Diversity in the Protein N-Glycosylation Pathways Within the Campylobacter Genus*[

Harald Nothaft†, Nichollas E. Scott‡§, Evgeny Vinogradov¶, Xin Liu¶‡‡, Rui Hu¶, Bernadette Beadle‡, Christopher Fodor‡, William G. Miller¶, Jianjun Li¶, Stuart J. Cordwell§, and Christine M. Szymanski‡**

The foodborne bacterial pathogen, Campylobacter jejuni, possesses an N-linked protein glycosylation (pgl) pathway involved in adding conserved heptasaccharides to asparagine-containing motifs of >60 proteins, and releasing the same glycan into its periplasm as free oligosaccharides. In this study, comparative genomics of all 30 fully sequenced Campylobacter taxa revealed conserved pgl gene clusters in all but one species. Structural, phylogenetic and immunological studies showed that the N-glycosylation systems can be divided into two major groups. Group I includes all thermotolerant taxa, capable of growth at the higher body temperatures of birds, and possesses an en bloc heptasaccharide (pgl)1 pathway was discovered in the epsilonproteobacterium Campylobacter jejuni (1). The organism transfers a conserved heptasaccharide to asparagine residues within the sequon D/E- X1-N-X2-S/T (X1, X2 ≠ P) of >60 glycoproteins (2–4). Furthermore, the pathway can be functionally transferred into Escherichia coli, and the oligosaccharyltransferase (OTase), PglB, is capable of adding foreign sugars to acceptor proteins (5–7). C. jejuni PglB also possesses hydrolase activity, influenced by the cellular growth phase and osmotic environment, releasing free oligosaccharides (fOS) into the periplasmic space in a 10:1 ratio relative to the amount of heptasaccharide N-linked to protein (8, 9).

In this study, we used phylogenetic, immunological, structural and glycoproteomic studies to compare the N-glycosylation systems of 29 Campylobacter species and identified unexpected variations. Thus, although the pathway is a common feature within this genus, variability in the N-glycans and fOS at the species level suggests that each species possesses a unique recognition to cancer development. The first general protein glycosylation pathway was discovered in the epsilonproteobacterium Campylobacter jejuni (1). The organism transfers a conserved heptasaccharide en bloc to asparagine residues within the sequon D/E- X1-N-X2-S/T (X1, X2 ≠ P) of >60 glycoproteins (2–4). Furthermore, the pathway can be functionally transferred into Escherichia coli, and the oligosaccharyltransferase (OTase), PglB, is capable of adding foreign sugars to acceptor proteins (5–7). C. jejuni PglB also possesses hydrolase activity, influenced by the cellular growth phase and osmotic environment, releasing free oligosaccharides (fOS) into the periplasmic space in a 10:1 ratio relative to the amount of heptasaccharide N-linked to protein (8, 9).

The N-glycosylation of the human gut commensal, C. hominis, which produces at least four different N-glycan structures, akin to the surface carbohydrate diversity observed in the well-studied commensal, Bacteroides. Both group I and II glycans are immunogenic and cell surface exposed, making these structures attractive targets for vaccine design and diagnostics. Molecular & Cellular Proteomics 11: 10.1074/mcp.M112.021519, 1203–1219, 2012.

In eukaryotes, glycosylated proteins are ubiquitous components of extracellular matrices and cellular surfaces. Their oligosaccharide moieties are implicated in a wide variety of essential cell-cell and cell-matrix processes ranging from immunological, developmental, and cell differentiation to cancer development. The first general protein glycosylation pathway was discovered in the epsilonproteobacterium Campylobacter jejuni (1). The organism transfers a conserved heptasaccharide en bloc to asparagine residues within the sequon D/E- X1-N-X2-S/T (X1, X2 ≠ P) of >60 glycoproteins (2–4). Furthermore, the pathway can be functionally transferred into Escherichia coli, and the oligosaccharyltransferase (OTase), PglB, is capable of adding foreign sugars to acceptor proteins (5–7). C. jejuni PglB also possesses hydrolase activity, influenced by the cellular growth phase and osmotic environment, releasing free oligosaccharides (fOS) into the periplasmic space in a 10:1 ratio relative to the amount of heptasaccharide N-linked to protein (8, 9).

The N-link glycosylation of the human gut commensal, C. hominis, which produces at least four different N-glycan structures, akin to the surface carbohydrate diversity observed in the well-studied commensal, Bacteroides. Both group I and II glycans are immunogenic and cell surface exposed, making these structures attractive targets for vaccine design and diagnostics. Molecular & Cellular Proteomics 11: 10.1074/mcp.M112.021519, 1203–1219, 2012.

From the †Alberta Glycomics Centre and Department of Biological Sciences, University of Alberta, Edmonton, AB, Canada; §School of Molecular and Microbial Biosciences, University of Sydney, Sydney, NSW, Australia; ¶Institute for Biological Sciences, National Research Council, Ottawa, ON, Canada; ||Produce Safety and Microbiology Research Unit, Agricultural Research Service, US Department of Agriculture, Albany, CA, USA

Received June 18, 2012, and in revised form, July 28, 2012

© 2012 by The American Society for Biochemistry and Molecular Biology, Inc. This paper is available on line at http://www.mcponline.org

1 The abbreviations used are: pgl, protein glycosylation; N-linked, asparagine-linked; fOS, free oligosaccharides; PEtN, phosphoethanolamine.
a unique array of glycosyltransferases, which correlate with their phylogenetic relatedness.

**EXPERIMENTAL PROCEDURES**

**Plasmids, Bacterial Strains, and Growth Conditions**—All bacterial strains and plasmids used in this study are listed in Table I. Campylobacter strains were grown on Mueller Hinton or Brain Heart Infusion supplemented with 5% horse blood under microaerobic (85% N₂, 10% CO₂, and 5% O₂) conditions. Alternatively, anaerobic strains were grown for 18–72 h at 37 °C in the following medium: 30 g/l Tryptic Soy Broth, 2 g/l yeast extract, 3.0 g/l Bacto-Peptone, 2.0 g/l NaCl, 3.0 g/l sodium formate, 4.0 g/l formic acid, 4.0 g/l asparagine monohydrate, 1.0 g/l cysteine hydrochloride, pH 7.8; supplemented with 5% (v/v) horse blood. For large scale anaerobic cultures, bacteria were grown in anaerobic medium (without horse blood) at 37 °C in ventilated 1 liter bottles in anaerobic jars using the AnaeroGen gas pack system (Oxoid). E. coli strains were routinely grown in 2× YT medium at 37 °C. Growth media were supplemented with chloramphenicol (Cm, 25 μg/ml) and/or kanamycin (Kan, 30 μg/ml), where appropriate.

**Phylogenetic Analysis**—The AtpA (ATP synthase subunit, housekeeping gene) protein sequence was used as the phylogenetic marker for trees constructed with the neighbor-joining algorithm of the MEGA3.1 software package (20) using a multiple alignment parameters gap opening penalty of 8, a gap extension penalty of 2, and the PAM protein weight matrix. To identify putative pgl genes, the predicted proteomes of 29 Campylobacter genomes, representing all validly described taxa, were compared with the predicted proteome of C. jejuni strain NCTC 11168 using BLASTP. Proteins demonstrating >35% amino acid identity to Pgl proteins from 11168 were scored as positive matches. Putative gene products within the 29 proposed pgl loci were identified by BLASTP comparison to proteins present within

---

**Table I**

**Plasmids and strains**

NCTC: National Collection of Type Cultures; CCUG: Culture Collection, University of Göteborg, Sweden; LMG: Belgian coordinated collection of microorganisms; ATCC: American Type Culture Collection.

| Strain/plasmid | Relevant genotype, description or source | Reference |
|---------------|----------------------------------------|-----------|
| **Plasmids:** | | |
| pACYC(pglmut) | Encodes the C. jejuni pgl cluster with mutations W458A and D459A in the oligosaccharyltransferase PglB, CmR | (5) |
| **Strains:** | | |
| E. coli K12 (BW25113) | Δ araD-araB)67, ΔlacZ4787::(rrnB-3), lambda-, rph-1, Δ(rhaD-rhaB)568, hsdR514 | (59) |
| K12 (BW25113) wzy::kan | E. coli K-12 (BW25113) O-antigen polymerase mutant | (60) |
| **Campylobacter** | | |
| C. jejuni NCTC 11168; wild-type, clinical isolate used for genome sequencing, human origin | (61) |
| C. jejuni ΔpgIB | NCTC 11168 pgIB::kan | (9) |
| C. fetus subsp. fetus | 82–40; brain of sheep fetus | (62) |
| C. hominis | ATCC BAA-381; feces, healthy adult male | (47) |
| C. rectus | ATCC 33238; human periodontal pocket | (63) |
| C. lari subsp. lari | RM2100; human feces | (64) |
| C. coli | RM2228; chicken carcass | (64) |
| C. upsaliensis | RM3195; feces, human (child) | (64) |
| C. fetus subsp. venerealis | NCTC 10354; vaginal mucus of heifer | (62) |
| C. peloridis | LMG 11251; feces | (65) |
| C. lari subsp. concheus | LMG 11760; human | (65) |
| C. sputorum bv. sputorum | RM3237; human | (66) |
| C. gracilis | ATCC 33236; human gingival sulcus | (63) |
| C. showae | CCUG 30254; human gingival crevice | (67) |
| C. lari (UPTC) | NCTC 11845; freshwater; river water | (68) |
| C. lanienae | NCTC 13004; isolated in human abattoir workers, asymptomatic | (69) |
| C. curvus | ATCC BAA-1459; human oral cavity | (63) |
| C. helveticus | CCUG 30566; cat feces, diarrhea | (70) |
| C. hyointestinalis subsp. hyointestinalis | LMG 9260; human feces, diarrhea | (71) |
| C. hyointestinalis subsp. lawsonii | CCUG 27631; porcine stomach | (72) |
| C. mucosalis | CCUG 21559; porcine, necrotic colitis | (73) |
| C. sputorum bv. paraeolyticus | LMG 11764; human stool | (66) |
| C. sputorum bv. fecalis | CCUG 20703; cow, feces | (74) |
| C. ureolyticus | R4126; human, feces | (75) |
| C. insulaeigraea | NCTC 12927; marine mammal | (76) |
| C. concisus | NCTC 13826; human feces | (63) |
| C. subantarcticus | LMG 24377; sea-bird | (77) |
| C. avium | LMG 24591; chicken, caecal content | (78) |
| C. cuniculorum | LMG 24588; rabbit, caecal content | (79) |
| C. volucris | LMG 24379; sea-bird | (80) |
the NCBI nonredundant (nr) database. Annotation of the pgl loci also used the Sanger Centre Pfam protein families database (v 26.0; Nov. 2011) (21) and the CAZy database (22) to identify additional putative glycosyltransferases.

**IgN Purification and NMR Analysis—**Campylobacter cells from 6 liter cultures were harvested (10 min, 10,000 × g, 4 °C) and digested with proteinase K at pH 8 at 37 °C for 48 h. Digests were separated on a Sephadex G-15 column (1.5 × 60 cm). Fractions were dried, resuspended in D2O and analyzed by 1H NMR. Fractions containing fOS were separated by anion exchange chromatography on a Hitrap Q column (5 ml size, Amersham Biosciences). Glycans were eluted with a linear gradient of NaCl (0–1 M, 1 h). Desalting was performed on Sephadex G15 prior to glycan analysis on a Varian INOVA 500 MHz (1H) NMR spectrometer with a 3 mm gradient probe at 25 °C with acetone as the internal reference (2.252 ppm for 1H, 31.45 ppm for 13C) using standard pulse sequences: DQCOFSY, TOCSY (mixing time 120 ms), ROESY (mixing time 500 ms), HSQC and HMBC (100 ms long range transfer delay). AQ time was kept at 0.8–1 s for H-H correlations and 0.25 s for HSQC, 256 increments was acquired for t1. Sugars and sugar linkages were identified by characteristic coupling constants and chemical shift patterns. For fOS of C. concisus, the absolute configuration of the glucolactilic acid carboxyethyl group was determined as described (23).

**HPAEC-PAD—Cff fOS were separated by high performance anion exchange chromatography with pulsed amperometric detection (HPAEC/PAD) using a semipreparative CarboPac® PA 100 (9 × 25 mm equipped with a Guard Column: 3 × 50 mm) and a fraction collector (DIONEX UltiMate 3000) under the following conditions: flow rate: 0.5 ml/min; eluent system: 50 mM sodium hydroxide; detection mode: pulsed amperometry, quadrupole waveform, Au electrode; AgCl reference electrode; and the ambient temperature was set to 30 °C. Elution fractions containing the single fOS forms were neutralized with equimolar amounts of 0.2 M HCl and 1/10 of each elution fraction was applied to an analytical column CarboPac® PA200 (3 × 250 mm CarboPac PA100 equipped with a Guard Column: 3 × 50 mm) under the same conditions to confirm proper separation of individual fOS forms. Individual fOS-containing fractions were combined, lyophilized and stored at −20 °C until further use.

**IOS Analysis Using Mass Spectrometry—**fOS were identified and analyzed as described (8, 24), with the difference that precursor ion scanning was carried out for the detection of the diagnostic oxonium ion fragments at m/z 204 (for HexNAc) and m/z 407 (HexNAz2).

**Preparation of Whole-cell Lysates and Immunoblotting—**Whole-cell lysates were prepared as described (8). Immunoblotting using R1, hR6 (25), GRPII-1 (anti-Cff-minor), GRPII-2 (anti-Cff-major), or alkaline phosphatase (AP)-conjugated wheat germ agglutinin (WGA) (Vector Laboratories, Canada, Inc.) was performed as described (24) with dilutions of 1:10,000 for each primary antisera, 1:2000 for WGA-AP and 1:2000 for the secondary AP-labeled antibody (Santa Cruz Biotechnology, Santa Cruz, CA).

**Generation of fOS-BSA Glycconjugates—**Individual Cff fOS preparations were conjugated to BSA by reductive amination as described (26). Formation of BSA-glycoconjugates was confirmed by Western blotting with alkaline phosphatase (AP)-conjugated WGA.

**Generation of Group I and Group II C. fetus Subgroup Glycer-specific Antibodies—**New Zealand White Rabbits were immunized with individual group II C. fetus subgroup I fOS-BSA conjugates using a 10 week immunization protocol (approved Animal Care Committee protocol No. 717). Blood sera were analyzed for the production of pgl glycan-specific antibodies by Western blotting against whole cell lysates from various Campylobacter species. The antisera against HexNAc-[Hex]-HexNAc2-5iNacBac-BSA was named GRPII-1, the antisera against HexNAc-[Hex-Nac]-HexNAc2-5iNacBac-BSA was named GRPII-2. To generate group I-N-glycan specific (R1) antisera, rabbits were immunized with formaldehyde-fixed (according to (27)) cells of C. coli K12 wzy::kan (pACYC184p136) expressing the C. jejuni N-glycan on their surface (17).

**Digestion of Whole-cell Lysates for ZIC-HILIC Enrichment—**One mg of proteins, determined using the Quibit™ kit (Invitrogen, Carlsbad CA), was resuspended in 6 × 1 ml urea, 2 × thiourea, and 40 mg NH4HCO3. The solubilized sample was then reduced for 1 h with 10 mg dithiothreitol, subsequently alkylated using 20 mg iodoacetamide for 1 h in the dark and was quenched using 10 mg dithiothreitol. Then, 1/200 (w/w) of endoproteinase Lys-C (Sigma-Aldrich) was added and digestion was allowed to proceed for 4 h at 25 °C. Samples were diluted 1:4 with 100 mM NH4HCO3 and digested with 1/50 (w/w) of porcine sequencing grade trypsin (Promega, Madison, WI) for 18 h at 25 °C. Samples were then dialyzed against ultra-pure water overnight using a Mini Dialysis Kit with a molecular mass cut off of 1000 Da (GE Healthcare) and lyophilized.

**Enrichment of Glycopeptides by Tandem Mass Tag (iTRAQ) labeling—**Enrichment of glycopeptides was performed using a Quanta BioSciences (CBM-40) kit according to the manufacturer’s instructions. iTRAQ labeled glycopeptides were separated using a 4.6 × 50 mm ZIC-HILIC resin (Sequant, Umeå, Sweden) and digested with 1/50 (w/w) excised C8 Empore™ disc (Sigma) were packed to a bed length of 0.5 cm. Before use, the columns were washed with ultra-pure water, followed by 95% acetonitrile (ACN) and then equilibrated with 80% ACN and 5% formic acid (FA). Samples were resuspended in 80% ACN, 5% FA and insoluble material removed by centrifugation at 20,000 × g for 5 min at 4 °C. Samples were adjusted to a concentration of 2 μg/ml and 100 μg of peptide material loaded onto a column and washed with 10 load volumes of 80% ACN, 5% FA. Peptides were eluted with three load volumes of ultra-pure water into low-bond tubes and concentrated using vacuum centrifugation. Triplicate enrichments were generated for each glycopeptide analysis.

**Identification of Glycopeptides Using Tandem Mass Tag (iTRAQ) labeling—**Identification of glycopeptides was performed using a Q Exactive Mass Spectrometer (Thermo Scientific, San Jose, CA). Peptides were eluted and separated using a 60 cm, 100 μm id, ReproSil - Pur C18 AQ 3 μm column using a 120 min gradient in which the proportion of phase A (0.1% FA) and phase B (0.1% FA, 80% ACN) were altered from 100% phase A to 40% phase B over 120 min at 200 nL/min in an LTQ-Orbitrap Velos mass spectrometer (Thermo Scientific, San Jose, CA). The LTQ-Orbitrap Velos was operated using Xcalibur v2.2 with a capillary temperature of 275 °C in a data-dependent mode automatically switching between MS and MS/MS with CID followed by HCD. For each MS scan, the three most abundant precursor ions were selected for fragmentation with CID (normalized collision energy 35, activation time 30 ms) followed by HCD (normalized collision energy 45, activation time 30 ms) to obtain MS/MS spectra. MS resolution was set to 60,000 with an ACG of 1e6, maximum fill time of 500 ms and a mass window of 600 to 2000 m/z. CID fragmentation was carried out with an ACG of 2e6, maximum fill time of 100 ms whereas HCD fragmentation was carried out with an ACG of 2e6, maximum fill time of 500 ms, resolution of 7500 and mass window 170 to 2000 m/z.

**Glycopeptide Data Processing—**The raw files were then processed within Proteome Discoverer version 1.0 Build 43 (Thermo Scientific, San Jose, CA) to identify possible glycopeptides within this list, a new feature within the database. Searches were exported to Excel (Microsoft, Redmond, WA). To identify possible glycopeptides within this list, a new feature within the database. Searches were exported to Excel (Microsoft, Redmond, WA).
identified all scan events within the generated mgfs containing the diagnostic oxonium 204.086 m/z ion. Using Excel, all scan events that were not matched by Sequest and contained a predicted marker of glycosylation were identified. These events were manually inspected and identified as possible glycopeptides based on the presence of the glycan fragment within the CID scan. To facilitate glycopeptide assignments from HCD scans, the ions below the mass of the predicted deglycosylated peptides were extracted with Xcalibur v2.2 using the Spectrum list function. Ions with a deconvoluted mass above the deglycosylated peptide and ions corresponding to known carbohydrate oxoniums were removed in a similar approach to post-spectral deglycosylation. MASCOT v2.2 searches were conducted via the Australasian Proteomics Computational Facility (www.apcf.edu.au) of the proteobacteria taxonomy of the Lugwig NR database (database release LugwigNR_Q410, composed of 13112897 entries available via ftp://ftp.ch.embnet.org/pub/databases/nr_prot/). Searches were carried out with a parent ion mass accuracy of 20 ppm and a product ion accuracy of 0.02 Da with no protease specificity as well as the fixed modification carboxymethyl (C) and variable modifications, oxidation (M) and deamidation (N). The instrument setting of MALDI-QIT-TOF was chosen because of previous studies showing quadrupole-like fragmentation within HCD spectra (30) (generating a, b, and y ions) and our observation of internal cleavage products that are all included in this setting. All spectra were searched with the decoy option enabled and no matches to this database were detected (FDR 0%). To further validate glycopeptide matches, all spectra were manually annotated to ensure all major peaks were match providing further confidence of identity and localization. All annotated spectra are provided in supplemental MS-Figures.

RESULTS

Campylobacters Divide into Two Major Groups Based on Phylogeny and Antigenicity—All Campylobacter species, except C. canadensis, possess a complete repertoire of pgl genes (Fig. 1). Phylogenetic comparisons showed that campylobacters divide into two major groups that comprise the thermotolerant (group I) and the nonthermotolerant species (group II) with C. lari and related species forming a subgroup (the C. lari subgroup), within group I (Fig. 1). When whole cell lysates of all 29 Campylobacter pgl-containing species were screened with C. jejuni N-glycan specific antisera, all group I organisms were recognized (Figs. 2A, 2B). No reactivity against proteins from the group II species was observed suggesting that these organisms are either incapable of synthesizing N-glycans or they produce different oligosaccharide structures compared with C. jejuni.

The Characterization of Group I Campylobacter fOS Structures—Because the C. jejuni fOS to N-glycan ratio is 10:1 under standard laboratory growth conditions (9), we screened species from each group for fOS by mass spectrometry using precursor ion scanning. In group I organisms, except the C. lari subgroup, a mass of m/z 1425.0 strongly suggested the presence of a heptasaccharide structure similar to C. jejuni (supplemental Fig. S1) (9). For fOS isolated from the C. lari subgroup, masses of m/z 1263.0, m/z 1244.6, and 1343.0 were detected (supplemental Fig. S1). In C. lari subsp. lari, MS² analysis of m/z 1263.0 and 1245.6 indicated a linear [HexNAc]₆-diNACBac oligosaccharide composition and its nonhydrated form (Figs. 3A, 3B) whereas m/z 1344.0 correlated to the hexasaccharide with phosphate attached, ([HexNAc]₆-diNACBac-P; Fig. 3C). The presence of phosphate was confirmed by the neutral loss of phosphoric acid (~98 m/z) from the 1344.0 m/z structure (31).

Because multiple glycoforms were identified from the C. lari subsp. lari fOS analysis, we used NMR to determine the phosphate position on the fOS and investigate the authenticity of the nonphosphorylated form (Fig. 6, supplemental Table S2). For the C. lari subsp. lari fOS, the NMR H-1 signal of diNACBac showed correlation to a phosphate signal at 1 ppm and had an additional coupling of 7 Hz because of phosphorylation. The C-1 signal was at 95.2 and no beta anomer was present. Therefore the phosphorylation was at the O-1 of diNACBac and the linear [GalNAc]₆-diNACBac-P structure was confirmed for this species with no evidence of dephosphorylated glycoforms detected. As a control, structural analysis of fOS from C. jejuni (Fig. 6, supplemental Table S3) confirmed the previously established N-linked glycan structure is identical to the fOS glycan.

The Characterization of Group II Campylobacter fOS Structures—As representatives of the group II species, we analyzed Cff, Cfv, C. hyointestinalis subsp. hyointestinalis (Chyo) and C. hyointestinalis subsp. lawsonii (further referred to as the C. fetus subgroup). Precursor ion scanning for the diagnostic oxonium ion m/z 204 (for HexNAc) identified two parent ions with m/z 1263.0 and m/z 1222.0 (supplemental Fig. S1). MS² analyses of these ions for Cff identified the monosaccharide compositions and sequence of both IOS species as HexNAc-[Hex]-HexNAc₃-diNACBac and HexNAc-[HexNAc]-HexNAc₃-diNACBac (Figs. 3D, 3E). NMR analysis of the fOS of Cff, Cfv and Chyo confirmed that these species produce two oligosaccharides, as suggested by MS analysis (Fig. 6, supplemental Table S1). Within these glycans the three sugars at the reducing end were identical to those of C. jejuni whereas the three sugars at the nonreducing end were: β-GlcNAc-3-[α-GlcNAc-6]-α-GlcNAc-4 and β-Glc-3-[α-GlcNAc-6]-α-GlcNAc-4.

The presence of two oligosaccharide variants in the C. fetus subgroup raised the question about their expression and antigenicity. Quantitative and preparative HPAEC-PAD analyses on purified C. fetus subgroup II glycans (Fig. 4) not only allowed the separation of the two Cff fOS species, but also demonstrated that the ratio of peak 1 (major)/peak 2 (minor) group II fOS was ~3.6:1 (Fig. 4). Immunofluorescence staining of intact cells of Cff, Cfv and Chyo with GRPII-1 and GRPII-2 antisera individually raised in rabbits against each of the two structures, showed that both glycans are antigenic and expressed on the bacterial cell surface (Fig. 4F). In contrast, the GRPII antisera did not react with C. jejuni (group I) cells (Fig. 4F). Similarly, group II species could not be labeled with group I antiserum whereas C. jejuni cells showed strong reactivity.

To confirm that these glycoforms are indeed N-linked to protein, Cff whole cell lysates were subjected to ZIC-HILIC
glycopeptide enrichment and MS-MS analysis. A total of 65 unique glycopeptides were identified within this species (supplemental Table S7) with both glycans detected on protein substrates. Forty-six peptides contained the HexNAc[HexNAc]-HexNAc3-diNAcBac glycan whereas 19 contained the HexNAc-[Hex]-HexNAc3-diNAcBac glycan with 14 glycosylation sites confidently assigned independent of sequon alone (supplemental Table S9) and included examples of both glycans occupying the same glycosylation site. The identification of glycopeptides containing both glycans, which co-eluted, enabled the comparison of the ratio of the HexNAc-[HexNAc]-HexNAc3-diNAcBac and HexNAc-[Hex]-HexNAc3-diNAcBac glycoforms of the 19 glycopeptide pairs detected (supplemental Table S8). An example of this is the glycopeptide $1^{135}$DTNETELSAQK$^{145}$ of the putative protein A0RN00_CAMFF, which demonstrates the ratio of $1^{135}$/H$^{145}$1 for the two glycoforms (Fig. 5) that is representative for the average, $1^{135}$/H$^{145}$3.7/1, observed for all confirmed glycopeptide pairs (supplemental Table S8). This ratio is in good agreement with the observed fOS ratio.

Group II Campylobacter Species Show Extensive Diversity and Split into Eight Subgroups—The generation of C. fetus specific glycans antibodies provided a rapid screening tool to investigate the attachment of similar glycans to protein substrates. Whole-cell lysates of all 29 Campylobacter species were probed with GRPII-1/GRPII-2 $\text{N}$-glycan specific antisera (Figs. 2C, 2D). Similar to the immunofluorescence results, proteins of Group I species did not react, whereas proteins...
from *C. fetus* sp. and *C. hyointestinalis* sp. strongly reacted with GRPII sera. The GRPII sera showed variable reactivity with other group II *Campylobacter* species, which ranged from moderate with *C. gracilis* proteins, weak with *C. sputorum* bv. *sputorum*, *C. hyointestinalis*; *C. lari* subsp. *venerealis*; *C. sputorum* bv. *paraureolyticus*; *C. hyointestinalis* subsp. *hyointestinalis*. Group I species are highlighted in red, group I-*C. lari* subgroup species in green, group II species are in black. Note: hR6 reacts against all group I species as it recognizes the linear GalNAc chain independent on the presence of the Glc branch (25) whereas R1 requires the Glc branch on the GalNAc chain for full reactivity and therefore is less reactive against the *C. lari* subgroup. Molecular weight markers (MW) in kilodaltons (kDA) are indicated on the left.

from *C. fetus* sp. and *C. hyointestinalis* sp. strongly reacted with GRPII sera. The GRPII sera showed variable reactivity with other group II *Campylobacter* species, which ranged from moderate with *C. gracilis* proteins, weak with *C. sputorum* bv. *sputorum* and *C. lanienae*, and no reactivity against *C. hominis*, *C. rectus*, *C. showae*, *C. mucosalis*, *C. concisus*, *C. curvus*, *C. ureolyticus*, *C. sputorum* bv. *fecalis* and *C. sputorum* bv. *paraureolyticus* (Figs. 2C, 2D). Interestingly *Campylobacter* species with a closer phylogenetic relationship demonstrated similar patterns of reactivity toward GRPII-1/GRPII-2 antisera suggesting that group II may contain subgroups that differ in their glycan structures. Because of the observed variations in antibody reactivity, we analyzed the fOS composition of the remaining group II strains by MS and MS². In the *C. sputorum* subgroup (including *C. lanienae* and *C. sputorum* bv. *sputorum*, *paraureolyticus*, and *fecalis*), precursor ion scanning identified masses at m/z 1384.4 and 1425.0 (supplemental Fig. S1) corresponding to Hex-HexNAc-[Hex]-HexNAc₂-diNAcBac and HexNAc₂-[Hex]-HexNAc₃-diNAcBac, respectively (Figs. 3F, 3G). The ion at m/z 528.2 (Fig. 3G) may represent [Hex]₂-HexNAc and thus suggests the branch to be terminal. For *C. ureolyticus* and *C. concisus*, m/z 1456.0 (supplemental Fig. S1) represents 234-HexNAc-[Hex]-HexNAc-HexNAc₂-diNAcBac (Fig. 3H) whereas in *C. showae* and *C. rectus*, precursor ion scans identified a parent ion at m/z 1560.0 (supplemental Fig. S1). MS² strongly suggested the composition [Hex]₄₂-[HexNAc]₃; [217]₁ and [diNAcBac]₁, but the full oligosaccharide sequence could not be solved (Fig. 3I). Similarly, in *C. hominis*, a parent ion at m/z 1497.5 (supplemental Fig. S1) most likely reflects a

**Fig. 2.** N-glycan-specific antiserum reactivity against whole cell lysates from 29 *Campylobacter* species. Immunoblots with (A) hR6 and (B) R1 antisera specific against the N-glycan of *C. jejuni*; immunoblots with (C) GRPII-1 and (D) GRPII-2 antisera specific against the individual *C. fetus* subgroup oligosaccharides are shown. Forty micrograms of whole cell lysates of the indicated strains were applied to 12.5% SDS-PAGE and transferred to a PVDF membrane. *Cff*, *C. fetus* subsp. *fetus*; *Csp*, *C. sputorum*; *Chyo*, *C. hyointestinalis*; *Cfv*, *C. fetus* subsp. *venerealis*; *Csp* bv. *para*, *C. sputorum* bv. *paraureolyticus*; *Chyo* subsp. *hyo*, *C. hyointestinalis* subsp. *hyointestinalis*. Group I species are highlighted in red, group I-*C. lari* subgroup species in green, group II species are in black. Note: hR6 reacts against all group I species as it recognizes the linear GalNAc chain independent on the presence of the Glc branch (25) whereas R1 requires the Glc branch on the GalNAc chain for full reactivity and therefore is less reactive against the *C. lari* subgroup. Molecular weight markers (MW) in kilodaltons (kDA) are indicated on the left.
Fig. 3. Mass-spectrometric identification and analyses of fOS. Breakdown products (MS²) of characteristic ions obtained by precursor ion scanning (supplemental Fig. S1): group I C. lari subgroup (C. lari subsp. lari) for (A) m/z 1263.7; (B) m/z 1244.6, and (C) m/z 1344.6; group II C. fetus subgroup (Cff) for (D) m/z 1263.8 and (E) m/z 1221.0; group II C. sputorum subgroup (C. lanienae) for (F) m/z 1344.6; group II C. concisae and C. ureolyticus (C. concisae) for m/z 1456.7: 234-HexNAc-[Hex]-HexNAc₂-dINAcBac; group II C. concisae and C. ureolyticus (C. concisae) for m/z 1456.7: 234-HexNAc-[Hex]-HexNAc₂-dINAcBac; group II C. rectus and C. showae (C. rectus) for m/z 1605: [Hex]₃-HexNAc₂-[217], dINAcBac; group II C. rectus and C. showae (C. rectus) for m/z 1605: [Hex]₃-HexNAc₂-[217], dINAcBac; group II C. hominis for m/z 1497.8: [HexNAc]₂, dINAcBac; group II C. concisae and C. mucosa (C. mucosa) for m/z 1440.5: Hex-HexNAc-[HexNAc]₁-HexNAc₂-217-HexNAc₂-dINAcBac; group II C. gracilis for m/z 1417.4: PE-HexNAc-[234]-HexNAc₂-dINAcBac.
sugar composition of [234], [HexNAc]5, and [diNAcBac]1, (Fig. 3J) whereas for C. curvus and C. mucosalis, identical masses of m/z 1439.6 (supplemental Fig. S1) represent HexNAc2-[Hex]-HexNAc-217-HexNAc-diNAcBac (Fig. 3K). For C. gracilis, a parent ion mass at m/z 1417.4 (supplemental Fig. S1) represents a hexasaccharide with the sequence 123-HexNAc- [234]-HexNAc2-diNAcBac (Fig. 3L). The fragment ions with m/z 626.8, 487.1 and 327.7 were indicative of a PETN residue linked to an acetylated amino sugar (HexNAc). Interestingly, fragmentation of this PETN containing glycan was dominated by B-ions compared with the typically Y-ion series dominant spectra seen for the other fOS glycans.

The Characterization of Group II Subgroups by the Identification of fOS Structures—Because fOS analysis identified the presence of unusual monosaccharides (m/z 234 and m/z 217) and we detected differences in glycoprotein reactivity with GRPII-1/GRPII-2 antisera within the C. sputorum subgroup (although MS2 suggests similar fOS sequences), we solved the representative fOS structures of each subgroup by NMR. Within C. concisus, the 234 Da component is glucolactic acid (Glc4RLac) and the structure of the fOS was assigned to be Glc4RLac-6-α-GalNAc-4-[α-Glc-6]-α-GalNAc-4-α-GalNAc-4-α-GalNAc-4-α-GalNAc-3-α,β-diNAcBac (Fig. 6, supplemental Table S6). The position of the glucolactic acid was determined based on strong NOEs between H-2 of the lactilic acid and H-4 of the glucose and on the low field chemical shift of the glucose C-4 (78.5 ppm) that was caused by alkylation of the Glc O-4 by C-2 of lactilic acid (data not shown). For C. sputorum bv. sputorum the structure was determined to be β-Glc-4-β-GlcNAc-3-[α-GlcNAc-6]-α-GlcNAc-4-α-GalNAc-4-α-GalNAc-4-α-GalNAc-3-α,β-diNAcBac (Fig. 6, supplemental Table S4), whereas the structure for C. sputorum bv. fecalis is β-Glc-4-β-GlcNAc-3-[α-GalNAc-6]-α-GlcNAc-4-α-GalNAc-4-α-GalNAc-3-α,β-diNAcBac (Fig. 6, supplemental Table S5). Interestingly, these fOS structures only differ by the configuration of the HexNAc branch. Although multiple attempts were made to analyze fOS from all representative subgroups, we were not able to obtain the amount of fOS that is required for NMR to elucidate the glycan structure of the remaining Campylobacter species.

Characterization of Group II Glycans by Glycopeptide Enrichment—Because we could not fully elucidate the fOS structure of certain Campylobacter group II species such as C. rectus and C. hominis, by MS or NMR, we employed ZIC-HILIC glycopeptide enrichment to further characterize these glycans and their composition. Enrichment of C. hominis glycopeptides confirmed that the fOS composition of [234], [HexNAc]5, [diNAcBac]1 was indeed found on protein substrates such as the preprotein translocase, SecG subunit.

Fig. 4. Generation of group II C. fetus subgroup-specific antibodies and detection of N-glycans on the bacterial cell surface. Elution profiles of group II-C. fetus subgroup fOS before and after separation by high performance anion exchange chromatography with pulsed amperometric detection (HPAEC/PAD) are shown. A, Approximately 0.5 nmol of group II C. fetus subgroup fOS and (B) the minor and (C) the major forms after separation were analyzed. Individual group II C. fetus subgroup fOS-glycoconjugates after coupling by reductive amination were separated by 12.5% PAGE and analyzed by (D) Coomassie staining and by (E) Western blotting with alkaline phosphatase-conjugated WGA. No reactivity against the lectin could be seen with unconjugated BSA (lane 1, 400 µg BSA fraction V). Both fOS-BSA conjugates (lane 2, 400 µg of group II HexNAc-[Hex]-HexNAc2-diNAcBac; lane 3, 400 µg of HexNAc-[HexNAc]-HexNAc2-diNAcBac coupled to BSA fraction V) migrated at the expected size of ~66 kDa and reacted strongly with WGA. In parallel, analysis of the same glycoconjugates by SDS-PAGE and Coomassie staining confirmed the integrity of the preparation. F, Fluorescent microscope images of whole cells of Cff, Cfv, C. hyointestinalis subsp. hyointestinalis (Chyo), and C. jejuni (Cj, negative control) probed with GRPII-1 or GRPII-2 as the primary antisera and a fluorescent-tagged secondary antibody (AlexaFluor546-antirabbit) are shown. R1 antisera in combination with Cj, Cj (pglB), or Cff cells were used as additional positive and negative controls.

1425.5 and (G) m/z 1384.3; group II C. concisus and C. ureolyticus (C. concisus is shown) for (H) m/z 1456.5; group II C. rectus and C. showae (C. rectus is shown) for (I) m/z 1560.0; group II C. hominis for (J) m/z 1497.5; group II C. curvus and C. mucosalis (C. curvus is shown) for (K) m/z 1440.5 and group II C. gracilis for (L) m/z 1417.4. The deduced monosaccharide composition or the oligosaccharide sequence and the designations of the corresponding breakdown ions are indicated.
and shown to be: HexNAc$_2$-[234]-HexNAc$_3$-diNAcBac (Figs. 7A, 7D). Additional variations in the glycopeptides were detected that varied in the composition of the two terminal sugars. HCD and CID fragmentation patterns clearly showed the presence of two variants that contain a unique mass of $m/z$ 245 (accurate mass analysis exhibited $MH^+ = 246.09506$): 245-HexNAc-[234]-HexNAc$_3$-diNAcBac and 245–245-[234]-HexNAc$_3$-diNAcBac (Figs. 7B, 7C). Within the glycopeptide analysis, isobaric glycoforms of the 4271.61 Da glycan were noted (Fig. 8A). This mass was found to correspond to two variants; 245-HexNAc-[234]-HexNAc$_3$-diNAcBac and HexNAc-245-[234]-HexNAc$_3$-diNAcBac spectra in which both Y-ions associated with the fragmentation of glycosidic linkages between the carbohydrate moieties and lower molecular mass oxonium ions (366, 407, 549, 610, and 813) could be observed. D, HCD fragmentation of novel glycopeptides, 135DTNETELSAQK identified, Mascot ion score 89 and 5 respectively, within both HexNAc-[HexNAc]-HexNAc$_3$-diNAcBac and HexNAc-[Hex]-HexNAc$_3$-diNAcBac spectra. Comparison of relative glycoform abundance based on MS analysis within the glycoproteome of Cff is shown in supplemental Table S8.

9EDQNKSEPFPVLVPDAK$_{109}$ and shown to be: HexNAc$_2$-[234]-HexNAc$_3$-diNAcBac (Figs. 7A, 7D). Additional variations in the glycopeptides were detected that varied in the composition of the two terminal sugars. HCD and CID fragmentation patterns clearly showed the presence of two variants that contain a unique mass of $m/z$ 245 (accurate mass analysis exhibited $MH^+ = 246.09506$): 245-HexNAc-[234]-HexNAc$_3$-diNAcBac and 245–245-[234]-HexNAc$_3$-diNAcBac (Figs. 7B, 7C). Within the glycopeptide analysis, isobaric glycoforms of the 4271.61 Da glycan were noted (Fig. 8A). This mass was found to correspond to two variants; 245-HexNAc-[234]-HexNAc$_3$-diNAcBac and HexNAc-245-[234]-HexNAc$_3$-diNAcBac (Figs. 8B, 8C), which differ in the position of the 245.09 carbohydrate moiety but are attached to the same protein substrate (foldase protein PrsA$_{263}$AANVQFTDTNK$_{265}$ Fig. 8D). In total 18 unique glycopeptides corresponding to 10 unique peptide sequences and five localized glycosylation sites were identified within C. hominis (supplemental Table S10). Comparison of relative glycoform abundance based on MS analysis within the glycoproteome of Cff is shown in supplemental Table S8.

Analysis of C. rectus glycopeptides confirmed the fOS mass of 1560.0 Da and also enabled the complete assignment of the oligosaccharide as Hex-HexNAc-[Hex-Hex]-HexNAc-217-HexNAc-diNAcBac (Fig. 9A). Within the generated HCD spectra, the unusual 217 Da carbohydrate moiety lead to the formation of an oxonium ion of mass $MH^+ =
Campylobacter N-Glycan Pathways

**Fig. 6.** NMR structures of select group I and group II fOS. 
$^1$H-NMR spectra for fOS of C. concisus, C. sputorum bv. fecalis, Cff (the insert shows an overlay of 2D HSQC for Cff (blue), and Cfv (purple) with $^1$H and $^1$C chemical shifts as X and Y axes), C. sputorum bv. sputorum, C. lari subsp. lari, and C. jejuni. For chemical shifts and structures, see supplemental Tables S1–S6.

218.06398 that strongly suggested an elemental composition of C6H10O5N4 (-2.710ppm) with the presence of the oxonium ion being diagnostic for C. rectus glycopeptides. A total of 25 unique glycopeptides were identified enabling the localization of 11 glycosylation sites across multiple proteins (supplemental Table S11) such as the M23 peptidase domain protein, $^{43}$KAEENLT$^{45}$ (Fig. 9B). The C. rectus N-glycan represents the first example of an octasaccharide within the Campylobacter genus.

Confirmation of a Noncarbohydrate Moiety in the C. gracilis Glycans by Glycopeptide Enrichment—Because our initial analysis of C. gracilis fOS lead to the identification of a glycan composition of 123-HexNAc-[234]-HexNAc$_3$diNAcBac with the identity of the 123 Da moiety being unassigned, we further characterized this moiety and verified the attachment of this unusual glycan to protein substrates by ZIC-HILIC glycopeptide enrichment. Multiple glycoforms were identified in C. gracilis that were characterized by the presence or absence of the noncarbohydrate modification PEtN. In total eight glycopeptides from seven glycoproteins were identified enabling the localization of four glycosylation sites (supplemental Table S12). Within this dataset, two proteins and three glycopeptides were demonstrated to contain PEtN on the glycan (supplemental Table S12). Examples of the identified C. gracilis glycopeptides were the PEtN-HexNAc-[234]-HexNAc$_3$-diNAcBac glycopeptide, $^{8}$VKIEANATK$^{35}$ of preprotein translo-
case SecG subunit (Fig. 9C, D) and the HexNAc-[234]-HexNAc$_3$-diNAcBac glycopeptide, $^{117}$TGADNGTLAEQR$^{129}$ of the periplasmic nitrate reductase (Figs. 9E, 9F). Within this dataset, the use of CID and HCD fragmentation led to the generation of internal PEtN fragments, including the loss of the diagnostic ethanamine group at 43.04 Da, which aided in the assignment of this noncarbohydrate modification.

The N-linked Glycan of C. lari subsp. lari Differs from Its fOS Counterpart—To determine whether the N-glycan of C. lari, like its fOS counterpart, is modified with a phosphate residue at the reducing end, ZIC-HILIC enrichment and a combination of MS HCD and ETD fragmentation was used to identify glycopeptides in C. lari subsp. lari (supplemental Table S13). Within the identified glycopeptides, such as $^{58}$AAVDAN-ASGSEK$^{59}$ from ubiquinol cytochrome c oxidoreductase, only [HexNAc]$_3$-diNAcBac glycans were detected (Figs. 9G, 9H) and there was no evidence of phosphate-containing N-glycan suggesting that the C. lari subsp. lari N-linked glycan is [HexNAc]$_3$-diNAcBac.

**DISCUSSION**

Phylogenetic analysis coupled with antibody screening divided the Campylobacter genus into two major groups: group I includes all thermotolerant species capable of living at the high body temperatures found within avian hosts, whereas group II contains the nonthermotolerant campylobacters. Interestingly, a similar grouping was also observed when the sequence of the oligosaccharyltransferase PglB, necessary for N-glycosylation, was used as the phylogenetic marker (12). The C. jejuni heptasaccharide is conserved in all group I species (excluding the C. lari subgroup), and identical fOS and N-glycan structures are produced. The C. lari subgroup lack the glucosyltransferase (PglI) responsible for the addition of the glucose branch (Fig. 1) and instead produce a novel heptasaccharide that is modified with a phosphate on the fOS, but not on the N-glycan. This is the only Campylobacter species that produces fOS structures that are different from the N-glycans. This linear hexasaccharide was also observed when the C. lari pgl cluster was transferred into the heterologous E. coli host (25). Remarkably, although Neisseria species O-glycosylate their proteins, they too possess enzymes for the synthesis of diNAcBac at the reducing end of the O-glycan and have recently been described to display regression of their glucosyltransferase enzyme (32). In the eukaryotic N-glycosylation system, a single glucose is recognized by lectins involved in the quality control of glycoprotein folding leading to transport to the Golgi or retrotranslocation to the cytoplasm; whereas in C. jejuni, we have previously demonstrated that mutation of the glucosyltransferase does not have a phenotype (2). It is interesting to speculate that the general O-glycoprotein pathway of Neisseria and the N-glycoprotein pathway of Campylobacter are exhibiting parallel bacterial evolution. Also, the fact that C. lari has lost several key metabolic enzymes and likely relies on metabolites provided by its restricted marine habitat will be interesting to investigate in future studies.
Fig. 7. N-glycan variation in C. hominis. A, B, and C show CID-associated fragmentation of three glycoforms that demonstrated the alteration of the distal moiety by the substitution of HexNAc for the 245 Da component. a, b, and c show extracted ion chromatograms and the corresponding high mass accuracy MS spectra of the associated glycoforms denoted as α, β, γ. Ion chromatograms showed an ~1 min
Campylobacter N-Glycan Pathways

We are currently trying to understand why the fOS in C. lari is phosphorylated. If C. lari fOS is involved in osmoregulation similar to C. jejuni (9), the negatively charged phosphate may assist in the retention of periplasmic fOS between the membranes, and similar modifications have been observed in other Gram-negative bacterial membrane-derived oligosaccharides (34). Alternatively, C. lari may have developed a more efficient method for recycling undecaprenylphosphate (UndP) by possessing an alternate enzyme capable of releasing P-fOS through cleavage of the pyrophosphate linkage of the lipid carrier (because the phosphate residue is alpha-linked to diNAcBac) or coupling UndP recycling with phosphorylation of PglB-released fOS.

Within group II, we identified several novel glycan structures. The C. fetus subgroup produces two different fOS and N-glycans that are both present in an 4:1 ratio (with GlcNAc or Glc branches). The reason the GlcNAc-containing hexasaccharide is preferred over the Glc-hexasaccharide is unknown, but may be caused by transferase efficiencies or nucleotide-donor availabilities. In the C. fetus subgroup, the narrower host range (35) may have resulted in the formation of two glycan structures as a potential mechanism to evade the host immune system by modulating the amount of N-glycan populations on their surface. A similar mechanism was found.

Fig. 8. The C. hominis isobaric glycoforms. The identification of isobaric glycans within C. hominis glycopeptides of the foldase protein, PrsA, is shown. A, An extracted ion chromatogram of 1450.62–1450.64 m/z demonstrates the present of two similarly eluting ions differing in elution by less than 1 min. CID of (B), the first peak, identified an alternated arrangement of the glycan compared with (C), the second peak, and HCD fragmentation of both peaks in (D) confirms the peptide component of these glycans as 246AANVQFTDTNATK265, Mascot ion score 73 and 55 respectively.

increase in elution time associated with substitution of the 245 Da moiety for a HexNAc. Inserts show high mass accuracy MS spectra of each glycoform demonstrating an increase in –42.02 Da associated with the β glycoform compared with the α glycoform, and the γ glycoform compared with β glycoform. D, HCD fragmentation of the β glycoform containing glycopeptide 99EDQNKEPFVLPDK109 of the preprotein translocase, SecG subunit (uniprot accession number A7I0C3_CAMHC; Mascot ion Score: 43).
Campylobacter N-Glycan Pathways

Molecular & Cellular Proteomics 11.11

1215
for the general O-glycosylation pathway in Neisseria species that was speculated to be driven by selection imposed by host adaptive and innate immunity (32, 36).

Because both C. fetus subgroup glycans are immunogenic, and the generated antisera do not cross react with group I species, this indicates that the terminal sugars are the immunodominant N-glycan epitopes on the bacterial cell surface. The split between terminal GalNAc residues in group I and predominantly terminal GlcNAc residues in group II is quite striking (Fig. 1). This is particularly noteworthy because it has been demonstrated that the human C-type lectin, MGL, binds terminal GalNAc residues on the C. jejuni N-glycan (17), but an MGL homolog in chickens has not been identified. Instead, the only characterized C-type lectin isolated from chicken hepatocytes was shown to recognize glycans terminating in GlcNAc residues (37), which are not present in any of the thermotolerant campylobacters. We are examining the C-type lectin repertoire on chicken dendritic cells to characterize the immune recognition of the pgl-derived glycans and to better understand the vaccine potential of these structures (38, 39).

Initial MS and MS² fOS analysis of the C. sputorum biovars and C. lanienae identified multiple unique glycans. At the MS level, two glycans were observed, one with a mass identical to that observed for group I whereas the other corresponded to a single glycan, 1.66TGADLNGTLAEQR129, Mascot ion score 49, of the periplasmic nitrate reductase was identified with the enzyme activity. Alkaline 284.05 ion corresponding to a HexNAc-P generated by the loss of ethanolamine. Conversely the nature of the 217 Da and 245 Da sugars remain obscure. A 217 Da monosaccharide is present in the N-glycan of H. pullorum (44) and identified within the capsule and LPS of Aeromonas salmonicida strain 80204-1 (45). Furthermore, high mass accuracy measurements of the 217 oxonium ion derived from C. rectus (m/z 218.06398) and that of the 217 oxonium ion of H. pullorum glycopeptides (m/z 218.06503) (data not shown) confirms these carbohydrates possess identical mass. This suggests the presence of acidic sugar residues in pgl glycans of specific group II species, which are, with the exception of C. hominis, predominantly found in the human oral cavity (46). The 245 Da component is unique to C. hominis, which is the only known human commensal among the taxa (47). C. hominis also exhibits the most extensive repertoire of N-glycan structures, with 3 glycans of differing mass detected of which one corresponds to two isobaric glycan forms. The 245 Da component may, based on its mass measurements, correspond to an O-acetylated HexNAc. Currently it is unknown if the observation of these multiple glycans provides insight into the true nonstoichiometric heterogeneity in the glycan or is a result of chemical loss of acetylation during glycopeptide preparation. It is interesting that commensal Bacteroides species also have an extensive repertoire of glycan structures on their surface that enable the organism to persist in the gut by modulating the host immune responses (48, 49). The reason why the group I species express a single, conserved N-glycan structure whereas group II

Additional variations in fOS composition were observed in group II. One striking feature was the presence of sugars with masses of 217 Da, 234 Da and 245 Da. We identified the 234 Da component as glucolactic acid, a sugar that was described as a component of the Shigella boydii type 17 (23) and E. coli O124 (43) lipopolysaccharides (LPS). However, the nature of the 217 Da and 245 Da sugars remain obscure. A 217 Da monosaccharide is present in the N-glycan of H. pullorum (44) and identified within the capsule and LPS of Aeromonas salmonicida strain 80204-1 as 2-acetamido-2-deoxy-D-galacturonic acid (GalNacA) (45). Furthermore, high mass accuracy measurements of the 217 oxonium ion derived from C. rectus (m/z 218.06398) and that of the 217 oxonium ion of H. pullorum glycopeptides (m/z 218.06503) (data not shown) confirms these carbohydrates possess identical mass. This suggests the presence of acidic sugar residues in pgl glycans of specific group II species, which are, with the exception of C. hominis, predominantly found in the human oral cavity (18, 46). The 245 Da component is unique to C. hominis, which is the only known human commensal among the taxa (47). C. hominis also exhibits the most extensive repertoire of N-glycan structures, with 3 glycans of differing mass detected of which one corresponds to two isobaric glycan forms. The 245 Da component may, based on its mass measurements, correspond to an O-acetylated HexNAc. Currently it is unknown if the observation of these multiple glycans provides insight into the true nonstoichiometric heterogeneity in the glycan or is a result of chemical loss of acetylation during glycopeptide preparation. It is interesting that commensal Bacteroides species also have an extensive repertoire of glycan structures on their surface that enable the organism to persist in the gut by modulating the host immune responses (48, 49). The reason why the group I species express a single, conserved N-glycan structure whereas group II
species express a range of N-glycan structures will be exciting to explore. The terminal phosphoethanolamine (PETN) modification identified on the C. gracilis IOS and glycoproteins was also unique and requires further study. PETN modifications present on the lipid A and FlgG protein of C. jejuni 81–176 are important for resistance against polymyxin B and for motility (50). In other bacteria, PETN modification of lipid A or LOS core sugars acts as a mechanism to increase antibiotic resistance (51, 52). Interestingly, isolates of C. gracilis show striking resistance against penicillin and various cephalosporins compared with other Campylobacter species (53). One of the striking features of the C. gracilis IOS analysis by MS was the dominant presence of B-ions compared with the Y-ions seen within the analysis of other IOS structures. This alteration in the dominant ion series is because of the proton sequestration ability of the basic ethanolamine group, akin to the alteration in the dominant ion series seen in LysN peptides or when C-terminal chemical derivatization is used to increase basicity (54, 55). This observation further supports the assignment of the PETN residue.

This work provides a comprehensive analysis of the N-glycosylation pathways of all known Campylobacter taxa through genome analysis, phylogenetic comparisons, antibody screening, NMR structural characterization and MS technologies through direct examination of both the N-glycosylated proteins and free oligosaccharides. These studies enabled us to confirm that this pathway is functional in all but one of these species and expands on recent studies examining N-glycosylation in H. pullorum and a subset of Campylobacter species examined in vitro (25, 44, 56). There are some discrepancies in structures observed in our study compared with those described by Jervis et al. (56). The difference in the N-glycan structures in C. concisus might be a result of intra-species diversity at the genome level (57), whereas the differences for C. laniariae are likely because of the methods used for analysis. In the latter study, the authors elucidated the N-glycan structures by glycosylating biotin-tagged peptides with membranes from the select Campylobacter species, but obtained yields too low to generate reliable mass-spectrometry data and therefore supplemented the reactions with the heterologous OTase from C. jejuni (56). In the case of C. fetus subsp. fetus, Jervis et al. observed one rather than the two glycan structures that we observed by NMR and MS at both the glycopeptide and fOS levels, but again this is likely caused by not analyzing this system directly in the host (56). Our studies suggest that protein N-glycosylation plays a fundamental role in Campylobacter biology because of the observed evolutionary pressure to conserve this pathway and the demonstration that glycosylation of specific proteins appears to be conserved across multiple species, such as the general secretory pathway protein, SecG, which we identified to be glycosylated in both C. hominis and C. gracilis and previously shown to be glycosylated in C. jejuni (58). Comparative studies with C. canadensis, the only Campylobacter species lacking this pathway, may help to elucidate the key function of protein glycosylation in campylobacters. The diversity in glycan structures among the species parallels the phylogenetic relatedness of the organisms and could potentially be associated with the natural hosts and niches that each species inhabits. Because the nonreducing sugar residues have been shown to be the determinants for adaptive and innate immune detection, host recognition mechanisms may in part be responsible for the pressures influencing N-glycan diversity within this genus and this diversity could be exploited to generate novel glycoconjugate vaccines and diagnostics.

Acknowledgments—We thank Tracy Raivio for access to the E. coli Keio collection, Markus Aebi for hR6 antiserum, Emma Yee for help with Campylobacter growth and Richard Fahlman for allowing the use of instrumentation.

* This work was supported by the Alberta Glycomics Centre. NES holds a CJ Martin Overseas Biomedical Postdoctoral Research Fellowship from the Australian National Health and Medical Research Foundation (APP1037373). CMS holds an Alberta Innovates Scholar award.

§ This article contains supplemental Fig. S1, Tables S1 to S13, and supplemental MS-Figures.

‡‡ Current address: College of Life Science and Technology, Hua- zhong University of Science and Technology, Wuhan, 430074, China.

** To whom correspondence should be addressed: Department of Biological Sciences, CCIS 6-061, University of Alberta, Edmonton, AB, T6G 2E9, Canada. Tel.: (780) 248-1234; Fax: (780) 492-9234; E-mail: cszymans@ualberta.ca.

REFERENCES

1. Szymanski, C. M., Yao, R., Ewing, C. P., Trust, T. J., and Guerry, P. (1999) Evidence for a system of general protein glycosylation in Campylobacter jejuni. Mol. Microbiol. 32, 1022–1030

2. Kelly, J., Jarrell, H., Millar, L., Tessier, L., Fiori, L. M., Lau, P. C., Allain, B., and Szymanski, C. M. (2006) Biosynthesis of the N-linked glycan in Campylobacter jejuni and addition onto protein through block transfer. J. Bacteriol. 188, 2427–2434

3. Kowark, M., Young, N. M., Numao, S., Schulz, B. L., Hug, I., Callewaert, N., Mills, D. C., Watson, D. C., Hernandez, M., Kelly, J. F., Wacker, M., and Aebi, M. (2006) Definition of the bacterial N-glycosylation site consensus sequence. EMBO J. 25, 1957–1966

4. Scott, N. E., Parker, B. L., Connolly, A. M., Paulech, J., Edwards, A. V., Crossett, B., Falconer, L., Kolarich, D., Djordjevic, S. P., Hojrup, P., Packer, N. H., Larsen, M. R., and Cordwell, S. J. (2010) Simultaneous glycans-peptide characterization using hydrophilic interaction chromatography and parallel fragmentation by CID, HCD and ETD-MS applied to the N-linked glycoproteome of Campylobacter jejuni. Mol. Cell. Proteomics 10, 10.1074/mcp.M000031

5. Wacker, M., Linton, D., Hitchen, P. G., Nita-Lazar, M., Haslam, S. M., North, S. J., Panico, M., Morris, H. R., Dell, A., Wren, B. W., and Aebi, M. (2002) N-linked glycosylation in Campylobacter jejuni and its functional transfer into E. coli. Science 298, 1790–1793

6. Feldman, M. F., Wacker, M., Hernandez, M., Hitchen, P. G., Marolda, C. L., Kowark, M., Morris, H. R., Dell, A., Valvano, M. A., and Aebi, M. (2005) Engineering N-linked protein glycosylation with diverse O-antigen lipopolysaccharide structures in Escherichia coli. Proc. Natl. Acad. Sci. U.S.A. 102, 3016–3021

7. Wacker, M., Feldman, M. F., Callewaert, N., Kowark, M., Clarke, B. R., Pohl, N. L., Hernandez, M., Vines, E. D., Valvano, M. A., Whitfield, C., and Aebi, M. (2006) Substrate specificity of bacterial oligosaccharyltrans-
ferase suggests a common transfer mechanism for the bacterial and eukaryotic systems. Proc. Natl. Acad. Sci. U.S.A. 103, 7088–7093

28. Good, D. M., Wenger, C. D., McAlister, G. C., Bai, D. L., Hunt, D. F., and Coon, J. J. (2009) Post-acquisition ETD spectral processing for increased peptide identifications. J. Am. Soc. Mass Spectrom. 20, 1435–1440

29. Qin, J., and Chait, B. T. (1997) Identification and characterization of post-translational modifications of MALDI ion trap mass spectrometry. Anal. Chem. 69, 4002–4009

30. Miller, W. G., Wang, G., Binnewies, T. T., and Parker, C. T. (2008) The chaperone gene nhaA in rabbits. World J. Gastroenterol. 14, 778–779

31. Inglis, G. D., Kalischuk, L. D., and Busz, H. W. (2003) A survey of Campylobacter species shed in faeces of beef cattle using polymerase chain reaction. Can. J. Microbiol. 49, 721–730

32. Wuhrer, M., Koeleman, C. A., and Deelder, A. M. (2009) Hexose rearrangement of the family Rhizobiaceae. Proc. Natl. Acad. Sci. U.S.A. 106, 9643–9648

33. Krinos, C. M., Coyne, M. J., Weinacht, K. G., Tzianabos, A. O., Kasper, J. L., Morris, H. R., Dell, A., Wren, B., and Linton, D. (2010) Characterisation of N-linked protein glycosylation in Helicobacter pullorum. J. Bacteriol. 192, 5228–5236

34. Wang, Z., Larocque, B., Vinogradov, E., and Brisson, J. R. (2003) Detection of conserved O antigen of Aeromonas salmonicida strain 80204-1 produced under in vivo growth conditions. Eur. J. Biochem. 271, 4507–4516

35. Lee, H., Hsu, F. F., Turk, J., and Groisman, E. A. (2004) The PmrA-regulated pmrC gene mediates phosphoethanolamine modification of lipid A and poly- myxin resistance in Salmonella enterica. J. Bacteriol. 186, 4124–4133
52. Cox, A. D., Wright, J. C., Li, J., Hood, D. W., Moxon, E. R., and Richards, J. C. (2003) Phosphorylation of the lipid A region of meningococcal lipopolysaccharide: identification of a family of transferases that add phosphoethanolamine to lipopolysaccharide. *J. Bacteriol.* **185**, 3270–3277.

53. Johnson, C. C., Reinhardt, J. F., Edelstein, M. A., Mulligan, M. E., George, W. L., and Finegold, S. M. (1985) Bacteroides gracilis, an important anaerobic bacterial pathogen. *J. Clin. Microbiol.* **22**, 799–802.

54. Hennrich, M. L., Boersema, P. J., van den Toorn, H., Mischerikow, N., Heck, A. J., and Mohammed, S. (2009) Effect of chemical modifications on peptide fragmentation behavior upon electron transfer induced dissociation. *Anal. Chem.* **81**, 7814–7822.

55. Hohmann, L., Sherwood, C., Eastham, A., Peterson, A., Eng, J. K., Eddes, A. J., and Mohammed, S. (2009) Effect of chemical modifications on peptide fragmentation behavior upon electron transfer induced dissociation. *Anal. Chem.* **81**, 7814–7822.

56. Baba, T., Ara, T., Miyano, K., Ishihama, Y., and Datsenko, K. A. (2002) In-frame, site-directed insertions in Escherichia coli K-12 (e15). *Proc. Natl. Acad. Sci. U.S.A.* **99**, 7154–7159.

57. Deshpande, N. P., Kaakoush, N. O., Mitchell, H., Janitz, K., Raftery, M. J., Ve´ron, M., and Chatelain, R. (1973) Taxonomic study of the genus Campylobacter. *Zbl. Bakt. Abt. I Orig.* **2170–2185.

58. Parkhill, J., Wren, B. W., Mungall, K., Ketley, J. M., Churcher, C., Basham, D., Chillingworth, T., Davies, R. M., Feltwell, T., Holroyd, S., Jagels, K., Jin, L., Li, S. S., and Wilkins, M. R. (2011) Sequencing and validation of the glycans of campylobacter species. *PLoS One.* **6**, e22170.

59. Hohmann, L., Sherwood, C., Eastham, A., Peterson, A., Eng, J. K., Eddes, A. J., and Mohammed, S. (2009) Effect of chemical modifications on peptide fragmentation behavior upon electron transfer induced dissociation. *Anal. Chem.* **81**, 7814–7822.

60. Baba, T., Ara, T., Miyano, K., Ishihama, Y., and Datsenko, K. A. (2002) In-frame, site-directed insertions in Escherichia coli K-12 (e15). *Proc. Natl. Acad. Sci. U.S.A.* **99**, 7154–7159.

61. Parkhill, J., Wren, B. W., Mungall, K., Ketley, J. M., Churcher, C., Basham, D., Chillingworth, T., Davies, R. M., Feltwell, T., Holroyd, S., Jagels, K., Jin, L., Li, S. S., and Wilkins, M. R. (2011) Sequencing and validation of the glycans of campylobacter species. *PLoS One.* **6**, e22170.

62. Ve´ron, M., and Chatelain, R. (1973) Taxonomic study of the genus Campylobacter. *Zbl. Bakt. Abt. I Orig.* **2170–2185.

63. Tanner, A. G. R., Badger, J. H., Fraser, D. A., Lai, C. H., Listgarten, M. A., Visconti, R. A., and Socransky, S. S. (1981) Wolinella recta sp. nov., a hippurate-positive, butyrate-negative anaerobic bacterial pathogen. *Int. J. Syst. Bacteriol.* **31**, 432–445.

64. Logan, J. M. J., Burnens, A., Linton, A., Lawson, A. J., and Stanley, J. (2000) Characterization of the structurally diverse N-linked glycans of Clostridium perfringens species. *J. Bacteriol.* **182**, 7154–7159.

65. Gorkiewicz, G., Feierl, G., Schober, C., Dieber, F., Kofer, J., Zechner, R., and Zechner, E. L. (2003) Species-specific identification of campylobacters by partial 16S rRNA gene sequencing. *J. Clin. Microbiol.* **41**, 2537–2546.

66. On, S. L., Atabay, H. I., Corry, J. E., Harrington, C. S., and Vandamme, P. (1998) Emended description of *Campylobacter jejuni* and *Campylobacter coli*. *Int. J. Syst. Bacteriol.* **48 Pt 1**, 195–206.

67. Etoh, Y., Dewhirst, F. E., Paster, B. J., Yamamoto, A., and Goto, N. (1993) *Campylobacter showae* sp. nov., isolated from the human oral cavity. *Int. J. Syst. Bacteriol.* **43**, 631–639.

68. Stanely, J., Burnens, A., Linton, A., Lawson, A. J., and Owen, R. J. (1992) *Campylobacter helveticus* sp. nov., a new thermophilic species from domestic animals: characterization, and cloning of a species-specific DNA probe. *Microbiology* **138**, 2293–2303.

69. Gebhart, C. J., Edmonds, P., Ward, G. E., Kurtz, H. J., and Brenner, D. J. (1985) “Campylobacter hyointestinalis” sp. nov.: a new species of Campylobacter found in the intestines of pigs and other animals. *J. Clin. Microbiol.* **21**, 715–720.

70. Churchill, G. A., and Steigerwalt, A. G., Lawson, P. A., Thorne, P., Byrer, D. E., Ross, H. M., Xeroy, J., Thompson, P. M., and Collins, M. D. (2004) *Campylobacter insulae* sp. nov., isolated from marine mammals. *Int. J. Syst. Evol. Microbiol.* **54**, 2369–2373.

71. Debruyne, L., Broman, T., Bergstrom, S., Olsen, B., On, S. L., and Vandamme, P. (2010) *Campylobacter jejuni* sp. nov., isolated from the porcine stomach, and an emended description of *Campylobacter jejuni*. *Int. J. Syst. Evol. Microbiol.* **60**, 2364–2369.

72. Foster, G., Holmes, B., Steigerwalt, A. G., Lawson, P. A., Thorne, P., Byrer, D. E., Ross, H. M., Xeroy, J., Thompson, P. M., and Collins, M. D. (2004) *Campylobacter insulae* sp. nov., isolated from marine mammals. *Int. J. Syst. Evol. Microbiol.* **54**, 2369–2373.

73. Zanoni, R. G., Debruyne, L., Rossi, M., Revez, J., and Vandamme, P. (2009) *Campylobacter subantarcticus* sp. nov., isolated from black-headed gulls (Larus ridibundus). *Int. J. Syst. Evol. Microbiol.* **60**, 1870–1875.