

# A functional *Campylobacter jejuni maf4* gene results in novel glycoforms on flagellin and altered autoagglutination behaviour

Lieke B. van Alphen,<sup>1†</sup> Manfred Wuhrer,<sup>2</sup> Nancy M. C. Bleumink-Pluym,<sup>1</sup> Paul J. Hensbergen,<sup>2</sup> André M. Deelder<sup>2</sup> and Jos P. M. van Putten<sup>1</sup>

## Correspondence

Jos P. M. van Putten  
j.vanputten@uu.nl

<sup>1</sup>Department of Infectious Diseases and Immunology, Utrecht University, Yalelaan 1, 3584 CL Utrecht, The Netherlands

<sup>2</sup>Biomolecular Mass Spectrometry Unit, Department of Parasitology, Leiden University Medical Center, Albinusdreef 2, 2300 RC Leiden, The Netherlands

Flagellin of *Campylobacter jejuni* is extensively modified with (derivatives of) pseudaminic acid. The flagellar glycosylation locus contains several genes with homopolymeric G-tracts prone to slipped-strand mispairing, some of which belong to the *maf* gene family. We investigated the function of the putative phase-variable *maf4* gene of *C. jejuni* strain 108. A constructed *maf4* mutant displayed unaltered flagella assembly and bacterial motility. 2D-PAGE analysis revealed that the flagellin of strain 108 migrated at a more acidic pI than the protein of the Maf4 mutant. MS-MS in combination with high-resolution matrix-assisted laser desorption/ionization Fourier transform ion cyclotron MS (MALDI-FT-ICR-MS) on flagellin-derived glycopeptides showed that the flagellins of the mutant lacked two previously unidentified modifications of pseudaminic acid. These glycoforms carried additional CO<sub>2</sub> and C<sub>2</sub>H<sub>2</sub>O<sub>2</sub> groups, consistent with the more acidic pI of the wild-type flagellin. Phenotypically, the *maf4* mutant displayed strongly delayed bacterial autoagglutination. Collectively, our results suggest that the presence of a functional Maf4 expands the flagellin glycan repertoire with novel glycoforms of pseudaminic acid and, in the event of phase variation, alters the population behaviour of *C. jejuni*.

Received 23 April 2008

Revised 8 July 2008

Accepted 16 July 2008

## INTRODUCTION

Flagella are important bacterial organelles that enable swimming of microbes through watery environments and may contribute to bacterial pathogenesis by facilitating tissue colonization (Caldwell *et al.*, 1985). The flagellar apparatus is composed of three basic elements: the basal body that is embedded in the membrane and contains the flagellar rotor, a short filamentous hook structure that protrudes into the extracellular environment, and a long polymeric filament that is mainly built up of thousands of flagellin subunits. Each flagellin molecule typically consists of two  $\alpha$ -helical structures formed by the N- and C-

terminal regions of the protein that are buried in the core of the filament, and a central hypervariable surface-exposed domain (Samatey *et al.*, 2001). In some bacterial species the flagellin protein undergoes post-translational modifications (for a review, see Logan, 2006). The function of these modifications is unknown, but they contribute to the serospecificity of the flagellin (Alm *et al.*, 1991).

Post-translational modification of flagellin was first demonstrated in *Campylobacter* species (Logan *et al.*, 1989). Further characterization of flagellin using periodate treatment, specific lectins (Doig *et al.*, 1996), and, at a later stage, state-of-the-art chemical analysis (Logan *et al.*, 2002; McNally *et al.*, 2006, 2007; Thibault *et al.*, 2001) indicates the presence of O-linked carbohydrate residues. To date, the flagellin of *Campylobacter jejuni* strain 81-176 is known to be decorated predominantly with 5,7 diacetamido-3,5,7,9 tetra-deoxy-L-glycero-L-manno-nonulosonic acid (pseudaminic acid, Pse5Ac7Ac), which is attached to up to 19 different Ser/Thr residues in the flagellin. Substitution of the acetamido group of pseudaminic acid with acetamidino groups (Pse5Am7Ac), hydroxypropionyl groups (Pse5Pr7Pr), N-acetylglucosamine (Pse5Am7AcOGlnAc) or acetyl groups (Pse5Am7Ac8OAc) (Schirm *et al.*, 2005) causes further heterogeneity in the carbohydrate moieties

† Present address: Department of Biological Sciences, M469, Biological Sciences building, University of Alberta, Edmonton, AB T6G 2E9, Canada.

**Abbreviations:** MALDI-FT-ICR-MS, matrix-assisted laser desorption/ionization Fourier transform ion cyclotron MS; Pse5Ac7Ac, 5,7 diacetamido-3,5,7,9 tetra-deoxy-L-glycero-L-manno-nonulosonic acid (pseudaminic acid); Pse5Am7Ac, pseudaminic acid with acetamido group substituted with acetamidino groups.

The GenBank/EMBL/DDBJ accession numbers for the *maf4* gene and the glycosylation locus of the *neuB3* to *maf2* sequences of *Campylobacter jejuni* strain 108WT are EU448320 and EU448321, respectively.

(Thibault *et al.*, 2001). This heterogeneity is evident from the migration of flagellin as separate glycoforms during IEF (Logan *et al.*, 2002). Recently, other glycosyl groups, namely legionaminic acid and derivatives thereof, have been detected on the flagella of *Campylobacter coli* VC167 (McNally *et al.*, 2007). Thus, *Campylobacter* appears to have evolved a considerable repertoire of carbohydrate structures to decorate its flagella.

The genes known to contribute to the glyco-modification of *Campylobacter* flagellin are clustered in a single locus, the flagellar glycosylation locus (Guerry *et al.*, 2006). The locus varies in size in different *Campylobacter* strains and contains genes involved in the synthesis of Pse5Ac7Ac, as well as a family of up to seven very homologous genes, termed the Cj1318 or *maf* gene family. The function of these genes is under active investigation. *maf2* (*pseD*), but not *maf3*, *maf6* or *maf7*, in strain 81-176 appears to be involved in flagellin glycosylation (Guerry *et al.*, 2006), while *maf5* (*pseE*) in strain 11168 is required for flagella assembly (Karlyshev *et al.*, 2002). Analysis of the flagellin in a *pseD* mutant has demonstrated that it contains Pse5Ac7Ac but lacks Pse5Am7Ac on flagellin (Guerry *et al.*, 2006). However, intracellular CMP-Pse5Am7Ac was detected, suggesting a role in glycosyl transfer to flagellin and not in monosaccharide biosynthesis. This is in contrast to a *pseA* mutant, which lacks intracellular CMP-Pse5Am7Ac, suggesting a defect in biosynthesis (McNally *et al.*, 2006).

Some of the *maf* genes including the (identical) *maf1* and *maf4* genes of strain 11168 contain a homopolymeric G-tract that is prone to slipped-strand mispairing, which may result in a shift of the ORF (Karlyshev *et al.*, 2002). The exact functions of these *maf* genes are unknown, as in most strains they appear to be in the 'off' state or are not present at all. In the present study we investigated the function of the potentially phase-variable *maf4* gene in *C. jejuni* strain 108. In this strain the gene is in the 'on' state, and thus encodes a full-length Maf4 protein. Targeted gene inactivation combined with phenotypic analysis with respect to flagellin glycosylation, flagella-mediated bacterial agglutination, and the ability to cause infection of host cells,

revealed that Maf4 is not required for flagellar assembly, but does alter flagellin glycosylation and bacterial agglutination behaviour. Thus, phase variation of *maf4* may serve to change the functional properties of *Campylobacter* flagellin.

## METHODS

**Bacteria and cell culture.** Wild-type *C. jejuni* 108 (129108, 108WT; Endtz *et al.*, 1993) and its derivatives were routinely grown on saponin agar plates with 4% lysed horse blood or in 5 ml heart infusion (HI) broth under microaerophilic conditions at 37 °C, unless indicated otherwise. Chang (CCL-20.2, ATCC) and INT-407 (CCL-6, ATCC) epithelial cells were routinely maintained in RPMI (Gibco) supplemented with 5% fetal bovine serum (FBS) or Dulbecco's minimal essential medium (DMEM; Gibco) supplemented with 10% FBS in a humidified 37 °C atmosphere of 5 or 10% CO<sub>2</sub>, respectively.

**Analysis of the *maf* glycosylation locus.** The sequence of the glycosylation locus of 108WT from *neuB3* to *maf2* (GenBank accession no. EU448321) was determined (Baseclear) after amplification by PCR using primer sets Maf1 Fwd+NeuC Rev and NeuC Fwd+Maf2 Rev (Table 1), and subsequent cloning of the PCR fragments into pGEM-T easy. The sequence of *maf4* of strain 108WT (GenBank accession no. EU448320) was determined after amplification of the corresponding gene by PCR with primer set Maf4 Fwd+Maf4 Rev using *Taq* polymerase with proofreading (Invitrogen) and cloning of the resulting PCR product into pGEM-T easy. The primer set Maf4 Fwd+FlaB Rev was used to verify that the *maf3-flaB* region was of the expected size.

**Construction of *C. jejuni* 108 *maf4* mutants.** Genetic inactivation of *maf4* was accomplished by amplifying the corresponding gene and the flanking region by PCR with the primers Maf4 Fwd and Maf4 Rev. PCR fragments were cloned into the vector pGEM-T easy and transformed into *Escherichia coli* DH5 $\alpha$ . Purified plasmid was used as a template in an outward-directed PCR with primers  $\Delta$ Maf4 Fwd and  $\Delta$ Maf4 Rev, which yielded products that lacked ~260 bp of coding sequence and carried *Bam*HI restriction sites at the ends. These sites served to introduce a chloramphenicol-resistance cassette derived from pAV35 (van Vliet *et al.*, 1998). The resulting vector, pG*maf4*::cat, with the chloramphenicol cassette in the same orientation as *maf4*, was introduced into *C. jejuni* 108 WT and 108P4 via electroporation, and chloramphenicol (20  $\mu$ g ml<sup>-1</sup>)-resistant transformants were selected after 48 h growth on saponin agar plates. Targeted gene disruption was verified by PCR.

**Table 1.** Primers used in this study

Primer	Sequence
Maf1 Fwd	5'-GCAAGAAGCTTATATGCAAGTAAGG-3'
NeuC Rev	5'-CCATTCAGCCCTTGTGCACTGAC-3'
NeuC Fwd	5'-GTCAGTGCAACAAGGGCTGAATGG-3'
Maf2 Rev	5'-GGTGTCTGTTGTAGGTGGTGAT-3'
Maf4 Fwd	5'-GCTATTTATTTTCATAACGAATGCG-3'
Maf4 Rev	5'-TCCCCGCGGTCACCTTCTCACCCCTTCCA-3'
FlaB Rev	5'-TGTTAAAGCAGCTGAATCAACCA-3'
$\Delta$ Maf4 Fwd	5'-CGGGATCCGAATATTACTTAAAGCCTCAAT-3'
$\Delta$ Maf4 Rev	5'-CGGGATCC AATTCCATTTCCAAAACCATA-3'

**Motility assay.** To assess bacterial migration in semi-solid agar, broth-grown *C. jejuni* 108WT, 108 $\Delta$ maf4 and the non-flagellated mutant 108 $\Delta$ rpoN (Wösten, 1997) were stabbed into semi-solid medium (thioglycollate medium containing 0.4% agar). Swarming was assessed after incubation under microaerobic conditions at 37 °C for 24 h.

**Infection and gentamicin protection assay.** Infection experiments were carried out as described previously (van Alphen *et al.*, 2008) with some modifications. Briefly, for microscopic analysis, epithelial cells (75% confluence) were rinsed twice with serum-free medium and placed in 1 ml of this medium in a microaerobic incubator at 30 min prior to infection. Bacteria grown in HI broth for 16 h (OD<sub>550</sub> 1.2) were collected by centrifugation (3000 g, 10 min, 20 °C), resuspended in Dulbecco's PBS (DPBS), and added to the cells at an m.o.i. of 200. At 2 h of incubation (microaerophilic conditions, 37 °C), the cells were rinsed three times with 1 ml DPBS and fixed (>1 h, 20 °C) in 1.5% formaldehyde in DPBS for microscopic analysis (van Putten *et al.*, 1994). Bacterial subvasion was scored by multiplanar microscopy with counting of the number of bacteria per cell for 10 randomly selected microscopic fields. Bacterial viability during infection was assessed using the gentamicin protection assay. In this assay, 2 h-infected cells were rinsed three times with 1 ml DPBS, incubated (for 3 h) in 0.5 ml medium containing 250  $\mu$ g gentamicin ml<sup>-1</sup>, rinsed three times with 1 ml DPBS, lysed with 250  $\mu$ l 0.1% Triton X-100 in DPBS (15 min, 20 °C), and plated at various dilutions to quantify the number of viable bacteria. No viable bacteria were recovered from the culture supernatant after gentamicin treatment. Experiments were performed in duplicate and the mean  $\pm$  SEM of three separate experiments are presented.

**Autoagglutination (AAG) assays.** For AAG assays, bacteria grown in HI broth (16 h, 37 °C) were collected by centrifugation (3000 g, 5 min) and resuspended in DPBS to a final OD<sub>550</sub> of 0.5. Bacterial suspensions were transferred to glass tubes and kept at room temperature. At 45 min intervals, photographs were taken and 0.5 ml samples of the top layer of suspension were collected for measurement of OD<sub>550</sub>.

**Isolation of bacterial outer membranes.** Bacteria grown on saponin agar plates (16 h, 42 °C) or in HI broth (16 h, 37 °C) were collected and resuspended in 2 ml 50 mM Tris-HCl, pH 7.5. After ultrasonic disruption (60 s on ice), crude fragments were removed by centrifugation (12 000 g, 30 min, 4 °C) and bacterial membranes were collected (50 000 g, 1 h, 4 °C). Outer membranes were obtained by dissolving the inner membranes with 1% Sarkosyl in Tris-HCl buffer (10 min, 20 °C), followed by centrifugation (50 000 g, 2 h, 4 °C). The outer-membrane fraction was resuspended in Tris-HCl buffer and stored at -20 °C.

**2D gel electrophoresis.** For 2D electrophoresis, 50  $\mu$ g protein from bacterial outer membranes was mixed with rehydration solution containing 7 M urea, 2 M thiourea, 4% (w/v) CHAPS, 0.5% IPG buffer, pH 4–7 (GE Healthcare) and 0.3% (w/v) DTT to a final volume of 250  $\mu$ l. Proteins were resolved in the first dimension by IEF in an IPGphor (GE Healthcare) using 13 cm long, pH 4–7, precast immobilized nonlinear pH gradient strips (GE Healthcare). The IEF parameters were as follows: rehydration was carried out for 11 h at 30 V, followed by 500 V for 1 h, 1000 V for 1 h, 8000 V for 5 h, and 500 V for 30 min. Then, the strips were equilibrated (for 15 min) in 5 ml buffer I (50 mM Tris-HCl, pH 8.8, 6 M urea, 2% SDS, 30% w/v, glycerol and 10 mg DTT ml<sup>-1</sup>), and subsequently for 15 min in 5 ml buffer II (50 mM Tris-HCl, pH 8.8, 6 M urea, 2% SDS, 30% w/v, glycerol and 25 mg iodoacetamide ml<sup>-1</sup>). Strips were placed on 12.5% polyacrylamide gels and run for 30 min at 15 mA and then at 30 mA until the front reached the bottom of the gel, using an LBK

Hoefler Scientific cell apparatus. Proteins were visualized by staining with silver (Shevchenko *et al.*, 1996).

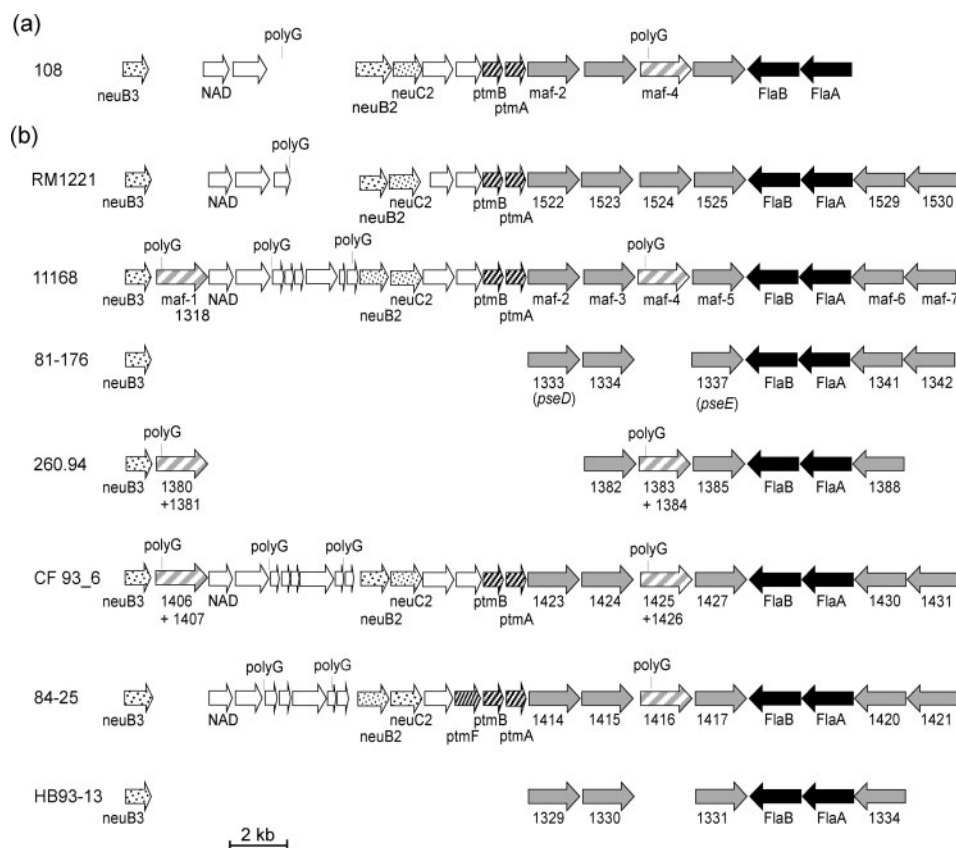
**Western blotting of flagellar proteins.** For Western blotting, the procedure described above for 2D gel electrophoresis was used, with modifications to suit a minigel system. Separated proteins were transferred to Hybond-C Extra nitrocellulose using a Bio-Rad miniblott system (GE Healthcare). Filters were incubated with anti-FlaAB antiserum and, subsequently, with horseradish-peroxidase-conjugated goat anti-rabbit IgG (Santa Cruz Biotechnology). Immune-reactive bands were visualized with Super Signal West Pico chemiluminescent substrate (Pierce).

**MS.** Coomassie-stained flagellin protein bands (64 kDa) were excised from gels after separation of outer-membrane proteins on 8% SDS-PAGE gels. Subsequently, flagellin proteins were reduced, alkylated and in-gel digested using trypsin (modified, sequencing grade, Promega) as described elsewhere (Steen *et al.*, 2002). After digestion, peptides were collected using two rounds of extraction with 0.1% trifluoroacetic acid and stored at -20 °C until further use. For the analysis of tryptic (glyco)peptides, samples were injected onto a Nano LC System (Ultimate, Dionex) equipped with a peptide trap column [Pepmap 100, 300  $\mu$ m inner diameter (i.d.)  $\times$  10 mm] and an analytical column (Pepmap 100, 75  $\mu$ m i.d.  $\times$  150 mm, Dionex). The mobile phases consisted of (A) 0.04% formic acid/0.4% acetonitrile and (B) 0.04% formic acid/90% acetonitrile. A 45 min linear gradient from 0 to 60% B was applied at a flow rate of 0.2  $\mu$ l min<sup>-1</sup>. The outlet of the LC system was coupled to an HCTUltra ion-trap mass spectrometer (Bruker Daltonics) using a nano-electrospray ionization source. The spray voltage was set at 1.2 kV and the temperature of the heated capillary was set to 165 °C. Eluting peptides were analysed in the data-dependent MS/MS mode over a 400–1600 *m/z* range. The five most abundant fragments in each MS spectrum were selected for MS/MS analysis by collision-induced dissociation. For an accurate mass analysis, the tryptic digest of bacterial flagellins of 108WT was separated by reverse-phase HPLC, and 3 min fractions were collected. Fractions were dried in a vacuum centrifuge, resuspended in 50  $\mu$ l water, and 1  $\mu$ l was spotted on a stainless steel target plate together with 1  $\mu$ l 2,5-dihydroxybenzoic acid (DHB; 20 mg ml<sup>-1</sup> in 30% acetonitrile), allowed to dry at ambient temperature and analysed by matrix-assisted laser desorption/ionization Fourier transform ion cyclotron MS (MALDI-FT-ICR-MS; ApexUltra equipped with a combisource, Bruker Daltonics). An internal calibration was performed for the spectrum obtained from fraction 5 using the theoretical masses of various non-glycosylated peptides detected in this fraction. Obtained accurate masses for the various glycoforms of peptide T<sub>465</sub>-K<sub>481</sub> of FlaA/FlaB were used for the calculation of the masses of the [M+H]<sup>+</sup> species of the glycan moieties. Suggestions for elemental compositions of the glycan moieties were obtained from these calculated masses using an online tool (<http://www.ch.cam.ac.uk/magnus/EadFormW.html>).

## RESULTS

### Identification of phase-variable *maf* genes

*In silico* analysis of seven sequenced *C. jejuni* flagellar glycosylation loci revealed the presence of a variable number of *maf* genes, with a minimum of four per strain (Fig. 1b). In four out of seven sequenced strains, one or two of the *maf* genes (named *maf1* and *maf4*) contain a homopolymeric G-tract prone to slipped strand mispairing. When all putative phase-variable *maf* genes were altered in their poly-G region to translate into the largest



**Fig. 1.** Schematic representation of the organization of the *maf* (1318) gene family within the flagellar glycosylation locus of (a) *C. jejuni* 108 and (b) sequenced *C. jejuni* strains. The numbers underneath the genes refer to the gene numbers in the corresponding strain. The bar represents 2.0 kb. polyG, location of a stretch of G residues. Variable shading of the arrows indicates differences in the predicted function and similarity of the genes: grey, *maf* family or related genes; hatched grey lines, phase-variable *maf* family genes; solid black, flagellin genes; hatched black lines, genes involved in flagellin modification; dots, putative *N*-acetylneuraminic acid synthetases; solid white, other genes.

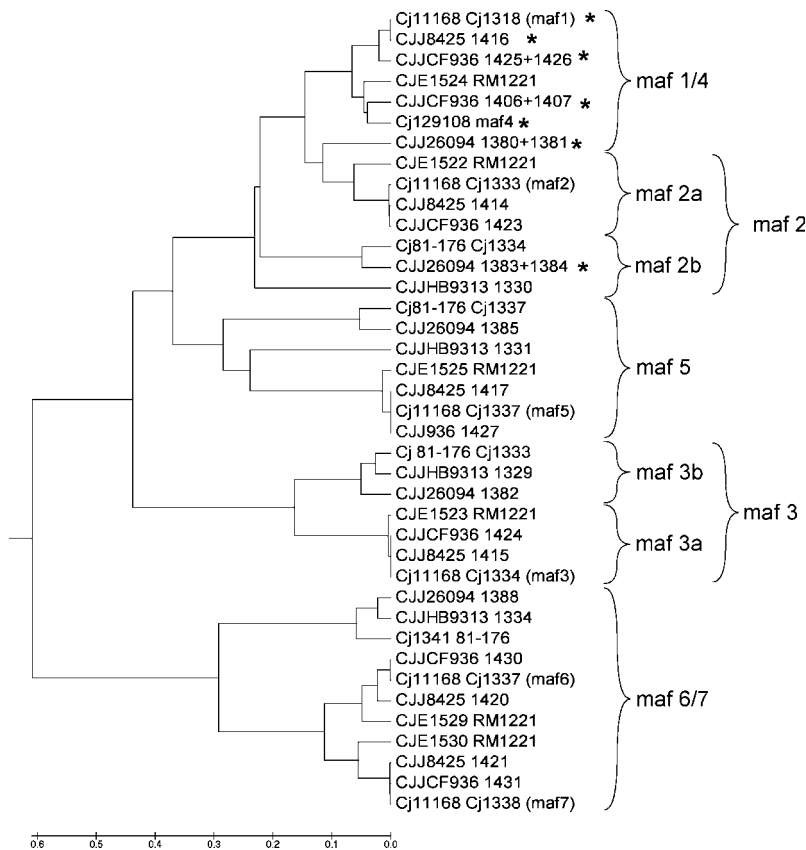
possible ORF, their alignment showed in all cases high similarity (63–99%) at the amino acid level to Maf1 of strain 11168. When two potentially phase-variable genes were present in one strain they had virtually identical sequences (with the exception of strain CF93\_6), suggesting that they evolved by gene duplication. Furthermore, the location of these *maf* genes within the glycosylation locus seemed conserved in all strains: either directly downstream of *neuB3*, which encodes the enzyme sialic acid synthase, or ~2000 bp downstream of the flagellin locus (Fig. 1b). Two exceptions in the characteristics of *maf1/4*-like genes were the presence of a phase-variable gene (Cj1383/1384) in strain 260.94 that was most similar to *maf3* rather than *maf1/4*, and the presence of a *maf4* gene with a seemingly deleted poly-G tract region in strain RM1221.

On the basis of phylogenetic and molecular evolutionary analyses conducted using MEGA version 3.1 (Kumar *et al.*, 2004), the *maf* genes fall into different clusters (Fig. 2) that sometimes do not follow the current gene nomenclature. For example, the *maf2* genes from strain 81-176 (Cj1333),

HB9313 (CJJ1329) and 260.94 (CJJ1382) appear to belong in one group with the *maf3* gene family of the other four strains, while the *maf3* genes of strains 81-176 (Cj1334), HB9313 (CJJ1330) and 260.94 (CJJ1383 + 1384) better fit the *maf2* gene cluster (Fig. 2). Furthermore, on the basis of sequence similarity, the *maf* genes located upstream of *flaA* form a distinct category and appear to belong to the *maf6* and *maf7* gene cluster (Fig. 2). However, for the potentially phase-variable *maf1/4* genes the current nomenclature was consistent with the MEGA analysis.

### Identification of *maf1/4*-like gene(s) in strain 108

PCR and sequence analysis of the glycosylation locus of *C. jejuni* strain 108WT downstream of *neuB3* demonstrated a similar organization to that in strain RM1221 (Fig. 1a). No *maf* gene was detected adjacent to *neuB3*, in contrast to the situation in several other strains including strain 11168 (Fig. 1b). A putative *maf4* gene was located at the same position in the glycosylation locus as in strains RM1221 and 11168. Sequencing of this gene indicated the presence of a poly-G



**Fig. 2.** Neighbour-joining tree of aligned *maf* genes of all sequenced strains. The phase-variable *maf* genes are indicated by asterisks. These *maf* genes form a separate cluster, with one exception (the *maf3* CjJ1383+1384 gene of strain 260.94). For strains in which *maf1* and *maf4* are gene duplications, only *maf1* is depicted.

tract in the 5'-terminal region of the gene. The gene contained an intact ORF that was highly similar at the amino acid level to the in-frame phase-variable *maf1/4* sequences of strain 11168, with the exception of two distinct regions of 12–50 amino acids in length (Fig. 3). These regions were most similar to the corresponding region of the potentially phase-variable *maf1* of strain 260.94 and the non-phase-variable *maf4* of strain RM1221 (Fig. 3). This architecture suggests that the Maf proteins may consist of variable peptide modules embedded in a relatively conserved protein backbone.

### Role of *maf4* in flagella assembly and bacterial motility

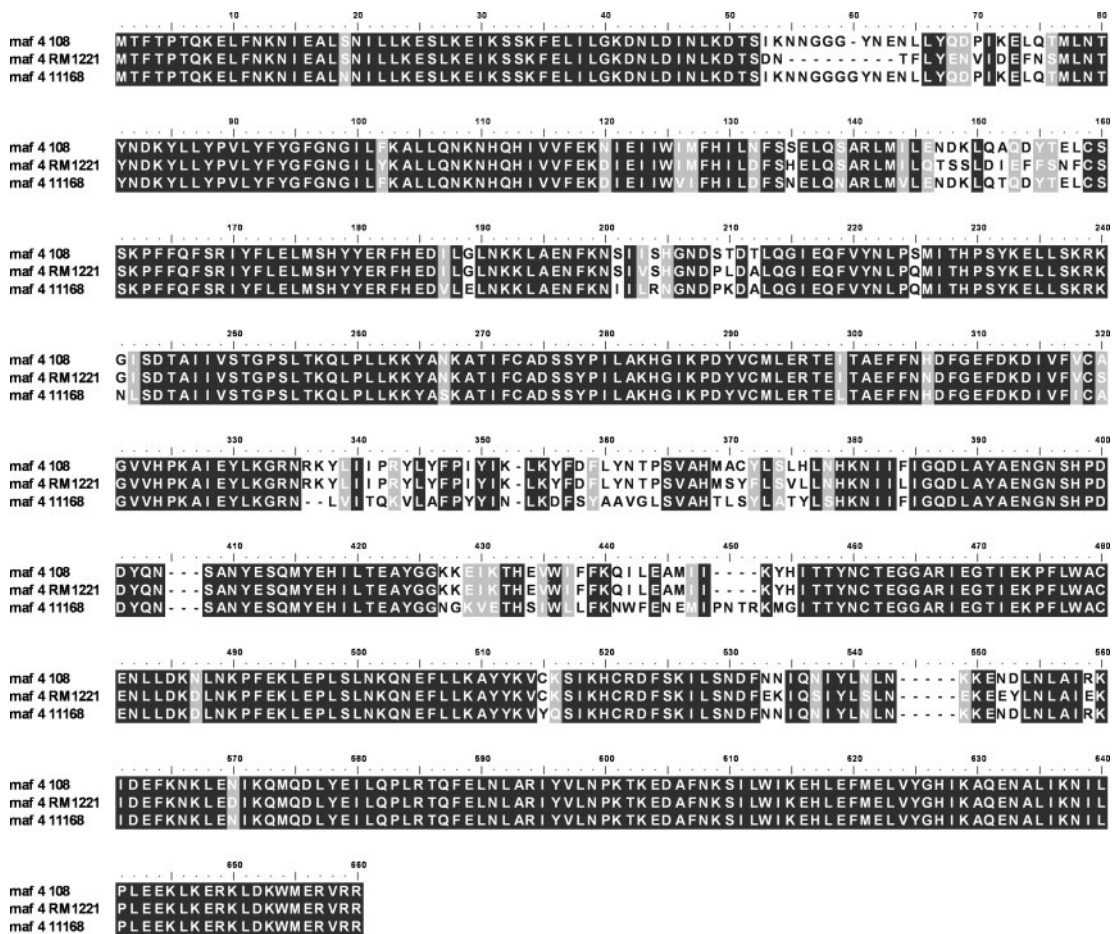
As strain 108 contained only a single *maf1/4* gene encoding a full-length protein (in contrast to strain 11168), we used this strain to investigate the function of the *maf4* gene in *C. jejuni* biology. Gene inactivation of *maf4* in strain 108 was accomplished by allelic replacement with a *maf4* gene copy in which 260 bp was deleted and replaced by a chloramphenicol-resistance cassette. PCR confirmed the correct insertion of the disrupted gene copy (not shown). Phenotypic analysis demonstrated that the 108*maf4*::*cat* mutant, designated 108Δ*maf4*, displayed similar swarming behaviour in semi-solid thioglycollate agar to the parent strain (Fig. 4a). These data indicate that Maf4 is not required either for flagella assembly or for bacterial motility, in contrast to Maf5 (Karlyshev *et al.*, 2002).

### Maf4 is involved in the glycosylation of *C. jejuni* flagellin

The finding that inactivation of *maf4* still allowed flagella formation enabled us to investigate the glycosylation status of the flagellin in the mutant strain. For this purpose, whole bacterial lysates of strain 108WT and 108Δ*maf4* were analysed by 2D-PAGE. This demonstrated that flagellin of strain 108WT was heterogeneous with respect to its pI, suggesting variable post-translational modifications. For the mutant strain a similar migration pattern was observed, except that the most dominant protein forms had shifted towards a more neutral pI (from 4.5–5.0 to 5.0–5.5) (Fig. 4b). This shift was more pronounced when bacteria were grown at 37 °C than at 42 °C (data not shown). No shift in apparent molecular mass of the flagellins was observed in the wild-type compared with the *maf4* mutant (Fig. 4b). Western blotting with a flagellin-specific mAb demonstrated that the shifted spots correspond to flagellin (data not shown). Furthermore, MS-MS of tryptic digests confirmed the spots to be FlaA and FlaB.

### Inactivation of *maf4* alters flagellin glycoforms

In order to elucidate molecular differences between the flagellins of strain 108WT and 108Δ*maf4*, outer-membrane fractions containing high amounts of flagellin were prepared and subjected to SDS-PAGE. Flagellin bands



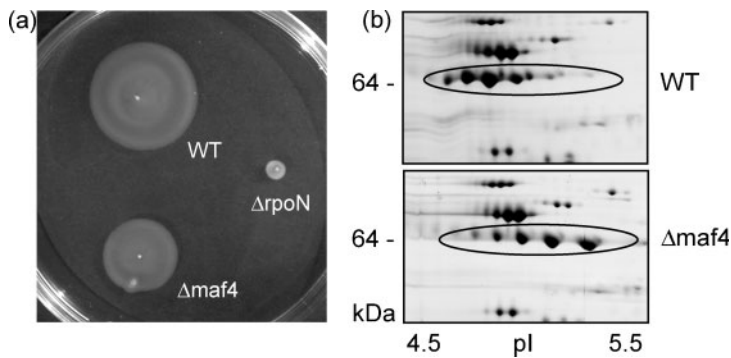
**Fig. 3.** Alignment of the Maf4 protein of strain 108 with the corresponding genes of strains RM1221 and 11168. Note that the highly conserved regions are interspersed by two highly variable regions, which show high similarity to *maf4* (CJE1524) of RM1221, but are different from *maf4* (Cj1318 or Cj1335+1336) of 11168. Identical amino acids are shaded black, similar amino acids grey.

were excised, in-gel-digested with trypsin, and subsequently analysed by nano reverse-phase HPLC ion-trap MS-MS (Fig. 5). The samples showed very similar patterns of both non-glycosylated peptides and glycopeptides. The obvious differences in the chromatogram (e.g. an additional peak for the mutant sample at 16 min; Fig. 5a) were caused by differences in the extent of methionine oxidation, most likely introduced during sample work-up, and could not explain the differences observed in isoelectric points observed in the 2D-PAGE gels.

Candidate glycopeptides were identified by their characteristic tandem mass spectra. Due to its pronounced lability the *O*-glycosidic linkage was predominantly cleaved on collisional activation, resulting in oxonium ions of the bacterial monosaccharides next to a prominent signal of the deglycosylated peptide moiety. A manual screening of the MS/MS dataset revealed many glycopeptides with one or two monosaccharides attached (Table 2). The prevailing monosaccharides were Pse5Ac7Ac and Pse5Am7Ac, which

were indicated by oxonium ions of  $m/z$  317 and  $m/z$  316, respectively.

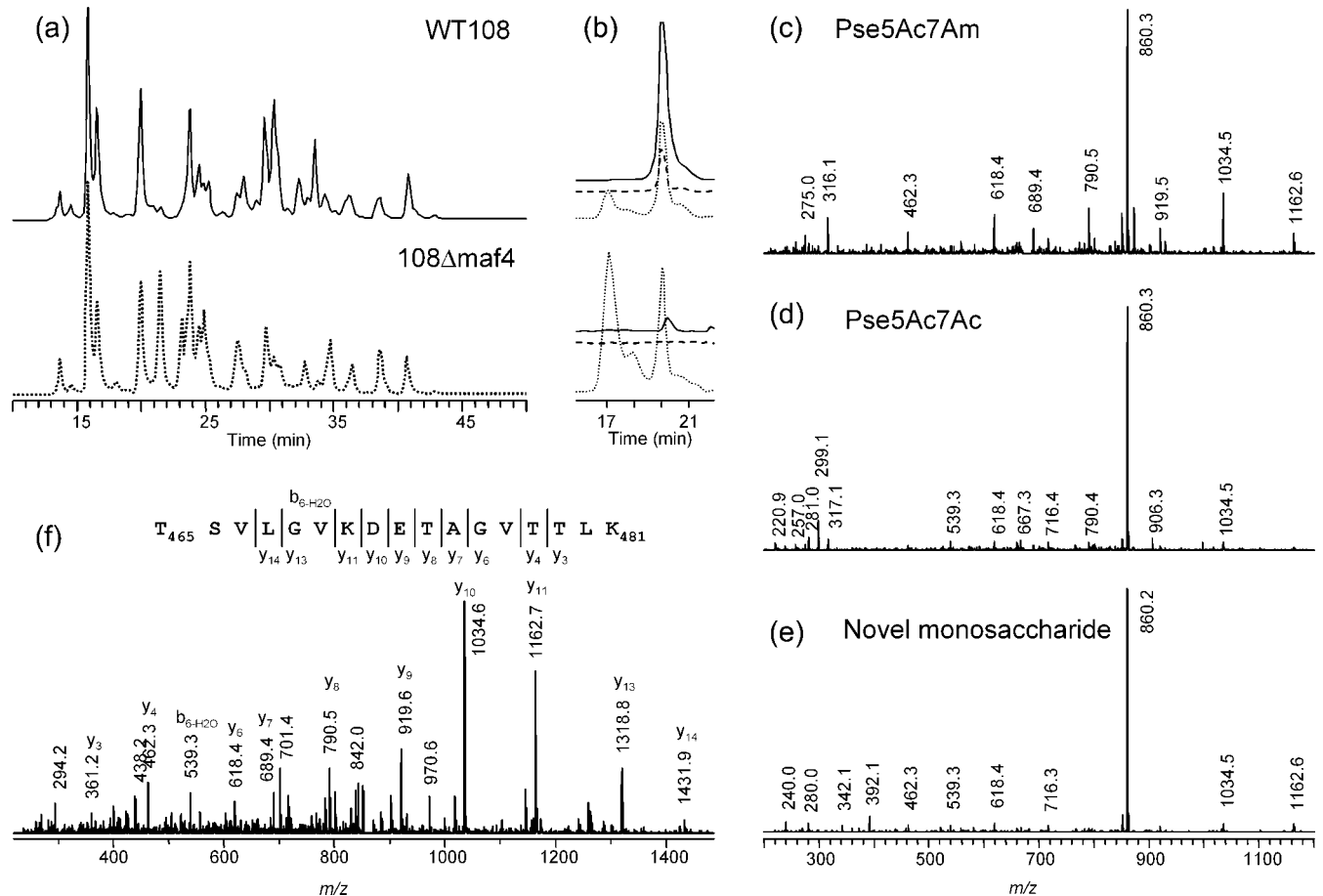
Screening of the LC-MS/MS dataset for possible differences between the flagellins of strains 108WT and 108 $\Delta$ maf4 indicated differential glycosylation of the peptide T<sub>465</sub>SVLGVK<sub>471</sub>, which is shared between FlaA and FlaB. In strain 108WT, this peptide stretch was in part occupied by Pse5Am7Ac and Pse5Ac7Ac, resulting in glycopeptide ions (peptide moiety T<sub>465</sub>-K<sub>481</sub>, one missed tryptic cleavage site) of  $m/z$  678.7 (at 17 min) and  $m/z$  679.1 (at 19 min), respectively. These glycoforms were also detected for 108 $\Delta$ maf4, yet in different ratios and at higher abundances, as evidenced by the extracted ion chromatograms shown in Fig. 5(b). Moreover, in strain 108WT two additional glycoforms of the peptide were observed which were absent in 108 $\Delta$ maf4. The prevailing glycoform ( $m/z$  693 [M+3H]<sup>3+</sup>, continuous line, and  $m/z$  1039 [M+2H]<sup>2+</sup>, dashed line; Fig. 5b) contained a carbohydrate residue, which, to the best of our knowledge, has not been



**Fig. 4.** (a) Swarming assay for strain 108WT, the 108 $\Delta maf4$  mutant ( $\Delta maf4$ ) and the non-flagellated mutant 108 $\Delta rpoN$  ( $\Delta rpoN$ ). (b) 2D analysis of outer-membrane protein spots of both strains as visualized with silver. Encircled spots represent flagellin. Note the shift of the flagellar spots of 108 $\Delta maf4$  to a less acidic pI.

described before. MS-MS of this glycopeptide led to an oxonium ion, which, after elimination of water, resulted in a signal at  $m/z$  342 (Fig. 5e). The deglycosylated peptide

was observed at  $m/z$  860 ( $[M+2H]^{2+}$ ), which was also the dominant fragment observed for the glycoforms with Pse5Am7Ac (Fig. 5c) and Pse5Ac7Ac (Fig. 5d). The



**Fig. 5.** Nano reverse-phase HPLC ion-trap MS-MS analysis of bacterial flagellins after tryptic digestion. Tryptic (glyco-) peptides from flagellins of strain 108WT and the 108 $\Delta maf4$  mutant were analysed by nano reverse-phase HPLC ion-trap MS-MS. (a) Base-peak chromatograms for wild type (upper panel) and mutant (lower panel). (b) Extracted-ion chromatograms of  $m/z$  693 (continuous line),  $m/z$  1039 (dashed line) and  $m/z$  678/679 (dotted line), representing three glycoforms of the glycopeptide  $T_{465}-K_{481}$  of Flag A. (c–e) Ion trap tandem mass spectra of triple-charged glycoforms of  $T_{465}-K_{481}$  of  $m/z$  678.7 (c),  $m/z$  679.1 (d) and  $m/z$  693.4 (e) observed for the wild-type sample. (f)  $MS^3$  spectrum obtained from the double-charged fragment at 860.2 from panel (e), confirming the identity of the  $T_{465}-K_{481}$  peptide moiety.

**Table 2.** Analysis of tryptic glycopeptides from flagellins by nano reverse-phase HPLC ion-trap MS-MS

Glycopeptide signal ( <i>m/z</i> )	Deduced [M+H] <sup>+</sup>	MS-MS glycan marker ions	Peptide moiety	Flagellin	Glycan B-ion
510.3 [M+2H] <sup>2+</sup>	1019.6	240.1, 281.1, 299.1, 317.1	703.1 [M+H] <sup>+</sup> T <sub>465</sub> SVLGVK <sub>471</sub> (confirmed with MS <sup>3</sup> )	A, B	317.1
531.8 [M+2H] <sup>2+</sup>	1062.6	240.1, 280.1, 324.1, 342.1, 360.2	703.1 [M+H] <sup>+</sup> T <sub>465</sub> SVLGVK <sub>471</sub>	A, B	<b>360.1</b>
509.3 [M+4H] <sup>4+</sup> ; 678.7 [M+3H] <sup>3+</sup> ; 1017.6 [M+2H] <sup>2+</sup>	2034.2	275.1, 316.1	860.0 [M+2H] <sup>2+</sup> T <sub>465</sub> SVLGVKDETAGVTTLK <sub>481</sub>	A, B	316.1
679.1 [M+3H] <sup>3+</sup> ; 1018.1 [M+2H] <sup>2+</sup>	2035.2	299.1, 317.1	860.1 [M+2H] <sup>2+</sup> T <sub>465</sub> SVLGVKDETAGVTTLK <sub>481</sub>	A, B	317.1
693.4 [M+3H] <sup>3+</sup> ; 1039.6 [M+2H] <sup>2+</sup>	2078.2	240.0, 280.1, 324.1, 342.1	860.1 [M+2H] <sup>2+</sup> T <sub>465</sub> SVLGVKDETAGVTTLK <sub>481</sub>	A, B	<b>360.1</b>
698.0 [M+3H] <sup>3+</sup> ; 1046.6 [M+2H] <sup>2+</sup>	2092.2	374.1	860.1 [M+2H] <sup>2+</sup> T <sub>465</sub> SVLGVKDETAGVTTLK <sub>481</sub>	A, B	<b>374.1</b>
532.2 [M+3H] <sup>3+</sup> ; 797.9 [M+2H] <sup>2+</sup>	1594.8	275.0, 299.1, 316.1, 334.1	1279.7 [M+H] <sup>+</sup> I <sub>180</sub> TSGGEVQFTLK <sub>191</sub>	A, B	316.1
532.6 [M+3H] <sup>3+</sup> ; 798.4 [M+2H] <sup>2+</sup>	1595.8	281.0, 299.0, 317.1	1279.7 [M+H] <sup>+</sup> I <sub>180</sub> TSGGEVQFTLK <sub>191</sub>	A, B	317.1
532.2 [M+3H] <sup>3+</sup> ; 753.8 [M+3H] <sup>3+</sup> ; 1130.1 [M+2H] <sup>2+</sup>	1594.6 2259.2	275.0, 299.1, 316.1, 334.1 275.1, 316.1, 334.1	1279.7 [M+H] <sup>+</sup> I <sub>180</sub> TSGGEVQFTLK <sub>191</sub> 1944.0 [M+H] <sup>+</sup> V <sub>204</sub> VISTSVGTGLGALADEINK <sub>223</sub>	A, B A	 316.1
754.1 [M+3H] <sup>3+</sup> ; 1130.6 [M+2H] <sup>2+</sup>	2260.2	281.0, 299.1, 317.1	972.6 [M+2H] <sup>2+</sup> ; 1944.0 [M+H] <sup>+</sup> V <sub>204</sub> VISTSVGTGLGALADEINK <sub>223</sub>	A	317.1
950.5 [M+4H] <sup>4+</sup> ; 1266.9 [M+3H] <sup>3+</sup>	3798.7	275.0, 299.0, 316.1,	Loss of one glycan: 1161.8 [M+3H] <sup>3+</sup> N <sub>335</sub> DGKDILVSGTGLTAAGFGANSFISQASISLR <sub>366</sub>	A, B	316.1, two gly- cans
950.8 [M+4H] <sup>4+</sup>	3802.2	275.0, 281.0, 299.0, 316.1, 317.1	Loss of one glycan: 1161.8 [M+3H] <sup>3+</sup> N <sub>335</sub> DGKDILVSGTGLTAAGFGANSFISQASISLR <sub>366</sub>	A, B	316.1, 317.1
951.0 [M+4H] <sup>4+</sup>	3803.0	281.0, 299.1, 317.1	Loss of one glycan: 1161.8 [M+3H] <sup>3+</sup> N <sub>335</sub> DGKDILVSGTGLTAAGFGANSFISQASISLR <sub>366</sub>	A, B	317.1, two gly- cans
846.9 [M+4H] <sup>4+</sup> ; 1128.9 [M+3H] <sup>3+</sup>	3384.7	275.0, 316.1	D <sub>339</sub> ILVSGTGLTAAGFGANSFISQASISLR <sub>366</sub>	A, B	316.1, two gly- cans
847.2 [M+4H] <sup>4+</sup> ; 1129.2 [M+3H] <sup>3+</sup>	3385.8	281.0, 299.1, 317.1	D <sub>339</sub> ILVSGTGLTAAGFGANSFISQASISLR <sub>366</sub>	A, B	317.1, two gly- cans
748.1 [M+3H] <sup>3+</sup> ; 561.3 [M+4H] <sup>4+</sup>	2242.3	275.0, 288.1, 299.0, 316.1	963.8 [M+2H] <sup>2+</sup> F <sub>174</sub> ETGGRITSGGEVQFTLK <sub>191</sub>	A, B	316.1
748.4 [M+3H] <sup>3+</sup>	2243.2	ND*	F <sub>174</sub> ETGGRITSGGEVQFTLK <sub>191</sub>	A, B	317.1
896.8 [M+3H] <sup>3+</sup>	2688.4	281.1, 299.0, 317.1	1186.6 [M+2H] <sup>2+</sup> V <sub>204</sub> VISTSVGTGLGALADEINK NADK <sub>227</sub>	A	317.1
758.8 [M+3H] <sup>3+</sup>	2274.4	281.1, 299.0, 317.1	979.5 [M+2H] <sup>2+</sup> V <sub>204</sub> VISTSVGTGLGALAEINK <sub>223</sub>	B	317.1

\*ND, Not determined



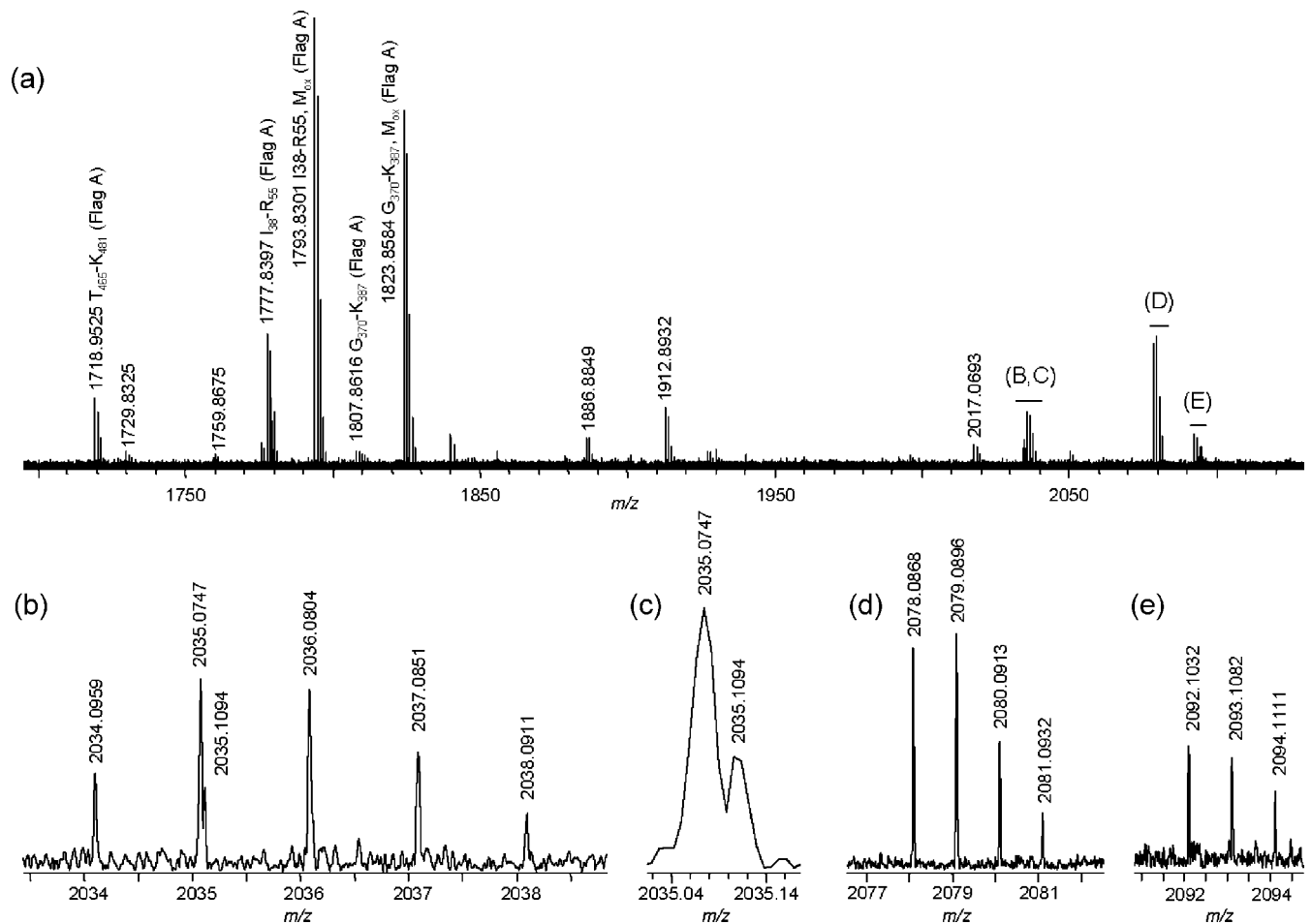
identity of the peptide was confirmed by performing an addition ion selection/fragmentation cycle with the fragment at  $m/z$  860 (MS<sup>3</sup> experiment), which provided a series of Y ions with peptide sequence information (Fig. 5f). Another, minor glycoform, which was likewise exclusively observed for the 108 parent strain, resulted in an oxonium ion at  $m/z$  374 upon MS-MS, next to the peptide signal ( $m/z$  860).

### High-resolution MS of the novel flagellin glycoforms

To characterize the novel flagellar monosaccharides in more detail the tryptic digest of the flagellin of strain 108WT was fractionated by reverse-phase HPLC and analysed by MALDI-FT-ICR-MS. The masses of the various glycoforms of T<sub>465</sub>-K<sub>481</sub> were determined with sub-p.p.m. mass accuracy (Fig. 6a, Table 3). The signals for

the glycopeptides with Pse5Am7Ac and Pse5Ac7Ac were overlapping (Fig. 6b): two peaks at  $m/z$  2035.0747 and  $m/z$  2035.1094 could be partially resolved, representing the first isotope peak of Pse5Am7Ac and the second isotope peak of Pse5Ac7Ac, respectively (Fig. 6c). The non-glycosylated form of peptide T<sub>465</sub>-K<sub>481</sub> was detected at  $m/z$  1718.9525. The two novel glycoforms were detected at  $m/z$  2078.0868 (Fig. 6d) and  $m/z$  2092.1032 (Fig. 6e). From these data the masses of the corresponding glycan moieties (oxonium ions) were calculated ( $m/z$  360.1406 and  $m/z$  374.1570; Table 3) and used to search for possible molecular compositions.

As the modification was expected to be a derivative of Pse5Ac/Am7Ac, only molecular compositions within the range C<sub>13–20</sub>H<sub>18–30</sub>N<sub>1–8</sub>O<sub>5–10</sub> were taken into account. Within the expected mass accuracy range of  $\pm 2$  mDa the only matching molecular composition for the glycan of  $m/z$



**Fig. 6.** MALDI-FT-ICR-MS analysis of the glycoforms of the tryptic peptide T<sub>465</sub>-K<sub>481</sub> of FlaA. The tryptic digest of bacterial flagellins was fractionated by reverse-phase HPLC. The fraction containing the glycoforms of the tryptic peptide T<sub>465</sub>-K<sub>481</sub> was analysed by MALDI-FT-ICR-MS (a). In addition to the glycoforms with Pse5Am7Ac ( $m/z$  2034.0959) and Pse5Ac7Ac (2035.0747) (b, c), two additional glycoforms were determined representing novel pseudaminic acid derivatives ( $m/z$  2078.0868 and  $m/z$  2092.1032) (d, e). The accurate masses were used to deduce the molecular composition of the attached glycan moieties (see Table 3).

**Table 3.** Analysis of the glycoforms of the differentially glycosylated peptide by MALDI-FT-ICR-MS

Registered masses [M + H] <sup>+</sup>	Deduced mass of modification [M + H] <sup>+</sup>	Compound	Theoretical mass [M + H] <sup>+</sup>	Mass accuracy
1718.9525	–	T <sub>465</sub> SVLGKDETAGVTTLK <sub>481</sub> (Flag A)	1718.9535	0.0010 mDa (0.6 p.p.m.)
2017.0693	299.1238	T <sub>465</sub> –K <sub>481</sub> with C <sub>13</sub> H <sub>18</sub> O <sub>6</sub> N <sub>2</sub>	2017.0700	0.0007 mDa (0.3 p.p.m.)
2034.0959	316.1497	T <sub>465</sub> –K <sub>481</sub> with C <sub>13</sub> H <sub>21</sub> O <sub>6</sub> N <sub>3</sub> (Pse5Am7Ac)	2034.0965	0.0007 mDa (0.3 p.p.m.)
2035.0747; 2038.0911 (fourth isotope peak)	317.1348	T <sub>465</sub> –K <sub>481</sub> with C <sub>13</sub> H <sub>20</sub> O <sub>7</sub> N <sub>2</sub> (Pse5Ac7Ac)	2035.0805; 2038.0906 (fourth isotope peak)	0.0005 mDa (0.2 p.p.m.)*
2078.0868	360.1406	T <sub>465</sub> –K <sub>481</sub> with C <sub>14</sub> H <sub>21</sub> O <sub>8</sub> N <sub>3</sub>	2078.0863	0.0005 mDa (0.2 p.p.m.)
2092.1032	374.1570	T <sub>465</sub> –K <sub>481</sub> with C <sub>15</sub> H <sub>23</sub> O <sub>8</sub> N <sub>3</sub>	2092.1020	0.0012 mDa (0.6 p.p.m.)

\*Calculated for the fourth isotope peak.

360.1406 was C<sub>14</sub>H<sub>21</sub>O<sub>8</sub>N<sub>3</sub>. The compositional difference between this glycan (C<sub>14</sub>H<sub>21</sub>O<sub>8</sub>N<sub>3</sub>) and Pse5Am7Ac (C<sub>13</sub>H<sub>21</sub>O<sub>6</sub>N<sub>3</sub>) is CO<sub>2</sub>, indicating that the 360 kDa glycoform may be Pse5Am7Ac substituted with additional carboxylic acid group. This modification fits with the lower average pI of 108WT flagellins compared with 108Δmaf4 flagellins in 2D gels (Fig. 4). For the 374 Da glycan (calculated oxonium ion of *m/z* 374.1570), two possible matches were found: C<sub>15</sub>H<sub>23</sub>O<sub>8</sub>N<sub>3</sub> with a calculated mass of *m/z* 374.1558 (Δ*m*=1.2 mDa), and C<sub>18</sub>H<sub>22</sub>O<sub>5</sub>N<sub>4</sub> with a calculated mass of *m/z* 374.1585 (Δ*m*=1.5 mDa). The first option (C<sub>15</sub>H<sub>23</sub>O<sub>8</sub>N<sub>3</sub>) seems the more probable one, as this would mean a compositional difference of CH<sub>2</sub> compared with the 360 Da glycan. Overall, the results indicate that Maf4 alters the glycosylation of the flagellin in the peptide region T<sub>465</sub>–K<sub>481</sub> by enabling the addition (or formation) of Pse5Am7Ac substituted with an additional carboxyl group, resulting in a more acidic glycopeptide.

### Maf4 influences *C. jejuni* autoagglutination kinetics

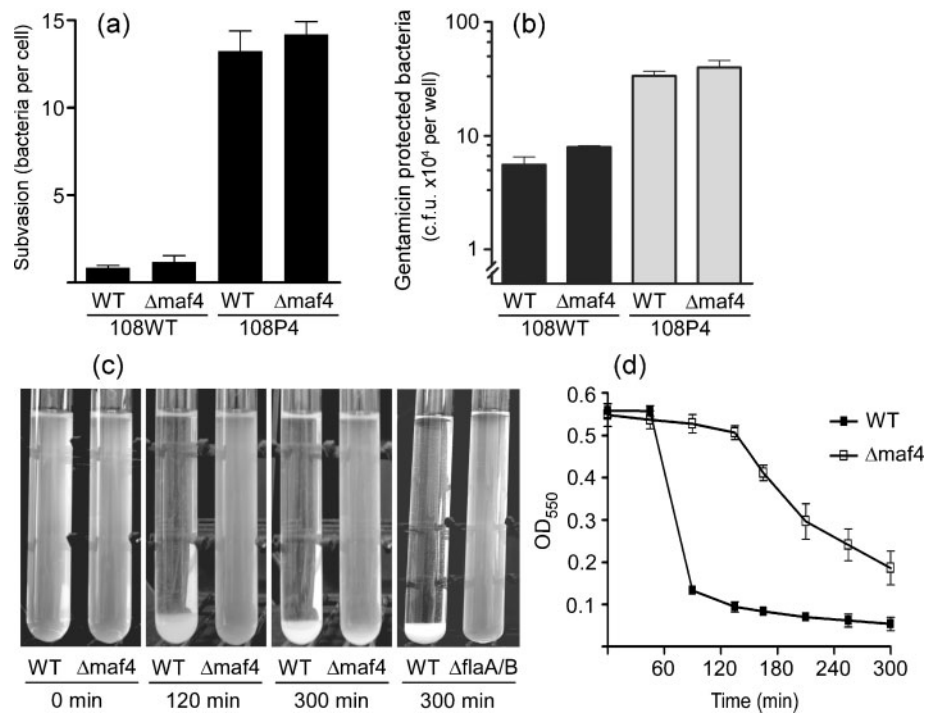
What are the functional consequences of the presence of a functional Maf4? This question was addressed first by testing the *in vitro* behaviour of a highly invasive derivative of strain 108WT (designated 108P4) (van Alphen *et al.*, 2008). Sequencing of the poly-G region of *maf4* of 108P4 confirmed that no slipped-strand mispairing had occurred during the selection of this strain and that the gene was in the ‘on’ state (data not shown). To analyse the role of *maf4* in strain 108P4 the gene was inactivated as described for 108WT. 2D-PAGE analysis of the flagellins of strains 108P4 and 108P4Δmaf4 demonstrated a similar shift in pI to that observed for strain 108WT and its *maf4*-negative derivative (data not shown), suggesting that the observed effects in both mutants are caused by inactivation of the *maf4* gene rather than by changes in other phase-variable genes. Infection assays with 108P4Δmaf4 revealed unaltered (high) subversion behaviour in Chang epithelial cells compared to 108WT (Fig. 7a). Gentamicin protection

assays on infected cells using INT-407 cells indicated a slight (but consistent) increase in the number of recovered bacteria for the *maf4*-negative derivatives of both the highly invasive strain 108P4 and the low invasive strain 108WT (Fig. 7b).

We also investigated the effect of Maf4 on the auto-agglutination of *C. jejuni*, as this trait has been associated with changes in flagellar glycosylation (Guerry *et al.*, 2006). As shown in Fig. 7(c), incubation (2 h) of a suspension of strain 108WT of OD 0.5 resulted in strong agglutination with the formation of a clear bacterial pellet at the bottom of the tube. In contrast, strain 108Δmaf4 was still largely in suspension at this time point (Fig. 7c). Measurement of the ODs of the bacterial suspension at different time points confirmed the altered agglutination behaviour of the *maf4* mutant (Fig. 7d). At 5 h of incubation both strain 108WT and 108Δmaf4 had agglutinated (Fig. 7d), in contrast to the strain lacking flagellin. These results suggest that the Maf4-induced post-translational modification influences the kinetics of the flagellin-mediated bacterial agglutination, and thus that the *maf4* gene in the ‘on’ state may alter the population behaviour of *C. jejuni*.

## DISCUSSION

In the present study we identified a functional, potentially phase-variable gene (*maf4*) involved in post-translational modification of the flagellin of the bacterial pathogen *C. jejuni*. Sequencing of *maf4* from strain 108WT indicated that it contained a homopolymeric stretch of nine G residues, resulting in an ORF encoding full-length protein in contrast to the homologous *maf1* of strain 11168. Evidence that Maf4 is involved in flagellin glycosylation includes the altered migration of flagellin isoforms in 2D gels after inactivation of *maf4*, and MS analysis of flagellin tryptic glycopeptides. Functional studies indicate that Maf4 functionality influences flagella-mediated bacterial auto-agglutination. This suggests that, in the event of phase variation, variable *maf4* expression alters the behaviour of *C. jejuni*, which may add to the virulence of the pathogen.



**Fig. 7.** Infection and autoagglutination assays. (a, b) Infection assays of epithelial cells. Subversion of Chang epithelial cells (a) and invasion of INT-407 epithelial cells (b) of strain 108WT, the highly subvasive/invasive strain 108P4, and their corresponding *maf4* mutants ( $\Delta maf4$ ). Data represent the mean  $\pm$  SEM of three experiments. (c, d) Autoagglutination assays. Suspensions of strain 108WT and 108 $\Delta maf4$  were incubated for up to 5 h at room temperature. (c) Autoagglutination at 0, 2 and 5 h of incubation is depicted to illustrate the difference in kinetics of pellet formation between the wild-type and the mutant strain. The non-flagellated RpoN mutant served as a control for flagella-dependent autoagglutination. (d) Kinetics of the autoagglutination as measured by the drop in OD<sub>550</sub> over a 5 h period. Values represent the mean  $\pm$  SEM of three separate experiments.

Comparative genome analysis indicated that *C. jejuni* carries up to seven different *maf* genes that are all part of the *C. jejuni* flagellar glycosylation locus. The variable number of *maf* genes in different *C. jejuni* strains suggests that the *maf* gene repertoire is highly diverse, although analysis based on amino acid similarity allows clustering into distinct groups (Fig. 2). In most cases, the different types of *maf* genes are located at the same position in the flagellar glycosylation locus. Comparison of related *maf* genes from different strains, such as *maf4* from strains 108 and 11168, suggests that they are relatively conserved except in distinct regions. This seemingly modular architecture resembles that of certain families of enzymes, such as glycosyltransferases (Kapitonov & Yu, 1999). In these cases, the variable domains contribute to the substrate specificity of the enzymes.

A search for *maf* genes carrying homopolymeric nucleotide tracts indicated that some, but not all, strains contain one or two potentially phase-variable genes. Interestingly, the homopolymeric stretch of nucleotides was in all except one strain located in the *maf1* and/or *maf4* gene, suggesting that these genes are particularly attractive loci for the bacterium to undergo phase variation. The finding that

*maf1/4* genes with both intact and truncated ORFs exist in nature suggests that slipped-strand mispairing occurs and is not harmful to the bacterium. Indeed, phase variation of *maf1* (which identical to *maf4*) has been demonstrated to occur for strain 11168 (Karlyshev *et al.*, 2002). Why *maf1/4* is phase variable in contrast to, e.g. *maf5*, is unclear, but may be related to the apparently essential function of *maf5* in flagella assembly (Karlyshev *et al.*, 2002).

The unchanged swarming behaviour of our *maf4* mutant in semisolid agar (Fig. 4a) indicates that the gene is not essential for motility of *C. jejuni*. This is surprising, as *maf* genes have been implicated as determinants of bacterial motility in strain 11168 (Karlyshev *et al.*, 2002), although inactivation of *maf2* and *maf3* does not result in loss of motility in strain 81-176 (Guerry *et al.*, 2006; McNally *et al.*, 2006). The finding that insertional inactivation of *maf4* still allowed flagella assembly enabled us to investigate the glycosylation status of the flagellin in the mutant strain. Comparison of the flagellins of 108WT and 108 $\Delta maf4$  by 2D-PAGE analysis demonstrated multiple flagellin spots with different pI values, consistent with the existence of multiple glycoforms in both strains (Thibault *et al.*, 2001). However, the protein spots from 108 $\Delta maf4$  flagellins

migrated on average toward the less acidic region of the IEF gel. This observation fits with MS results regarding the apparent decoration of the flagellin of strain 108WT, but not that of 108 $\Delta$ maf4, with Pse5Am7Ac possessing an additional CO<sub>2</sub> or C<sub>2</sub>H<sub>2</sub>O<sub>2</sub> moiety. The more acidic glycoforms were found for the peptide region T<sub>465</sub>–K<sub>481</sub>, which is shared between FlaA and FlaB. As only part of the bacterial O-glycosylation sites was monitored in our LC-MS approach, additional sites carrying these glycans may exist, which likewise may contribute to the differences in pI of the bacterial flagellins of the two strains. Considering the nature of the modification of Pse5Am7Ac in the Maf4-positive strain, it is possible that Maf4 facilitates additional carboxylation of nucleotide-activated Pse5Am7Ac. Alternatively, the protein may have no role in the biosynthesis but may specifically transfer the novel Pse5Am7Ac derivative(s) to the flagellin protein. Also, insertional inactivation of the *maf2* (*pseD*) gene does not interfere with the biosynthesis of pseudaminic acid precursors (McNally *et al.*, 2006) but results in a specific lack of Pse5Am7Ac but not Pse5Ac7Ac glycans on flagellin (Guerry *et al.*, 2006), consistent with substrate-specific transfer of the glycans.

The exact function of flagellar glycosylation in *C. jejuni* biology is not well understood. Our results indicate a clear difference in flagella-mediated autoagglutination between the flagellated strain 108WT and its *maf4*-defective derivative (Fig. 7). Autoagglutination of *C. jejuni* has been reported to influence flagella typing methods based on lectin interactions (Wong *et al.*, 1985), co-agglutination (Wong *et al.*, 1985) and serospecificity (Lior *et al.*, 1981). The self-aggregation is associated with the presence of flagella and with cell surface hydrophobicity (Misawa & Blaser, 2000). The role of flagella has been confirmed in a random transposon mutagenesis screen on *C. jejuni* 480 that showed autoagglutination to be influenced by genes in the flagellar glycosylation locus (Golden & Acheson, 2002). Because of the absence of autoagglutination after a loss of flagellar Pse5Am7Ac (but not Pse5Ac7Ac) residues (Guerry *et al.*, 2006), it has been proposed that interactions between both glycan forms across flagellar filaments are crucial in the aggregative interaction. However, it has been difficult to study this interaction, as bacteria unable to decorate their flagellins fail to build a functional flagellum (Karlyshev *et al.*, 2002). Our results indicate that the addition of CO<sub>2</sub> (*m/z* 360.1) or C<sub>2</sub>H<sub>2</sub>O<sub>2</sub> (*m/z* 374.1) groups to Pse5Ac7Ac strongly accelerates bacterial autoagglutination, suggesting a perhaps even more intense interaction of the more acidic Pse5Ac7Ac flagellin with other glycoforms. In other studies, a correlation has been found between autoagglutination and bacterial invasion of INT-407 cells (Guerry *et al.*, 2006). In our hands, similar numbers of gentamicin-protected bacteria were recovered from INT-407 cells for parental strains 108WT and 108P4 and their *maf4*-negative derivatives. Furthermore, no differences in subvase behaviour (van Alphen *et al.*, 2008) could be detected between the parental and Maf4-mutant strains.

Here, we demonstrated for the first time, to our knowledge, that the presence of a functional *maf4* gene of *C. jejuni* results in a novel glycan modification of the flagellin and enhanced bacterial autoagglutination compared with a *maf4*-negative strain. On the basis of these results it can be expected that variable functionality of Maf4, as occurs in the event of phase variation due to the changes in the poly-G tract in the *maf4* gene, influences the glycosylation status of flagellin and alters the population behaviour of *C. jejuni*.

## ACKNOWLEDGEMENTS

The authors thank Carolien A. M. Koeleman and Irina Dragan for expert technical assistance.

## REFERENCES

- Alm, R. A., Guerry, P., Power, M. E., Lior, H. & Trust, T. J. (1991). Analysis of the role of flagella in the heat-labile Lior serotyping scheme of thermophilic *Campylobacters* by mutant allele exchange. *J Clin Microbiol* **29**, 2438–2445.
- Caldwell, M. B., Guerry, P., Lee, E. C., Burans, J. P. & Walker, R. I. (1985). Reversible expression of flagella in *Campylobacter jejuni*. *Infect Immun* **50**, 941–943.
- Doig, P., Kinsella, N., Guerry, P. & Trust, T. J. (1996). Characterization of a post-translational modification of *Campylobacter* flagellin: identification of a sero-specific glycosyl moiety. *Mol Microbiol* **19**, 379–387.
- Endtz, H. P., Giesendorf, B. A., van Belkum, A., Lauwers, S. J., Jansen, W. H. & Quint, W. G. (1993). PCR-mediated DNA typing of *Campylobacter jejuni* isolated from patients with recurrent infections. *Res Microbiol* **144**, 703–708.
- Golden, N. J. & Acheson, D. W. (2002). Identification of motility and autoagglutination *Campylobacter jejuni* mutants by random transposon mutagenesis. *Infect Immun* **70**, 1761–1771.
- Guerry, P., Ewing, C. P., Schirm, M., Lorenzo, M., Kelly, J., Pattarini, D., Majam, G., Thibault, P. & Logan, S. (2006). Changes in flagellin glycosylation affect *Campylobacter* autoagglutination and virulence. *Mol Microbiol* **60**, 299–311.
- Kapitonov, D. & Yu, R. K. (1999). Conserved domains of glycosyltransferases. *Glycobiology* **9**, 961–978.
- Karlyshev, A. V., Linton, D., Gregson, N. A. & Wren, B. W. (2002). A novel paralogous gene family involved in phase-variable flagella-mediated motility in *Campylobacter jejuni*. *Microbiology* **148**, 473–480.
- Kumar, S., Tamura, K. & Nei, M. (2004). MEGA3: integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinform* **5**, 150–163.
- Lior, H., Woodward, D. L., Edgar, J. A. & LaRoche, L. J. (1981). Serotyping by slide agglutination of *Campylobacter jejuni* and epidemiology. *Lancet* **2**, 1103–1104.
- Logan, S. M. (2006). Flagellar glycosylation – a new component of the motility repertoire? *Microbiology* **152**, 1249–1262.
- Logan, S. M., Trust, T. J. & Guerry, P. (1989). Evidence for posttranslational modification and gene duplication of *Campylobacter* flagellin. *J Bacteriol* **171**, 3031–3038.
- Logan, S. M., Kelly, J. F., Thibault, P., Ewing, C. P. & Guerry, P. (2002). Structural heterogeneity of carbohydrate modifications affects serospecificity of *Campylobacter* flagellins. *Mol Microbiol* **46**, 587–597.

- McNally, D. J., Hui, J. P., Aubry, A. J., Mui, K. K., Guerry, P., Brisson, J. R., Logan, S. M. & Soo, E. C. (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., Aubrey, A. J., Hui, J. P., Khieu, N. H., Whitfield, D., Ewing, C. P., Guerry, P., Brisson, J. R., Logan, S. M. & Soo, E. C. (2007). Targeted metabolomics analysis of *Campylobacter coli* VC167 reveals legionaminic acid derivatives as novel flagellar glycans. *J Biol Chem* **282**, 14463–14475.
- Misawa, N. & Blaser, M. J. (2000). Detection and characterization of autoagglutination activity by *Campylobacter jejuni*. *Infect Immun* **68**, 6168–6175.
- Samatey, F. A., Imada, K., Nagashima, S., Vonderviszt, F., Kumasaka, T., Yamamoto, M. & Namba, K. (2001). Structure of the bacterial flagellar protofilament and implications for a switch for supercoiling. *Nature* **410**, 331–337.
- Schirm, M., Schoenhofen, I. C., Logan, S. M., Waldron, K. C. & Thibault, P. (2005). Identification of unusual bacterial glycosylation by tandem mass spectrometry analyses of intact proteins. *Anal Chem* **77**, 7774–7782.
- Shevchenko, A., Wilm, M., Vorm, O. & Mann, M. (1996). Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. *Anal Chem* **68**, 850–858.
- Steen, H., Kuster, B., Fernandez, M., Pandey, A. & Mann, M. (2002). Tyrosine phosphorylation mapping of the epidermal growth factor receptor signaling pathway. *J Biol Chem* **277**, 1031–1039.
- Thibault, P., Logan, S. M., Kelly, J. F., Brisson, J. R., Ewing, C. P., Trust, T. J. & Guerry, P. (2001). Identification of the carbohydrate moieties and glycosylation motifs in *Campylobacter jejuni* flagellin. *J Biol Chem* **276**, 34862–34870.
- van Alphen, L. B., Bleumink-Pluym, N. M. C., Rochat, K. D., Van Balkom, B. W. M., Wösten, M. M. S. M. & van Putten, J. P. M. (2008). Active migration into the subcellular space precedes *Campylobacter jejuni* invasion of epithelial cells. *Cell Microbiol* **10**, 53–66.
- van Putten, J. P. M., Weel, J. F. & Grassme, H. U. (1994). Measurements of invasion by antibody labeling and electron microscopy. *Methods Enzymol* **236**, 420–437.
- van Vliet, A. H., Wooldridge, K. G. & Ketley, J. M. (1998). Iron-responsive gene regulation in a *Campylobacter jejuni* *fur* mutant. *J Bacteriol* **180**, 5291–5298.
- Wong, K. H., Skelton, S. K., Patton, C. M., Feeley, J. C. & Morris, G. (1985). Typing of heat-stable and heat-labile antigens of *Campylobacter jejuni* and *Campylobacter coli* by coagglutination. *J Clin Microbiol* **21**, 702–707.
- Wösten, M. M. S. M. (1997). *Initiation of transcription and gene organization in Campylobacter jejuni*. PhD thesis, Utrecht University.

Edited by: P. H. Everest