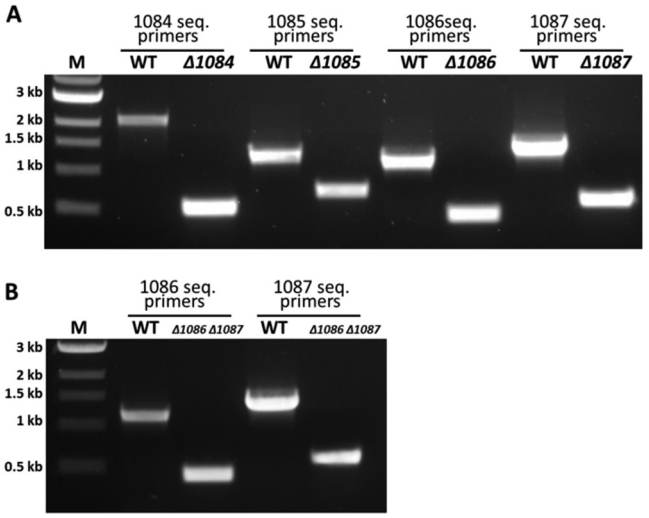

FIG 3 Confirmation of in-frame deletions of Δ*mmp1084*, Δ*mmp1085*, Δ*mmp1086*, and Δ*mmp1087* single mutants and the Δ*mmp1086* Δ*mmp1087* double mutant by PCR. (A) Washed whole cells of each deletion strain, as well as wild-type (WT) cells, were used as a template for the PCR confirmation of each in-frame deletion, using corresponding sequencing primers as marked in the figure. (B) For the Δ*mmp1086* Δ*mmp1087* double mutant, the in-frame deletions of *mmp1086* and *mmp1087* were determined individually. Washed whole cells of the Δ*mmp1086* Δ*mmp1087* mutant were used as a template with sequencing primers for both *mmp1086* and *mmp1087* mutants. In all cases, amplicons of the expected size from both the wild type and deletion mutants were obtained. M, 1-kb DNA ladder.

the opposite direction (Fig. 2B). PCRs using the same primers and the total RNA that had not undergone the reverse transcriptase reaction did not yield any amplification products, demonstrating that the RNA was free of contaminating DNA. Standard PCRs using Mm900 genomic DNA and the same primers were also conducted as positive controls to confirm the size of the amplicons and the specificity of the primers.

Generation of Δ*mmp1084*, Δ*mmp1085*, Δ*mmp1086*, and Δ*mmp1087* single mutants and the Δ*mmp1086* Δ*mmp1087* double mutant. To identify the possible involvement of *mmp1084* to *mmp1087* in the N-glycosylation system in *M. maripaludis* S2, in-frame deletions of each gene were generated and confirmed by PCR of transformant cells using primers that amplified across the targeted gene (Fig. 3). In each case, Mm900 genomic DNA (wild type) was also used as a template with the same primer pair as the corresponding deletion mutant to confirm the specificity of the primers and to provide a comparison wild-type size of the target gene. In the PCR amplification using cells of each deletion mutant as a template, the size of the PCR product was smaller than that from wild-type genomic DNA by the length predicted from the size of the deletion made in the target gene. In each case, sequencing of the PCR product confirmed that each deletion made was in frame. In the case of the double mutant Δ*mmp1086* Δ*mmp1087*, each gene deletion was confirmed individually by PCR and sequencing. As shown in Fig. 3B, the amplicons obtained with the Δ*mmp1086* Δ*mmp1087* deletion strain as a template using the *mmp1086* and *mmp1087* screening primers were smaller than those from the wild-type genomic DNA by the expected lengths, illustrating that the Δ*mmp1086* Δ*mmp1087* double mutant was successfully generated.

Western blot analysis of the archaellins from the mutants. To examine whether the deletion of any of the *mmp1084* to *mmp1087* genes had any detectable effect on the glycan N-linked to the *M. maripaludis* archaellins, the migration of the archaellin reporter glycoprotein FlaB2 in each deletion strain was initially examined by Western blots probed with antibody specific for this archaellin (34). It has been shown previously that even small truncations to the glycan have effects on archaellin migration that are detectable by Western blots (17, 34). As shown in Fig. 4, FlaB2 migrated as a slightly smaller protein in both Δ*mmp1084* and Δ*mmp1085* mutants compared to the wild-type cells, suggesting a possible effect on the glycan. In contrast, no change in the migration of FlaB2 in either Δ*mmp1086* or Δ*mmp1087* mutants compared to wild-type cells was evident. However, the 56% identity between MMP1086 and MMP1087 suggests that they might compensate for each other. To exclude the possibility that in either the Δ*mmp1086* or the Δ*mmp1087* single mutant, the remaining gene product complements the function of the in-frame deletion, FlaB2 from the Δ*mmp1086* Δ*mmp1087* double mutant was also analyzed by Western blot analysis, as shown in Fig. 4. The apparent molecular mass of FlaB2 from the Δ*mmp1086* Δ*mmp1087* double mutant was the same as that from wild-type cells, indicating that at least under the culture conditions used neither *mmp1086* nor *mmp1087* appeared to be involved in the N-glycosylation pathway. Since the roles of *aglY* and *aglZ* in N-glycosylation were not evident when cells are grown in the presence of ammonia (34), the

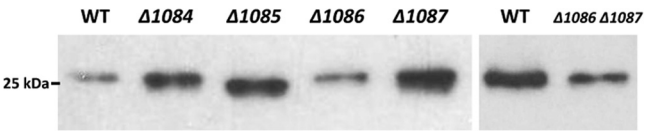

FIG 4 Western blot of whole-cell lysates of the wild type, Δ*mmp1084*, Δ*mmp1085*, Δ*mmp1086*, and Δ*mmp1087* single mutants, and the Δ*mmp1086* Δ*mmp1087* double mutant probed with anti-FlaB2 antibodies. The decrease in apparent molecular mass of FlaB2 in the Δ*mmp1084* and Δ*mmp1085* mutants was the initial evidence of a truncated N-glycan in these mutants. No change in the migration of FlaB2 was evident in the Δ*mmp1086* mutant, the Δ*mmp1087* mutant, or the Δ*mmp1086* Δ*mmp1087* double mutant.

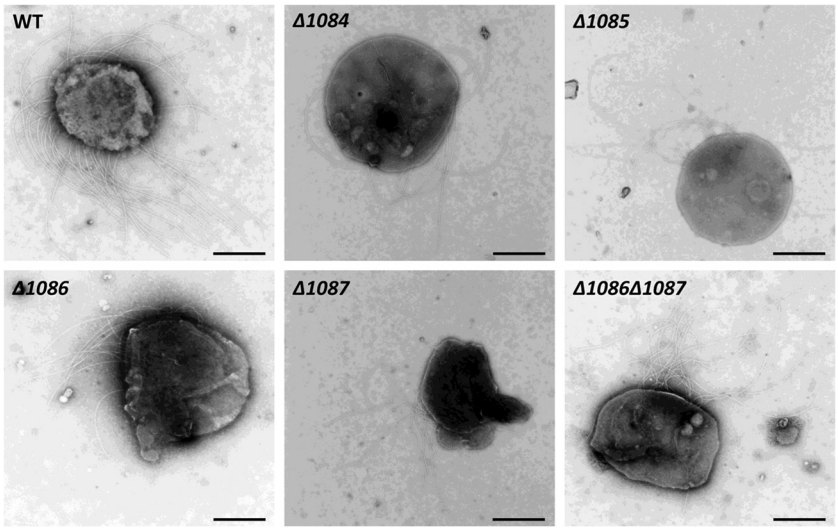

FIG 5 Electron microscopy of wild-type and mutant cells obtained in the present study. Numerous archaella are observed attached to the cell surface of the Δ*mmp1084*, Δ*mmp1085*, Δ*mmp1086*, and Δ*mmp1087* single mutants and the Δ*mmp1086* Δ*mmp1087* double mutant, as well as the wild-type cell. Scale bars, 500 nm.

size of FlaB2 in the Δ*mmp1086* Δ*mmp1087* double mutant was also examined under ammonia-limiting conditions by growing cells in nitrogen-free medium supplemented with alanine. No reduction in the size of FlaB2 in the double mutant was observed in Western blots under these conditions either (data not shown).

In-frame deletions of *mmp1084* to *mmp1087* do not interfere with archaellation. In order to determine whether proteins encoded by *mmp1084* to *mmp1087* are involved in archaella assembly, the Δ*mmp1084*, Δ*mmp1085*, Δ*mmp1086*, and Δ*mmp1087* deletion mutants, as well as the Δ*mmp1086* Δ*mmp1087* double mutant, were examined by electron microscopy. As shown in Fig. 5, numerous archaella were observed on the surfaces of all of the mutant cells, as well as the wild-type cell, suggesting that deletion of these genes had no gross effect on archaella assembly, number, or cellular location.

Semisolid swarm motility plate assay. Examination of wild-type cells and the various mutants by light microscopy demonstrated that all strains were motile. However, since the motility of even wild-type cells is weak (46), motility was also examined using semisolid agar. Here, a motility similar to that of wild-type cells was observed for the Δ*mmp1084*, Δ*mmp1085*, Δ*mmp1086*, Δ*mmp1087*, and Δ*mmp1086* Δ*mmp1087* deletion mutants (data not shown).

Mass spectrometry analysis of archaellins. Mass spectroscopy of isolated archaella confirmed the Western blot results that showed effects on glycosylation in the Δ*mmp1084* and Δ*mmp1085* deletions but not in Δ*mmp1086*, Δ*mmp1087*, or Δ*mmp1086* Δ*mmp1087* deletion strains (Fig. 6). Tryptic glycopeptides from the Δ*mmp1084* archaellin were predominantly modified with the tetrameric glycan designated 203Da-258Da-257Da-217Da (Fig. 6b) corresponding to the loss of the threonine moiety from the third sugar in the wild-type glycan (Fig. 6a). A minor trimeric glycan species, 203Da-258Da-257Da, was also observed in the Δ*mmp1084* mutant strain. Archaellins from the Δ*mmp1085* mutant were modified with a different tetrameric glycan 203Da-258Da-358Da-203Da, indicating that the methyl group on the fourth sugar was absent (Fig. 6c). Archaella from the Δ*mmp1086*, Δ*mmp1087*, and Δ*mmp1086* Δ*mmp1087* deletion strains were modified with the wild-type glycan, 203Da-258Da-358Da-217Da (Fig. 6d).

Complementation of *mmp1084* and *mmp1085*. To confirm that deletions of *mmp1084* and *mmp1085* are the only contributors to the defects observed by mass spectrometry, Δ*mmp1084* and Δ*mmp1085* mutants were complemented with plasmids bearing the pertinent wild-type copy of the deleted gene under an inducible *nif* promoter, which is induced under ammonia-limiting conditions (as when L-alanine, for example, is used as the sole nitrogen source) while repressed in the presence of ammonia (37, 47). When cells were grown with alanine as the sole nitrogen source the apparent molecular mass of FlaB2 was restored to wild-type size in both *mmp1084* and *mmp1085* complementations (Fig. 7), indicating that the molecular mass loss in FlaB2 in each mutant was caused by the in-frame deletion of that single gene. When ammonia was used as the nitrogen source, the apparent mass of FlaB2 in both complementations was also restored. This effect has been previously observed in the complementation of certain enzymes involved in the N-linked glycosylation process, likely due to a low basal level of expression from the *nif* promoter even in the presence of ammonia (17). This low level of expression may be sufficient to provide enough product to still complement the deletion.

DISCUSSION

Using archaellins and S-layer proteins as the reporter proteins, significant progress has been made in the study of N-glycosylation pathways in several model archaeal organisms, such as *H. volcanii*, *S. acidocaldarius*, *M. voltae*, and *M. maripaludis* (2, 9, 48). In these systems, numerous *agl* genes in the N-glycosylation pathways have been identified encoding the oligosaccharyltransferase *aglB*, glycosyltransferases and enzymes involved in the biosynthesis of component sugars in the N-glycan (2, 9, 12, 14, 19, 20, 22–26, 33, 34, 49–53). The viability of an *aglB* deletion mutant shows that the N-glycosylation of proteins is not essential for *M. voltae* (14), *M. maripaludis* (17), or *H. volcanii* (22). However, disruption of the system leads to defects in archaellation in both methanogens (14, 17) and *H. volcanii* (29) and the stability of the S layer in the

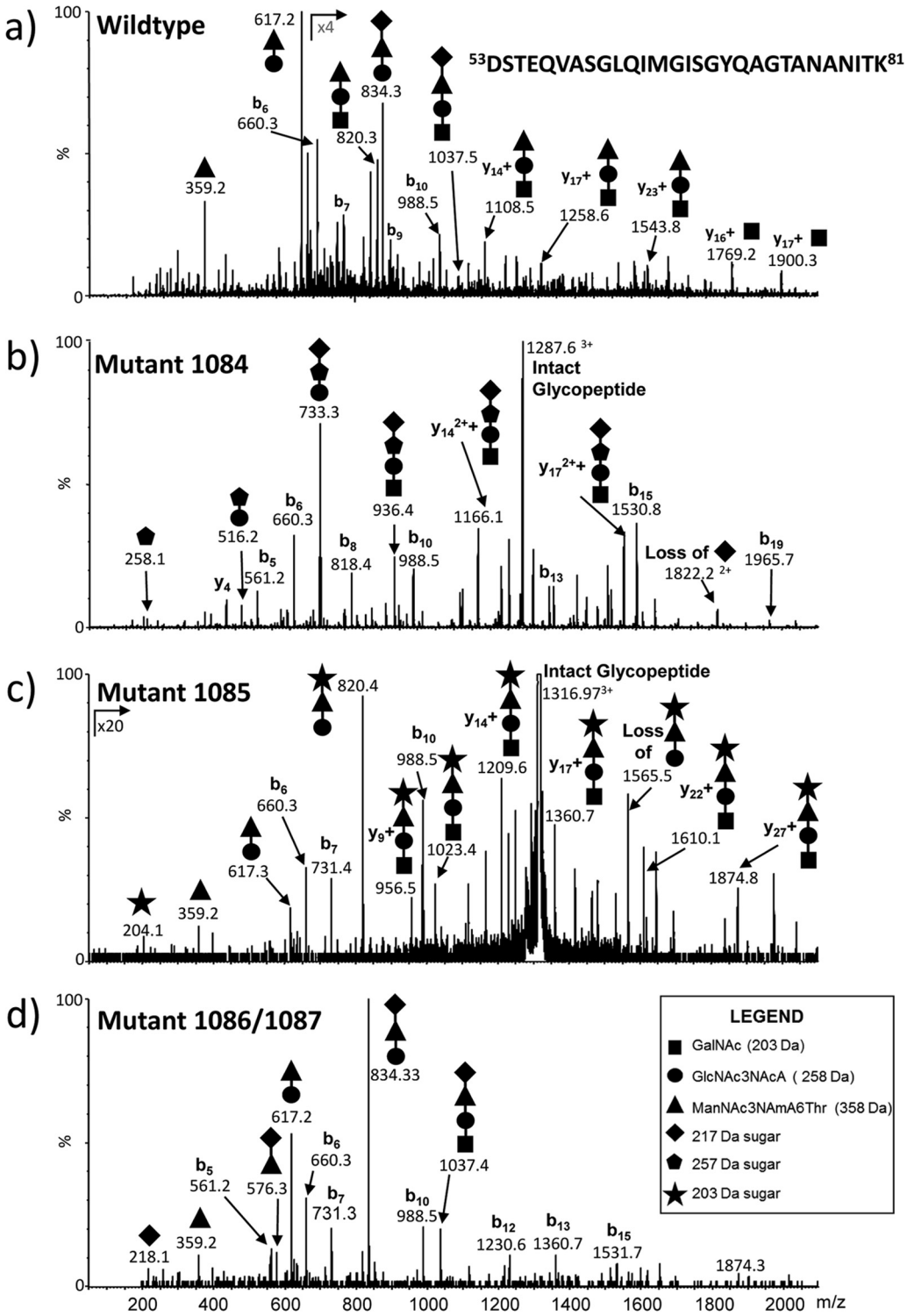

FIG 6 Nano-LC-MS/MS analysis of the FlaB2 tryptic glycopeptide, T^{53-81}, from the wild-type and mutant strains investigated in the present study. The FlaB2 tryptic peptide T^{53-81} contains one site of N-glycosylation (DSTEQVASGLQIMGISGYQAGTANANITK). The major carbohydrate oxonium ions are identified in the MS/MS spectra using symbols to indicate the sugar residues present as follows: ■, GalNAc; ●, GlcNAc3NAcA; ▲, ManNAc3NAmA6Thr; solid pentagon, ManNAc3NAmA; ◆, (5*S*)-2- acetamido-2,4-dideoxy-5-*O*-methyl-α-L-erythro-hexos-5-ulo-1,5-pyranose (217-Da sugar); and ★, previous sugar lacking a methyl group (203-Da sugar). The b and y ions arising from fragmentation of the peptide bonds are also shown. (a) The structure of tetrameric wild-type glycan (MS/MS spectrum acquired in an earlier study) has been described previously (16). (b) The glycopeptide from the Δ*mmp1084* mutant strain was modified with a tetrasaccharide lacking the threonine modification that is linked to the third sugar residue of the wild-type glycan. (c) The same tryptic peptide from the Δ*mmp1085* mutant was also modified with a tetrasaccharide but one whose terminal sugar is 14 Da less than that in the wild-type glycan. (d) The equivalent glycopeptide from the Δ*mmp1086* Δ*mmp1087* double mutant was modified with wild-type glycan.

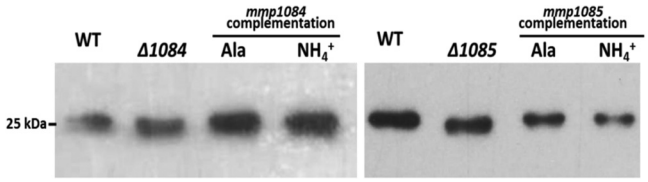

FIG 7 Western blot analysis of complementation experiments with Δ*mmp1084* and Δ*mmp1085* mutants. A Western blot of whole-cell lysates from Δ*mmp1084* and Δ*mmp1085* mutants harboring respective complementation vectors cultured in nitrogen-free medium supplemented by either alanine (Ala, promoter on) or ammonia (NH_4^+, promoter off) was probed with anti-FlaB2 antibodies. Wild-type (WT) and deletion mutants were used as controls to show the apparent molecular masses of FlaB2 attached to wild-type and truncated N-glycans.

halophile, which results in growth defects at elevated NaCl concentrations (23). In contrast to the situation in these euryarchaeotes, an *aglB* deletion could not be created in the crenarchaeote *S. acidocaldarius*, indicating an essential role for this protein modification in this organism (9). Interference with N-glycosylation to the point where truncated glycans are still attached in *S. acidocaldarius* leads to both nonarchaellation and growth defects at higher salinities (9, 19, 20).

In *M. maripaludis*, genes encoding the oligosaccharyltransferase and several glycosyltransferases were identified (17), and it was shown that interference in the N-glycosylation pathway resulted in adverse effects on archaella formation and function but not on pilus assembly (14, 17). Recently, proteins involved in the biosynthesis of the glycan component sugars have been identified (33, 34, 54, 55). We expand here on these findings and identify two genes whose products are involved in the synthesis of the final two sugars of the glycan: *mmp1084* is involved in the threonine modification of the third sugar, and *mmp1085* encodes the methyltransferase required for the biosynthesis of the terminal sugar.

In *M. maripaludis* S2, an in-frame deletion of *mmp1084* resulted in a truncated N-glycan which was shown by mass spectrometry analyses to be missing only the threonine modification of the third sugar. Possible roles for MMP1084 include involvement in either the biosynthesis of the threonine or the transfer of this amino acid onto the lipid-linked tetrasaccharide. *In silico* analysis shows that MMP1084 belongs to the Gn_AT_II superfamily and has sequence similarity to the putative asparagine synthase AsnB in many sequenced genomes. However, in the genome of *M. maripaludis* S2, *mmp0918* has been annotated as *asnB*, and it is located with adjacent genes whose products are annotated to be involved in amino acid metabolism, suggesting that *mmp0918* is more likely to encode the true asparagine synthase. Genes homologous to *mmp0918* are found in a conserved genomic island in numerous members of the Methanococcales, including in the sequenced genomes of four other *M. maripaludis* strains (C5, C6, C7, and X1), indicating that *mmp0918* is a housekeeping gene. On the other hand, homologues of *mmp1084* are rare and are not found in other sequenced *M. maripaludis* strains, although they can be identified in *M. voltae* strains PS and A3. In keeping with the assigned function of MMP1084, the structure of the N-linked glycan attached to *M. voltae* PS archaellins and S layer protein includes a threonine attached to the C-6 of a mannuronic acid as the third sugar (13). The lack of an *mmp1084* homolog in other sequenced *M. maripaludis* strains suggest that these strains do not have a glycan carrying a threonine-modified mannuronic acid. In addition, the growth of Δ*mmp1084* cells is not impaired in nitrogen-free medium supplemented with NH_4Cl as the sole nitrogen source, which means the *de novo* synthesis pathway of amino acids is not affected in the mutant. This would seem to rule out a role for MMP1084 in threonine biosynthesis.

Further analysis of MMP1084 supports a role for this protein in the transfer of threonine to the third sugar. MMP1084 has 22% identity and 39% similarity with the product of *orf3* in the K40 gene cluster in *E. coli* O8:K40, a rare example of an identified gene that is responsible for transferring an amino acid, serine, onto the carboxyl group at the C-6 position of a hexuronic acid (56). Like MMP1084, ORF3 also possesses a conserved domain found in members of the Gn_AT_II superfamily. In *E. coli* O8:K40, the K40 CPS repeat unit is a trisaccharide (6GlcNAc-4-β-GlcA6Ser-4-α-GlcNAc) in which the serine is linked to the second sugar, and deletion of *orf3* does not block the addition of the last sugar onto the CPS repeat unit. Similarly, in *M. maripaludis*, in-frame deletion of *mmp1084* leads to a truncated N-glycan that is only missing the threonine linked to the third sugar residue but retains the full tetrasaccharide backbone. These results suggest that the amino acid modification may be the last step in the glycan synthesis pathway, occurring only after the fourth sugar is added to the growing glycan. This would be consistent with previous observations on the glycan of an *aglL* (encoding the fourth glycosyltransferase) mutant in which not only is the last sugar missing but also the threonine of the third sugar (17). Sequence analysis of MMP1084 does not show any evidence of a signal peptide or transmembrane domain, and it is predicted to be a cytoplasmic protein. It may function in the cytoplasm after the action of the glycosyltransferase for the last sugar but before the flipping of the lipid-linked N-glycan precursor to the external face of the cytoplasmic membrane. However, it is also possible that MMP1084 could be membrane associated if it interacts with integral membrane proteins, perhaps in a complex devoted to glycan formation.

A mutant carrying an in-frame deletion of *mmp1085* had a truncated N-glycan lacking only the methyl group O-linked to the C5 hydroxyl group on the unique fourth sugar residue. This is consistent with the identification of a conserved domain in MMP1085 associated with SAM-dependent methyltransferases and with BLAST search results that show significant sequence similarity of MMP1085 to SAM-dependent methyltransferases. O-methylation is a relatively common glycan modification and has been studied in another archaeal N-glycan attached to both S-layer protein and archaellins in *H. volcanii*. In *H. volcanii*, the C-4 position in the hexuronic acid fourth sugar residue in the pentasaccharide N-glycan is methylated through the action of AglP, also identified as a SAM-dependent methyltransferase (49). Several bacterial O-antigenic polysaccharides also contain a variety of different sugars with O-methylation modifications at various positions (C-2, C-3, C-4, or C-6) catalyzed by SAM-dependent methyltransferases (57–60). However, little sequence similarity is shared between the bacterial methyltransferases and the archaeal AglP and MMP1085. Thus, *mmp1085* is the first gene identified to be involved in the biosynthesis pathway of the unique terminal sugar, (5*S*)-2-acetamido-2,4-dideoxy-5-*O*-methyl-α-L-erythro-hexos-5-ulo-1,5-pyranose. Since deletion of *mmp1085* resulted in the loss of the methyl group of the terminal sugar, we assume that the methyl group is added onto the second hydroxyl group on the C-5 after the formation of the dialdose.

The products of the last two genes in the operon, *mmp1086* and *mmp1087*, did not have any demonstrable effect on the N-linked

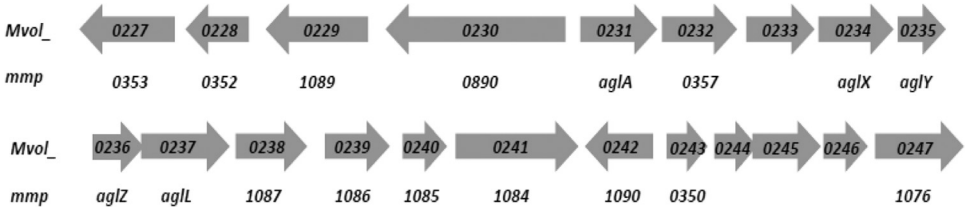

FIG 8 Gene locus from *Mvol_0225* to *Mvol_0247* in *M. voltae* A3. Many homologues of genes involved in the N-glycosylation pathway in *M. maripaludis* could be found in this gene island, including *mmp1084* to *mmp1087*, *aglA*, *aglL*, *aglXYZ*, *mmp1076*, and *mmp1090*. The *mmp* numbers or *agl* gene designations under the *M. voltae* genes indicate the respective homologues in *M. maripaludis* S2. No homologues of *Mvol_0233* or *Mvol_0244* to *Mvol_0246* could be found in the *M. maripaludis* S2 genome.

glycan. There was no detectable difference in the size of FlaB2 in Δ*mmp1086* and Δ*mmp1087* mutants compared to the wild-type version in Western blots, and mass spectrometry analysis of the purified archaella showed that a wild-type glycan was attached. Since MMP1086 and MMP1087 share 56% amino acid sequence identity and both contain an N-terminal radical-SAM domain and a C-terminal DUF4008 domain, it was considered that a loss of one of the two proteins might be compensated by the other. However, Western blot and mass spectrometry analyses of the Δ*mmp1086* Δ*mmp1087* double mutant also showed wild-type patterns. These data indicate that neither *mmp1086* nor *mmp1087* are involved in the N-glycosylation system under normal growth conditions. However, since the roles in N-glycosylation of the neighboring genes *aglY* and *aglZ* were not evident under normal growth conditions but only when the cells were ammonia limited (34), the Δ*mmp1086* Δ*mmp1087* mutant was also examined by Western blotting after growth under these conditions. However, again, only a wild-type size of FlaB2 was observed. These findings do not rule out the possibility that *mmp1086* and *mmp1087* might be involved in the N-glycosylation pathway under different growth conditions.

Recent analysis of sequenced archaeal genomes indicated that in many cases, especially among haloarchaea, there is a clustering of genes putatively involved in N-glycosylation in the immediate vicinity of *aglB* (8). However, this is not true of most methanogens, including *M. voltae* and *M. maripaludis* or the crenarchaeon *S. acidocaldarius* (8). Nonetheless, gene clusters involved in N-glycosylation not linked to *aglB* also exist. In *M. maripaludis*, a group of genes involved in the assembly or biosynthesis of the archaellin N-glycan are located in the genomic region from *mmp1076* to *mmp1094* (Fig. 1). Included in this region are the three glycosyltransferase genes *aglO* (*mmp1079*), *aglA* (*mmp1080*), and *algL* (*mmp1088*) (17), *mmp1076* (which encodes GlcN-1-P uridylyltransferase/acetyltransferase), *mmp1077* (which encodes a phosphomutase that converts α-D-glucosamine-6-phosphate to α-D-glucosamine-1-phosphate) (33), *mmp1090* (which encodes a UDP-Gal/GalNAc 4-epimerase [Ding et al., unpublished]), *aglXYZ* (*mmp1081-mmp1083*; responsible for the acetamidino modification of the third sugar) (34) and, in the present study, *mmp1084* and *mmp1085*. In addition to this genetic locus, another cluster of genes (*mmp0350* to *mmp0359*) has been identified and shown by deletion analysis and mass spectrometry to include genes that encode enzymes involved in biosynthesis of the second and third sugars of the glycan (54; S. Siu, K. F. Jarrell, S. M. Logan, and J. Kelly, unpublished data). Interestingly, in a related mesophilic methanococcus, *M. voltae* A3, homologues of both gene clusters are all located in a single large gene locus (from *Mvol_0230* to *Mvol_0247*) (Fig. 8).

In the present study, we examined four genes found in a single operon—*mmp1084*, *mmp1085*, *mmp1086*, and *mmp1087*—for their potential role in the N-linked glycosylation process in *M. maripaludis* S2. We were able to assign roles to two gene products. MMP1084 is the putative threonine transferase for the third sugar which functions after AglL has added the fourth sugar to the lipid carrier, and MMP1085 is the methyltransferase which transfers the methyl group to the second hydroxyl group on C-5 position in the terminal sugar. No role for MMP1086 or MMP1087 in N-glycan biosynthesis could be identified. Because of their demon-

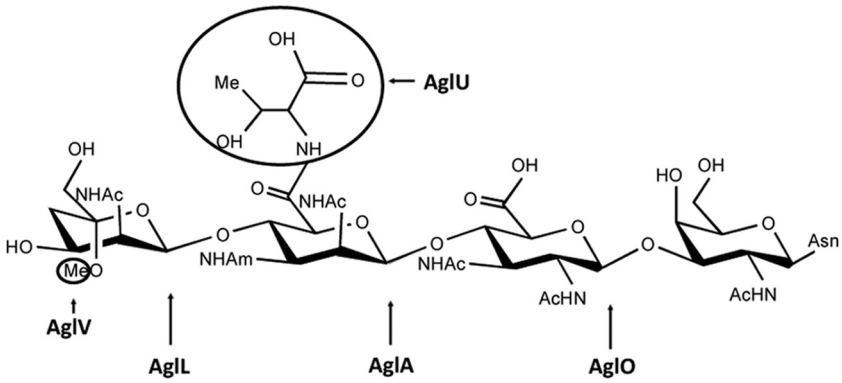

FIG 9 Summary of the function of AglU and AglV in the N-glycosylation pathway in *M. maripaludis*. MMP1084 is designated as AglU, and MMP1085 as AglV. AglU is the putative threonine transferase for the threonine modification of the third sugar. AglV is the putative methyltransferase which transfers the methyl group to the second hydroxyl group on C-5 position in the terminal sugar. The previously identified oligosaccharyltransferase AglB and glycosyltransferases AglA, AglO and AglL are also marked in this model (17).

strated involvement in the N-glycosylation pathway, we designate *mmp1084* as *aglU* and *mmp1085* as *aglV*, in keeping with the nomenclature scheme for genes involved in N-linked glycosylation in *Archaea* first proposed by Chaban et al. (14) (Fig. 9).

ACKNOWLEDGMENTS

This study was funded by the National Research Council of Canada (S.M.L. and J.K.) and by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (to K.F.J.). Y.D. is sponsored by the China Scholarship Council (grant 2010622028).

REFERENCES

1. **Calo D, Kaminski L, Eichler J.** 2010. Protein glycosylation in *Archaea*: sweet and extreme. Glycobiology **20:**1065–1076.
2. **Eichler J.** 2013. Extreme sweetness:protein glycosylation in *Archaea*. Nat. Rev. Microbiol. **11:**151–156.
3. **Schwarz F, Aebi M.** 2011. Mechanisms and principles of N-linked protein glycosylation. Curr. Opin. Struct. Biol. **21:**576–582.
4. **Jarrell KF, Jones GM, Kandiba L, Nair DB, Eichler J.** 2010. S-layer glycoproteins and flagellins: reporters of archaeal posttranslational modifications. Archaea doi:10.1155/2010/612948.
5. **Larkin A, Imperiali B.** 2011. The expanding horizons of asparagine-linked glycosylation. Biochemistry **50:**4411–4426.
6. **Nothaft H, Liu X, McNally DJ, Szymanski CM.** 2010. N-linked protein glycosylation in a bacterial system. Methods Mol. Biol. **600:**227–243.
7. **Magidovich H, Eichler J.** 2009. Glycosyltransferases and oligosaccharyltransferases in Archaea: putative components of the N-glycosylation pathway in the third domain of life. FEMS Microbiol. Lett. **300:**122–130.
8. **Kaminski L, Lurie-Weinberger MN, Allers T, Gophna U, Eichler J.** 2013. Phylogenetic- and genome-derived insight into the evolution of N-glycosylation in *Archaea*. Mol. Phylogenet. Evol. **68:**327–329.
9. **Meyer BH, Albers SV.** 2013. Hot and sweet: protein glycosylation in *Crenarchaeota*. Biochem. Soc. Trans. **41:**384–392.
10. **Guan Z, Naparstek S, Kaminski L, Konrad Z, Eichler J.** 2010. Distinct glycan-charged phosphodolichol carriers are required for the assembly of the pentasaccharide N-linked to the *Haloferax volcanii* S-layer glycoprotein. Mol. Microbiol. **78:**1294–1303.
11. **Calo D, Guan Z, Naparstek S, Eichler J.** 2011. Different routes to the same ending: comparing the N-glycosylation processes of *Haloferax volcanii* and *Haloarcula marismortui*, two halophilic archaea from the Dead Sea. Mol. Microbiol. **81:**1166–1177.
12. **Jarrell KF, Jones GM, Nair DB.** 2010. Role of N-linked glycosylation in cell surface structures of *Archaea* with a focus on flagella and S layers. Int. J. Microbiol. doi:10.1155/2010/470138.
13. **Voisin S, Houliston RS, Kelly J, Brisson JR, Watson D, Bardy SL, Jarrell KF, Logan SM.** 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.
14. **Chaban B, Voisin S, Kelly J, Logan SM, Jarrell KF.** 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.
15. **Ng SYM, Wu J, Nair DB, Logan SM, Robotham A, Tessier L, Kelly JF, Uchida K, Aizawa S, Jarrell KF.** 2011. Genetic and mass spectrometry analysis of the unusual type IV-like pili of the archaeon *Methanococcus maripaludis*. J. Bacteriol. **193:**804–814.
16. **Kelly J, Logan SM, Jarrell KF, Vandyke DJ, Vinogradov E.** 2009. A novel N-linked flagellar glycan from *Methanococcus maripaludis*. Carbohydr. Res. **344:**648–653.
17. **Vandyke DJ, Wu J, Logan SM, Kelly JF, Mizuno S, Aizawa SI, Jarrell KF.** 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*. Mol. Microbiol. **72:**633–644.
18. **Peyfoon E, Meyer B, Hitchen PG, Panico M, Morris HR, Haslam SM, Albers S-., Dell A.** 2010. The S-layer glycoprotein of the crenarchaeote *Sulfolobus acidocaldarius* is glycosylated at multiple sites with the chitobiose-linked N-glycans. Archaea doi:10.1155/2010/754101.
19. **Meyer BH, Peyfoon E, Dietrich C, Hitchen P, Dell A, Albers SV.** 2013. Agl16, a thermophilic glycosyltransferase, mediating the last step of the N-glycan biosynthesis in the thermoacidophilic crenarchaeon *Sulfolobus acidocaldarius*. J. Bacteriol. **195:**2177–2186.
20. **Meyer BH, Zolghadr B, Peyfoon E, Pabst M, Panico M, Morris HR, Haslam SM, Messner P, Schäffer C, Dell A, Albers SV.** 2011. Sulfoquinovose synthase: an important enzyme in the N-glycosylation pathway of *Sulfolobus acidocaldarius*. Mol. Microbiol. **82:**1150–1163.
21. **Eichler J, Maupin-Furlow J.** 2012. Post-translation modification in *Archaea*: lessons from *Haloferax volcanii* and other haloarchaea. FEMS Microbiol. Rev. doi:10.1111/1574-6976.12012.
22. **Abu-Qarn M, Eichler J.** 2006. Protein N-glycosylation in *Archaea*: defining *Haloferax volcanii* genes involved in S-layer glycoprotein glycosylation. Mol. Microbiol. **61:**511–525.
23. **Abu-Qarn M, Yurist-Doutsch S, Giordano A, Trauner A, Morris HR, Hitchen P, Medalia O, Dell A, Eichler J.** 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.
24. **Abu-Qarn M, Giordano A, Battaglia F, Trauner A, Hitchen PG, Morris HR, Dell A, Eichler J.** 2008. Identification of AglE, a second glycosyltransferase involved in N-glycosylation of the *Haloferax volcanii* S-layer glycoprotein. J. Bacteriol. **190:**3140–3146.
25. **Kaminski L, Abu-Qarn M, Guan Z, Naparstek S, Ventura VV, Raetz CR, Hitchen PG, Dell A, Eichler J.** 2010. AglJ adds the first sugar of the N-linked pentasaccharide decorating the *Haloferax volcanii* S-layer glycoprotein. J. Bacteriol. **192:**5572–5579.
26. **Kaminski L, Guan Z, Abu-Qarn M, Konrad Z, Eichler J.** 2012. AglR is required for addition of the final mannose residue of the N-linked glycan decorating the *Haloferax volcanii* S-layer glycoprotein. Biochim. Biophys. Acta **1820:**1664–1670.
27. **Jarrell KF, Walters AD, Bochiwal C, Borgia JM, Dickinson T, Chong JPJ.** 2011. Major players on the microbial stage: why archaea are important. Microbiology **157:**919–936.
28. **Jarrell KF, Albers SV.** 2012. The archaellum: an old motility structure with a new name. Trends Microbiol. **20:**307–312.
29. **Tripepi M, You J, Temel S, Ö. Önder Brisson D, Pohlschröder M.** 2012. N-glycosylation of *Haloferax volcanii* flagellins requires known Agl proteins and is essential for biosynthesis of stable flagella. J. Bacteriol. **194:**4876–4887.
30. **Zahringer U, Moll H, Hettmann T, Knirel YA, Schafer G.** 2000. Cytochrome b558/566 from the archaeon *Sulfolobus acidocaldarius* has a unique Asn-linked highly branched hexasaccharide chain containing 6-sulfoquinovose. Eur. J. Biochem. **267:**4144–4149.
31. **Kandiba L, Aitio O, Helin J, Guan Z, Permi P, Bamford DH, Eichler J, Roine E.** 2012. Diversity in prokaryotic glycosylation: an archaeal-derived N-linked glycan contains legionaminic acid. Mol. Microbiol. **84:**578–593.
32. **Chaban B, Ng SY, Kanbe M, Saltzman I, Nimmo G, Aizawa SI, Jarrell KF.** 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.
33. **Namboori SC, Graham DE.** 2008. Acetamido sugar biosynthesis in the *Euryarchaea*. J. Bacteriol. **190:**2987–2996.
34. **Jones GM, Wu J, Ding Y, Uchida K, Aizawa S, Robotham A, Logan SM, Kelly J, Jarrell KF.** 2012. Identification of genes involved in the acetamidino group modification of the flagellin N-linked glycan of *Methanococcus maripaludis*. J. Bacteriol. **194:**2693–2702.
35. **Moore BC, Leigh JA.** 2005. Markerless mutagenesis in *Methanococcus maripaludis* demonstrates roles for alanine dehydrogenase, alanine racemase, and alanine permease. J. Bacteriol. **187:**972–979.
36. **Balch WE, Fox GE, Magrum LJ, Woese CR, Wolfe RS.** 1979. Methanogens: reevaluation of a unique biological group. Microbiol. Rev. **43:**260–296.
37. **Lie TJ, Wood GE, Leigh JA.** 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.
38. **Tumbula DL, Makula RA, Whitman WB.** 1994. Transformation of *Methanococcus maripaludis* and identification of a PstI-like restriction system. FEMS Microbiol. Lett. **121:**309–314.
39. **Laemmli UK.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:**680–685.
40. **Towbin H, Staehelin T, Gordon J.** 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. U. S. A. **76:**4350–4354.
41. **Bardy SL, Mori T, Komoriya K, Aizawa S, Jarrell KF.** 2002. Identification and localization of flagellins FlaA and FlaB3 within flagella of *Methanococcus voltae*. J. Bacteriol. **184:**5223–5233.

42. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
43. **Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R, Lopez R.** 2005. InterProScan: protein domains identifier. Nucleic Acids Res. **33:**W116–W120.
44. **Bagos PG, Tsirigos KD, Plessas SK, Liakopoulos TD, Hamodrakas SJ.** 2009. Prediction of signal peptides in archaea. Protein Eng. Des. Sel. **22:**27–35.
45. **Yu NY, Laird MR, Spencer C, Brinkman FS.** 2011. PSORTdb: an expanded, auto-updated, user-friendly protein subcellular localization database for *Bacteria* and *Archaea*. Nucleic Acids Res. **39:**D241–D244.
46. **Jones WJ, Paynter MJB, Gupta R.** 1983. Characterization of *Methanococcus maripaludis* sp. nov., a new methanogen isolated from salt marsh sediment. Arch. Microbiol. **135:**91–97.
47. **Lie TJ, Leigh JA.** 2002. Regulatory response of *Methanococcus maripaludis* to alanine, an intermediate nitrogen source. J. Bacteriol. **184:**5301–5306.
48. **Kelly JF, Jarrell KF.** 2012. Protein glycosylation in the third domain of life: the *Archaea*, p 108–126. *In* Reid CW, Twine SM, Reid AN (ed), Bacterial glycomics. Caister Academic Press, Cumberland, RI.
49. **Magidovich H, Yurist-Doutsch S, Konrad Z, Ventura VV, Dell A, Hitchen PG, Eichler J.** 2010. AglP is an *S*-adenosyl-L-methionine-dependent methyltransferase that participates in the N-glycosylation pathway in *Haloferax volcanii*. Mol. Microbiol. **76:**190–199.
50. **Yurist-Doutsch S, Magidovich H, Ventura VV, Hitchen PG, Dell A, Eichler J.** 2010. N-glycosylation in *Archaea*: on the coordinated actions of *Haloferax volcanii* AglF and AglM. Mol. Microbiol. **75:**1047–1058.
51. **Yurist-Doutsch S, Abu-Qarn M, Battaglia F, Morris HR, Hitchen PG, Dell A, Eichler J.** 2008. *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.
52. **Chaban B, Logan SM, Kelly JF, Jarrell KF.** 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.
53. **Shams-Eldin H, Chaban B, Niehus S, Schwarz RT, Jarrell KF.** 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.
54. **VanDyke DJ, Wu J, Ng SY, Kanbe M, Chaban B, Aizawa SI, Jarrell KF.** 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.
55. **Namboori SC, Graham DE.** 2008. Enzymatic analysis of uridine diphosphate *N*-acetyl-D-glucosamine. Anal. Biochem. **301:**94–100.
56. **Amor PA, Yethon JA, Monteiro MA, Whitfield C.** 1999. Assembly of the K40 antigen in *Escherichia coli*: identification of a novel enzyme responsible for addition of L-serine residues to the glycan backbone and its requirement for K40 polymerization. J. Bacteriol. **181:**772–780.
57. **Stadthagen G, Sambou T, Guerin M, Barilone N, Boudou F, Korduláková J, Charles P, Alzari PM, Lemassu A, Daffé M, Puzo G, Gicquel B, Rivière M, Jackson M.** 2007. Genetic basis for the biosynthesis of methylglucose lipopolysaccharides in *Mycobacterium tuberculosis*. J. Biol. Chem. **282:**27270–27276.
58. **Ojeda KJ, Box JM, Noel KD.** 2010. Genetic basis for *Rhizobium etli* CE3 O-antigen O-methylated residues that vary according to growth conditions. J. Bacteriol. **192:**679–690.
59. **Nakata N, Fujiwara N, Naka T, Yano I, Kobayashi K, Maeda S.** 2008. Identification and characterization of two novel methyltransferase genes that determine the serotype 12-specific structure of glycopeptidolipids of *Mycobacterium intracellulare*. J. Bacteriol. **190:**1064–1071.
60. **Duelli DM, Tobin A, Box JM, Kolli VS, Carlson RW, Noel KD.** 2001. Genetic locus required for antigenic maturation of *Rhizobium etli* CE3 lipopolysaccharide. J. Bacteriol. **183:**6054–6064.