# Glycosylate and move! The glycosyltransferase Maf is involved in bacterial flagella formation

Gerlind Sulzenbacher, Veronique Roig-Zamboni, Régine Lebrun, Yann Guérardel, Dorothée Murat, Pascal Mansuelle, Nao Yamakawa, Xin-Xin Qian, Renaud Vincentelli, Yves Bourne, et al.

# Glycosylate and move! The glycosyltransferase Maf is involved in bacterial flagella formation

Gerlind Sulzenbacher[1], Véronique Roig-Zamboni[1], Régine Lebrun[2], Yann Guérardel[3], Dorothée Murat[4,5,6] Pascal Mansuelle[2], Nao Yamakawa[3], Xin-Xin Quian[4,5] Renaud Vincentelli[1], Yves Bourne[1], Long-Fei Wu[4,5] and François Alberto[4,5*]

1. Aix Marseille Univ, CNRS, AFMB UMR7257, Marseille 13288, France

2. Plate-forme Protéomique, Institut de Microbiologie de la Méditerranée, FR3479 Aix-Marseille Université and Centre National de la Recherche Scientifique, 13402, Marseille, France

3. Unité de Glycobiologie Structurale et Fonctionnelle, UMR 8576 Université de Lille and Centre National de la Recherche Scientifique, 59000, Lille, France

4. Aix Marseille Univ, CNRS, LCB UMR7283, Marseille 13402, France

5. International Associated Laboratory of Evolution and Development of Magnetotactic Organisms (LIA-MagMC), Centre National de la Recherche Scientifique, 13402, Marseille, France

6. Present address : Aix Marseille Univ, CNRS, IGS UMR7256, Marseille 13288, France

Running title: Role of Maf flagellin glycosyltransferase

*Correspondence to Francois Alberto, falberto@imm.cnrs.fr, Laboratoire de Chimie Bactérienne, UMR7283 AMU CNRS, 31 chemin Joseph Aiguier, 13402 Marseille, France

Telephone number: +33 491 16 41 21

Fax number: +33 491 71 89 14

**Originality-Significance Statement**

Flagellin glycosylation with a nonulosonic sugar is required for the formation of the flagellum on the cell surface of *Magnetospirillum magneticum* AMB-1. This work describes the structural and functional characterization of the bacterial glycosyltranferase Maf. Because flagellin glycosylation is primordial for the motility and the virulence of many pathogens, and since many nonulosonic acid derivatives and the glycosyltransferases responsible for their transfer onto flagellins are unique to bacteria, we believe that these glycosyltransferases represent attractive targets for the development of new antimicrobial agents.


## Abstract

The flagella of various Gram-negative bacteria are decorated with diverse glycan structures, amongst them nonulosonic acids related to the sialic acid family. Although nonulosonic sugar biosynthesis pathways have been dissected in various pathogens, the enzymes transferring the sugars onto flagellin are still poorly characterized. The deletion of genes coding for Motility associated factors (Mafs) found in many pathogenic strains systematically gives rise to non-flagellated bacteria lacking specific nonulosonic sugars on the flagellins, therefore relating Maf function to flagellin glycosylation and bacterial motility. We investigated the role of Maf from our model organism, *Magnetospirillum magneticum* AMB-1, in the glycosylation and formation of the flagellum. Deletion of the gene *amb0685* coding for Maf produced a non-flagellated bacterium where the flagellin was still produced but no longer glycosylated. Our X-ray structure analysis revealed that the central domain of Maf exhibits similarity to sialyltransferases from *Campylobacter jejuni*. Glycan analysis suggested that the nonulosonic carbohydrate structure transferred is pseudaminic acid or a very close derivative. This work describes the importance of glycosylation in the formation of the bacterial flagellum and

provides the first structural model for a member of a new bacterial glycosyltransferase family involved in nonulosonic acids transfer onto flagellins.

## Introduction

Over the last few decades bacterial glycosylation has received considerable attention, since it had been well established that bacteria, like eukaryotes, possess both *N*-linked and *O*-linked glycosylation pathways (Nothaft *et al.,* 2010). The *O*-glycosylation pathway has mainly been described for flagellin (the individual subunits of the flagellar filament) and pilin monomers (Logan, 2006; Ng *et al.,* 2006; Nothaft *et al.,* 2010). Bacterial flagellin glycosylation is now well-documented for pathogenic bacteria, including *Campylobacter jejuni*, *Helicobacter pylori*, *Pseudomonas aeruginosa*, *Legionella pneumophila*, *Aeromonas caviae*, *Clostridium difficile*, or *Listeria monocytogenes* (Schirm *et al.,* 2003; Schirm *et al.,* 2004b; Schirm *et al.,* 2004a; Logan, 2006; Twine *et al.,* 2009; Hopf *et al.,* 2011; Parker *et al.,* 2012; Iwashkiw *et al.,* 2013).-Flagellin O-glycosylation takes place on serine and/or threonine residues, often near a hydrophobic region, but so far a general consensus amino acid sequence has not been identified. The glycosylated amino acids reside in the central and variable region of the flagellin monomer, the region that is located at the surface of the flagellar filament (Bubendorfer *et al.,* 2013; Poweleit *et al.,* 2016). Bacterial flagellin glycosylation features a remarkable variety in the number of acceptor sites, the type and size of glycan structures and in some cases even strain-to-strain glycan heterogeneity (Thibault *et al.,* 2001; Schirm *et al.,* 2005).

Among the flagellin sugar modifications, pseudaminic acid (Pse5Ac7Ac) has been identified as a specific bacterial sialic acid related nonulosonic acid sugar (Zunk *et al.*, 2014), essential for flagellar assembly in *Campylobacter* spp., *A. caviae*, or *H. pylori* (Goon *et al.,*

2003). Since motility is required for colonization and the glycosylated extremities of the flagellins mediate host interactions (Zunk *et al.*, 2014), the presence of the Pse5Ac7Ac has been related to the virulence capacities of these bacteria. The complete biosynthesis of Pse5Ac7Ac has been characterized in *C. jejuni* and *H. pylori* (McNally *et al.*, 2006; Schoenhofen *et al.*, 2006). Within these two bacteria CMP-activated Pse5Ac7Ac is generated from UDP-N-acetylglucosamine in a six-step pathway involving six different enzymes (Morrison *et al.*, 2014).

Despite the thorough characterizations of the biosynthetic pathway of Pse5Ac7Ac, very little is known about the glycosyltransferases that transfer the sugar onto the flagellin. When the genome of *C. jejuni* was published, authors described a number of putative flagella-related genes similar to genes of *Escherichia coli* and other bacteria (Parkhill *et al.*, 2000). The genes involved in flagellin glycosylation are part of a large cluster containing genes involved in sugar biosynthesis and seven closely related genes of unknown function. These genes with unknown function were termed as the motility accessory factor (*maf*) family of flagellin associated genes (Karlyshev *et al.*, 2002). In *H. pylori* two of these genes code for HP0114 and HP0465, respectively (Parkhill *et al.*, 2000). Inactivation of the gene coding for HP0114 displayed a non-motile phenotype. As HP0114 does not appear to be involved in the biosynthesis of CMP-Pse5Ac7Ac but clearly plays a role in flagellar assembly, it was suggested that HP0114 is involved in the transfer of Pse5Ac7Ac from CMP-Pse5Ac7Ac to the flagellin monomer (Schirm *et al.*, 2003). Following deletion of the *maf* genes found in *Aeromonas hydrophila* AH-3, *maf1 and maf2,* and *A. caviae*, *maf1*, the flagellin monomers were still produced within the cell, along with CMP-Pse5Ac7Ac (Canals *et al.*, 2006; Canals *et al.*, 2007; Parker *et al.*, 2012), but were no longer glycosylated, suggesting that Maf proteins are glycosyltransferases targeting flagellins (Parker *et al.*, 2014). Furthermore, the authors of this work showed that the Maf-dependent glycosylation takes place before

interaction of the unfolded flagellin with its specific chaperone and subsequent export *via* the T3SS secretion system (Parker *et al.*, 2014).

We have performed genomic, genetic, biochemical, proteomic, and structural analyses to understand the structure and function of Maf in our magnetotactic model strain *Magnetospirillum magneticum* AMB-1 (AMB-1). We found that deletion of the *maf* gene affected glycosylation and assembly of Fla, the sole AMB-1 flagellin, into functional flagellar filaments. Glycan and proteomic analyses showed that the flagellin monomers are decorated with Pse5Ac7Ac or a close derivative - but only in the presence of Maf. Finally, we solved the 3D structure of Maf, which provides a structural template for the Maf family and the framework for the development of novel antibiotic agents specifically targeting this family.

## Results

### The AMB-1 flagellin is produced as a glycosylated protein.

Shearing-prepared AMB-1 flagellin fractions were resolved on SDS-PAGE gels and Coomassie staining revealed major bands around 70, 55, and 40 kDa (Fig. 1A). Polyclonal antibodies raised against the AMB-1 flagellin Fla, of expected theoretical mass of 54137 Da, (Murat *et al.*, 2015) hybridized with both the 70 and the 55 kDa bands (Fig. 1B). The two bands were analyzed in a bottom-up proteomics workflow and showed the presence of Fla. The coverage obtained for the sequence of the 70 and the 55 kDa bands was 80% and 65%, respectively. Among the other polypeptides identified in the 55 kDa band, many have a cytoplasmic location, thus the presence of Fla in this band is most likely due to cytoplasmic protein contamination. Indeed, immunoblot analysis with the anti-flagellin antibody of a concentrated fraction of AMB-1 culture supernatant revealed only the presence of the 70 kDa

band (Fig. S1A). Analysis of protein glycosylation by the periodate/biotin hydrazine method revealed the presence of a glycoprotein in the 70 kDa band (Fig. 1C).

**Deletion of *maf* prevents the glycosylation of the flagellin of AMB-1.**

Downstream of the *fla* gene coding for the AMB-1 flagellin is located *maf* (*amb0685*), coding for a 75 kDa protein of unknown function. The Maf protein displays sequence identity with hypothetical proteins from other Gram negative bacteria like *Magnetospira* QH2 (33%), *Alteromonas australica* (30%), and numerous *Aeromonas* species: *A. schubertii* (32%), *A. piscicola* (32%), *A. hydrophila* (31%), *A. caviae* (29%), or *C. jejuni* PseE (22%). The homologue from *A. caviae*, Maf1, had been shown to transfer Pse5Ac7Ac onto the flagellins FlaA/B (Parker *et al.*, 2014) and the same role was attributed to PseE in *C. jejuni* (McNally *et al.*, 2006). The central part of all the above homologous proteins bears a MAF_flag10 domain (accession number: pfam01973) of unknown function. HHpred analysis (Meier *et al.*, 2015) pointed to a structural similarity between this central domain of Maf and *C. jejuni* sialyltransferases Cst-I and Cst-II. These enzymes, members of the CAZy (http://www.cazy.org/) glycosyltransferase family GT42 (Lombard *et al.*, 2014), use CMP-N-acetylneuraminic acid (CMP-Neu5Ac) as the universal donor sugar and transfer a sialic acid moiety onto cell surface lipo-oligosaccharides (Chiu *et al.*, 2004a; Chiu *et al.*, 2007). The sequence similarities with *A. caviae* Maf1 and the predicted structural similarity to the well-characterized sialyltransferases Cst-I and Cst-II advocated for a possible involvement of Maf in the transfer of a nonulosonic acid onto Fla.

To test this hypothesis, we prepared a Δ*maf* deletion mutant in which the genetic regions flanking *maf* where cloned into an AMB-1 non-replicative vector (pAK31), where *maf* was replaced by a short linker in order to avoid the deletion from having a polar effect. The Δ*maf* mutant cells were non-motile and electron microscopy confirmed the absence of flagella (Fig.

1E) as compared to the wild-type strain (Fig. 1D). Expression of *maf in trans* in the Δ*maf* mutant restored the wild type phenotype, that is to say, cell motility and the presence of a functional flagellar filament outside the cell (Fig. 1F). To verify whether the deletion of *maf* had a deleterious effect on flagellin production, we hybridized the anti-flagellin antibody against total protein extracts of wild type AMB-1, the Δ*maf* mutant and the complemented Δ*maf* mutant. As expected, the Western blot showed that hybridizations occurred at 70 and 55 kDa for the wild-type total protein extract (Fig. 1G), but only the 55 kDa band could be revealed in total protein extract of the deletion mutant, indicating that the flagellin was still produced in Δ*maf*, but not glycosylated (Fig. 1H). Immunoblotting of the complemented Δ*maf* mutant protein extract revealed the 70 kDa band, indicating that flagellin glycosylation was restored by the complementation (Fig. 1I). Immunoblotting of a concentrated fraction of the supernatant of the Δ*maf* mutant culture revealed no band at all (Fig.S1B), indicating that the non-glycosylated protein is not exported.

**The flagellin of AMB-1 is glycosylated by pseudaminic acid or a close derivative.**

In a first approach, we tried to identify flagellin-associated glycans by purifying potential nonulosonic acids from flagellins. To accomplish that, we released nonulosonic acids from purified flagellin by mild acid hydrolysis. Free potential nonulosonic acids were then derivatized by 1,2-diamino-4,5-methylenedioxybenzene (DMB), a fluorogenic reagent for 2-keto-acids that shows high specificity for all derivatives of nonulosonic acids, and analyzed by reverse phase HPLC coupled to mass spectrometry (LC-$MS^n$) . The analysis of the hydrolysis product of flagellin by LC-MS showed a MS signal at *m/z* 451.25 corresponding to the $[M+H]^+$ adduct of a compound with a molecular mass of 450.25 Da. This value could correspond to the expected mass of DMB-derivatized Pse5Ac7Ac or DMB-derivatized Leg5Ac7Ac (legionaminic acid) that both exhibit a $[M+H]^+$ value at *m/z* 451 (Lewis *et al.*, 2009; Ricaldi *et al.*, 2012). The attribution to Pse5Ac7Ac or Leg5Ac7Ac was further

confirmed by $MS^2$ analysis of the parent ion at *m/z* 451. Indeed, recurrent losses of $H_2O$ and *N*-acetyl groups at *m/z* 433.24 and 415.23, respectively, show similarities with the fragmentation pattern of standard Neu5Ac sialic acid (Fig. 2A and 2B), in accordance with a previous study (Lewis *et al.,* 2009).

A BLAST search (Altschul *et al.,* 1990) using the sequences of the enzymes involved in nonulosonic acid biosynthesis pathways in various bacterial species (*C. jejuni*, *H. pylori*, *or A. caviae)* against the AMB-1 protein database identified hits for all of the enzymes involved in the pseudaminic acid biosynthesis pathway and the corresponding cluster from genes *amb0713* to *amb0718*. Thus, PseB of *H. pylori* is homologous to Amb0716 (77% sequence similarity), PseC to Amb0717 (55%), PseG to Amb0714 (45%), PseH to Amb0715 (43%), PseI to Amb0713 (68%), and PseF to Amb0718 (42%). Transcriptomic analysis showed that all these genes are expressed in the wild-type AMB-1 strain (Fig. S2B). Genetic deletion of *amb0715* led to the same phenotype as the Δ*maf* mutant, that is to say the absence of flagellum and the disappearance of the 70 kDa band as revealed by Western blotting (Fig. S3A and S3C). Complementation of the Δ*amb0715* mutant by a plasmid carrying *amb0715* restored flagellum assembly and the presence of the glycosylated form of Fla (Fig. S3B and S3D). A BLAST search (Altschul *et al.,* 1990) for the biosynthetic pathway of legionaminic acid in the AMB-1 protein database gave no conclusive result. In other bacterial species the number of *maf* homologues is correlated with the number of nonulosonic acids decorating the flagellin (Merino *et al.,* 2014). Amb0685 presents no other homologue in the AMB-1 genome, suggesting that Pse5Ac7Ac (or a close Pse5Ac7Ac derivative) is the only nonulosonic sugar synthesized and transferred onto the flagellin.

Tandem mass spectrometry data from enzymatically digested Fla were analysed searching for Pse5Ac7Ac (316.13 Da) as a variable modification on serine and threonine residues. The sequence of Fla was covered at 80% by unique peptides and four among them displayed one

or two modifications by Pse5Ac7Ac (Fig. S4A). For example, the parent ion at $[M+H]^+$ 3192.624 Da corresponding to the sequence -GLALNAQSASSSQIGSLVQQFNALR-, fitted with a mass accommodating two groups of Pse5Ac7Ac on serine residues (Fig. S4B). MS/MS fragmentation of this modified peptide in the HCD (Higher-energy Collisional Dissociation) cell displayed the nonulosonic acid specific signature as described previously (Thibault *et al.*, 2001), in the lowest m/z range: 299, 281, 257, 239, 221, 180, and 134 (Fig. S4C). Conversely, tandem mass spectrometry analysis of digested flagellin produced by the Δ*maf* mutant did not reveal any Pse5Ac7Ac modification on any peptide (Fig. S4D). Altogether these data strongly suggest that the nonulosonic acid found on the AMB-1 flagellin is Pse5Ac7Ac or a Pse5Ac7Ac derivative.

**Overall structure of full-length Maf.**

In order to gain insight into structure/function relationships, the crystal structure of Maf was determined taking advantage of multiple wavelength anomalous diffraction (MAD) data recorded at the Yb edge and refined at a resolution of 2.3 Å to $R_{work}$ of 18.94% and $R_{free}$ of 22.20% (Table S1). The final model of good stereo-chemical quality encompasses residues 7–664, but two peptide regions, spanning from Ser86 to His89 and from Lys422 to Gly443, respectively, could not be modelled because of lacking electron density.

*In cristallo*, Maf adopts a helical shape and is organized into three domains (Fig. 3). At the N-terminus (Ile7-Ser81 and Ser119-Ile237) a degenerated Rossmann-like domain, consisting of a four-stranded parallel β-sheet, flanked by five α-helices, is preceded by three additional α-helices and a β-hairpin. A structural homology search with the DALI server (Holm *et al.*, 1993) did not identify significant structural resemblance with coordinates from the Protein Data Bank, except for weak similarities limited to the Rossmann-like motifs as found in nucleotide- and dinucleotide-binding enzymes and S-adenosyl-L-methionine (AdoMet)-

dependent methyltransferases. A Gly-rich motif is located in a position structurally corresponding to the Gly-rich motif I in Mtases (Malone *et al.*, 1995), but where class I Mtases present a cavity accommodating the methyl donor, bulky side-chains occupy the equivalent space in Maf.

A long loop with an α-helical insert connects the N-terminal domain to the central MAF_flag10 domain (Thr273-Arg530), displaying a modified glycosyltransferase GT-A topology as found in family GT29 and GT42 sialyltransferases (Chiu *et al.*, 2004b; Chiu *et al.*, 2007). Briefly, a central parallel seven-stranded twisted β-sheet (topology β8, β7, β1, β2, β3, β4, β5) is bordered by four (αA, αB, αC, αD) and three (αE, αF, αG) α-helices on either sides and capped by a last α-helix (αH) at the N-terminal end of the β-strands. Between β–strand β6 and helix αG a long loop, comprising the disordered region Lys422-Gly443, projects away from the α/β domain and packs against the N-terminal domain of a symmetry-related molecule.

A long loop comprising an α-helix connects the α/β domain to the C-terminal domain (Trp546-Asp664), composed of a mostly antiparallel six-helix bundle. One of the α-helices making up this helical bundle originates from the N-terminal part of Maf (Leu92-Asp107). Structural homology searches with DALI (Holm *et al.*, 1993) and PDBeFold (Krissinel *et al.*, 2004), taking into account the connectivity of the peptide chain of the C-terminal domain, pointed to weak structural similarity to the N-terminal domain of a fungal translation initiation factor (PDB 5DBO, Z-score of 8.0) and the *E. coli* histidine kinase NarX sensor domain (PDB 3EZH, Z-score of 7.6). Interestingly, weaker structural similarities (Z-scores in the range of 5 to 6) can be found between the Maf C-terminal domain and flagellin export chaperones and flagellins, like *Aquifex aeolicus* FliS (PDB 1ORJ) and FliC from *Burkholderia pseudomallei* (PDB 4CFI).

In the crystal, Maf appears to form a dimer, generated by a crystallographic two-fold axis, with a buried surface of 8420 $Å^2$ per monomer as calculated by the PISA server (Krissinel *et al.*, 2007) (Fig. S5). However, this is undoubtedly a crystallization artefact, as size-exclusion chromatography coupled with multi-angle laser light scattering (SEC-MALLS) (Fig. S6) clearly showed that Maf is a monomer in solution.

**Comparison of the central MAF_flag10 domain with sialyltransferases and identification of functional residues.**

The closest structural homologue of the central domain of Maf is *C. jejuni* sialyltransferase Cst-II (PDB 1RO7) belonging to CAZy (Lombard *et al.*, 2014) family GT42, with a DALI Z-score of 9.8 and an rmsd of 3.6 Å for 158 equivalent Cα-atoms (this similarities is comparable with that to Cst-I, but for simplicity only Cst-II will be discussed). More distant similarities can be found to members of glycosyltransferase family GT29 and of the thiamine pyrophosphokinase family. The central β-sheet and α-helices αA, αF, and αG are structurally well conserved between Maf and Cst-II (Fig. 4), but larger structural divergence can be observed for the remaining α-helices and especially for the last α-helix of the domain, αH, which in Cst-II lies on the opposite end of the β-sheet when compared to Maf. Family GT42 sialyltransferases catalyze the transfer of a sialic acid moiety from CMP-Neu5Ac onto the terminal position of glycoconjugates with overall inversion of configuration of the anomeric carbon. Their active site is located between the edge of the central β-sheet and a small lid-like structure, which undergoes a conformational change, becomes ordered and folds over the active site upon substrate binding. The general base has been identified as being His188 in Cst-II (Chan *et al.*, 2009). A superposition of the Maf coordinates with the structures of GT42 members in complex with glycosyl donors permitted identification of a putative active site for Maf. Interestingly, the disordered loop region of Maf has its boundaries in positions

equivalent to the boundaries of the lid-like domain in the GT42 members. Maf has no residue with a functional group that could act as reaction catalyst in a position corresponding to His188 in Cst-II. By docking we produced a model of CMP-Pse5Ac7Ac bound to Maf, where the ligand adopted a position very similar to that occupied by CMP-3FNeu5Ac in the Cst-II sugar donor complex (PDB 1RO7) (Fig. S7). Intriguingly the phosphate group of the modelled CMP-Pse5Ac7Ac overlaps precisely with a sulfate ion present in the crystal structure of Maf, indicating that the sulfate might mimic binding of a genuine phosphate group present in sugar donors and corroborating the validity of the docking result. In the model of Maf with docked CMP-Pse5Ac7Ac Glu324 is located on the C-terminal end of β-strand β3 in a position equivalent to Asn51 in Cst-II, a residue shown to be crucial for acceptor binding and catalysis. Other features in the substrate binding region common to Cst-II and Maf are conserved residues located in the loop regions following β-strands β1 (Gly279, Gly281, Pro282, Ser283) and β6 (Gly413, Asp415) as well as a Thr preceding helix αF (Fig. 4). In the model of the Maf/CMP-Pse5Ac7Ac complex this latter Thr392 is at hydrogen-bonding distance from the ribosyl O3 hydroxyl group.

A sequence alignment of the MAF_flag10 domain of AMB-1 with the corresponding domains of closest Maf homologues, displaying around 70% of sequence identity, and of more distant homologues (~35% of sequence identity), representing different genera (Fig. 5A) highlighted fully conserved residues with functional groups that could potentially steer catalysis: Ser283, Asp318, Glu324, Asp415, and Asn491. Point mutations S283A, D318A, and N491A on the pAK22-*maf* plasmid we used to complement the Δ*maf* strain did not affect AMB-1 motility and the synthesis of the 70 kDa band, as verified under an optical microscope and revealed by immunoblotting, respectively. Conversely, point mutations E324A and D415A on the complementation plasmid affected directly Maf activity, as bacterial motility was not restored and the glycosylated form of the flagellin could not be detected (Fig. 5B). As mentioned

before, Glu324 is located in proximity of the putative reaction center (Fig. 4), and could thus be directly involved in catalysis. Asp415 corresponds to Asp154 in Cst-II, where this residue structures a loop forming the binding pocket of the nucleotide portion of the sugar-donor and Asp415 might exercise the same role in Maf.

## Discussion

In this work we show that the deletion of the gene *amb0685* (*maf*) in AMB-1 prevents the glycosylation of the flagellin with pseudaminic acid or a close derivative. The flagellin is still produced but the individual monomers are not exported outside the cell to form a functional flagellum, suggesting that flagellin glycosylation by the nonulosonic acid is an essential prerequisite for flagellum export and assembly. By sequence threading we found that the central MAF_flag10 domain common to all Maf proteins is distantly related to family GT42 sialyltransferases and this prediction has been corroborated with the X-ray structure of Maf. The β-strand core of the MAF_flag10 domain is structurally well conserved between Maf and GT42 members, although more variability can be found for the position of the surrounding α-helices. Furthermore, the boundaries of the disordered segment between strand β6 and helix αG of the MAF_flag10 domain overlap with the boundaries of the lid-like domain, which is present in all bacterial (GT42) and mammalian (GT29) sialyltransferases and participates in the structuring of the active site. It is not surprising that the highest structural conservation between the α/β domain of Maf and GT42 members can be found in the donor binding site, accommodating in both cases activated nonulosonic acids sugars, whereas more structural divergence occurs in the acceptor binding site, which in the case of the sialyltransferases accommodates a saccharide, galactose or sialic acid, as opposed to Maf, which accommodates a polypeptide chain. It should be noted that in *C. jejuni* the role of Cst-I and Cst-II is to

transfer Neu5Ac onto cell wall lipo-oligosaccharide outer core structures, whereas Pse5Ac7Ac is predicted to be transferred onto the *C. jejuni* flagellin by the gene product of *pseE* (*maf 5*) (Merino *et al.*, 2014).

In the artificial dimeric arrangement within the crystal the active site of Maf is shielded from the solvent and consequently we were not able to obtain structures of complexes with a donor substrate by soaking experiments. All attempts to obtain by cocrystallization a complex of Maf with a carbohydrate donor related to pseudaminic acid, CMP-Neu5Ac failed, indicating that ligand binding is governed by a fine substrate recognition mechanism. Unquestionably, further biochemical and structural studies are warranted to dissect the precise mechanism by which Maf transfers pseudaminic acid or a close derivative onto the flagellin and to ascertain the identity of the catalytic residue, which at the time being we can only hypothetically assign to Glu324.

Even though we provide strong evidence that the Maf-10 domain of Maf is a catalytic module responsible for the transfer of a nonulosonic acid onto the flagellin, the function of the N- and C-terminal modules remains elusive. Degenerated Rossmann-like motifs are found in a large number of proteins of diverse function and the similarity with methyltransferases is too distant to allow assignment of a precise function to the N-terminal domain. Six methylated lysine were observed by mass spectrometry analysis of the flagellin. Since there is a gene (*amb0693*) encoding for a putative methyltranferase in close vicinity of the flagellin gene, further work would be needed to determine whether Maf or Amb0693 is responsible for flagellin methylation. Concerning the C-terminal α-helical bundle, its intriguing resemblance to flagellar chaperones and flagellins points towards a putative role in flagellin recognition or export.

Maf proteins have been most extensively studied in gram-negative pathogens like *C. jejuni* (Thibault *et al.,* 2001; Karlyshev *et al.,* 2002; Logan *et al.,* 2002; Schirm *et al.,* 2005; McNally *et al.,* 2007), *H. pylori* (Schirm *et al.,* 2003), *A. hydrophila* (Gavin *et al.,* 2002), and *A. caviae* (Rabaan *et al.,* 2001), where the *maf* genes are often, though not systematically, associated with genes coding for the flagellins and the enzymes involved in nonulosonic acid biosynthesis. The genes encoding other structural flagella elements, like the hook and the basal body, are found elsewhere on the genome. The number of *maf* genes can vary from one organism or strain to another. In AMB-1 the only *maf* gene (*amb0685*) and the adjacent flagellin are distant from the gene cluster comprising Pse5Ac7Ac biosynthesis genes.

It has been shown by others and within this work that flagellin glycosylation can be essential for export and assembly of the bacterial flagellum, but glycosylation appears also to be involved in adhesion and regulation of the host immune response (Chaban *et al.,* 2015). The highly modified prokaryote-specific nonulosonic acid sugars can be considered as a virulence factor (Thibault *et al.,* 2001; Logan *et al.,* 2002; Goon *et al.,* 2003), and the biosynthetic pathways underlying their synthesis and transfer onto the flagellin may thus represent appealing targets for the development of new antibiotics. With the first three-dimensional structure of a Maf glycosyltransferases family member, we deliver the structural framework that should enable screening for novel compounds to address the growing challenge associated to antibiotic resistance.

## Experimental Procedures

### Bacterial strains and growth conditions.

*Magnetospirillum magneticum* AMB-1 was grown in MG (Magnetospirillum Growth) media as previously described (Murat *et al.,* 2010) and *E. coli* in LB media. *E. coli* DH5αλpir strain

was used for cloning. The *E. coli* WM3064 strain was used as donor strain for conjugation. For *E. coli* strains antibiotics were used as follows: 50 µg/ml for kanamycine and 100 µg/ml for ampicilline. For AMB-1, kanamycine was used at 15 µg/ml on plates and 10 µg/ml in liquid.

**Construction of the *maf* deletion mutant and complementation experiments.**

The AMB-1 *maf* deletion mutant (Δ*maf*) was created as described previously (Murat *et al.*, 2010). PCR runs were performed using GoTaq green mix (Promega). Briefly, the flanking regions of *maf* were amplified using AMAF2/BMAF and CMAF3/DMAF2 primers (Table S2). Then the two fragments were fused by PCR using AMAF2/DMAF2 primers. PCR cycles were : 95°C for 2 min, 30 cycles of as follows : 95°C for 10 s, 55°C for 30 s, 72°C for 90 s or 200 s (for the PCR fusion), and 72°C for 10 min. The resulting fragment of PCR fusion was cloned into the pAK31 vector (Komeili *et al.*, 2004) using the In-Fusion HD Cloning Kit method (Clontech) to create the pAK31-DELMAF plasmid. This plasmid was transformed in donor strain *E. coli* WM3064, and conjugation was done in AMB-1 as previously described (Murat *et al.*, 2010). Transformed AMB-1 cells were plated on MG agar supplemented with kanamycine and let grow for six days at 30°C in a jar to maintain microaerophilic conditions (0.4% oxygen). Some of the hereby obtained colonies were grown in 1.5 ml of MG supplemented with kanamycine for 48 h at 30°C. Then, 10 ml of MG were inoculated at 1/100 with one transconjugant and let grow at 30°C for 48 h. One, two, and six hundred microliters of this culture were used to inoculate MG plates containing sucrose (2%). The plates were incubated 8 days at 30°C in a jar under microaerophilic conditions. Colonies obtained were screened by PCR using AMAF2/DMAF2 primers to check the deletion of the *maf* gene.

For complementation experiments, the *maf* gene was amplified by PCR using GoTaq green mix (PROMEGA) and the MAFINFPAK22F/MAFINFPAK22R primer couples (Table S2).

The PCR cycle consisted of 98°C for 2 min and 35 cycles as follows: 98°C for 30 s, 55°C for 30 s, 72°C for 120 s, and 72°C for 10 min. Then the fragment generated was cloned in the AMB-1 replicative pAK22 plasmid (Murat *et al.,* 2010) using the In-Fusion HD Cloning Kit method (Clontech). The plasmid was transferred by conjugation into the Δ*maf* AMB-1 strain using *E. coli* WM3064 strain as donor. Transconjugants were plated on MG supplemented with kanamycine and incubated in a jar at 30°C for 6 days under microaerophilic conditions. Colonies were picked and grown in 1.5 ml MG supplemented with kanamycine at 30°C for 48 h. The same protocol was used to generate the AMB-1 *amb0715* deletion mutant (Δ*amb0715*). Primers 0715A/0715B and 0715C/0715D were used to generate the amplicon cloned into the deletion plasmid and primers 0715PAK22F/0715PAK22R were used to generate the amplicon cloned into the complementation plasmid (Table S2).

Complementation of the Δ*maf* strain with mutated *maf* was obtained by generating point mutations in the pAK22-*maf* complementation plasmid by PCR. 100 ng of plasmid were amplified with 0.4 µM of each primer couple (S283AF2/S83AR2, D318AF/D318AR, E324AF/E324AR, D415AF/D415AR, and N491AF/N491AR) (Table S2), 0.3 mM of dNTP, 2 U of PfuTurbo DNA polymerase (Agilent Technologies) in 1X Buffer. PCR was carried out as follows: 60 sec at 92°C, 16 x (30 sec at 92°C, 60 sec at 55°C, 1200 sec at 68°C). Then 40 U of DpnI (New English Biolabs) were added to 50 µl of PCR volume and incubated at 37°C for 16 h. Five microliters of the total digestion volume were used to transform competent *E. coli* DH5α cells. Some colonies, which were able to grow on LB agar plates supplemented with 50 µg/ml of kanamycine were sequenced to confirm the point mutation (Genewiz).

**RNA preparation and Reverse Transcription for transcriptional analyses.**

RNAs were prepared from AMB-1 cultures (50 ml) in late exponential growth phase ($OD_{600}$~0.2). Cells were harvested and frozen at -80°C. Total RNAs were isolated from the

pellet using the Maxwell® 16 LEV simplyRNA Blood Kit (Promega) according to the manufacturer's instructions and an extra TURBO DNase digestion step was performed to eliminate the contaminating DNA. RNA quality was assessed by an Experion chip (Bio-Rad), and the absence of DNA contamination was confirmed by PCR (Fig. 2SA). RNA was quantified spectrophotometrically at 260 nm (NanoDrop 1000, Thermo Fisher Scientific). For cDNA synthesis, 1 µg of total RNA and 0.5 µg of random primers (Promega) were used with the GoScript™ Reverse transcriptase (Promega) according to the manufacturer instruction. Templates for the amplifications were cDNA, diluted at 1/20 (Fig. 2SB) and water (as a negative control of the PCR, Fig. 2SC). PCR runs were performed using Platinum SuperFi DNA Polymerase (Invitrogen) and primers 0713RTF/0713RTR, 0714RTF/0714RTR, 0715RTF/0715RTR, 0716RTF/0716RTR, 0717RTF/0717RTR, 0718RTF/0718RTR (Table S2). The PCR cycle consisted of 98°C for 30 s and 25 cycles as follows: 98°C for 10 s, 64°C for 10 s, 72°C for 10 s, and 72°C for 5 min.

**Detection of protein glycosylation.**

Proteins were transferred from the SDS-PAGE gel to a nitrocellulose membrane. The membrane was incubated 30 min in 15 mM sodium periodate and 50 mM sodium acetate, then washed 4 times in PBS (Phosphate Buffer Saline) and incubated 1 h in 5 mM biotin hydrazine. The membrane was blocked overnight in TBS (Tris Buffer Saline) containing 5% BSA (Bovine Serum Albumin), and then incubated 1 h in alkaline phosphate buffer. After washing in TBS, sugar revelation was performed using the ECL Western Blotting Substrate (Thermo Scientific).

**Characterization of flagellin and its glycosylation sites by tandem mass spectrometry**.

Flagellins were isolated from cells using the shearing method (Murat *et al.,* 2015). Samples were resuspended in a minimal volume of running buffer and submitted to SDS-PAGE

analyses. The protein-containing bands were cut out from SDS-PAGE gels and treated as follows: washed with 100 mM acetonitrile/ammonium bicarbonate pH 7.5, reduced by 10 mM dithiothreitol in 100 mM ammonium bicarbonate pH 7.5, alkylated by 55 mM iodoacetamide in 100 mM ammonium bicarbonate pH 7.5, digested in a single or two steps by trypsin from porcine pancreas (Promega) followed or not by α-chymotrypsin (Promega). Each protease was used at 10 ng/µL and incubations were performed overnight at 37°C in 100 mM ammonium bicarbonate pH 7.5. Digested peptide solutions were dried and then solubilized in 8 µL of 0.05% TFA and 2% acetonitrile in water. Aliquots of 2 µl were analyzed on a ESI-Q-Exactive Plus (ThermoFisher) mass spectrometer coupled to nanoliquid chromatography (Ultimate 3000, Dionex). Peptides were first desalted on a nano-trap (Acclaim PepMap100, 300 µm i.d.x 5 mm, 5 µm, 100 Å, Dionex), mounted in a 6-port valve, before elution onto a C18 column (Acclaim PepMap RSLC, 75 µm i.d. x 150 mm, 2 µm, 100 Å, Dionex). Peptides were eluted with a linear gradient from 6% to 40% of mobile phase B (80% acetonitrile and 0.1% formic acid in water) in mobile phase A (0.1% formic acid in water) for 52 min. The peptides were detected in the mass spectrometer in a positive ion mode using a Top 10 data dependent workflow. One scan event full MS in the Orbitrap (resolution 70 000 at m/z 400), in a 350-1900 range (Auto Gain Control target at $3 \times 10^6$), was followed by fragmentation MS/MS of the top 10 ions, in the Higher Energy Collisional Dissociation cell (HCD) set at a normalized value of 27. Analysis of the MS/MS fragments in the Orbitrap was performed at resolution 17 500 at m/z 400 and "no fixed first mass" set (Auto Gain Control target at $1 \times 10^5$). The spectra were analyzed with the Proteome Discoverer software (ThermoFisherScientific, version: 1.4.0.288) using the following parameters: (1) AMB-1 proteins database (4561 entries); (2) enzymes used: trypsin/α-chymotrypsin (5 missed cleavages); (3) fixed modification: carbamidomethyl (Cys) (4) variable modifications including: oxidation (Met), methylation (Lys) and pseudaminic acid (Ser, Thr); (5) mass

values specific for monoisotopic; (6) precursor mass tolerance: ± 10 ppm; (7) fragment mass tolerance: ± 0.02 Da. A high confidence filter was applied with validation on q-Value (Strict Target FDR: 0.01) and Maximum Delta Cn: 0.05.

**Secretion assays.**

Five hundred mL of a log-phase culture of wild-type AMB-1 and Δ*maf* mutant were filtered onto a 0.22 µm membrane (Millipore) to remove the cells. Secreted fractions were generated by precipitating culture supernatant in ten volumes of ice-cold ethanol and incubated overnight at -20°C. Samples were centrifuged at 24,610 g for 30 min and protein pellet was washed in 70% ice-cold ethanol, dried, resuspended in 8 M urea, 50 mM Tris-HCl pH 7.5 and incubated 1 h at 30°C in loading buffer. Samples were analyzed by Western blotting as described previously (Murat *et al.,* 2015).

**Expression and purification of Maf.**

The coding region of Maf was amplified from *Magnetospirillum magneticum* AMB-1 DNA using *Pfu* polymerase (Invitrogen) and the following PCR reaction primers: AMB0685FW/AMB0685RV (Table S2). The amplified PCR product was inserted into a pGEX20 vector, enabling fusion of a thioredoxin/polyhistidine-tag, using the Gateway cloning technology (Invitrogen). The recombinant vectors were checked by double-stranded DNA sequencing. Constructs were synthesized from *E. coli* Rosetta pLys cells, grown in ZYP5052 medium (Studier, 2005) at 37°C, until exponential growth phase was reached, and then transferred to 17°C for overnight protein production. Cells were harvested by centrifugation at 5,000 *g* and stored at -80°C. Dry cell pellet was resuspended in lysis buffer (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 5 mM DTT, 1% (*w/v*) Triton X-100, 1 mM PMSF, 0.25 mg/ml DNase, 10 mM $MgCl_2$, 0.25 mg/ml lysozyme) and sonicated at 4°C. After centrifugation at 16,000 *g*, the soluble fraction was loaded onto a HisTrap™ 5 ml column (GE

Healthcare) and eluted with a 10-500 mM gradient of imidazole in 50 mM Tris-HCL pH 8.0 and 300 mM NaCl. Eluted fractions were pooled and desalted in a buffer containing 50 mM Tris-HCL pH 8.0, 300 mM NaCl and 10 mM imidazole. The thioredoxin/6xHis-tag was removed by overnight incubation with tobacco etch virus protease at 4°C. The reaction mixture was loaded onto a HisTrap™ 5 ml column and the flow-through was concentrated using a Vivaspin 30 kDa cutoff centrifugal filter unit (GE Healthcare). A final purification step was carried out by size exclusion chromatography using a HiLoad™ 26/60 Superdex™ 200 column (GE Healthcare) with 20 mM Hepes buffer pH 7.0 and fractions containing the desired protein were pooled and concentrated as above. The oligomerization state of Maf was investigated by Size-Exclusion Chromatography coupled with Multi-Angle Laser Light Scattering (SEC-MALLS, Wyatt technology).

**Crystallization and structure solution.**

Crystals were grown at room temperature by the vapour diffusion method in sitting drops consisting of equal volumes (100 + 100 nl) of Maf solution at 7 mg/ml and a crystallization solution composed of 1.5-1.7 M $(NH_4)_2SO_4$, 5% (v/v) polyethylene glycol 400 and 0.1 M MES buffer, pH 6.5-7.0. Crystals belong to space group $P2_12_12$ with unit cell dimensions of 117x127x65 Å and one molecule per asymmetric unit. Because crystals were extremely fragile, a heavy atom derivative was prepared by adding with a cryo-loop step-wise infinitely small amounts of a 100 mM $YbCl_3$ solution to a drop containing the crystals. Crystals were flash-cooled in liquid nitrogen without prior addition of a cryo-protectant. A 3-wavelength multiple anomalous dispersion data set for an Yb-derivatized crystal and a dataset for a crystal of native Maf were collected on the ESRF beamlines (Grenoble, France) ID29 (de Sanctis *et al.,* 2012) and ID30A-1 (Svensson *et al.,* 2015), respectively. Diffraction data were indexed with XDS (Kabsch, 2010) and scaled with the program AIMLESS (Evans *et al.,* 2013). The structure of Maf was solved with the CRANK software suite (Pannu *et al.,* 2011), which

delivered an almost complete (95%) structural model. Model completion and adjustment, using the native data set, were carried out with the programs Refmac (Murshudov *et al.*, 1997) and Coot (Emsley *et al.*, 2010), respectively. A random set of 5% of reflections was set aside for cross-validation purposes. The model quality was assessed using the Molprobity server (Chen *et al.*, 2010). The coordinates for CMP-Pse5Ac7Ac and its library used for docking where prepared with UCSF Chimera (Pettersen *et al.*, 2004). Docking was performed with the SwissDock web service (Grosdidier *et al.*, 2011). Figures representing structural renderings where generated with the PyMOL Molecular Graphics System (DeLano, W.L. The PyMOL Molecular Graphics on http://www.pymol.org/). Atomic coordinates and structure factors have been deposited within the Protein Data Bank http://www.rcsb.org (Berman *et al.*, 2000).

**Identification of sialic acids.**

Intact sialic acids were liberated by mild hydrolysis in 0.1 N TFA at 80°C and coupled to 1,2-diamino-4,5-methylenedioxybenzene dihydrochloride (DMB). For DMB derivatization, sialic acids reacted with a volume of DMB reagent at 50°C for 2h30 min. Sialic acid derivatives were first observed by fluorescence detection after HPLC separation. To do that, compounds were separated isocratically on a C18 reverse phase HPLC column (250 x 4.6 mm, 5 micron, Vydac) by a solvent mixture of acetonitrile:methanol:water (7:9:84, v/v) and identified by referring to the elution positions of standard Neu5Ac, Neu5Gc, and Kdn derivatives (Hara *et al.*, 1987). The sialic acid derivatives were then identified by LC/$MS^n$ analysis in positive ion mode on an amaZon speed ETD ion trap mass spectrometer equipped with the standard electrospray ionization (ESI) ion source and controlled by Hystar (ver. 3.2) software (Bruker Daltonics). DMB-coupled sialic acid separation was achieved on a micro LC system (Prominence LC-20AB, Shimadzu). Samples were diluted 5-fold with 0.1% formic acid and 5 µL of dilution were applied to the analytical column (luna 3u C18 (2) 100 Å 150 x 1.00 mm,

Phenomenex) with an isocratic elution of acetonitrile: methanol: water (4:6:90, v/v) at a flow rate of 60 µL/min.

**ACCESSION NUMBERS:** Coordinates and structure factors have been deposited in the Protein Data Bank with accession number **5MU5**.

The authors declare no conflict of interest.

**Funding Information**

This work was supported by l'Agence Nationale de la Recherche (ANR-2010-BLAN-1320-01 to L.-F.W.) and the French Infrastructure for Integrated Structural Biology (FRISBI) ANR-10-INSB-05-01.

**Acknowledgments**

We thank the European Synchrotron Radiation Facility (ESRF) and Synchrotron Soleil for beam time allocation, and the staff of beam lines Proxima1, Proxiam2, ID29, and ID30A-1 for assistance with data collection. We are grateful to Artemis Kosta and Yann Denis (IMM, CNRS, Marseille) for electron micrographs and transcriptomic experiments, respectively, and Bernard Henrissat for helpful discussion. We also thank the Plateforme d'Analyse des Glycoconjugués (PAGés, http://plateforme-pages.univ-lille1.fr/) and Frabio FR 3688 CNRS for providing access to the instrumental facilities for carbohydrate analysis.

**References**

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic Local Alignment Search Tool. J Mol Biol 215: 403-410.
Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H. *et al.* (2000) The Protein Data Bank. Nucleic Acids Res 28: 235-242.
Bubendorfer, S., Ishihara, M., Dohlich, K., Heiss, C., Vogel, J., Sastre, F. *et al.* (2013) Analyzing the Modification of the Shewanella oneidensis MR-1 Flagellar Filament. PLoS ONE 8: e73444.
Canals, R., Vilches, S., Wilhelms, M., Shaw, J.G., Merino, S., and Tomas, J.M. (2007) Non-structural flagella genes affecting both polar and lateral flagella-mediated motility in Aeromonas hydrophila. Microbiology 153: 1165-1175.
Canals, R., Ramirez, S., Vilches, S., Horsburgh, G., Shaw, J.G., Tomas, J.A., and Merino, S. (2006) Polar flagellum biogenesis in Aeromonas hydrophila. J Bacteriol 188: 542-555.
Chaban, B., Hughes, H.V., and Beeby, M. (2015) The flagellum in bacterial pathogens: For motility and a whole lot more. Semin Cell Dev Biol 46: 91-103.
Chan, P.H., Lairson, L.L., Lee, H.J., Wakarchuk, W.W., Strynadka, N.C., Withers, S.G., and McIntosh, L.P. (2009) NMR spectroscopic characterization of the sialyltransferase CstII from Campylobacter jejuni: histidine 188 is the general base. Biochemistry 48: 11220-11230.
Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J. *et al.* (2010) MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr 66: 12-21.
Chiu, C.P.C., Lairson, L.L., Gilbert, M., Wakarchuk, W.W., Withers, S.G., and Strynadka, N.C.J. (2007) Structural analysis of the alpha-2,3-sialyltransferase cst-i from Campylobacter jejuni in apo and substrate-analogue bound forms. Biochemistry 46: 7196-7204.
Chiu, C.P.C., Gilbert, M., Lairson, L.L., Watts, A., Wakarchuk, W.W., Withers, S.G., and Strynadka, N.C.J. (2004a) Towards the understanding of the catalytic mechanism and substrate specificities of sialyltransferases from Campylobacter jejuni. Glycobiology 14: 1123-1123.
Chiu, C.P.C., Watts, A.G., Lairson, L.L., Gilbert, M., Lim, D., Wakarchuk, W.W. *et al.* (2004b) Structural analysis of the sialyltransferase CstII from Campylobacter jejuni in complex with a substrate analog. Nat Struct Mol Biol 11: 163-170.
de Sanctis, D., Beteva, A., Caserotto, H., Dobias, F., Gabadinho, J., Giraud, T. *et al.* (2012) ID29: a high-intensity highly automated ESRF beamline for macromolecular crystallography experiments exploiting anomalous scattering. J Synchrotron Radiat 19: 455-461.
Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010) Features and development of Coot. Acta Crystallogr D Biol Crystallogr 66: 486-501.
Evans, P.R., and Murshudov, G.N. (2013) How good are my data and what is the resolution? Acta Crystallogr D Biol Crystallogr 69: 1204-1214.
Gavin, R., Rabaan, A.A., Merino, S., Tomas, J.M., Gryllos, I., and Shaw, J.G. (2002) Lateral flagella of Aeromonas species are essential for epithelial cell adherence and biofilm formation. Mol Microbiol 43: 383-397.
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.
Grosdidier, A., Zoete, V., and Michielin, O. (2011) SwissDock, a protein-small molecule docking web service based on EADock DSS. Nucleic Acids Res 39: W270-277.
Hara, S., Takemori, Y., Yamaguchi, M., Nakamura, M., and Ohkura, Y. (1987) Fluorometric high-performance liquid chromatography of N-acetyl- and N-glycolylneuraminic acids and its application to their microdetermination in human and animal sera, glycoproteins, and glycolipids. Anal Biochem 164: 138-145.
Holm, L., and Sander, C. (1993) Protein structure comparison by alignment of distance matrices. J Mol Biol 233: 123-138.
Hopf, P.S., Ford, R.S., Zebian, N., Merkx-Jacques, A., Vijayakumar, S., Ratnayake, D. *et al.* (2011) Protein Glycosylation in Helicobacter pylori: Beyond the Flagellins? PLoS ONE 6: e25722.

Iwashkiw, J.A., Vozza, N.F., Kinsella, R.L., and Feldman, M.F. (2013) Pour some sugar on it: the expanding world of bacterial protein O-linked glycosylation. Mol Microbiol 89: 14-28.
Kabsch, W. (2010) XDS. Acta Crystallogr D Biol Crystallogr 66: 125-132.
Karlyshev, A.V., Linton, D., Gregson, N.A., and Wren, B.W. (2002) A novel paralogous gene family involved in phase-variable flagella-mediated motility in Campylobacter jejuni. Microbiology 148: 473-480.
Komeili, A., Vali, H., Beveridge, T.J., and Newman, D.K. (2004) Magnetosome vesicles are present before magnetite formation, and MamA is required for their activation. Proc Natl Acad Sci USA 101: 3839-3844.
Krissinel, E., and Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr D Biol Crystallogr 60: 2256-2268.
Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. J Mol Biol 372: 774-797.
Lewis, A.L., Desa, N., Hansen, E.E., Knirel, Y.A., Gordon, J.I., Gagneux, P. *et al.* (2009) Innovations in host and microbial sialic acid biosynthesis revealed by phylogenomic prediction of nonulosonic acid structure. Proc Natl Acad Sci U S A 106: 13552-13557.
Logan, S., Kelly, J.F., Thibault, P., Ewing, C.P., and Guerry, P. (2002) Structural heterogeneity of carbohydrate modifications affects serospecificity of Campylobacter flagellins. Mol Microbiol 46: 587-597.
Logan, S.M. (2006) Flagellar glycosylation - a new component of the motility repertoire? Microbiology 152: 1249-1262.
Lombard, V., Ramulu, H.G., Drula, E., Coutinho, P.M., and Henrissat, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res 42: D490-D495.
Malone, T., Blumenthal, R.M., and Cheng, X. (1995) Structure-guided analysis reveals nine sequence motifs conserved among DNA amino-methyltransferases, and suggests a catalytic mechanism for these enzymes. J Mol Biol 253: 618-632.
McNally, D.J., Hui, J.P.M., Aubry, A.J., Mui, K.K.K., Guerry, P., Brisson, J.R. *et al.* (2006) Functional characterization of the flagellar glycosylation locus in Campylobacter jejuni 81-176 using a focused metabolomics approach. J Biol Chem 281: 18489-18498.
McNally, D.J., Aubry, A.J., Hui, J.P.M., Khieu, N.H., Whitfield, D., Ewing, C.P. *et al.* (2007) Targeted metabolomics analysis of Campylobacter coli VC167 reveals legionaminic acid derivatives as novel flagellar glycans. J Biol Chem 282: 14463-14475.
Meier, A., and Soding, J. (2015) Automatic Prediction of Protein 3D Structures by Probabilistic Multi-template Homology Modeling. PLoS Comput Biol 11: e1004343.
Merino, S., and Tomas, J.M. (2014) Gram-negative flagella glycosylation. International Journal of Molecular Sciences 15: 2840-2857.
Morrison, M.J., and Imperiali, B. (2014) The Renaissance of Bacillosamine and Its Derivatives: Pathway Characterization and Implications in Pathogenicity. Biochemistry 53: 624-638.
Murat, D., Quinlan, A., Vali, H., and Komeili, A. (2010) Comprehensive genetic dissection of the magnetosome gene island reveals the step-wise assembly of a prokaryotic organelle. Proc Natl Acad Sci USA 107: 5593-5598.
Murat, D., Hérissé, M., Espinosa, L., Bossa, A., Alberto, F., and Wu, L.-F. (2015) Opposite and Coordinated Rotation of Amphitrichous Flagella Governs Oriented Swimming and Reversals in a Magnetotactic Spirillum. J Bacteriol 197: 3275.
Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr 53: 240-255.
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.
Nothaft, H., and Szymanski, C.M. (2010) Protein glycosylation in bacteria: sweeter than ever. Nat Rev Microbiol 8: 765-778.

Pannu, N.S., Waterreus, W.J., Skubak, P., Sikharulidze, I., Abrahams, J.P., and de Graaff, R.A. (2011) Recent advances in the CRANK software suite for experimental phasing. Acta Crystallogr D Biol Crystallogr 67: 331-337.
Parker, J.L., Day-Williams, M.J., Tomas, J.M., Stafford, G.P., and Shaw, J.G. (2012) Identification of a putative glycosyltransferase responsible for the transfer of pseudaminic acid onto the polar flagellin of Aeromonas caviae Sch3N. MicrobiologyOpen 1: 149-160.
Parker, J.L., Lowry, R.C., Couto, N.A.S., Wright, P.C., Stafford, G.P., and Shaw, J.G. (2014) Maf-dependent bacterial flagellin glycosylation occurs before chaperone binding and flagellar T3SS export. Mol Microbiol 92: 258-272.
Parkhill, J., Wren, B.W., Mungall, K., Ketley, J.M., Churcher, C., Basham, D. *et al.* (2000) The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hypervariable sequences. Nature 403: 665-668.
Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem 25: 1605-1612.
Poweleit, N., Ge, P., Nguyen, H.H., Loo, R.R., Gunsalus, R.P., and Zhou, Z.H. (2016) CryoEM structure of the Methanospirillum hungatei archaellum reveals structural features distinct from the bacterial flagellum and type IV pili. Nat Microbiol 2: 16222.
Rabaan, A.A., Gryllos, I., Tomas, J.M., and Shaw, J.G. (2001) Motility and the polar flagellum are required for Aeromonas caviae adherence to HEp-2 cells. Infect Immun 69: 4257-4267.
Ricaldi, J.N., Matthias, M.A., Vinetz, J.M., and Lewis, A.L. (2012) Expression of sialic acids and other nonulosonic acids in Leptospira. BMC Microbiol 12: 161.
Schirm, M., Schoenhofen, I.C., Logan, S.M., Waldron, K.C., and Thibault, P. (2005) Identification of unusual bacterial glycosylation by tandem mass spectrometry analyses of intact proteins. Anal Chem (Wash) 77: 7774-7782.
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., Kalmokoff, M., Aubry, A., Thibault, P., Sandoz, M., and Logan, S.M. (2004a) Flagellin from Listeria monocytogenes is glycosylated with beta-O-linked N-acetylglucosamine. J Bacteriol 186: 6721-6727.
Schirm, M., Arora, S.K., Verma, A., Vinogradov, E., Thibault, P., Ramphal, R., and Logan, S.M. (2004b) Structural and genetic characterization of glycosylation of type a flagellin in Pseudomonas aeruginosa. J Bacteriol 186: 2523-2531.
Schoenhofen, I.C., McNally, D.J., Brisson, J.R., and Logan, S.M. (2006) Elucidation of the CMP-pseudaminic acid pathway in Helicobacter pylori: synthesis from UDP-N-acetylglucosamine by a single enzymatic reaction. Glycobiology 16: 8C-14C.
Studier, F.W. (2005) Protein production by auto-induction in high-density shaking cultures. Protein Expression Purif 41: 207-234.
Svensson, O., Malbet-Monaco, S., Popov, A., Nurizzo, D., and Bowler, M.W. (2015) Fully automatic characterization and data collection from crystals of biological macromolecules. Acta Crystallogr D Biol Crystallogr 71: 1757-1767.
Thibault, P., Logan, S.M., Kelly, J.F., Brisson, J.R., Ewing, C.P., Trust, T.J., and Guerry, P. (2001) Identification of the carbohydrate moieties and glycosylation motifs in Campylobacter jejuni flagellin. J Biol Chem 276: 34862-34870.
Twine, S.M., Reid, C.W., Aubry, A., McMullin, D.R., Fulton, K.M., Austin, J., and Logan, S.M. (2009) Motility and Flagellar Glycosylation in Clostridium difficile. J Bacteriol 191: 7050-7062.
Zunk, M., and Kiefel, M.J. (2014) The occurrence and biological significance of the alpha-keto-sugars pseudaminic acid and legionaminic acid within pathogenic bacteria. RSC Advances 4: 3413-3421.

**Fig. 1.** Shearing-prepared surface fractions containing flagellin from wild-type AMB-1. The shearing-prepared protein fractions were submitted to SDS-PAGE and the gel was Coomassie stained (A), hybridized with polyclonal antibodies raised against the AMB-1 flagellin (B), or submitted to sugar detection by the periodate/biotin hydrazine method (C). Lane MW represents the Molecular Weight Marker. Electron micrographs of the flagella of wild-type AMB-1(D), the Δ*maf* mutant (E) and the complemented Δ*maf* mutant (F), and the corresponding hybridization with polyclonal anti-flagellin antibody of total protein extract from wild-type AMB-1 (G), the Δ*maf* mutant (H), and the complemented Δ*maf* mutant (I).

**Fig. 2.** Analysis of DMB-coupled sialic acids by LC-$MS^n$. (A) $MS^3$ fragmentation pattern of $[M-18+H]^+$ ion at *m/z* 408 of standard DMB-Neu5Ac (Inset Neu5Ac); (B) $MS^2$ fragmentation pattern of $[M+H]^+$ ion at m/z 451 of a nonulosonic acid purified from AMB-1 flagellin (Inset Pse5Ac7Ac). On the right panel, the fragment ions of Neu5Ac and Pse5Ac7Ac were deduced from previous studies on multiple sialic acids.

**Fig. 3.** Overall structure of Maf. Ribbon view of Maf colored by domains: the N-terminal domain is shown in blue, the central α/β domain in green, and the C-terminal domain in orange. The boundaries of the disordered region corresponding to the lid-like domain of family GT41 sialyltransferases are depicted in yellow. The helix Leu92-Ser108 participating to the C-terminal domain is shown in ruby, and helices inserted into the inter-domain linkers are in grey. The secondary structure elements of the central domain, β1–β7, and αA–αH are labelled, and the N- and C-termini are marked by labelled circles.

**Fig. 4.** Comparison of Maf and *C. jejuni* Cst-II. Superposition of the central α/β domain of Maf, color-coded as in Fig. 3, and *C. jejuni* Cst-II (PDB 1RO7), colored in light-blue. A-helices of Maf and the C-terminal α-helix of Cst-II are labelled in yellow and pink, respectively. A model of Maf-bound CMP-Pse5Ac7Ac obtained by docking is represented in

sticks, with carbon atoms colored in pink, and oxygen and nitrogen atoms colored in red and blue, respectively. A sulfate ion present in the crystal structure of Maf and overlapping with the phosphate group of docked CMP-Pse5Ac7Ac is shown in orange sticks. Conserved residues, the catalytic base His188 of Cst-II and the putative catalytic base Glu324 of Maf are highlighted and more detailed views are provided in the insets. Carbon atoms of Maf are colored in green and carbon atoms of Cst-II are shown in light-blue, whereas oxygen and nitrogen atoms of both Maf and Cst-II are shown in red and blue, respectively.

**Fig. 5.** (A). Sequence alignment of the MAF_flag10 domain (from amino acids T273 to R530) of Maf from AMB-1 (M.mag, NCBI accession WP_011383128.1) with homologous proteins from: M.QH2, *Magnetospira* sp. QH2 (WP_046021671.1), A.cav, *Aeromonas caviae* (WP_042015967.1), G.psy, *Glaciecola psychrophila* (WP_007637865.1), P.atl, *Pseudoalteromonas atlantica* (WP_011575799.1), M.med, *Marinomonas mediterranea* (WP_013662468.1), P.mar, *Prochlorococcus marinus* (WP_011129470.1), R.cen, *Rhodospirillum centenum* (WP_012568882.1), H.pyl, *Helicobacter pylori* (WP_001116302.1), W.suc, *Wolinella succinogenes* (WP_011139562.1), C.jej, *Campylobacter jejuni* (WP_015016479.1). Amino acids that were mutated to alanine on the complementation plasmid pAK22-*maf* are marked with a blue asterisk. (B) Western blots of total protein extracts from the Δ*maf* mutant complemented with pAK22-*maf* plasmids containing the point mutations: S283A, D318A, E324A, D415A, N494A. Polyclonal antibodies raised against the AMB-1 flagellin were used.

Figure 1

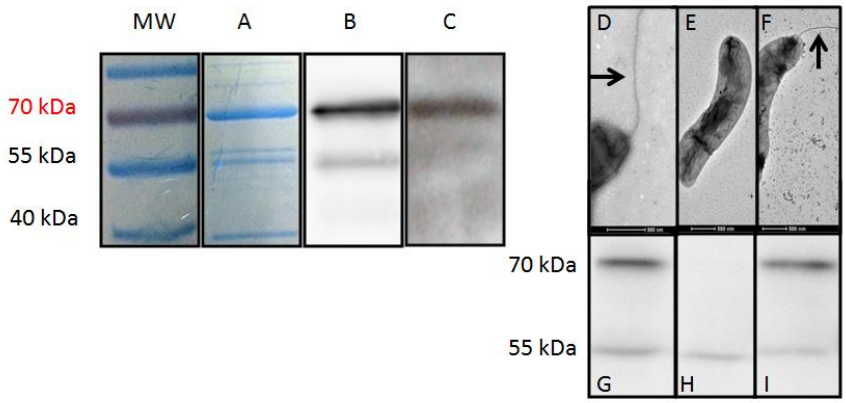


Figure 2

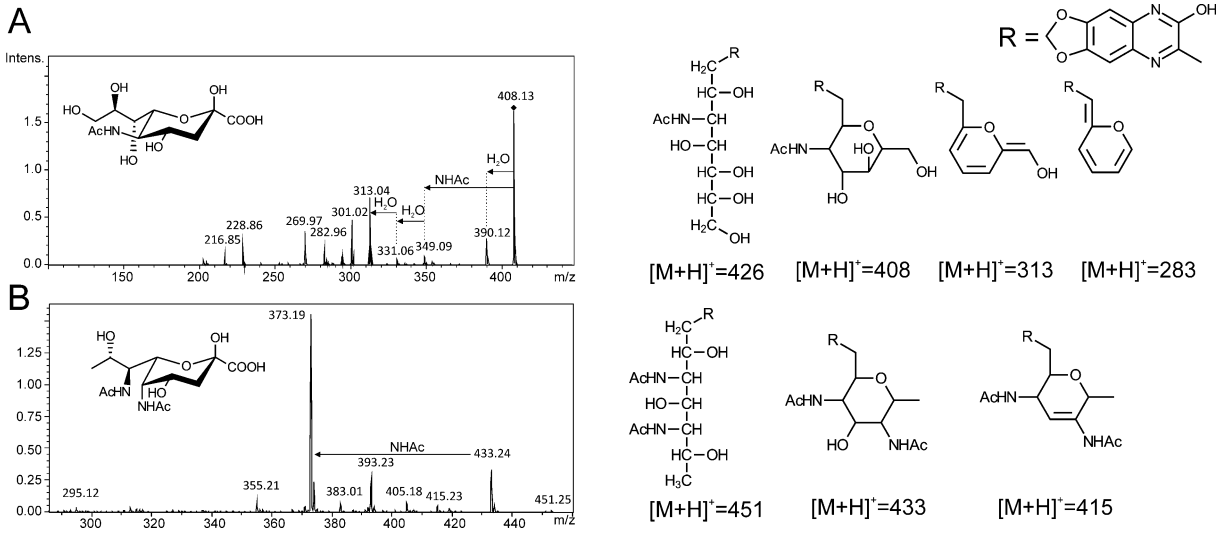

A
Intens.
408.13
313.04
301.02
269.97
282.96
228.86
216.85
331.06
349.09
390.12
NHAc
m/z
B
373.19
NHAc
433.24
393.23
295.12
355.21
383.01
405.18
415.23
451.25
m/z
R =
[M+H]+=426
[M+H]+=408
[M+H]+=313
[M+H]+=283
[M+H]+=451
[M+H]+=433
[M+H]+=415

Figure 3

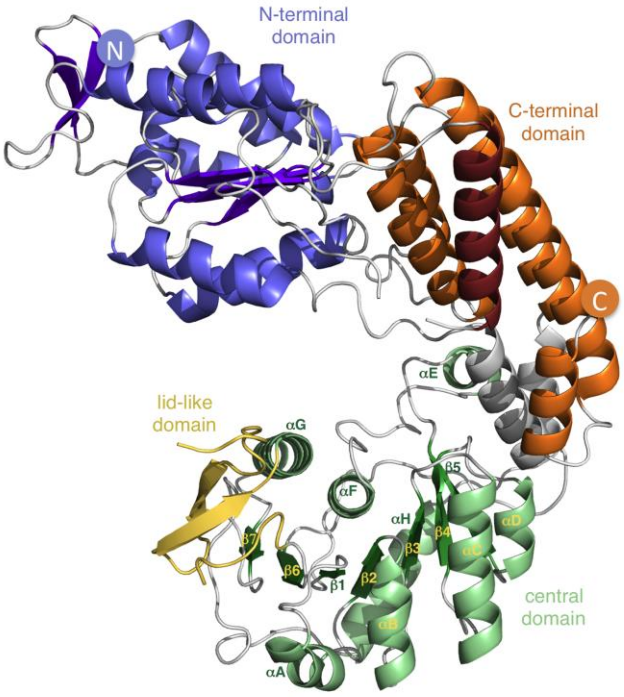

N-terminal
domain
N
C-terminal
domain
C
αE
lid-like
domain
αG
αF
β5
αH
β4
αD
β7
β3
αC
β6
β1
β2
central
domain
αB
αA

Figure 4

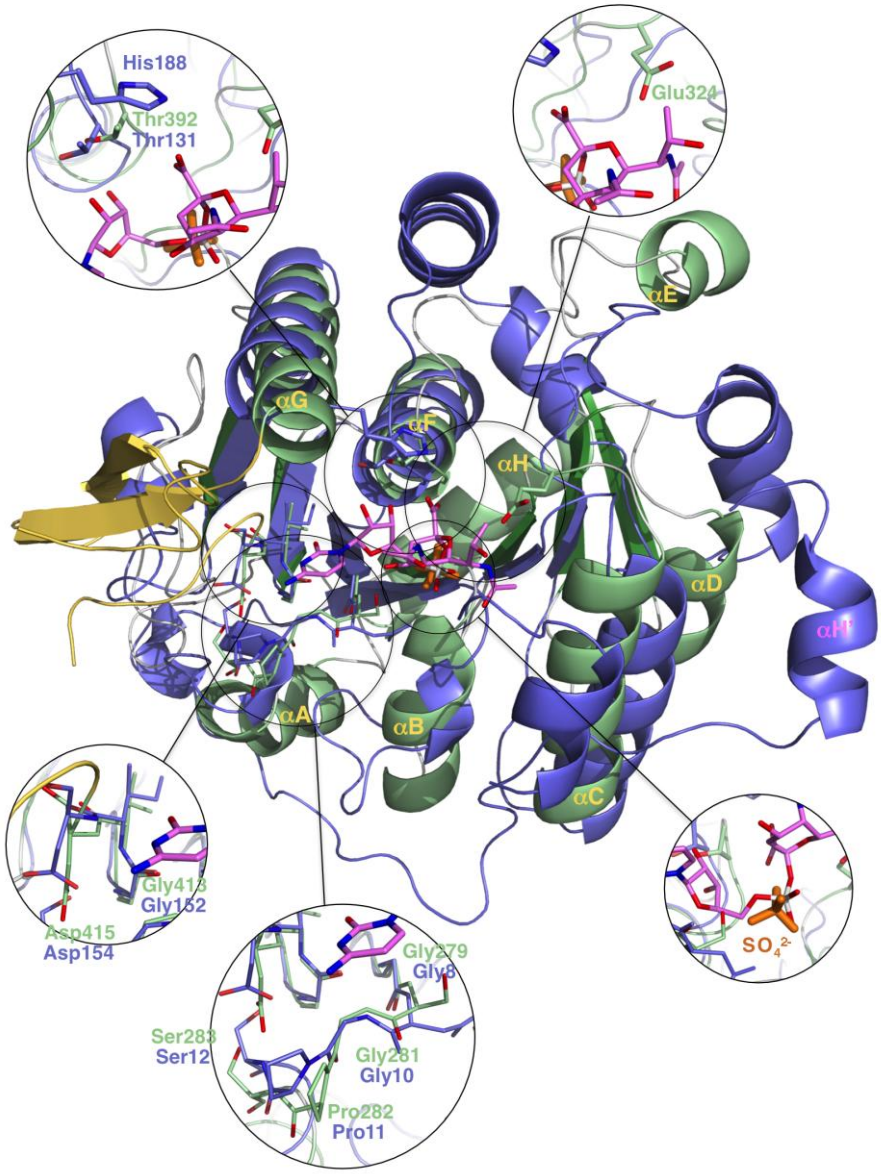
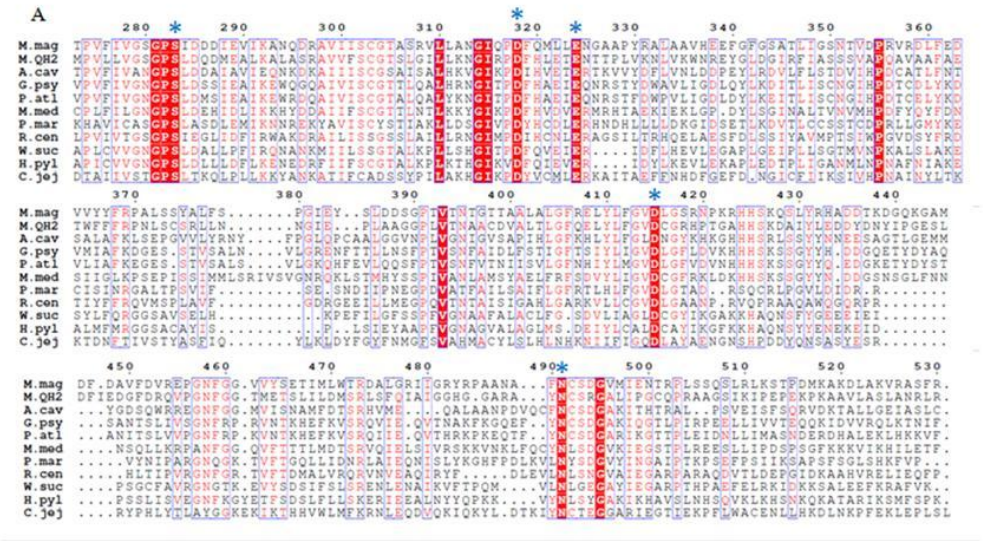
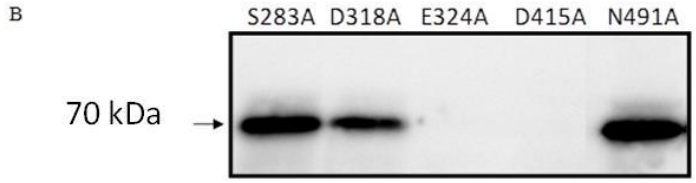

His188
Thr392
Thr131
Glu324
αE
αG
αF
αH
αD
αH'
αA
αB
αC
Gly413
Gly152
Asp415
Asp154
Gly279
Gly8
Ser283
Ser12
Gly281
Gly10
Pro282
Pro11
$SO_4^{2-}$

Figure 5


A
M.mag
M.QH2
A.cav
G.psy
P.atl
M.med
P.mar
R.con
W.suc
H.pyl
C.jej


B
S283A
D318A
E324A
D415A
N491A
70 kDa

**Supplemental Figures**

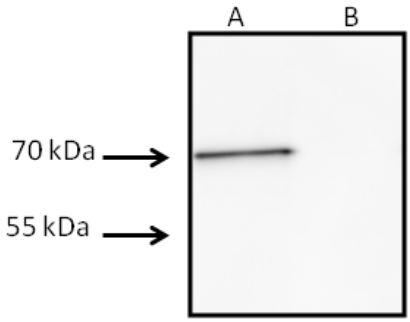



**Fig. S1.** Western blotting of secreted fractions from cultures of (A) wild-type AMB-1 and (B) the Δ*maf* mutant. Hybridization was done with a polyclonal antibody raised against the AMB-1 flagellin.

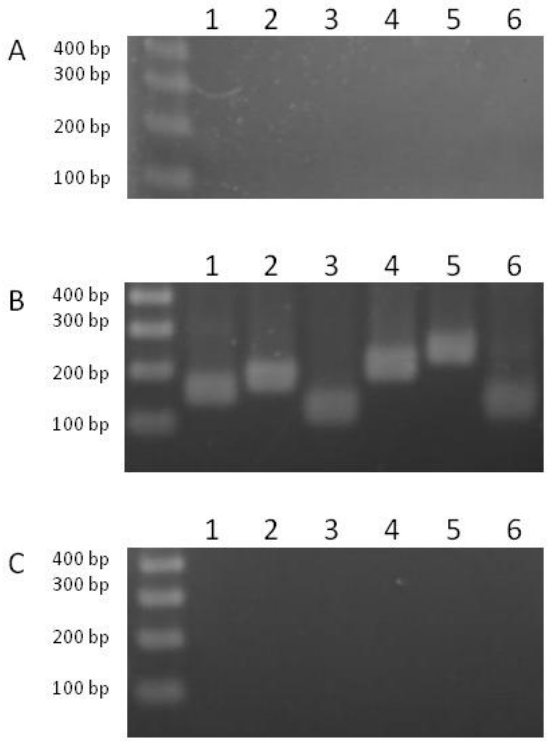



**Fig. S2.** Expression of genes from the cluster putatively involved in Pse5Ac7Ac cluster biosynthesis. The agarose gels presented in the figure show the result of PCR using primers for the expression of *amb0713* (lane 1), *amb0714* (lane 2), *amb0715* (lane 3), *amb0716* (lane 4), *amb0717* (lane 5), and *amb0718* (lane 6) utilizing (A) total extracted RNA, (B) cDNA, and (C) water as templates.

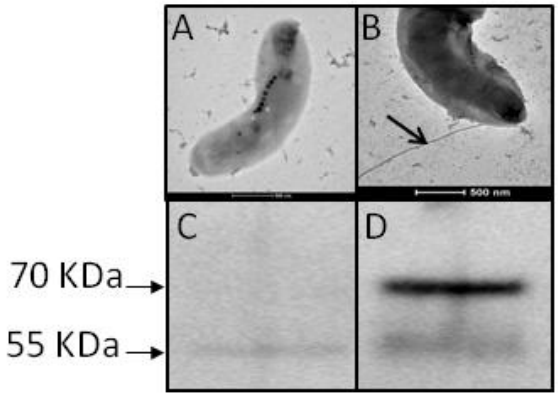
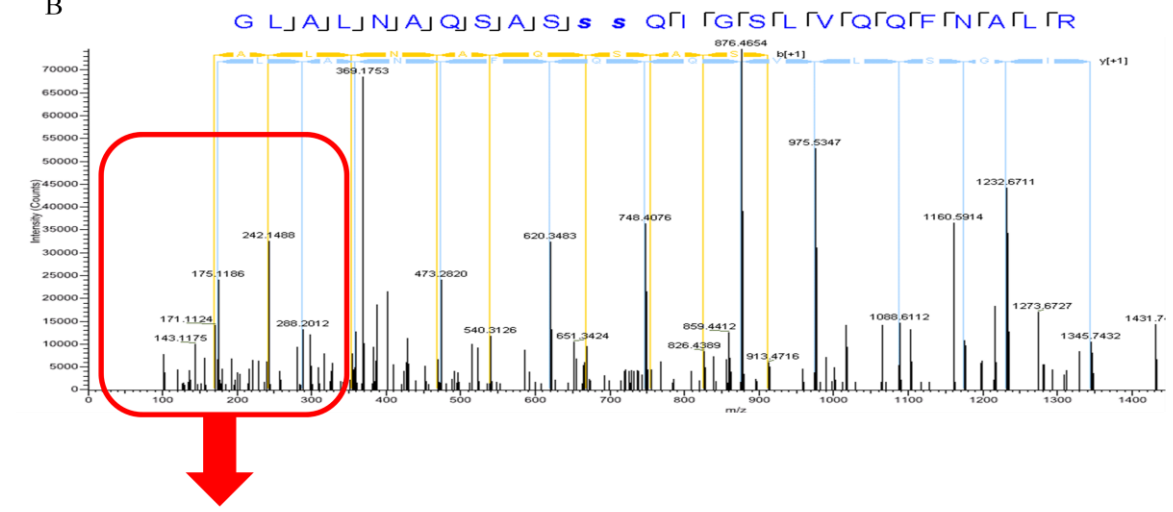
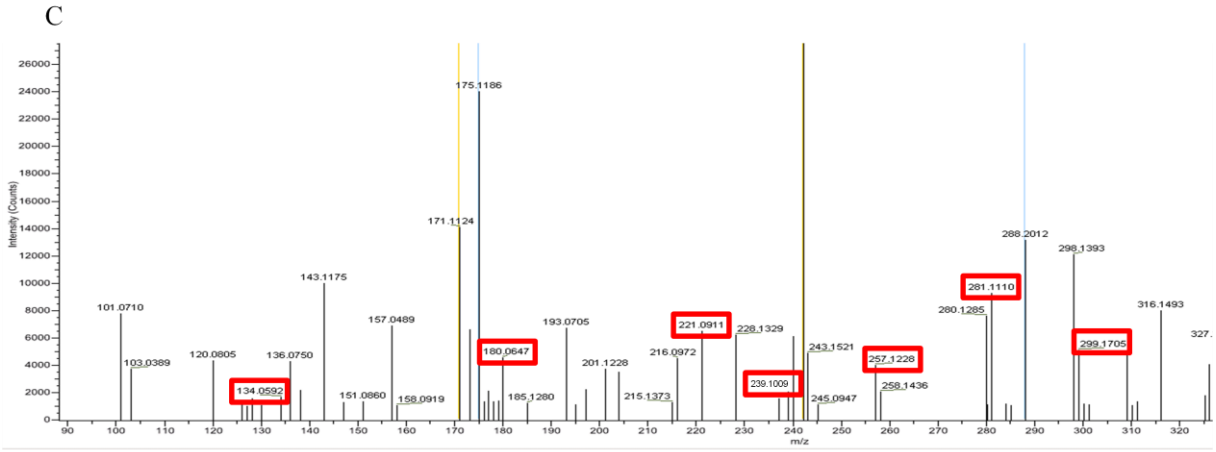



**Fig. S3.** Electron micrographs of AMB-1 (*A*) Δ*0715* mutant, (B) the complemented Δ*amb0715* mutant, and the corresponding hybridization with polyclonal anti-flagellin antibody of total protein extract from (C) the Δ*amb0715* mutant, and (D) the complemented Δ*amb0715* mutant.

A

| Sequence | Modifications | Charge | $[M+H]^+$ [Da] | ΔM [ppm] |
|---|---|---|---|---|
| SHAQNLGTNVALLNtRLDF | T15(Pse5Ac7Ac) | 3 | 2400,2121 | -1,36 |
| GQTLTVsFsNLTGSTLEVTSVDVR | S7 and S9 (Pse5Ac7Ac) | 3 | 3143,5660 | 5,76 |
| VAGKGQTLTVSFsNLTGSTLEVTSVDVR | S13(Pse5Ac7Ac) | 3 | 3182,6231 | -6,18 |
| GLALNAQSASssQIGSLVQQFNALR | S11 and S12(Pse5Ac7Ac) | 3 | 3192,6240 | 6,89 |



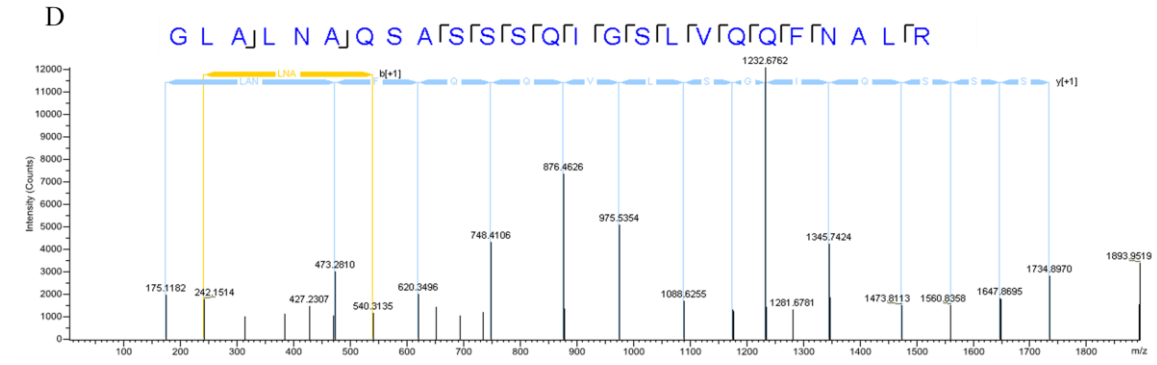




**Fig. S4.** Results of tandem mass spectrometry analysis (A) List of Pse5Ac7Ac modified peptides of digested wild-type AMB-1 flagellin identified by LC-MS/MS. (B) Nano-electrospray MS/MS spectrum in the range m/z 100-1800, of the peptide "$^{93}$GLALNAQSAS**ss**QIGSLVQQFNALR$^{117}$", including the double Pse5Ac7Ac modification (the two Ser residues displaying the Pse5Ac7Ac modification are in lowercase letters and in bold), $[M+H]^+$ 3192.624 Da (charge state +3, monoisotopic m/z 1064.880 (Δm 6.9ppm)). (C) Zoom in the low range m/z 90-330 of the spectrum in (B) to show the Pse5Ac7Ac signature (specific m/z in red boxes). (D) MS/MS spectrum of the same peptide as in B, from the Δ*maf* type flagellin, $[M+H]^+$ 2560.355 Da (charge state +2, monoisotopic m/z 1280.681 (Δm 2.78ppm). b and y ions from the MS/MS fragmentation are represented in yellow and blue, respectively, in all spectra.

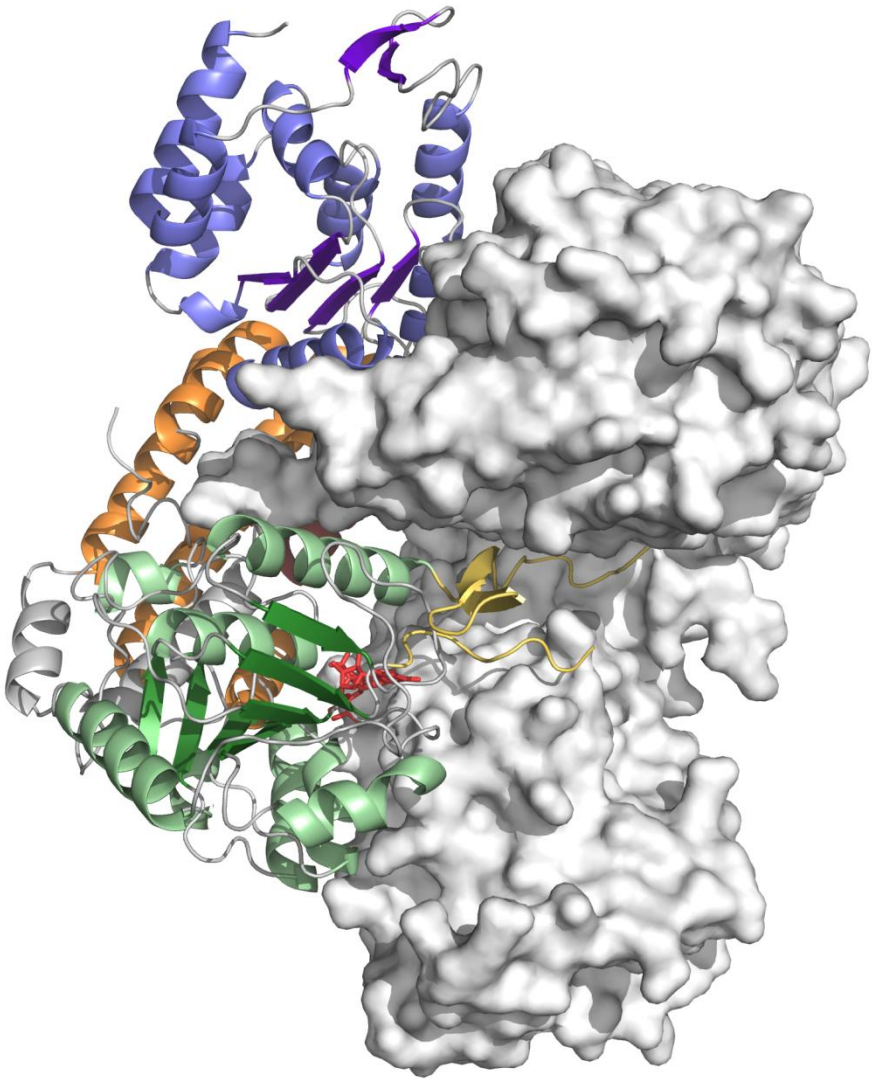


**Fig. S5.** The artificial dimer of Maf as observed in the crystal structure. A ribbon diagram of Maf colored as in Fig. 3 interacting with a symmetry-related molecule (grey surface) via the boundaries of the lid-like domain. CMP-Pse5Ac7Ac obtained by docking is shown in red sticks.

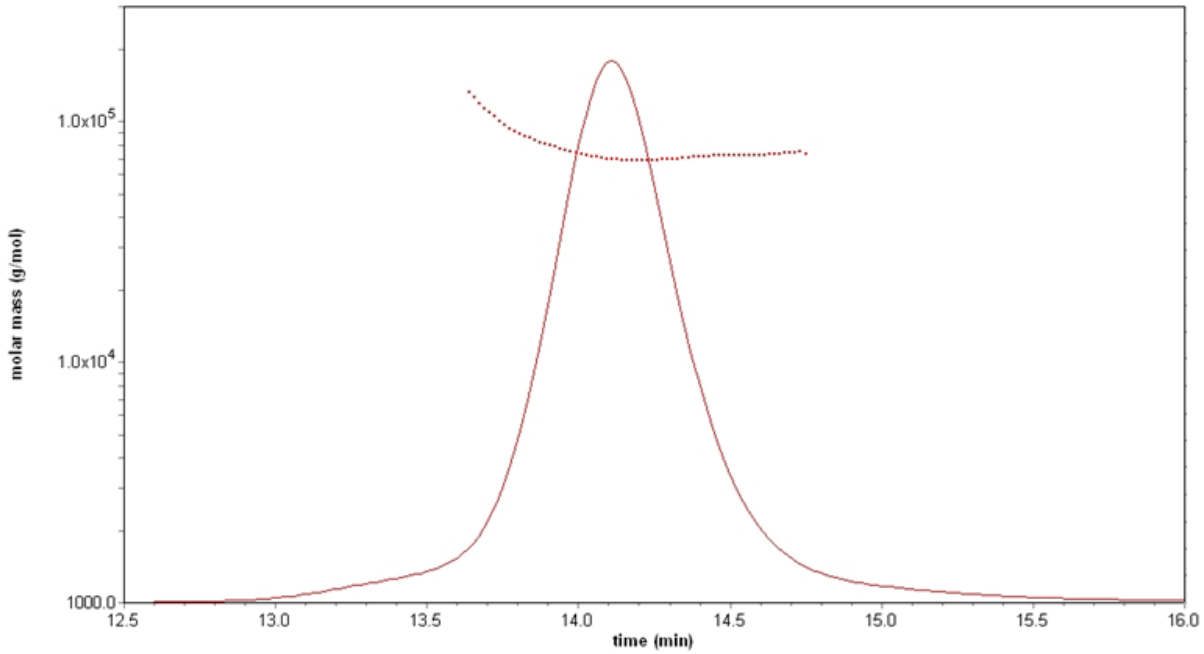



**Fig. S6.** Light scattering (SEC-MALLS) analysis of Maf. The chromatogram from the size-exclusion experiment (solid line) and the calculated molar mass of the protein (dotted line) are shown.

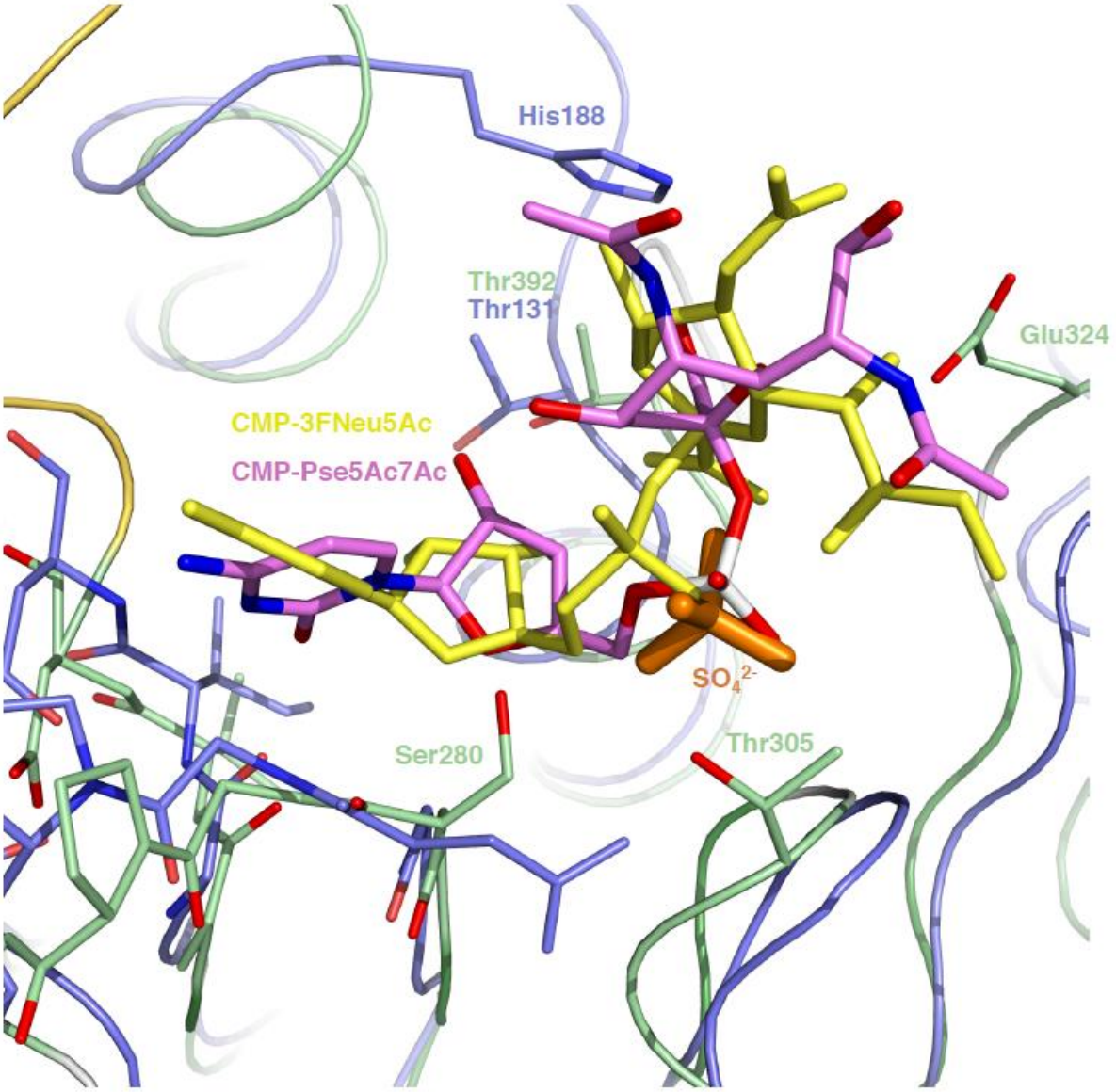



**Fig. S7.** Donor sugar binding to Maf and *C. jejuni* Cst-II. Superposition of the central α/β domain of Maf, color-coded as in Fig. 3, and *C. jejuni* Cst-II (PDB 1RO7), colored in light-blue. A model of CMP-Pse5Ac7Ac obtained by docking is represented in sticks, with carbon atoms colored in pink, and oxygen and nitrogen atoms colored in red and blue, respectively. A sulphate ion present in the crystal structure of Maf and overlapping with the phosphate group of modelled CMP-Pse5Ac7Ac is shown in orange sticks. CMP-3FNeu5Ac bound to Cst-II is shown in yellow sticks. Conserved residues, the catalytic base His188 of Cst-II and the putative catalytic base Glu324 of Maf are shown in sticks, with carbon atoms of Maf colored in green and carbon atoms of Cst-II shown in light-blue, whereas oxygen and nitrogen atoms of both Maf and Cst-II are shown in red and blue, respectively. Maf residues Thr305 and Ser280 might have a role in binding of the activated sugar-donor.

## Supplemental Tables

**Table S1:** Crystallographic data collection and refinement statistics

| | **Native** | **Yb-derivative** | | |
|---|---|---|---|---|
| | | Peak | Inflection | Remote |
| **Data collection** | | | | |
| Wavelength | 0.9650 | 1.38519 | 1.38578 | 1.38056 |
| Resolution range [a] | 45.26-2.30 (2.42-2.30) | 45.23-2.80 (2.95-2.80) | 45.27-2.80 (2.95-2.80) | 45.34-2.80 (2.95-2.80) |
| Redundancy | 5.1 (5.2) | 5.7 (5.7) | 5.7 (5.7) | 5.6 (5.7) |
| Anomalous redundancy | 2.7 (2.7) | 2.9 (2.9) | 2.9 (2.9) | 2.9 (2.9) |
| Completeness (%) | 99.7 (99.5) | 99.9 (99.9) | 99.9 (100) | 99.9 (100) |
| Anomalous completeness (%) | 99.8 (99.2) | 99.1 (99.3) | 99.2 (99.3) | 99.2 (99.3) |
| No. of unique reflections | 43504 | 24525 | 24620 | 24689 |
| $R_{merge}$ [b] | 0.085 (0.790) | 0.077 (1.035) | 0.057 (0.575) | 0.080 (1.502) |
| CC1/2 | 0.998 (0.769) | 0.996 (0.719) | 0.998 (0.885) | 0.998 (0.546) |
| Mean I/□(I) | 12.8 (2.4) | 13.9 (1.4) | 16.6 (2.7) | 12.6 (1.0) |
| B-factor from Wilson plot (Å$^2$) | 40.93 | 72.77 | 69.46 | 82.78 |
| **Refinement** | | | | |
| $R_{cryst}$ (%) [d] | 18.67 (28.0) | | | |
| $R_{free}$ (%) | 22.02 (31.30) | | | |
| No. of free reflections | 2224 | | | |
| Protein atoms | 5014 | | | |
| Ligand atoms | 46 | | | |
| Solvent atoms | 98 | | | |
| *r.m.s. deviations from target values* | | | | |
| Bond lengths (Å) | 0.009 | | | |
| Bond angles (°) | 1.389 | | | |
| Chiral volumes (Å$^3$) | 0.079 | | | |
| *Average B-factors (Å$^2$)* | | | | |
| Main/side chains | 47.80/53.59 | | | |
| Ligands/solvent | 67.50/47.81 | | | |
| *r.m.s. deviations on B-factors (A$^2$)* | | | | |
| Main chain | 1.663 | | | |
| Side chain | 2.660 | | | |
| *Ramachandran plot statistics (%)*[e] | | | | |
| Residues in favoured regions | 98.56 | | | |
| Residues in allowed regions | 1.44 | | | |

[a] Throughout the table the values in parentheses apply for the outermost resolution shell

[b] $R_{merge} = \sum_{hkl} \sum_i ( I_{i(hkl)} - \langle I_{hkl} \rangle ) / \sum_{hkl} \sum_i I_{i(hkl)}$, where I is an individual reflection measurement and <I> is the mean intensity for symmetry related reflection.

[d] $R_{cryst} = \sum_{hkl} | | F_o| - |F_c| | / \sum_{hkl} | F_o |$, where $F_o$ and $F_c$ are observed and calculated structure factors, respectively.

[e] There are no Ramachandran outliers

**Table S2:** Sequences of primers used in this study

| Primers name | Sequence | Description |
|---|---|---|
| **AMAF2** | CGGGGGATCCACTAGTAGTGTGTGGTCAAGGCCACCTAC | Generation of AB fragment for deletion of *amb0685* in AMB-1 |
| **BMAF** | CCCATCCACTAAATTTAAATAGCCCGACTCCCTTTTCCCAC | Generation of AB fragment for deletion of *amb0685* in AMB-1 |
| **CMAF3** | TATTTAAATTTAGTGGATGGGAGTCGGACGAGGAACTGTTCG | Generation of CD fragment for deletion of *amb0685* in AMB-1 |
| **DMAF2** | CCGCTCTAGAACTAGTCGAGGTGACCCAGTGGAC | Generation of CD fragment for deletion of *amb0685* in AMB-1 |
| **MAFINFPAK22F** | ACAGGAAACAGAATTCATGGCCGCAGCGCCGCAGATC | Amplification of *amb0685* to clone in pAK22 for complementation |
| **MAFINFPAK22R** | CCGCTCTAGAACTAGTTCAGTCATAATCCCACCCCGTC | Amplification of *amb0685* to clone in pAK22 for complementation |
| **0715A** | CCTGCAGCCCGGGGGATCCACTAGTAGACGATCAGTCCCAGTCC | Generation of AB fragment for deletion of *amb0715* in AMB-1 |
| **0715B** | CCCATCCACTAAATTTAAATACACTGCCCTGCGGAACTCCAT | Generation of AB fragment for deletion of *amb0715* in AMB-1 |
| **0715C** | TATTTAAATTTAGTGGATGGGGACGGTGGGCGCATGCCATGA | Generation of AB fragment for deletion of *amb0715* in AMB-1 |
| **0715D** | GCGGTGGCGGCCGCTCTAGAACTAGTACGTCGGTCAGCACCACT | Generation of AB fragment for deletion of *amb0715* in AMB-1 |
| **0715PAK22F** | ACAGGAAACAGAATTCATGGAGTTCCGCAGGGCA | Amplification of *amb0715* to clone in pAK22 for complementation |
| **0715PAK22R** | CCGCTCTAGAACTAGTTCATGGCATGCGCCCACC | Amplification of *amb0715* to clone in pAK22 for complementation |
| **S283AF2** | CGTGGGGTCCGGTCCCGCCATCGATGACGAC | Mutagenesis S283A in plasmid pAK22-amb0685 |
| **S283AR2** | GTCGTCATCGATGGCGGGACCGGACCCCACG | Mutagenesis S283A in plasmid pAK22-amb0685 |
| **D318AF** | ACGGCATTCAGCCTGCCTTCCAGATGCTGCTGG | Mutagenesis D318A in plasmid pAK22-amb0685 |
| **D318AR** | CCAGCAGCATCTGGAAGGCAGGCTGAATGCCGT | Mutagenesis D318A in plasmid pAK22-amb0685 |
| **E324AF** | TTCCAGATGCTGCTGGCAAATGGCGCCGCCCC | Mutagenesis E324A in plasmid pAK22-amb0685 and pETG-20A-amb0685 |
| **E324AR** | GGGGCGGCGCCATTTGCCAGCAGCATCTGGAA | Mutagenesis E324A in plasmid pAK22-amb0685 and pETG-20A-amb0685 |
| **D415AF** | CCTGTTCGGCGTCGCCCTGGGCAGCCGCAAC | Mutagenesis D415A in plasmid pAK22-amb0685 |
| **D415AR** | GTTGCGGCTGCCCAGGGCGACGCCGAACAGG | Mutagenesis D415A in plasmid pAK22-amb0685 |
| **N491AF** | CGGCCAACGCCTTCGCTTGCAGCGACGGGGT | Mutagenesis N491A in plasmid pAK22-amb0685 |
| **N491AR** | ACCCCGTCGCTGCAAGCGAAGGCGTTGGCCG | Mutagenesis N491A in plasmid pAK22-amb0685 |
| **0713RTF** | TCCCATTTCTCGCTGGAACC | Amplification of gene *amb0713* in qRT-PCR experiments |
| **0713RTR** | GTCCTCGACCACATAGAGCG | Amplification of gene *amb0713* in qRT-PCR experiments |
| **0714RTF** | GGTGATCGACCATTACGGCA | Amplification of gene *amb0714* in qRT-PCR experiments |
| **0714RTR** | GGAATTGTGGGGGACCAGTT | Amplification of gene *amb0714* in qRT-PCR experiments |
| **0715RTF** | GGGTCAAGCAAAGCAATCCG | Amplification of gene *amb0715* in qRT-PCR experiments |
| **0715RTR** | GTCTCCACAAGAAGGGGCAA | Amplification of gene *amb0715* in qRT-PCR experiments |
| **0716RTF** | CAAGCCGCACAAGATCATCG | Amplification of gene *amb0716* in qRT-PCR experiments |
| **0716RTR** | TCCGAGCCATAGCGGAAATC | Amplification of gene *amb0716* in qRT-PCR experiments |
| **0717RTF** | GATTCGGCTGTCTTCGTGGA | Amplification of gene *amb0717* in qRT-PCR experiments |
| **0717RTR** | ATGGGCTTTACCTTGGGAGC | Amplification of gene *amb0717* in qRT-PCR experiments |
| **0718RTF** | GACTACCCGGAGGATTACGC | Amplification of gene *amb0718* in qRT-PCR experiments |
| **0718RTR** | TGGCGTTGATATCGGCCAG | Amplification of gene *amb0718* in qRT-PCR experiments |
| **16SRTF** | GCTAATACCGCATACGCCCT | Amplification of gene *amb_r1* in qRT-PCR experiments |
| **16SRTR** | CAGACCAGCTACCGATCGTC | Amplification of gene *amb_r1* in qRT-PCR experiments |
| **AMB0685FW** | GGGGACAAGTTTGTACAAAAAAGCAGGCTTAGAAAACCTGTACTTCCA GGGTGCCGCAGCGCCGCAGATC | Amplification of *amb0685* to clone in pETG-20A for over-expression in *E.coli* |
| **AMB0685RV** | GGGGACCACTTTGTACAAGAAAGCTGGGTCTTATTAGTCATAATCCCA CCCCGTCACCTTATC | Amplification of *amb0685* to clone in pETG-20A for over-expression in *E.coli* |