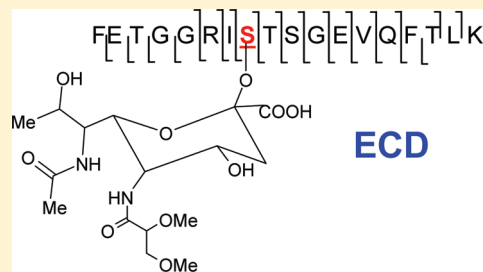


Novel Glycosylation Sites Localized in *Campylobacter jejuni* Flagellin FlaA by Liquid Chromatography Electron Capture Dissociation Tandem Mass Spectrometry

Cleidiane G. Zampronio, Gemma Blackwell, Charles W. Penn, and Helen J. Cooper*

School of Biosciences, College of Life and Environmental Sciences, University of Birmingham, Edgbaston, Birmingham, B15 2TT, U.K.

ABSTRACT: Glycosylation of flagellin in *Campylobacter jejuni* is essential for motility and virulence. It is well-known that flagellin from *C. jejuni* 81–176 is glycosylated by pseudaminic acid and its acetamidino derivative, and that *Campylobacter coli* VC167 flagellin is glycosylated by legionaminic acid and its derivatives. Recently, it was shown, by use of a metabolomics approach, that *C. jejuni* 11168 is glycosylated by dimethyl glyceric acid derivatives of pseudaminic acid, but the sites of glycosylation were not confirmed. Here, we apply an online liquid chromatography electron capture dissociation (ECD) tandem mass spectrometry approach to localize sites of glycosylation in flagellin from *C. jejuni* 11168. Flagellin A is glycosylated by a dimethyl glyceric acid derivative of pseudaminic acid at Ser181, Ser207 and either Thr464 or Thr 465; and by a dimethyl glyceric acid derivative of acetamidino pseudaminic acid at Ser181 and Ser207. For comparison, on-line liquid chromatography collision-induced dissociation of the tryptic digests was performed, but it was not possible to assign sites of glycosylation by that method.



KEYWORDS: glycosylation, flagellin, *Campylobacter*, ECD, mass spectrometry, FT-ICR, FT-MS

INTRODUCTION

The food-borne pathogen *Campylobacter jejuni* is one of the leading causes of bacterial gastroenteritis worldwide. *Campylobacter* are characterized by a rapid, darting motility, mediated by bipolar flagella, which is essential for virulence. The flagellum comprises a basal body linked to the flagellar filament, which consists of thousands of copies of the flagellin proteins FlaA and FlaB, with FlaA being the major component. Glycosylation of flagellins appears to be essential for flagellar biogenesis, as mutations in Cj1293, encoding a putative UDP-GlcNAc C6-dehydratase/C4-reductase involved in flagellin glycosylation, were aflagellate and nonmotile.¹ The Cj1293 orthologue in *Campylobacter coli* VC167 was not essential for flagellin glycosylation or flagellin biogenesis but a double mutant in this gene and *ptmD* involved in pseudaminic acid biosynthesis was also aflagellate and nonmotile.¹ These observations indicate that flagellin glycosylation with pseudaminic acid and derivatives is essential for targeting and/or secretion of flagellin in *C. jejuni* and *C. coli*. More recently, Howard et al.² have demonstrated the importance of specific structural modifications to the flagellin glycoform in the biological fitness of *C. jejuni* in colonization of chickens.

In 2001, Thibault et al.³ analyzed intact flagellins from three strains of *C. jejuni* (81–176, NCTC 11168 and OH4384) and one strain of *C. coli* (VC167 T2) by mass spectrometry. They showed that, in each case, the mass of the monomeric glycoprotein was ~6 kDa greater than that predicted from the primary sequence. Liquid chromatography mass spectrometry (LC–MS) of tryptic peptides of *C. jejuni* 81–176 flagellin FlaA revealed the presence of seven glycopeptides. The nature of the glycans was

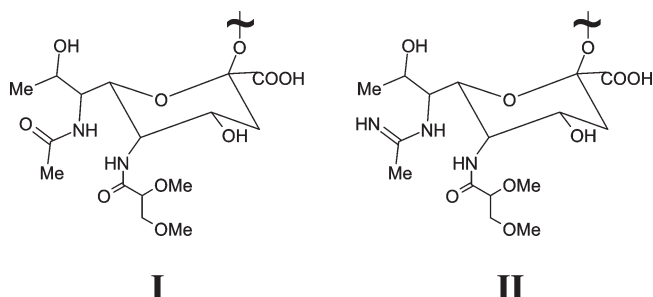
assigned following collision-induced dissociation (CID) tandem mass spectrometry (MS/MS) of HPLC fractions containing the suspected glycopeptides. The major glycosylation component was pseudaminic acid (Pse5Ac7Ac), with 5-acetamidino pseudaminic acid (Pse5Am7Ac) and 5,7-*N*-(2,3-dihydroxypropionyl)-pseudaminic acid (Pse5Pr7Pr) also present. Sites of glycosylation were determined by β -elimination reactions and subsequent MS/MS. In total, 19 serine and threonine glycosylation sites were identified on *C. jejuni* 81–176 FlaA. In further work, Thibault and co-workers applied a top-down approach for the identification of glycosylation in *C. jejuni* 81–176 flagellin.⁴ Again, the major glycans were Pse5Ac7Ac and Pse5Am7Ac. In addition, novel glycans, Pse5Am7Ac8GlnAc and Pse5Ac7Ac8OAc, were identified and localized to tryptic peptide [390–422]. The exact sites of the novel modification, however, were not determined.

Soo and co-workers⁵ developed a targeted metabolomics approach for the analysis of glycosylation in *C. jejuni* 81–176, confirming the presence of pseudaminic acid and its acetamidino derivative. Cell lysates were investigated for the presence of sugar-nucleotide metabolites by combining hydrophilic interaction liquid chromatography (HILIC) with precursor ion scanning mass spectrometry and the structures of the metabolites were confirmed by NMR spectroscopy. The method was subsequently applied to *C. coli* VC167.⁶ In addition to pseudaminic acid, acetamidino legionaminic acid (Leg5Am7Ac) and *N*-methylacetimidoyl legionaminic (Leg5-AmNMe7Ac) were identified as flagellar glycans.

Received: October 11, 2010

Published: December 15, 2010

Scheme 1. Structures of I and II



Most recently, the metabolomics approach was applied to the analysis of *C. jejuni* 11168 flagellin.⁷ Although 11168 was the first *C. jejuni* strain to have its genome sequenced,⁸ the least is known about its flagellin glycosylation. Moreover, its glycosylation locus is far more complex than either *C. jejuni* 81–176 or *C. coli* VC167. Logan et al.⁷ identified two novel glycans, see Scheme 1, corresponding to dimethylglyceric acid derivatives of pseudaminic acid and 7-acetamidino pseudaminic acid, hereafter referred to as I and II, respectively. Tryptic peptide [204–223] was shown to be modified by either I or II but the site(s) of glycosylation were not confirmed. Tryptic peptide [464–480] was shown to be modified by I. Again, the site of modification was not localized. Hitchen et al.⁹ used a ‘knock-in’ approach to demonstrate that gene Cj1295 is responsible for the presence of the dimethylglyceric acid derivative of pseudaminic acid.

Here, we have applied on-line reversed-phase liquid chromatography electron capture dissociation (ECD) tandem mass spectrometry¹⁰ to the characterization of glycosylation of flagellin A from *C. jejuni* 11168. ECD^{11–13} is a radically driven fragmentation technique which results in cleavage of the peptide backbone N–C α to produce *c* and *z* \bullet (or *c* \bullet and *z*) fragments. Unlike the more commonly used collision-induced dissociation (CID), labile post-translational modifications, such as glycosylation, are retained on the peptide backbone fragments following ECD.^{14,15} It is possible, therefore, to localize sites of modification unambiguously. We show that flagellin A is glycosylated by I at Ser181, Ser207 and either Thr464 or Thr465, and by II at Ser181 and Ser207. Glycosylation of Ser181 of flagellin A from *C. jejuni* 11168 has not been observed previously, although LC–MS/MS analysis of flagellins from the knockout strain 11168 Cj1295::aphA and the ‘knock-in’ strain 11168 Cj1295::aphA cmCj1295 revealed the tryptic peptide [179–190] to be modified.⁹ Interestingly, both Ser181 and Ser207 fall within the sequence motif ISTS (S is modified) suggesting the possibility of a consensus sequence for targeting of glycosylation. The equivalent site (also Ser207) has been shown to be modified by the glycans Pse5Ac7Ac and Pse5Am7Ac in flagellin A from *C. jejuni* strain 81–176.³ Finally, ECD revealed that flagellin A is glycosylated by I at Thr464 or Thr465. The equivalent residues in flagellin A from *C. jejuni* strain 81–176 have not been observed to be glycosylated.³ There was no evidence for modification of these residues by glycan II. Additionally, there was no evidence for the modification of flagellin A from strain 11168 by any of the glycans known for other strains of *Campylobacter*, for example, pseudaminic acid, legionaminic acid. For comparison, on-line liquid chromatography collision-induced dissociation tandem mass spectrometry was performed. Although the glycopeptides were observed, it was not possible to localize the sites of modification due to the preferential loss of the glycan modification.

EXPERIMENTAL SECTION

Preparation of Flagellin A

The method was based on that of Power et al.¹⁶ Cultures were grown in Mueller–Hinton broth for 24 h in an atmosphere of 5% O₂, 10% CO₂, and 85% N₂, with shaking at 37 °C. Cells were recovered from approximately 200 mL of culture by centrifugation at 6000g for 20 min at 4 °C and resuspended to a calculated OD of 20. Suspensions were homogenized at 20 500 rpm for 3 min using an Ultra-Turrax T-25 high speed homogenizer (IKA Labortechnik) before removal of cells and debris by centrifugation at 6000g for 20 min at 4 °C. The supernatant was then subjected to ultracentrifugation at 100 000g for 60 min at 5 °C. The pellet was resuspended in 1% (w/v) sodium dodecyl sulfate in distilled water to solubilize membrane fragments, incubated at 37 °C for 2 h, and again subjected to ultracentrifugation as above. The cycle of SDS washing and ultracentrifugation was repeated three times and purity checked by SDS-PAGE and electron microscopy before suspending in water and ultracentrifugation as above for three more cycles to remove SDS. Flagellar fragments were stored in water at –20 °C until required for analysis.

Flagellar proteins isolated from *C. jejuni* strain 11168 were separated by polyacrylamide gel (12%) electrophoresis and visualized with Coomassie stain R250 (BioRad). Bands with a molecular mass of approximately 60 kDa were cut out of the gel and stored at –20 °C until required for trypsinisation.

Trypsin Digestion

Digestion was performed by use of a Qiagen 3000 robot (Qiagen Sample & Assay Technologies, German). The gel plugs were destained and dehydrated using pure acetonitrile from J.T. Baker (Philipsburg, NJ) and 100 mM ammonium bicarbonate, (Sigma-Aldrich, St. Louis, MO). Cysteines were reduced with 10 mM dithiothreitol (Sigma-Aldrich, St. Louis, MO) and alkylated with 50 mM iodoacetamide (Sigma-Aldrich, St. Louis, MO). Samples were digested overnight with Trypsin Gold (Promega, Madison, WI), 12.5 ng/ μ L in 25 mM ammonium bicarbonate, at 37 °C.

Liquid Chromatography Tandem Mass Spectrometry (LC–MS/MS)

On-line reversed-phase liquid chromatography was performed using a Micro AS autosampler and Surveyor MS pump (Thermo Fisher Scientific, Germany). Peptides were separated using a binary solvent system consisting of (A) water (J.T. Baker, Holland) and 0.1% formic acid and (B) acetonitrile (J.T. Baker, Holland) and 0.1% formic acid. Peptides were loaded onto a 75 μ m (internal diameter) Integragrit (New Objective) C8 resolving column (length 10 cm) and separated over a linear 40 min gradient from 5% to 40% (B). Peptides eluted directly (\sim 300 nL min^{–1}) via a TriVersa Nanomate nanospray source (Advion, NY) into a 7 T LTQ FT mass spectrometer (Thermo Fisher Scientific).

ECD. The mass spectrometer alternated between a full FT-MS scan (m/z 380–2000) and a subsequent ECD MS/MS scan of the most abundant ions. Survey scans were acquired in the ICR cell with a resolution of 100 000 at m/z 400. Precursor ions were isolated in the ion trap and transferred to the ICR cell. Isolation width was 6 m/z and only multiply charged precursor ions were selected for MS/MS. Dynamic exclusion was used with a repeat count of 1 and exclusion duration of 180 s. Automatic gain control (AGC) was used to accumulate sufficient precursor ions (target value 1×10^6 , maximum fill time 4 s). The electrons for

MGFR**INTNVA** **ALNAKANADL** **NSKSLDASLS** **RLSSGLRINS** **AADDASGMAI** **ADSLRSQANT** 60
LQQATSNQND **ALGILQTDK** **AMDEQLK** **ILD** **TIKTKATQAA** **QDGQSLK** **TRT** **MLQADINRLM** 120
EELDNIA **NTT** **SFNGKQLLSG** **NFINQEFQIG** **ASSNQTVKAT** **IGATQSSKIG** **LTRFETGGRI** 180
STSGEVQFTL **KNYNGIDDFQ** **FQKVVISTSV** **GTGLGALADE** **INKNADKTGV** **RATFTVETRG** 240
IAAVRAGATS **DTFAINGVKI** **GKVDYKDGDA** **NGALVAAINS** **VKDTTGVEAS** **IDANGQLLLT** 300
SREGRGIKID **GNIGGGAFIN** **ADMKENYGR** **SLVKNDGKDI** **LISGSNLSSA** **GFGATQFISQ** 360
ASVSLRESKG **QIDANIADAM** **GFGSANKGVV** **LGYSVVSAY** **MSSAGSGFSS** **GSYVSGSGK** 420
NYSTGFANAI **AISAASQLST** **VYNVSAGSGF** **SSGSTLSQFA** **TMKTTAFGVK** **DETAGVTTLK** 480
GAMAVMDIAE **TAITNLDQIR** **ADIGSVQNVQV** **TSTINNITVT** **QVNVKAAESQ** **IRDVDFAAES** 540
ANYSKANILA **QSGSYAMAQA** **NSVQQNVLR** **LQ**

Figure 1. Protein sequence coverage obtained following LC-ECD-MS/MS analysis of tryptic digest of flagellin A from *C. jejuni* 11168. Identified peptides are shown in inverse bold.

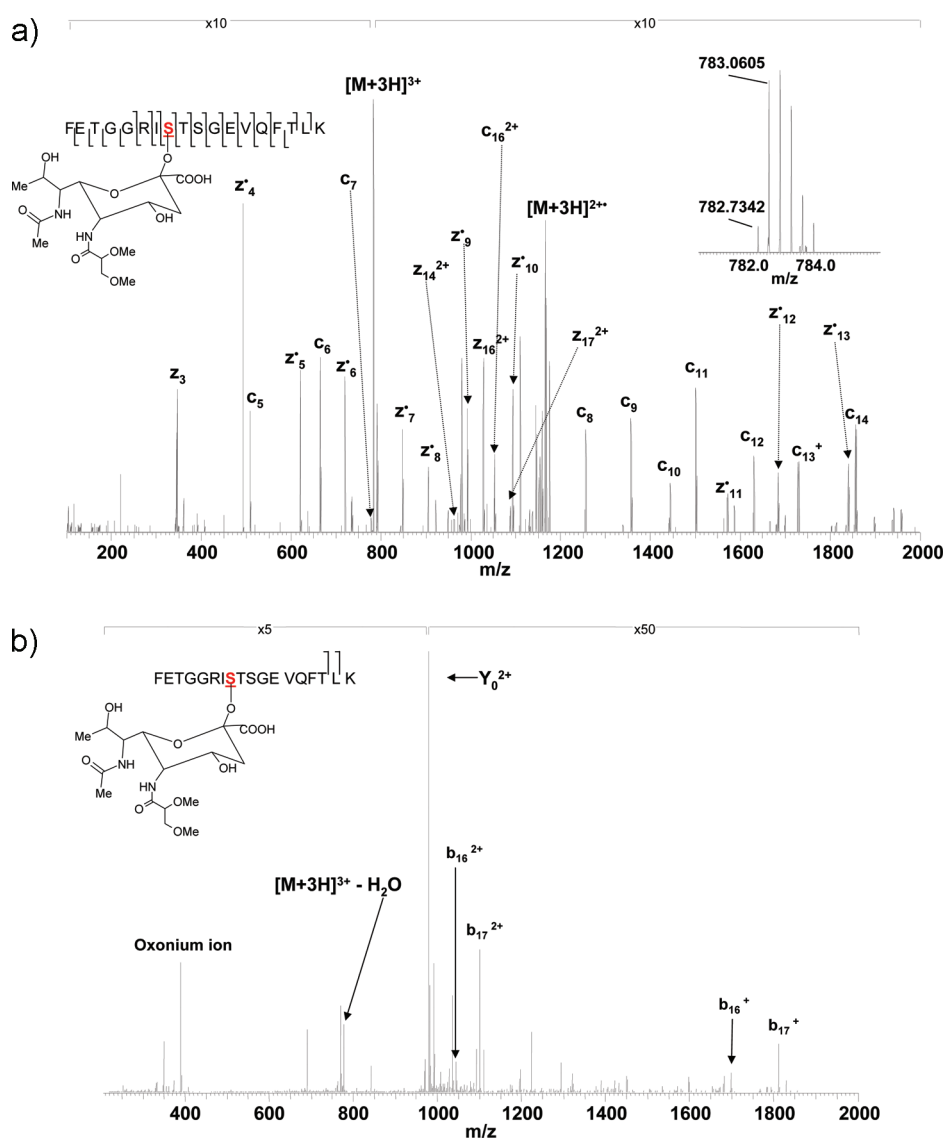
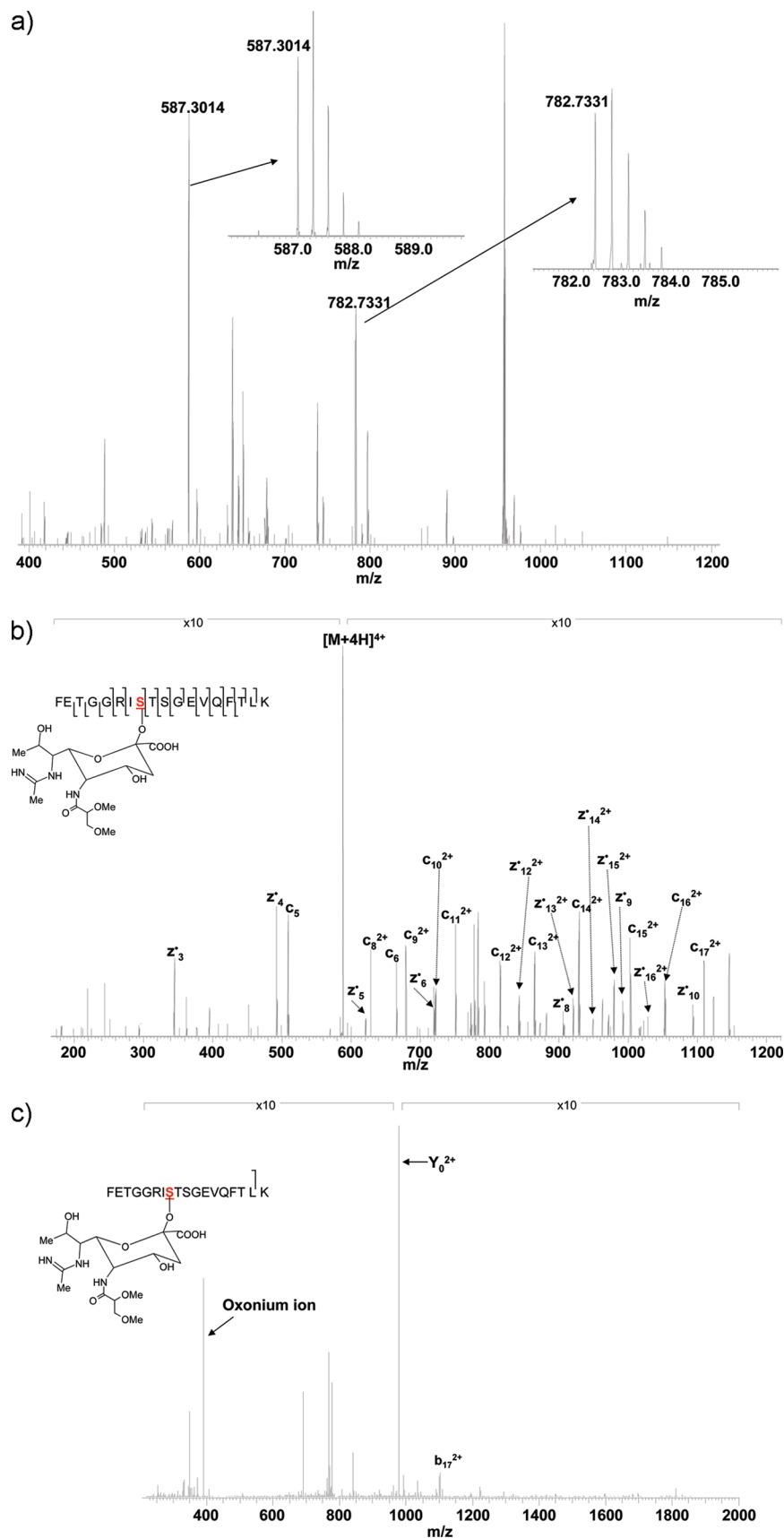


Figure 2. (a) ECD mass spectrum of 3+ ions of I-modified glycopeptide FETGGRISTSQGEVQFTLK [174–191]. Inset: isolated precursor ions showing measured monoisotopic m/z and peptide sequence coverage obtained. (b) CID mass spectrum of 3+ ions of I-modified glycopeptide FETGGRISTSQGEVQFTLK [174–191]. Inset: peptide sequence coverage obtained.



MGFRINTNVA	ALNAKANSDL	NAKSLDASLS	RLSSGLRINS	AADDASGMAI	ADSLRSQANT	60
MGFRINTNVA	ALNAKANADL	NSKSLDASLS	RLSSGLRINS	AADDASGMAI	ADSLRSQANT	60
LGQAISNGND	ALGILQTADK	AMDEQLKILD	TIKTKATQAA	QDGQSLKTRT	MLQADINKLM	120
LGQAISNGND	ALGILQTADK	AMDEQLKILD	TIKTKATQAA	QDGQSLKTRT	MLQADINRLM	120
EELDNIANTT	SFNGKQLLSG	NFTNQEFQIG	ASSNQTVKAT	IGATQSSKIG	VTRFETGAQS	180
EELDNIANTT	SFNGKQLLSG	NFINQEFQIG	ASSNQTVKAT	IGATQSSKIG	LTRFETGGRI	180
FTSGVVGLTI	KNYNGIEDFK	FDNVVISSTV	GTGLGALAE	INKSADKTGV	RATYDVKTG	240
<u>STSGEVQFTL</u>	KNYNGIDDFQ	FQKVVISSTV	GTGLGALAE	INKNADKTGV	RATFTVETRG	240
VYAIKEGTTT	QEFAINGVTI	GKIEYKDGDG	NGSLISAINA	VKDTTGVQAS	KDENGKLVLT	300
IAAVRAGATS	DTFaingvki	GKVDYKDGDA	NGALVAAINS	VKDTTGVVEAS	IDANGQLLLT	300
SADGRGIKIT	GDIGVGSgil	ANQKENYGRl	SLVKNDGRDI	<u>NI</u> <u>SGTNLS</u> <u>AI</u>	GMGTDMISQ	360
SREGRIKID	GNIGGGAFIN	ADMKENYGRl	SLVKNDGKDI	LISGSNLSSA	GFGATQFISQ	360
SSVSLRESKQ	QISATNADAM	GFNSYKGGGK	FVFTQNVSSI	<u>S</u> <u>AFMS</u> <u>AQGS</u> <u>S</u> <u>G</u>	FSRGSFSGSVG	420
ASVSLRESKQ	QIDANIADAM	GFGSANKGVV	LGGYSSVSAY	MSSAGSGFSS	SGSYVSGSK	420
SGKNLSVGLS	QGIQIISAA	<u>S</u> <u>MSNTYVV</u> <u>S</u> <u>A</u>	<u>G</u> <u>SGFS</u> <u>SG</u> <u>SGN</u>	<u>S</u> <u>QFAALK</u> <u>TTA</u>	ANTTDETAGV	480
NYSTGFANAI	AI.....SAAS	QLSTVYNVSA	GSGFSSGSTL	SQFATMK <u>TTA</u>	FGVKDETAGV	476
<u>T</u> <u>T</u> <u>L</u> <u>K</u> <u>G</u> <u>A</u> <u>M</u> <u>A</u> <u>V</u> <u>M</u>	DIAETAITNL	DQIRADIGSI	QNQVTSTINN	ITVTQVNVKA	AESQIRDVDF	540
TTLKGAMAVM	DIAETAITNL	DQIRADIGSV	QNQVTSTINN	ITVTQVNVKA	AESQIRDVDF	536
ASESANYSKA	NILAQSGSYA	MAQANSSQQN	VLRLLQ	576		
AAESANYSKA	NILAQSGSYA	MAQANSVQQN	VLRLLQ	572		

Figure 4. Comparison of the O-glycosylation sites identified in flagellin A from *C. jejuni* 81–176³ (top, blue typeface) and those identified in flagellin A from *C. jejuni* 11168 in the current study (bottom, black typeface). O-glycosylation sites are shown in red, bold, and underlined.

ECD were produced by an indirectly heated barium tungsten cylindrical dispenser cathode (5.1 mm diameter, 154 mm from the cell, 1 mm off-axis). The current across the electrode was ~1.1 A. Ions were irradiated for 60 ms at 5% energy (corresponding to a cathode potential of -2.775 V). ECD mass spectra were acquired with a resolution of 25 000 at m/z 400. Each ECD scan comprised 6 co-added microscans.

CID. The mass spectrometer alternated between a full FT-MS scan (m/z 380–2000) and subsequent CID MS/MS scans of the three most abundant ions. Survey scans were acquired in the ICR cell with a resolution of 100 000 at m/z 400. Precursor ions were isolated and subjected to CID in the linear ion trap. Automatic gain control (AGC) was used to accumulate sufficient precursor ions (target value 5×10^5 , maximum fill time 200 ms). Collisional activation was performed using helium gas at normalized collision energy 35% and $q_{\text{excite}} = 0.25$. The width of the precursor isolation window was 2 m/z and only multiply charged precursor ions were selected for MS/MS. Dynamic exclusion was used with a repeat count of 1 and exclusion duration of 180 s. Each CID scan comprised three co-added microscans.

Data acquisition was controlled by Xcalibur 2.0 software.

Data Analysis

Data were analyzed manually using Xcalibur 2.1 software (Thermo Fisher Scientific). *In silico* digests were generated by use of ProteinProspector v5.6.1 (<http://prospector.ucsf.edu/prospector/mshome.htm>) (University of California, San Francisco).

RESULTS AND DISCUSSION

The tryptic digests of the flagellin A isolated from *C. jejuni* 11168 were analyzed in duplicate by LC-ECD-MS/MS. The data

were analyzed manually by comparison with an *in silico* digest of flagellin A. The *in silico* digest included methionine oxidation and asparagine/glutamine deamidation as variable modifications. In addition, the data were searched for the presence of possible glycopeptides. The following potential glycans were considered: I (Δm 390.1638, C₁₆H₂₆N₂O₉); II (Δm 389.1798, C₁₆H₂₇N₃O₈); multiples and combinations of I and II, up to two glycans; pseudaminic acid, Pse5Ac7Ac (Δm 316.1270, C₁₃H₂₀N₂O₇); acetamidino pseudaminic acid, Pse5Am7Ac (Δm 315.1430, C₁₃H₂₁N₃O₆); Pse5Pr7Pr (Δm 408.1380, C₁₅H₂₄N₂O₁₁); Pse5Am7Ac8GlnAc (Δm 485.2122, C₂₀H₃₁N₅O₉); Pse5Ac7Ac8OAc (Δm 358.1376, C₁₅H₂₂N₂O₈); Pse5Ac7Ac8GlnAc (Δm 486.1962, C₂₀H₃₀N₄O₁₀); Leg5Am7Ac (Δm 315.1430, C₁₃H₂₁N₃O₆) [Note: this glycan has the same mass shift as acetamidino pseudaminic acid and the two are indistinguishable by the present methodology]; and Leg5AmNme7Ac (Δm 329.1587, C₁₄H₂₃N₃O₆).

Figure 1 shows the sequence coverage obtained (68.4%) following LC-ECD-MS/MS of flagellin. Twenty-nine unmodified peptides and five glycopeptides were observed in various charge states. Unmodified analogues of the glycopeptides were not observed in any case. Three peptides were modified by I and two peptides were modified by II. Figure 2a shows the ECD mass spectrum of triply charged ions of the tryptic peptide FETGGRISTSGEVQFTLK [174–191] modified by I, recorded at retention time 41.68 min. The measured monoisotopic m/z was 783.0605 (m/z_{calc} 783.0602, Δ 0.4 ppm). (The isolated precursor, shown in Figure 2a inset, also shows a minor peak corresponding to the glycopeptides modified with II). Complete sequence coverage was obtained for this glycopeptide, see Figure 2a inset. The peptide contains five possible sites of O-glycosylation

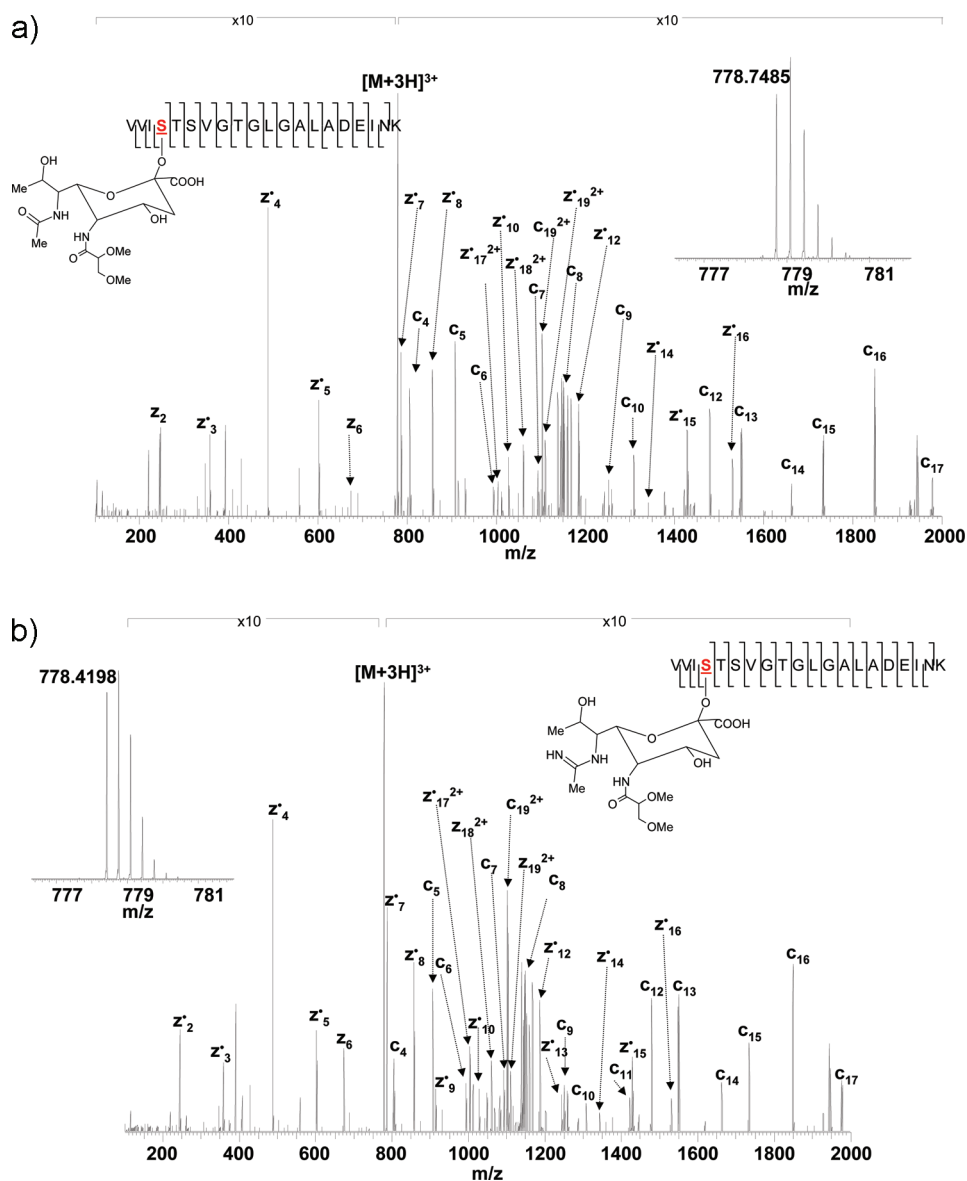


Figure 5. (a) ECD mass spectrum of 3+ ions of I-modified glycopeptide VVISTSVGTGLGALADEINK [204–223]. Inset: isolated precursor ions showing measured monoisotopic m/z and peptide sequence coverage obtained. (b) ECD mass spectrum of 3+ ions of II-modified glycopeptide VVISTSVGTGLGALADEINK [204–223]. Inset: isolated precursor ions showing measured monoisotopic m/z and peptide sequence coverage obtained.

(Ser/Thr). As labile modifications are retained on the peptide backbone fragments following ECD, the glycosylation site can be confirmed unequivocally as Ser181. There was no evidence for glycosylation at Thr176, Thr182, Ser183, or Thr189. For comparison, the CID mass spectrum of the $[M + 3H]^{3+}$ ions of the I-modified glycopeptide is shown in Figure 2b. The dominant fragment peak in that spectrum corresponds to the loss of the glycan from the doubly charged precursor peptide (the Y_0^{2+} fragment according to Domon and Costello's nomenclature¹⁷). Only two backbone fragments were observed: b_{16} and b_{17} . It is not possible to assign the site of glycosylation from these fragments.

The equivalent peptide modified by II was observed in the 3+ (m/z_{meas} 782.7331, m/z_{calc} 782.7323, Δ 1.0 ppm) and 4+ (m/z_{meas} 587.3014, m/z_{calc} 587.3010, Δ 0.7 ppm) charge states. The survey mass spectrum at retention time 40.02 min is shown in Figure 3a. Although both the 3+ and 4+ charge states of this

glycopeptides were observed, only the 4+ ions were subjected to ECD. That illustrates one of the disadvantages of ECD, namely, that the relatively long time scale of the ECD event reduces the total number of peptide ions analyzed.¹⁸ The ECD mass spectrum of the 4+ ions of tryptic peptide FETGGRISS-GEVQFTLK [174–191] modified by II, recorded at retention time 40.15 min, is shown in Figure 3b. Virtually complete sequence coverage is obtained and it is possible to assign the site of glycosylation as Ser181. Note that cleavage of the N-C α directly N-terminal of Ser181 residue was not observed. In part that is due to the m/z of the singly charged c_7 fragment overlapping with the triply charged ion arising from the loss of ammonia, a common and abundant fragment in ECD. Nevertheless, N-C α cleavages are observed between each of the possible sites of glycosylation allowing unambiguous localization. No evidence for glycosylation at the other serine/threonine

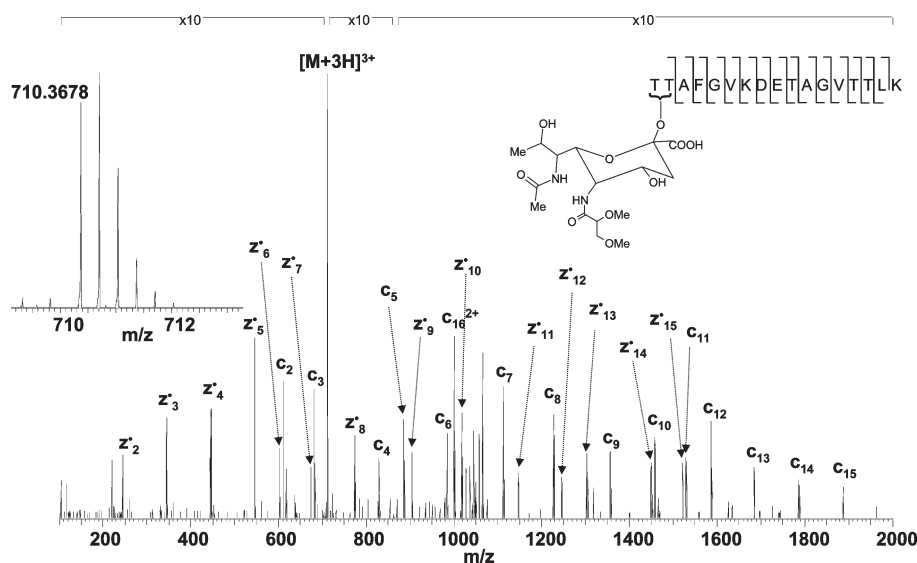


Figure 6. ECD mass spectrum of 3+ ions of I-modified glycopeptide TTAFGVKDETAGVTLK [464–480]. Inset: isolated precursor ions showing measured monoisotopic m/z and peptide sequence coverage obtained.

residues was obtained. For comparison, the CID mass spectrum of the $[M + 3H]^{3+}$ ions of the II-modified glycopeptide is shown in Figure 3c. Again, it is not possible to assign the site of glycosylation, nor obtain sequence information, from this mass spectrum.

Ser181 constitutes a novel glycosylation site in flagellin A. The equivalent amino acid residue in flagellin A from *C. jejuni* strain 81–176 is phenylalanine, see Figure 4. Glycosylation so far toward the N-terminus of the flagellin protein has not previously been reported and is potentially significant biologically. First, it questions the likely surface accessibility of glycan residues, previously assumed to be confined to the presumed surface-exposed, less conserved central region of the flagellin.³ Second and potentially more significantly, this region of the protein is very close to, if not a part of, the β -hairpin region recently implicated in TLR-5 mediated activation of the innate immune response in mammals when substituted with a Salmonella-derived peptide.¹⁹ This region may therefore be involved in modulation of host–pathogen interaction at some level.

The tryptic peptide [204–223] VVISTSVGTGLGALADEINK was observed to be glycosylated by both I and II. Figure 5a shows the ECD mass spectrum of triply charged ions of the tryptic peptide VVISTSVGTGLGALADEINK [204–223] modified by I, recorded at retention time 47.60 min. The measured monoisotopic m/z was 778.7485 (m/z_{calc} 778.7477, Δ 1.0 ppm). The isolated precursor peak, Figure 5a inset, clearly shows that the mass spectrum is obtained from the I-modified glycopeptide only, that is, there is no peak corresponding to the II-modified glycopeptide at this retention time. Peptide [204–223] contains four possible sites of O-glycosylation (two serine and two threonine residues). ECD of the glycopeptide results in complete sequence coverage, that is, all N–C α are cleaved, and the site of glycosylation is assigned as Ser207. There is no evidence for modification of the remaining serine/threonine residues. The I-modified glycopeptide was also observed in low abundance in the 2+ charge state; however, insufficient ions were available for ECD and a single backbone fragment was observed. Figure 5b shows the ECD mass spectrum of triply charged ions of the tryptic peptide VVISTSVGTGLGALA-

DEINK [204–223] modified by II, recorded at retention time 46.88 min. The measured monoisotopic m/z was 778.4198 (m/z_{calc} 778.4197, Δ 0.1 ppm), see Figure 5b inset. Again, ECD resulted in complete sequence coverage and the site of glycosylation is confirmed as Ser207. (The dominant fragment observed following CID of the I- and II-modified [204–223] glycopeptide corresponded to loss of the glycan and it was not possible to determine the site of glycosylation from that spectrum (data not shown).)

Glycosylation of the equivalent site in flagellin A, also Ser207, see Figure 4, from *C. jejuni* strain 81–176 has also been observed.³ In the case of *C. jejuni* 81–176, the glycans were pseudaminic acid and acetamidino pseudaminic acid, whereas the glycans observed in flagellin A from strain 11168 are the dimethylglyceric acid derivatives. Interestingly, both Ser207 and Ser181 (flagellin A, 11168) fall within the motif ISTS suggesting that a consensus sequence might exist for glycosylation in flagellin A.

Finally, tryptic peptide [464–480] (TTAFGVKDETAGVTLK) was observed to be glycosylated by I. There was no evidence for glycosylation of this peptide by glycan II. The I-glycopeptide was observed in the 3+ charge state with measured monoisotopic m/z 710.3678 (m/z_{calc} 710.3669, Δ 1.3 ppm), and the ECD mass spectrum of these ions is shown in Figure 6. There are five potential sites of glycosylation in this peptide, all threonine residues. Virtually complete sequence coverage is obtained, localizing the site of glycosylation to either Thr464 or Thr465. Unfortunately, no N–C α was observed between these two amino acid residues and it is not possible to confirm absolutely which site (or both) are modified. There was no evidence for glycosylation at the remaining threonine residues. The equivalent residues in flagellin A from *C. jejuni* 81–176 (Thr468 and Thr469, see Figure 4) have not previously been observed to be modified by any glycan. As in the case of the glycosylation of Ser181, this newly observed glycan-modified residue is located toward the C-terminal end of the known glycosylated, central variable region, but is not the most distal as modification of Thr483 had been observed in the work of Thibault et al.³

It is noteworthy that the region identified as heavily glycosylated in flagellin A from *C. jejuni* 81–176³ was not detected in the present analysis of the protein from *C. jejuni* 11168. The region contains many serine and threonine residues and it may be that this region is glycosylated in 11168 flagellin A. Three tryptic peptides are predicted from this region: [339–366], [388–420], and [421–463], containing 8, 11, and 14 potential O-glycosylation sites, respectively, all of which have high molecular mass, thus, increasing the likelihood that they will not be observed in the bottom-up LC–MS/MS approach. These peptides were also not observed in the LC–CID–MS/MS analysis. Note that the latter two (388–420 and 421–463) contain the bulk of the glycosylated residues observed by Thibault et al.³

CONCLUSION

We have applied LC–ECD–MS/MS to the analysis of glycosylation of flagellin A from *C. jejuni* 11168. The results confirm the presence of dimethylglyceric acid derivatives of pseudaminic acid and acetamidino pseudaminic acid. The glycosylation sites were assigned as Ser181, Ser207, and Thr464 or Thr465. This work is the first observation of Ser181 as a site of glycosylation. Ser181 was shown to be glycosylated by both I and II, and this modification may be significant biologically in view of its proximity to a region of the protein that was recently implicated in TLR-5-mediated interaction with the innate immune system. The equivalent amino acid in flagellin A from 81–176 is phenylalanine which clearly cannot be glycosylated. Ser207 was also shown to be glycosylated by both I and II. The equivalent site in flagellin A from 81–176, also Ser207, has been shown to be glycosylated by pseudaminic acid.³ Finally, Thr464 or Thr465 (or both, but exclusively) have been shown to be glycosylated by I in flagellin A from 11168. There was no evidence for glycosylation of these residues by II. The equivalent sites in flagellin A from 81–176 (Thr468 and Thr469) have not been observed to be modified by any glycan.

AUTHOR INFORMATION

Corresponding Author

*School of Biosciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, U.K. Telephone: +44 (0)121 414 7527. Fax: +44 (0)121 414 5925. E-mail: H.J.Cooper@bham.ac.uk.

ACKNOWLEDGMENT

H.J.C. gratefully acknowledges the Wellcome Trust (074131) for funding. G.B. was in receipt of a BBSRC research studentship.

REFERENCES

- (1) Goon, S.; Kelly, J. F.; Logan, S. M.; Guerry, P. Pseudaminic acid, the major modification on *Campylobacter* flagellin, is synthesized via the Cj1293 gene. *Mol. Microbiol.* **2003**, *50* (2), 659–671.
- (2) Howard, S. L.; Jagannathan, A.; Soo, E. C.; Hui, J. P. M.; Aubry, A. J.; Ahmed, I.; Karlyshev, A.; Kelly, J. F.; Jones, M. A.; Stevens, M. P.; Logan, S. M.; Wren, B. W. *Campylobacter jejuni* glycosylation island important in cell charge, legionaminic acid biosynthesis and colonization of chickens. *Infect. Immun.* **2009**, *77* (6), 2544–2556.
- (3) Thibault, P.; Logan, S. M.; Kelly, J. F.; Brisson, J. R.; Ewing, C. P.; Trust, T. J.; Guerry, P. Identification of the carbohydrate moieties and glycosylation motifs in *Campylobacter jejuni* flagellin. *J. Biol. Chem.* **2001**, *276* (37), 34862–34870.
- (4) Schirm, M.; Schoenhofen, I. C.; Logan, S. M.; Waldron, K. C.; Thibault, P. Identification of unusual bacterial glycosylation by

tandem mass spectrometry analyses of intact proteins. *Anal. Chem.* **2005**, *77*, 7774–7782.

- (5) McNally, D. J.; Hui, J. P.; Aubry, A. J.; Mui, K. K.; Guerry, P.; Brisson, J. R.; Logan, S. M.; Soo, E. C. Functional characterization of the flagellar glycosylation locus in *Campylobacter jejuni* 81–176 using a focused metabolomics approach. *J. Biol. Chem.* **2006**, *281* (27), 18489–18498.

- (6) McNally, D. J.; Aubry, A. J.; Hui, J. P. M.; Khieu, N. H.; Whitfield, D.; Ewing, C. P.; Guerry, P.; Brisson, J.-R.; Logan, S. M.; Soo, E. C. Targeted metabolomics analysis of *Campylobacter coli* VC167 reveals legionaminic acid derivatives as novel flagellar glycans. *J. Biol. Chem.* **2007**, *282* (19), 14463–14475.

- (7) Logan, S. M.; Hui, J. P. M.; Vinogradov, E.; Aubry, A. J.; Melanson, J. E.; Kelly, J. F.; Nothaft, H.; Soo, E. C. Identification of novel carbohydrate modifications in *Campylobacter jejuni* 11168 flagellin using metabolomics-based approaches. *FEBS J.* **2009**, *276*, 1014–1023.

- (8) Parkhill, J.; Wren, B. W.; Mungall, K.; Ketley, J. M.; Churcher, C.; Basham, D.; Chillingwirth, T.; Davies, R. M.; Feltwell, T.; Holroyd, S.; Jagels, K.; LKarlyshev, A. V.; Moule, S.; Pallen, M. J.; Penn, C. W.; Quail, M. A.; Rajandream, M. A.; Rutherford, K. M.; van Vliet, A. H. M.; Whitehead, S.; Barrell, B. G. The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. *Nature* **2000**, *403*, 665–668.

- (9) Hitchen, P.; Brzostek, J.; Panico, M.; Butler, J. A.; Morris, H. R.; Dell, A.; Linton, D. Modification of the *Campylobacter jejuni* flagellin glycan by the product of the Cj1295 homopolymeric-tract-containing gene. *Microbiology* **2010**, *156*, 1953–1962.

- (10) Cooper, H. J.; Akbarzadeh, S.; Heath, J. K.; Zeller, M. Data-dependent ECD FT-ICR mass spectrometry for proteomic analyses. *J. Proteome Res.* **2005**, *4* (5), 1538–1544.

- (11) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. ECD of multiply charged protein cations. A non-ergodic process. *J. Am. Chem. Soc.* **1998**, *120*, 3265–3266.

- (12) Cooper, H. J.; Hakansson, K.; Marshall, A. G. The role of electron capture dissociation in biomolecular analysis. *Mass Spectrom. Rev.* **2005**, *24*, 201–222.

- (13) Tsybin, Y. O.; Hakansson, P.; Budnik, B. A.; Haselmann, K. F.; Kjeldsen, F.; Gorshkov, M.; Zubarev, R. A. Improved low-energy electron injection systems for high rate electron capture dissociation in Fourier transform ion cyclotron resonance mass spectrometry. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 1849–1854.

- (14) Hakansson, K.; Cooper, H. J.; Emmett, M. R.; Costello, C. E.; Marshall, A. G.; Nilsson, C. L. ECD and IRMPD MS/MS of an N-glycosylated tryptic peptide to yield complementary sequence information. *Anal. Chem.* **2001**, *73*, 4530–4536.

- (15) Mirgorodskaya, E.; Roepstorff, P.; Zubarev, R. A. Localization of O-glycosylation sites in peptides by ECD in a Fourier transform mass spectrometer. *Anal. Chem.* **1999**, *71*, 4431–4436.

- (16) Power, M. E.; Guerry, P.; McCubbin, W. D.; Kay, C. M.; Trust, T. J. Structural and antigenic characteristics of *Campylobacter coli* FlaA flagellin. *J. Bacteriol.* **1994**, *176* (11), 3303–3313.

- (17) Domon, B.; Costello, C. E. A systematic nomenclature for carbohydrate fragmentations in FAB–MS/MS spectra of glycoconjugates. *Glycoconjugate J.* **1988**, *5* (4), 397–409.

- (18) Creese, A. J.; Cooper, H. J. Liquid chromatography electron capture dissociation tandem mass spectrometry (LC–ECD–MS/MS) versus liquid chromatography collision-induced dissociation tandem mass spectrometry (LC–CID–MS/MS) for the identification of proteins. *J. Am. Soc. Mass Spectrom.* **2007**, *18* (5), 891–897.

- (19) de Zoete, M. R.; Keestrs, A. M.; Wagenaar, J. A.; van Putten, J. P. Reconstitution of a functional Toll-like receptor 5 binding site in *Campylobacter jejuni* flagellin. *J. Biol. Chem.* **2010**, *285* (16), 12149–12158.