JB
Journals.ASM.org

Characterization of the Structurally Diverse N-Linked Glycans of *Campylobacter* Species

Adrian J. Jervis,[a,b] Jonathan A. Butler,[a] Andrew J. Lawson,[c] Rebecca Langdon,[b] Brendan W. Wren,[b] and Dennis Linton[a]

Faculty of Life Sciences, University of Manchester, Manchester, United Kingdom[a]; Pathogen Molecular Biology Unit, London School of Hygiene and Tropical Medicine, London, United Kingdom[b]; and Laboratory of Gastrointestinal Pathogens, Centre for Infections, Health Protection Agency, London, United Kingdom[c]

The Gram-negative bacterium *Campylobacter jejuni* encodes an extensively characterized N-linked protein glycosylation system that modifies many surface proteins with a heptasaccharide glycan. In *C. jejuni*, the genes that encode the enzymes required for glycan biosynthesis and transfer to protein are located at a single *pgl* gene locus. Similar loci are also present in the genome sequences of all other *Campylobacter* species, although variations in gene content and organization are evident. In this study, we have demonstrated that only *Campylobacter* species closely related to *C. jejuni* produce glycoproteins that interact with both a *C. jejuni* N-linked-glycan-specific antiserum and a lectin known to bind to the *C. jejuni* N-linked glycan. In order to further investigate the structure of *Campylobacter* N-linked glycans, we employed an *in vitro* peptide glycosylation assay combined with mass spectrometry to demonstrate that *Campylobacter* species produce a range of structurally distinct N-linked glycans with variations in the number of sugar residues (penta-, hexa-, and heptasaccharides), the presence of branching sugars, and monosaccharide content. These data considerably expand our knowledge of bacterial N-linked glycan structure and provide a framework for investigating the role of glycosyltransferases and sugar biosynthesis enzymes in glycoprotein biosynthesis with practical implications for synthetic biology and glycoengineering.

The asparagine-linked glycosylation of proteins is common in eukaryotes and is involved in protein folding, quality control, secretion, and sorting (10). Glycosylation involves the transfer of an oligosaccharide from a lipid carrier to asparagine residues within the consensus sequence N-X-S/T (X is any residue except proline) and takes place in the endoplasmic reticulum (ER) lumen as nascent polypeptides are secreted across the ER membrane (26). Analogous protein N-glycosylation reactions also occur in a restricted set of species from the archaeal and bacterial domains (3, 37). In eukaryotes, protein N-glycosylation usually involves a multiprotein complex (34), but only a single subunit, the STT3 protein of the yeast complex, has bacterial/archaeal orthologues, namely, PglB and AglB, respectively. The STT3, PglB, and AglB proteins all have multiple transmembrane domains within the amino terminus and a soluble carboxy terminus (3, 31, 32, 34, 37). The STT3 protein is the catalytic subunit of the yeast oligosaccharyltransferase (OTase) responsible for transferring oligosaccharide from the lipid carrier onto the target protein, but PglB and AglB act as single-subunit OTases (1, 18, 50).

The prototypical bacterial N-linked OTase is the *Campylobacter jejuni* PglB protein. The *C. jejuni pglB* gene is located within a cluster of genes termed the protein glycosylation (*pgl*) locus (45). This locus encodes all of the necessary enzymes for N-linked protein glycosylation with a heptasaccharide consisting of GalNAc–α1,4-GalNAc–α1,4(Glcβ1,3)-GalNAc–α1,4-GalNAc–α1,4-GalNAc–α1,3-Bac (Bac is bacillosamine or 2,4,-diacetamido-2,4,6-trideoxyglucose). Functional analysis of the genes from the *pgl* locus has revealed that the heptasaccharide is synthesized by five glycosyltransferases encoded by the *pglA*, *pglC*, *pglH*, *pglI*, and *pglJ* genes (6, 7, 30, 38). The lipid-linked heptasaccharide is transferred across the bacterial inner membrane and into the periplasm by a "flippase," PglK (2), where it is covalently linked to asparagine residues within an extended D/E-X-N-X-S/T sequon by the PglB protein (4, 22, 23).

More than 50 *C. jejuni* N-linked glycoproteins with diverse functions have been identified (29, 35, 42, 51). Mutants deficient in N-linked glycosylation are less competent (28), have a reduced capacity to invade *in vitro*-cultured epithelial cells, and are less able to colonize chicks and mice (11, 16, 17, 43). It has also been proposed that PglB-mediated release of free oligosaccharide in the periplasm plays a role in osmotic protection (36). Finally N-linked-heptasaccharide-specific recognition by macrophages is thought to have an immunosuppressive role (46). From these findings, it is clear that the role of N-linked protein glycosylation in campylobacters and related species is likely to be both significant and complex.

The *C. jejuni pgl* locus is fully functional when transferred to *Escherichia coli*, allowing glycosylation with heptasaccharide of both *C. jejuni* and non-*C. jejuni* proteins that contain the N-linked glycosylation sequon (50). When structurally variant lipid-linked glycans are made available, the *C. jejuni* PglB OTase is able to transfer many, but not all, of these onto proteins (5, 49). Such flexibility makes this an attractive approach for the development of an *in vivo* glycoconjugate synthesis system (14). The key structural requirement for PglB-mediated glycan transfer is the presence of an acetamido group on the C-2 carbon of the reducing-end sugar (5, 49).

Recently released bacterial genome sequence data indicate that many members of the *Deltaproteobacteria-Epsilonproteobacteria* subdivision of the *Proteobacteria* contain *pglB* orthologues (13, 15, 25, 37, 44). Prominent genera within the *Epsilonproteobacteria* subdivision include *Arcobacter*, *Campylobacter*, and *Helicobacter*.

Received 12 January 2012 **Accepted** 21 February 2012
Published ahead of print 2 March 2012
Address correspondence to Dennis Linton, james.d.linton@manchester.ac.uk.

doi:10.1128/JB.00042-12

Species of the *Arcobacter* genus lack *pglB* orthologues, and within the *Helicobacter* genus, only three closely related species (*H. pullorum*, *H. canadensis*, and *H. winghamensis*) contain *pglB* orthologues; we have characterized the *H. pullorum*-associated N-linked glycan (15). In contrast, all *Campylobacter* species with available genome sequence data contain one or, in a small number of cases, two *pglB* orthologues. Aside from *C. jejuni* and, to a much lesser extent, *Campylobacter lari* (41), other *Campylobacter*-encoded N-linked protein glycosylation systems are as yet completely uncharacterized.

In this study, we set out to provide a broader perspective on N-linked protein glycosylation within the *Campylobacter* genus as a whole. We employed an *in vitro* peptide N-glycosylation assay combined with mass spectrometry (MS) to rapidly characterize N-linked glycan structures from nine *Campylobacter* species. This considerably expands our knowledge of bacterial N-linked protein glycosylation systems, providing a further set of glycoactive enzymes for potential biotechnological exploitation.

MATERIALS AND METHODS

Bacterial strains and plasmids. All *E. coli* strains were grown in Luria-Bertani (LB) broth or on LB agar plates including antibiotic supplements as appropriate, and all *Campylobacter* species were grown on Columbia agar (Oxoid Ltd.) containing 5% defibrinated horse blood (TCS Biosciences). *C. jejuni* NCTC 11168H, *Campylobacter coli* RM2228, *C. lari* RM2100, and *Campylobacter upsaliensis* NCTC11541 were all grown at 42°C in a VA500 workstation (Don Whitley Ltd.) maintaining a microaerobic atmosphere of 85% N_2, 10% CO_2, and 5% O_2. *Campylobacter helveticus* NCTC 12470, *Campylobacter hyointestinalis* subsp. *hyointestinalis* NCTC 11608, *Campylobacter fetus* subsp. *fetus* NCTC 10842, and *Campylobacter lanienae* NCTC 13004 were grown at 42°C in gas jars under microaerobic conditions generated using the Campygen system (Oxoid Ltd.) supplemented with ~10% H_2. *Campylobacter concisus* NCTC 11485 was grown in the same manner but at 37°C, and *Campylobacter sputorum* biovar *sputorum* NCTC 11528 was grown anaerobically in a gas jar with H_2 at 42°C.

Insertional inactivation of *C. fetus pglB*. A 2.7-kb PCR product was amplified from *C. fetus* genomic DNA using primers CfetpglB-F (5′-CTT TTGATATTAGGCGAGGG-3′) and CfetpglB-R (5′-TACACCTTAGCG CTTGGATC-3′). The resultant PCR product, from ~420 bp upstream of the predicted start codon of the *pglB* gene to 12 bp upstream of the predicted stop codon, was ligated into pGEMTEasy (Stratagene) and digested with XbaI, removing a 1.2-kb internal region of the *pglB* gene. A kanamycin resistance cassette (*aphA*) lacking a transcriptional terminator was ligated into the XbaI-digested vector in the same transcriptional orientation as the *pglB* gene. The resultant plasmid was electroporated into electrocompetent *C. fetus* cells as described previously (47). Kanamycin-resistant transformants were isolated, and integration by a double-crossover event was verified by PCR. Approximately 1 in 20 recombination events was a double crossover, as previously reported (9).

Glycan detection with hR6 antiserum and soybean agglutinin (SBA) lectin. Campylobacters were grown as described above, and cells were harvested from agar plates, standardized by optical density at 600 nm, and heated to 95°C for 5 min in SDS-PAGE loading buffer. These whole-cell lysates were subjected to 12% SDS-PAGE, transferred to polyvinylidene difluoride membrane, and probed either with hR6 antiserum (S. Amber and M. Aebi, unpublished data) followed by an IRDye 800CW goat anti-rabbit IgG secondary antibody (Li-Cor) or with biotin-conjugated lectin SBA (Vector Labs) followed by IRDye 800CW streptavidin (Li-Cor). Blots were imaged using a Li-Cor Odyssey Infrared Imaging System.

***In vitro* OTase assays.** *In vitro* OTase assays were carried out as previously described (15, 21). In brief, a standard *in vitro* OTase assay mixture contained 26 μl of a Triton X-100-solubilized membrane preparation, 3 μl of 30 μM peptide (6-carboxyfluorescein–ADQNATA–NH_2), and 1 μl of 150 mM $MnCl_2$. When the glycopeptide yield was low, reaction mixtures were supplemented with 3 μl of a Triton X-100-solubilized membrane preparation from *E. coli* SCM7 (2) expressing the *C. jejuni pglB* gene from plasmid pMAF10 (5). Assay mixtures were incubated overnight at 30°C with agitation, and reaction products were analyzed by Tris-Tricine SDS-PAGE with a 16.5% resolving gel containing 6 M urea and a 10% spacer gel (39). Gels were visualized using a Typhoon Trio Plus Imager (GE Healthcare) with excitation at 532 nm (green laser) and a 526-nm short-pass emission filter.

Purification and matrix-assisted laser desorption ionization–time of flight (MALDI-TOF) MS analysis of biotinylated peptides. Biotinylated peptides (fluorescein isothiocyanate [FITC]-ADQNATAK-biotin) were purified from OTase assay mixtures using streptavidin-coated magnetic beads (Pierce) as previously described (15) but with the following modifications. When efficiency of transfer was low, some assays were carried out in larger quantities and/or with an increased amount of membrane preparation relative to that of peptide (up to 10-fold). In such cases, the quantity of magnetic beads was increased in accordance with the volume of the membrane preparation used. Samples were eluted from streptavidin-coated beads by incubation in 10 μl 70% acetonitrile–5% formic acid–125 μM biotin for 30 min at room temperature and desalted on a C_{18} ZipTip (Millipore). MALDI-TOF MS and MALDI-LIFT-TOF/TOF MS spectra were acquired by laser-induced dissociation (LID) using a Bruker Ultraflex II mass spectrometer in the positive-ion reflection mode with 50 mM α-cyano-4-hydroxycinnamic acid (in 80% acetonitrile) as the matrix. Data were analyzed with FlexAnalysis 3.0 software (Bruker Daltonics).

RESULTS

Diversity of *Campylobacter pgl* gene loci. Available genome sequence data indicate that all *Campylobacter* species encode a *pglB* orthologue located within a cluster of genes many of which are orthologues of genes from the well-characterized *C. jejuni pgl* gene locus (25, 37) (Fig. 1). This gene cluster in *C. coli*, the species most closely related to *C. jejuni*, has an essentially identical overall arrangement of genes, and individual gene products show significant levels of amino acid sequence identity (82 to 97%) to the *C. jejuni pgl* locus-encoded gene products (data not shown). Similarly, the *pglB* gene-containing locus from *C. lari*, also closely related to *C. jejuni*, is very similar in overall organization but with lower levels of orthologue sequence identity (56 to 83%; data not shown). The *C. lari* locus also lacks an orthologue of the *C. jejuni pglI* gene encoding a glucosyltransferase involved in heptasaccharide biosynthesis (30, 38, 41). More significant differences in *pgl* locus gene content and arrangement are evident in *Campylobacter* species more distantly related to *C. jejuni* (Fig. 1; data not shown). For example, in *C. upsaliensis* and *C. fetus*, although the overall arrangement of the putative *pgl* locus is still similar to that of the *C. jejuni* locus, there are insertions of 21 and 24 open reading frames (ORFs), respectively, between the putative *pglD* and *pglE* genes (Fig. 1). These insertions are apparently unrelated to each other and encode putative proteins with no similarity to enzymes involved in sugar biosynthesis or transport and are therefore likely not involved in N-linked protein glycosylation (data not shown). Comparison of the *C. fetus* and *C. jejuni* loci reveals further divergence, with additional ORFs encoding putative glycosyltransferases and sugar biosynthesis enzymes upstream of *pglB* (Fig. 1). Loci from *C. gracilis*, *C. hominis*, *C. showae*, and *C. curvus* also contain similarly located additional ORFs encoding putative glycosyltransferases and enzymes involved in sugar biosynthesis. Several *Campylobacter* species (*C. concisus*, *C. curvus*, and *C. gracilis*) possess two *pglB* orthologues, and these can be located within the

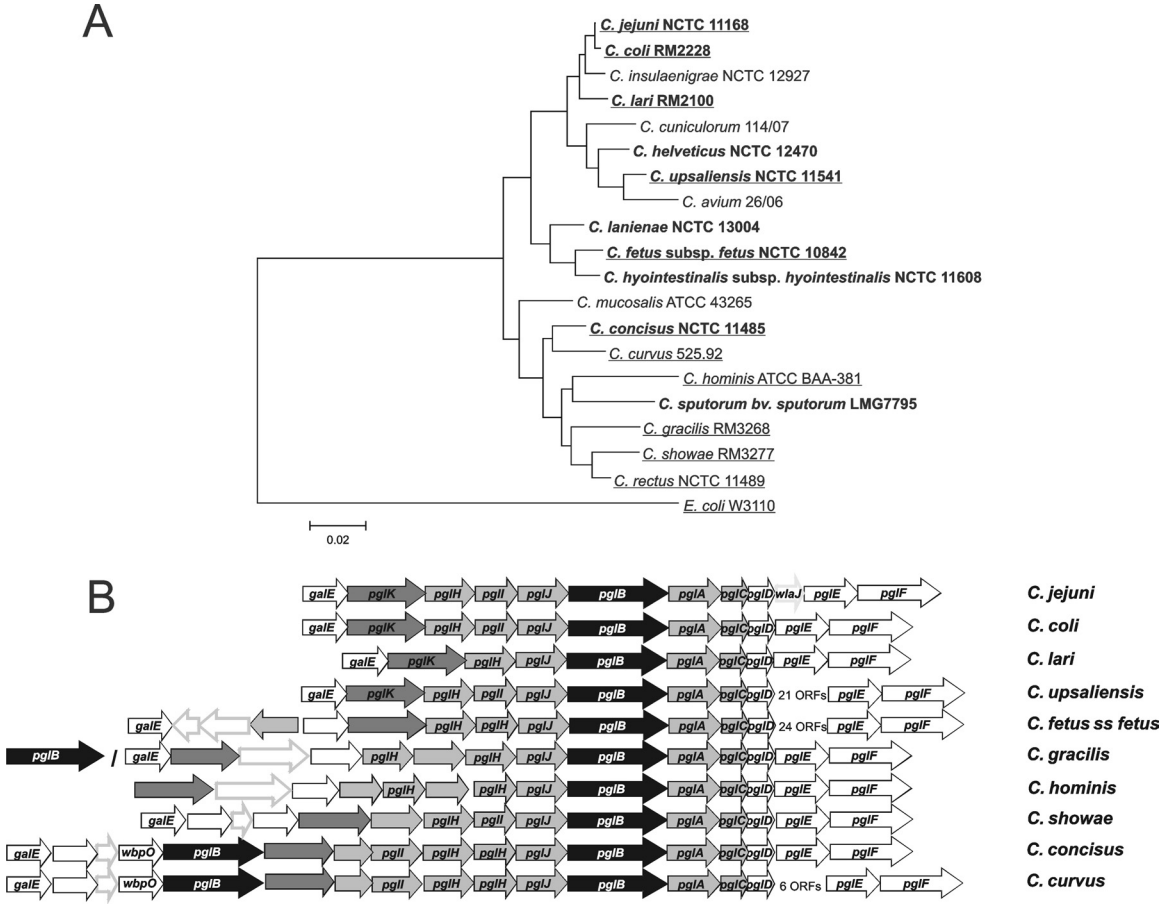

FIG 1 Inferred phylogeny (A) and putative *pgl* loci (B) of *Campylobacter* species. (A) A phylogenetic tree was generated using approximately 1,300 bases of 16S rRNA gene sequence from *Campylobacter* species and *E. coli* K-12 strain W3110 as an outlier. Sequences were aligned using ClustalW2 (27) and a neighbor-joining phylogenetic tree produced using MEGA5 software (24). Species in bold type were analyzed further in this study, and the genome sequences of those underlined are available. (B) The *C. jejuni* NCTC 11168 PglB amino acid sequence was used as the query in a BLAST search against the nonredundant protein database to identify further *Campylobacter pglB* genes. For species whose genomes have been sequenced, predicted amino acid sequences from ORFs surrounding the *pglB* genes were used as queries in BLAST searches to identify potential functions. Where BLAST searches resulted in hits with *C. jejuni* Pgl proteins, a Needleman-Wunsch pairwise global alignment (NCBI) of the query and the corresponding *C. jejuni* NCTC 11168 orthologue was performed. Where the amino acid identity was >30% (gaps, <10%), ORFs were labeled according to the *C. jejuni pgl* nomenclature (data not shown). Those with predicted functions involved in polysaccharide biosynthesis are identified as follows: putative N-linked OTases or PglB proteins, black shading; sugar biosynthesis enzymes, no shading; glycosyltransferases, light shading; transporters, dark shading; unrelated genes or genes of unknown function, grey outline with no shading. Excluded are *Campylobacter rectus*, which has sequencing errors in the *pglB* gene (37), and *C. fetus* subsp. *venerealis*, which has the same gene arrangement in this locus as *C. fetus* subsp. *fetus*.

pgl locus (*C. concisus*, *C. curvus*) or one of the two may be located outside the *pgl* locus (*C. gracilis*).

Reactivity of diverse *Campylobacter* species with *C. jejuni* glycan interacting reagents. The *C. jejuni N*-linked heptasaccharide can be detected on glycoproteins via interaction with lectins such as SBA (29) which bind to terminal GalNAc residues or a polyclonal antiserum (hR6) specific for the terminal two or three GalNAc residues of the *C. jejuni* N-linked glycan (Amber and Aebi, unpublished). When Western blot assays of whole-cell lysates of *Campylobacter* species were probed with hR6 antiserum, those from *C. jejuni*, *C. coli*, *C. lari*, *C. helveticus*, and *C. upsaliensis* were strongly immunoreactive (Fig. 2A). The dependence of hR6 reactivity on the presence of N-linked glycan was confirmed by the complete lack of signal (Fig. 2A) in a previously characterized *C. jejuni pglB* knockout mutant (30, 51). Interestingly, whole-cell extracts from *C. hyointestinalis*, *C. fetus*, *C. lanienae*, *C. sputorum*, and *C. concisus* were unreactive with the hR6 antiserum, suggesting structural differences in any N-linked glycans produced by these species. The lectin SBA produced essentially the same pattern and specificity of binding, although for some species there were additional bands that we attribute to nonspecific interactions (Fig. 2B). These data suggest that the N-linked glycans from the phylogenetically closely related species *C. jejuni*, *C. coli*, *C. lari*, *C. helveticus*, and *C. upsaliensis* all have structural features in common. In contrast, the species more distantly related to this group either do not produce N-linked glycoproteins or produce structurally diverse N-linked glycans that do not interact with either antiserum hR6 or lectin SBA. The latter explanation is strongly favored, at least for *C. fetus* and *C. concisus*, given the presence of putative N-linked glycosylation loci in their respective genome sequences (Fig. 1). Furthermore, these loci differ significantly in gene content from the *C. jejuni* prototype (Fig. 1), indicating potential for the biosynthesis of structurally distinct N-linked glycans that may well not interact

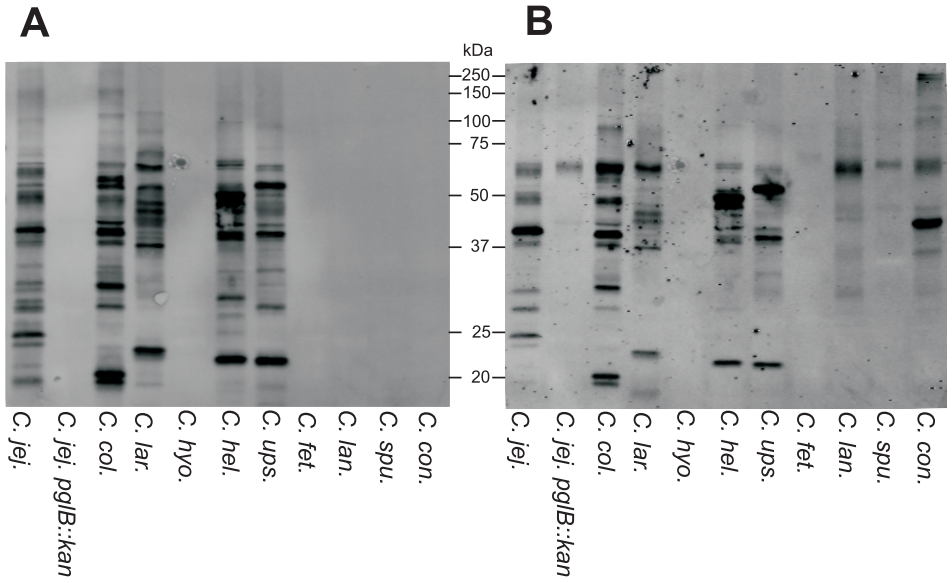

FIG 2 Reactivity of *Campylobacter* proteins with *C. jejuni* N-linked-glycan-specific reagents. Western blot assays of whole-cell lysates from *C. jejuni* (*C. jej.*) and the corresponding isogenic *pglB* knockout mutant (*C. jej. pglB*::*kan*) along with *C. coli* (*C. col.*), *C. lari* (*C. lar.*), *C. hyointestinalis* (*C. hyo.*), *C. helveticus* (*C. hel.*), *C. upsaliensis* (*C. ups.*), *C. fetus* (*C. fet.*), *C. lanienae* (*C. lan.*), *C. sputorum* (*C. spu.*), and *C. concisus* (*C. con.*) were probed with *C. jejuni* N-linked-glycan-specific hR6 antiserum (A) or SBA lectin (B).

with reagents with some degree of specificity for the *C. jejuni* N-linked glycan.

***In vitro* N-linked glycosylation in *Campylobacter* species.** We recently employed a modification of an *in vitro* OTase assay (21) to investigate *N*-linked glycosylation processes in *H. pullorum* (15), and we have now used a similar approach to characterize N-linked protein glycosylation in nine *Campylobacter* species. Detergent-solubilized membrane preparations from individual *Campylobacter* species were incubated with a fluorescently labeled peptide, ADQNATA, incorporating a *C. jejuni*-optimized N-glycosylation acceptor sequon (4). Products were separated by Tris-Tricine SDS-PAGE in which glycopeptides displayed reduced electrophoretic mobility compared to that of the unmodified peptide (Fig. 3). Each of the *Campylobacter* species produced at least one band with reduced electrophoretic mobility, although some are relatively faint, indicating varying efficiencies of individual reactions (Fig. 3A). In order to improve the efficiency of glycan transfer, reaction mixtures were supplemented with a membrane preparation from *E. coli* SCM7 expressing the *C. jejuni pglB* gene (see Materials and Methods for details). When resultant products were analyzed by gel electrophoresis, the overall banding pattern was essentially the same as that produced by reaction mixtures containing only *Campylobacter*-derived membrane preparations (compare Fig. 3A and B), but for a number of species (*C. hyointestinalis*, *C. helveticus*, *C. upsaliensis*, *C. lanienae*, and *C. concisus*), the fluorescence intensity of the derived glycopeptide bands was significantly increased, in-

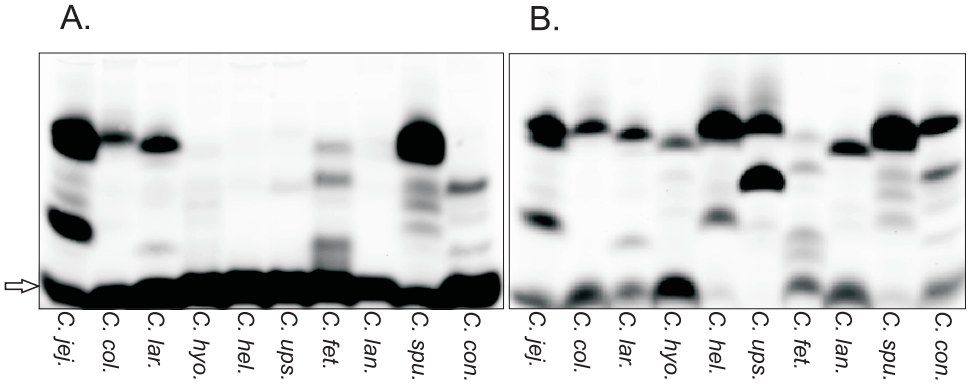

FIG 3 OTase activities of *Campylobacter* species. *Campylobacter* membrane preparations were solubilized in detergent and assayed for N-linked OTase activity using a fluorescently labeled peptide (ADQNATA). Reaction products were separated by Tris-Tricine SDS-PAGE. The unmodified peptide (arrow) runs toward the bottom of the gel, while derived glycosylated peptides display reduced mobility. For reference, the lowest-mobility product produced by *C. jejuni* membranes corresponds to the peptide modified with a heptasaccharide (15). The OTase activity displayed by membrane preparations from a range of *Campylobacter* species (A) was, in a number of cases, significantly increased when membranes prepared from *E. coli* expressing *C. jejuni pglB* were added to reaction mixtures (B). The species tested were *C. jejuni* (*C. jej*), *C. coli* (*C. col.*), *C. lari* (*C. lar.*), *C. hyointestinalis* (*C. hyo.*), *C. helveticus* (*C. hel.*), *C. upsaliensis* (*C. ups.*), *C. fetus* (*C. fet.*), *C. lanienae* (*C. lan.*), *C. sputorum* (*C. spu.*), and *C. concisus* (*C. con.*).

dicating enhanced transfer of glycan, presumably due to *C. jejuni* PglB activity.

The differing mobilities of the presumed glycopeptide products (Fig. 3) indicated potential structural differences in N-linked glycans, and these were further structurally characterized (see below). In parallel experiments employing synthetic peptides with either an Asn-to-Gln substitution or an Asp-to-Ala substitution at the −2 position of the N-glycosylation sequon, bands with reduced electrophoretic mobility were not generated by any of the *Campylobacter* species (unpublished data). The requirement for an Asn residue is consistent with an N-linked modification (structurally confirmed [see below]), and the requirement for an Asp residue suggests that all *Campylobacter* PglB proteins require a negatively charged residue at the −2 position, as demonstrated for *C. jejuni* PglB (23, 35).

Structural diversity in *Campylobacter* N-linked glycans. Using a biotin-tagged sequon-containing peptide (FITC-ADQNATAK-biotin) that can be readily purified from OTase reaction mixtures prior to analysis by MS (15), we attempted to structurally characterize N-linked glycans of glycopeptides produced by diverse *Campylobacter* species. For a number of species (*C. upsaliensis*, *C. helveticus*, *C. lanienae*, and *C. fetus*), the low yield of glycopeptide produced was insufficient to generate reliable MS data and two distinct approaches were employed to increase glycopeptide production. The first involved combining membrane preparations from an *E. coli* strain expressing the *C. jejuni pglB* gene with *Campylobacter* membrane preparations as described above, and this modification significantly increased the glycopeptide yield for the majority of the species. For *C. fetus*, an alternative approach involving the construction of a *C. fetus pglB* insertional knockout mutant was required (see Materials and Methods). When membrane preparations from the *C. fetus pglB*::*aphA* insertional knockout mutant and the corresponding wild-type strain were combined, significantly increased glycopeptide product was observed (data not shown), presumably due to the accumulation of lipid-linked oligosaccharide (LLO) on membranes derived from the *pglB* mutant strain (15). By these approaches, we were able to produce sufficient glycopeptides from the set of nine *Campylobacter* species for subsequent MS characterization.

During MALDI-TOF MS analysis, the reporter peptide produces a number of predominant peaks, including *m*/*z* 1,589 (theoretical mass, 1,588 Da), 1,556, and 1,199, the latter of which is due to the total loss of FITC (15). Purified glycopeptides produced by *Campylobacter* species were characterized by MALDI-LIFT-TOF/TOF MS, and representative spectra are presented (Fig. 4). For each species, a single glycopeptide was detected with glycan fragmentation patterns containing a peak with an *m*/*z* value of either 1,199 or 1,556 at the lower-mass end of the spectrum. Peaks with *m*/*z* values of 1,182 and 1,539 (17 Da less than 1,199 and 1,556, respectively) were also observed, characteristic of the fragmentation of a side-chain amide bond of an N-linked glycan (Fig. 4). Therefore, most glycopeptide parent ions consisted of intact peptide plus glycan, and under tandem MS analysis, the FITC moiety fragmented (spectra containing an *m*/*z* 1,199 peak), while some did not lose the whole moiety (spectra containing an *m*/*z* 1,556 peak). In the case of *C. lari*, the FITC was not present in the parent ion.

The MALDI-LIFT-TOF/TOF MS fragmentation patterns produced were consistent with sequential loss of sugar residues due to cleavage at glycosidic bonds, as indicated in Fig. 4. All *Campylobacter* N-linked oligosaccharides were composed of residues with masses of 228 Da (diacetamidotrideoxyhexose [DATDH]), 203 Da (*N*-acetylhexosamine [HexNAc]), and 162 Da (hexose [Hex]), apart from a single residue of 217 Da present in the *C. concisus* glycan. Comparison of the spectra shows that a variety of N-linked glycans are generated among *Campylobacter* species (Fig. 4). Key conserved features include the presence of the 228-Da residue as the reducing-end sugar of a HexNAc-rich oligosaccharide. More variable features include the number of sugar residues present in the N-linked oligosaccharide, with *C. coli*, *C. upsaliensis*, *C. helveticus*, C. *sputorum*, and *C. concisus* generating N-linked heptasaccharides; *C. lari*, *C. hyointestinalis*, and *C. fetus* generating hexasaccharides; and *C. lanienae* generating a pentasaccharide. Also, seven species (*C. coli*, *C. upsaliensis*, *C. helveticus*, *C. hyointestinalis*, *C. fetus*, *C. sputorum*, and *C. concisus*) produced oligosaccharides with a single branched sugar, a hexose residue in all species except for *C. sputorum*, with a branching HexNAc residue. Linear N-linked oligosaccharides were produced by *C. lari* and *C. lanienae*.

Overall, we have demonstrated experimentally that N-linked glycosylation is a conserved feature of *Campylobacter* species. It is also evident from MS data that the differing reactivities of whole-cell lysates (Fig. 2) and mobilities of *in vitro*-generated glycopeptides (Fig. 3) are a reflection of structural diversity in *Campylobacter* N-linked glycans, particularly at the nonreducing end (Fig. 4).

DISCUSSION

The extensively characterized *C. jejuni* N-linked protein glycosylation system is now established as both a model for understanding protein glycosylation and a biotechnological tool for *in vivo* generation of glycoconjugates (5, 12, 14, 50). However, genome sequence data indicate that not only *C. jejuni* but all *Campylobacter* species contain a gene cluster encoding a putative N-linked protein glycosylation system. In this study, we have demonstrated the functionality of these *Campylobacter* N-linked protein glycosylation systems and structurally characterized the associated N-linked glycans, thereby highlighting their conserved and more variable features.

In order to investigate N-linked glycosylation processes in multiple *Campylobacter* species, we employed a simple *in vitro* peptide glycosylation assay requiring only a source of bacterial membranes containing both an LLO and the integral membrane OTase protein along with a target peptide with the appropriate glycosylation sequon (21). An important assumption with this assay is that only LLOs that are involved in N-linked protein glycosylation will be used as donors by the OTase rather than other LLOs involved in glycolipid biosynthesis, for example. Though we have not formally excluded this possibility, the data obtained strongly suggest that this is a valid assumption. With this approach, we have characterized *in vitro*-generated, N-linked glycan from nine *Campylobacter* species. The observed varying levels of peptide glycosylation activity with membranes from different *Campylobacter* species (Fig. 3A) may have a number of explanations, but given that for most species this activity could be significantly increased by the addition of membranes derived from an *E. coli* strain expressing *C. jejuni pglB* (Fig. 3B), we suggest that either a lack of sufficient native PglB or the *C. jejuni*-optimized peptide employed (4) is not optimal for all *Campylobacter* PglB proteins. Using this assay, it was also clear that an acidic residue at the −2 position of the N-glycosylation sequon required for glycosylation

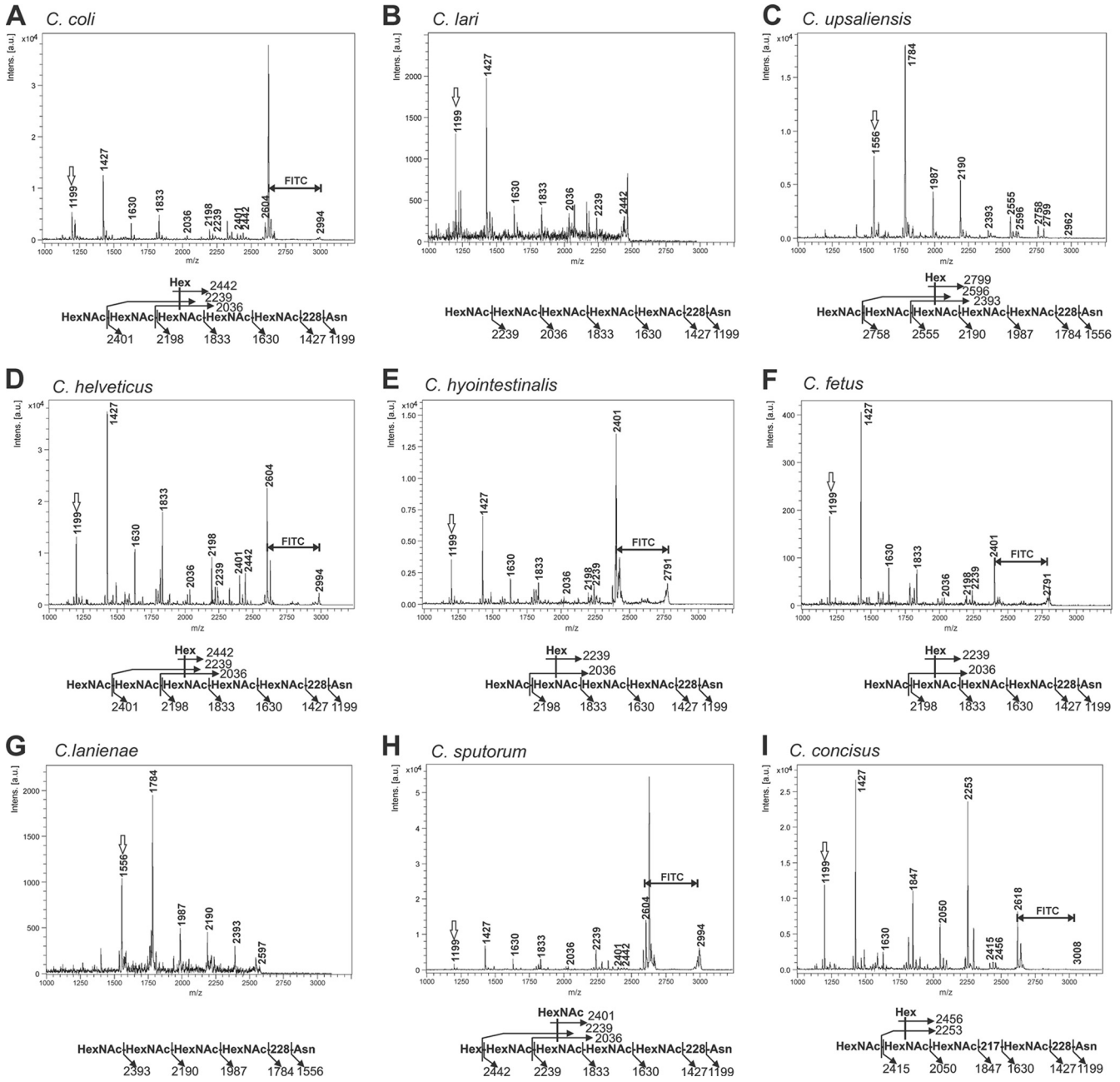

FIG 4 MS analysis of *Campylobacter in vitro*-generated N-linked glycans. *In vitro* glycosylation of a biotin-labeled fluorescent peptide (FITC-ADQNATAK-biotin) was analyzed by MALDI-TOF MS. Panels A through to I are MALDI-LIFT-TOF/TOF MS-generated spectra derived from glycopeptides produced by the indicated *Campylobacter* species. Below the spectra are the corresponding inferred N-linked glycan structures. Fragments corresponding to peptides with no glycan (arrows) generated peaks with an *m*/*z* value of either 1,556 or 1,199, depending on the presence or absence of the labile FITC moiety. The *m*/*z* values of peaks corresponding to fragment ions resulting from the sequential loss of sugar residues are indicated in the spectra. a.u., arbitrary units.

in *C. jejuni* (23) is also necessary for glycosylation in all other *Campylobacter* species (data not shown), although this was not absolutely required by *C. lari in vivo* (41). The principal limitation of the peptide glycosylation assay as employed herein is that the analytical scale of the *in vitro* reactions limits further structural investigation of N-linked glycan by nuclear magnetic resonance approaches; however, in all cases, sufficient glycopeptide was generated for structural analysis by MS.

MS-based analyses indicated that a conserved feature of all of the *Campylobacter* N-linked glycan structures elucidated was the presence of a 228-Da residue, assumed to be DATDH, as the reducing-end sugar. In *C. jejuni*, the 228-Da residue is the DATDH sugar bacillosamine, synthesized from GlcNAc by-products of the *pglDEF* genes and transferred onto a lipid carrier by the glycosyltransferase PglC (38, 40). Orthologues of the *C. jejuni pglCDEF* genes are present in all *Campylobacter* genome sequences (Fig. 1), suggesting that the 228-Da DATDH residue in these species is also bacillosamine, although this requires verification. At the nonre-

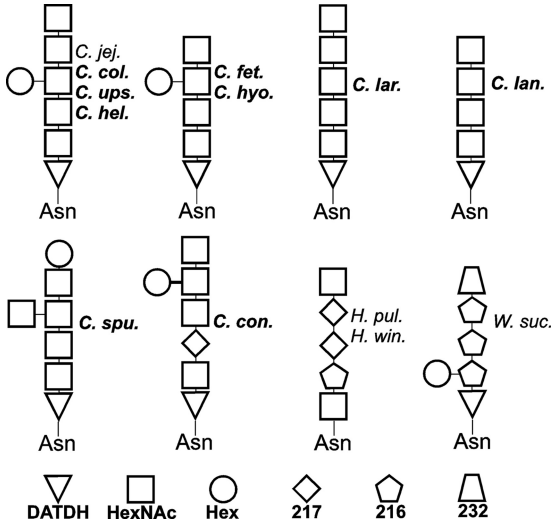

FIG 5 Schematic overview of epsilonproteobacterial *N*-linked glycan structures. The *C. jejuni* N-linked glycan structure is from reference 51, the *C. lari* structure is from reference 41 and this study, the *H. pullorum* structure is from reference 15, and the *H. winghamensis* and *W. succinogenes* structures are unpublished data from our laboratory. All other structures are from this study. The abbreviated species names *C. jej.*, *C. col.*, *C. ups.*, *C. hel.*, *C. fet.*, *C. hyo.*, *C. lar.*, *C. lan.*, *C. spu.*, *C. con.*, *H. pul.*, *H. win.*, and *W. suc.* correspond to *C. jejuni*, *C. coli*, *C. upsaliensis*, *C. helveticus*, *C. fetus*, *C. hyointestinalis*, *C. lari*, *C. lanienae*, *C. sputorum*, *C. concisus*, *H. pullorum*, *H. winghamensis*, and *W. succinogenes*, respectively. Unusual sugar residues are indicated by their respective masses in daltons.

ducing end of the *in vitro*-generated, N-linked glycans, MS data demonstrated structural variation that correlated to some degree with inferred evolutionary relationships of the corresponding bacterial species (Fig. 1A and 5). Thus, N-linked glycans from relatively closely related *Campylobacter* species (*C. jejuni*, *C. coli*, *C. upsaliensis*, and *C. helveticus*) were indistinguishable by MS while those from species more distantly related to this group (*C. fetus*, *C. hyointestinalis*, *C. lanienae*, *C. sputorum*, and *C. concisus*) displayed differing arrangements of Hex and HexNAc sugars at the nonreducing terminus, with the *C. concisus* N-linked glycan containing a residue of 217 Da that is not found in other *Campylobacter* species (Fig. 5). The N-linked glycans from species of the genera *Helicobacter* and *Wolinella* (Fig. 5) display further structural diversity, including the presence of unusual sugar residues of 216, 217, and 232 Da (15). A sugar residue of 216 Da was detected in the O-antigens of *Francisella tularensis* (48), *Pseudomonas aeruginosa* O6 (20), and *E. coli* O121 (19) and is known to be an *N*-acetylhexuronamide (HexNAcAN). In *P. aeruginosa*, HexNAcAN is synthesized from nucleotide-activated HexNAc through conversion by WbpO to an *N*-acetylhexuronic acid (HexNAcA) of 217 Da, followed by conversion to the 216-Da HexNAcAN residue by the amidotransferase WbpS (19, 33, 52). While we have no further structural data on the 216- and 217-Da sugars from *H. pullorum* and *W. succinogenes* N-linked glycans, both species possess orthologues of the *P. aeruginosa wbpO* and *wbpS* genes located close to orthologues of *C. jejuni pgl* genes. Furthermore, the presence of a *wbpO*, but not a *wbpS*, orthologue in *C. concisus* (Fig. 1B) is consistent with a 217-Da (presumably HexNAcA) but not a 216-Da residue in the N-linked glycan.

Structural data presented herein indicate that *Campylobacter* N-linked glycan structures vary at the species level and above. There is also little or no evidence from the now considerable number of *C. jejuni* genome sequences of genetic variation at the *pgl* locus, suggesting intraspecies conservation of the N-linked heptasaccharide. This is in contrast to the intraspecies and even intrastrain structural variation of glycan components of capsular polysaccharide, lipooligosaccharide (LOS), and the flagellin glycoprotein in *C. jejuni* (8). This intraspecies structural variation is due to differences in gene content among strains of *C. jejuni* within the specific loci responsible for the biosynthesis of LOS, capsule, and the glycan component of flagellin glycoprotein. These loci also contain many homopolymeric-tract-containing, phase-variable genes encoding products involved in the biosynthesis of these structures, so generating intrastrain structural variation. There is no evidence of phase-variable genes within *Campylobacter pgl* loci. This relatively limited structural diversity of *Campylobacter* N-linked glycans indicates that they are not under the same degree of selective pressure as capsule, LOS, and flagellin glycan. This relative structural conservation also makes N-linked glycans potential subunit vaccine components and targets for antibody-based detection methods. For these approaches, the multiplicity of N-linked glycan-containing proteins and their location at the bacterial surface are advantageous. Despite this relative conservation of *Campylobacter* N-linked glycosylation systems, there appears to be limited conservation among target glycoproteins. Thus, BLAST searches with each of the 55 known *C. jejuni* NCTC 11168 glycoprotein amino acid sequences against predicted proteomes from available genome sequences of *C. coli*, *C. lari*, *C. upsaliensis*, *C. fetus*, and *C. concisus* identified 39 homologues present in all of the species. However, only 15 of these proteins also have an extended N-glycosylation sequon of D/E-X-N-X-S/T within the predicted amino acid sequences from all six *Campylobacter* species. This suggests that there is only partial overlap in the glycoproteomes of *Campylobacter* species.

The major biotechnological application of the *C. jejuni* N-linked protein glycosylation technology is in developing systems for *in vivo* glycoconjugate synthesis (12, 14, 31). The *C. jejuni* OTase PglB is central to this approach, and our data demonstrate that there are other PglB proteins from non-*C. jejuni* campylobacters that may be similarly useful. This *in vivo* glycoconjugate synthesis technology is limited primarily by the restricted capacity of the *C. jejuni* PglB OTase to transfer only glycans with specific sugars at the reducing end (5, 49). Our data on the structure of *Campylobacter* N-linked glycans, specifically, the universal presence of a 228-Da sugar residue at the reducing end, suggest that all *Campylobacter* PglB proteins may have similar limited substrate specificity. Thus, characterization of N-linked OTases present in species of *Deltaproteobacteria* and *Epsilonproteobacteria* but more distantly related to the *Campylobacter* genus may be more productive in the search for alternate substrate specificities of biotechnological potential. However, it may well be worthwhile to evaluate individual *Campylobacter* N-linked OTases for increased activity or yield in specific biotechnological applications.

ACKNOWLEDGMENTS

This work was funded by UK Biotechnology and Biological Science Research Council (BBSRC) grants BB/F009321/1 (B.W. and D.L.) and BB/H017542/1 (D.L. and B.W.).

We thank David Knight and staff of the Biomolecular Analysis Core Research Facility, Faculty of Life Sciences, University of Manchester, for

assistance with MS. We also thank Anne Dell and Paul Hitchen for advice on the interpretation of MS data. Antiglycan antiserum was kindly provided by Markus Aebi.

REFERENCES

1. **Abu-Qarn M, Eichler J.** 2006. Protein N-glycosylation in archaea: defining *Haloferax volcanii* genes involved in S-layer glycoprotein glycosylation. Mol. Microbiol. **61**:511–525.
2. **Alaimo C, et al.** 2006. Two distinct but interchangeable mechanisms for flipping of lipid-linked oligosaccharides. EMBO J. **25**:967–976.
3. **Calo D, Kaminski L, Eichler J.** 2010. Protein glycosylation in archaea: sweet and extreme. Glycobiology **20**:1065–1076.
4. **Chen MM, Glover KJ, Imperiali B.** 2007. From peptide to protein: comparative analysis of the substrate specificity of N-linked glycosylation in *C. jejuni*. Biochemistry **46**:5579–5585.
5. **Feldman MF, et al.** 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.
6. **Glover KJ, Weerapana E, Chen MM, Imperiali B.** 2006. Direct biochemical evidence for the utilization of UDP-bacillosamine by PglC, an essential glycosyl-1-phosphate transferase in the *Campylobacter jejuni* N-linked glycosylation pathway. Biochemistry **45**:5343–5350.
7. **Glover KJ, Weerapana E, Imperiali B.** 2005. In vitro assembly of the undecaprenylpyrophosphate-linked, heptasaccharide for prokaryotic N-linked glycosylation. Proc. Natl. Acad. Sci. U. S. A. **102**:14255–14259.
8. **Guerry P, Szymanski CM.** 2008. *Campylobacter* sugars sticking out. Trends Microbiol. **16**:428–435.
9. **Guo P, Lin J, Reynolds DL, Zhang Q.** 2010. Contribution of the multidrug efflux transporter CmeABC to antibiotic resistance in different *Campylobacter* species. Foodborne Pathog. Dis. **7**:77–83.
10. **Helenius A, Aebi M.** 2004. Roles of N-linked glycans in the endoplasmic reticulum. Annu. Rev. Biochem. **73**:1019–1049.
11. **Hendrixson DR, Dirita VJ.** 2004. Identification of *Campylobacter jejuni* genes involved in commensal colonization of the chick gastrointestinal tract. Mol. Microbiol. **52**:471–484.
12. **Hug I, Feldman MF.** 2011. Analogies and homologies in lipopolysaccharide and glycoprotein biosynthesis in bacteria. Glycobiology **21**:138–151.
13. **Ielmini MV, Feldman MF.** 2011. *Desulfovibrio desulfuricans* PglB homolog possesses oligosaccharyltransferase activity with relaxed glycan specificity and distinct protein acceptor sequence requirements. Glycobiology **21**:734–742.
14. **Ihssen J, et al.** 2010. Production of glycoprotein vaccines in *Escherichia coli*. Microb. Cell Fact. **9**:61.
15. **Jervis AJ, et al.** 2010. Characterization of N-linked protein glycosylation in *Helicobacter pullorum*. J. Bacteriol. **192**:5228–5236.
16. **Jones MA, et al.** 2004. Adaptation of *Campylobacter jejuni* NCTC11168 to high-level colonization of the avian gastrointestinal tract. Infect. Immun. **72**:3769–3776.
17. **Karlyshev AV, et al.** 2004. The *Campylobacter jejuni* general glycosylation is important for attachment to human epithelial cells and in the colonization of chicks. Microbiology **150**:1957–1964.
18. **Kelleher DJ and Gilmore R.** 2006. An evolving view of the eukaryotic oligosaccharyltransferase. Glycobiology **16**:47R–62R.
19. **King JD, Vinogradov E, Tran V, Lam JS.** 2010. Biosynthesis of uronamide sugars in *Pseudomonas aeruginosa* O6 and *Escherichia coli* O121 O antigens. Environ. Microbiol. **12**:1531–1544.
20. **Knirel YA, Bystrova OV, Kocharova NA, Zaehringer U, Pier GB.** 2006. Conserved and variable structural features in the lipopolysaccharide of *Pseudomonas aeruginosa*. J. Endotoxin Res. **12**:324–336.
21. **Kohda D, Yamada M, Igura M, Kamishikiryo J, Maenaka K.** 2007. New oligosaccharyltransferase assay method. Glycobiology **17**:1175–1182.
22. **Kowarik M, et al.** 2006. N-linked glycosylation of folded proteins by the bacterial oligosaccharyltransferase. Science **314**:1148–1150.
23. **Kowarik M, et al.** 2006. Definition of the bacterial N-glycosylation site consensus sequence. EMBO J. **25**:1957–1966.
24. **Kumar S, Nei M, Dudley J, Tamura K.** 2008. MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. Brief. Bioinform. **9**:299–306.
25. **Langdon RH, Cuccui J, Wren BW.** 2009. N-linked glycosylation in bacteria: an unexpected application. Future Microbiol. **4**:401–412.
26. **Larkin A, Imperiali B.** 2011. The expanding horizons of asparagine-linked glycosylation. Biochemistry **50**:4411–4426.
27. **Larkin MA, et al.** 2007. Clustal W and Clustal X version 2.0. Bioinformatics **23**:2947–2948.
28. **Larsen JC, Szymanski C, Guerry P.** 2004. N-linked protein glycosylation is required for full competence in *Campylobacter jejuni* 81-176. J. Bacteriol. **186**:6508–6514.
29. **Linton D, Allan E, Karlyshev AV, Cronshaw AD, Wren BW.** 2002. Identification of *N*-acetylgalactosamine-containing glycoproteins PEB3 and CgpA in *Campylobacter jejuni*. Mol. Microbiol. **43**:497–508.
30. **Linton D, et al.** 2005. Functional analysis of the *Campylobacter jejuni* N-linked protein glycosylation pathway. Mol. Microbiol. **55**:1695–1703.
31. **Lizak C, Gerber S, Numao S, Aebi M, Locher KP.** 2011. X-ray structure of a bacterial oligosaccharyltransferase. Nature **474**:350–U377.
32. **Maita N, Nyirenda J, Igura M, Kamishikiryo J, Kohda D.** 2010. Comparative structural biology of eubacterial and archaeal oligosaccharyltransferases. J. Biol. Chem. **285**:4941–4950.
33. **Miller WL, et al.** 2008. Flagellin glycosylation in *Pseudomonas aeruginosa* PAK requires the O-antigen biosynthesis enzyme WbpO. J. Biol. Chem. **283**:3507–3518.
34. **Mohorko E, Glockshuber R, Aebi M.** 2011. Oligosaccharyltransferase: the central enzyme of N-linked protein glycosylation. J. Inherit. Metab. Dis. **34**:869–878.
35. **Nita-Lazar M, Wacker M, Schegg B, Amber S, Aebi M.** 2005. The N-X-S/T consensus sequence is required but not sufficient for bacterial N-linked protein glycosylation. Glycobiology **15**:361–367.
36. **Nothaft H, Liu X, McNally DJ, Li J, Szymanski CM.** 2009. Study of free oligosaccharides derived from the bacterial N-glycosylation pathway. Proc. Natl. Acad. Sci. U. S. A. **106**:15019–15024.
37. **Nothaft H, Szymanski CM.** 2010. Protein glycosylation in bacteria: sweeter than ever. Nat. Rev. Microbiol. **8**:765–778.
38. **Olivier NB, Chen MM, Behr JR, Imperiali B.** 2006. In vitro biosynthesis of UDP-*N*,*N*′-diacetylbacillosamine by enzymes of the *Campylobacter jejuni* general protein glycosylation system. Biochemistry **45**:13659–13669.
39. **Schägger H.** 2006. Tricine-SDS-PAGE. Nat. Protoc. **1**:16–22.
40. **Schoenhofen IC, et al.** 2006. Functional characterization of dehydratase/aminotransferase pairs from *Helicobacter* and *Campylobacter*: enzymes distinguishing the pseudaminic acid and bacillosamine biosynthetic pathways. J. Biol. Chem. **281**:723–732.
41. **Schwarz F, et al.** 2011. Relaxed acceptor site specificity of bacterial oligosaccharyltransferase *in vivo*. Glycobiology **21**:45–54.
42. **Scott NE, et al.** 2011. Simultaneous glycan-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**:M000031–MCP201.
43. **Szymanski CM, Burr DH, Guerry P.** 2002. *Campylobacter* protein glycosylation affects host cell interactions. Infect. Immun. **70**:2242–2244.
44. **Szymanski CM, Wren BW.** 2005. Protein glycosylation in bacterial mucosal pathogens. Nat. Rev. Microbiol. **3**:225–237.
45. **Szymanski CM, Yao RJ, Ewing CP, Trust TJ, Guerry P.** 1999. Evidence for a system of general protein glycosylation in *Campylobacter jejuni*. Mol. Microbiol. **32**:1022–1030.
46. **van Sorge NM, et al.** 2009. N-glycosylated proteins and distinct lipooligosaccharide glycoforms of *Campylobacter jejuni* target the human C-type lectin receptor MGL. Cell. Microbiol. **11**:1768–1781.
47. **Van Vliet AHM, Wood AC, Henderson J, Wooldridge K, Ketley JM.** 1998. Genetic manipulation of enteric *Campylobacter* species. Methods Microbiol. **27**:407–419.
48. **Vinogradov EV, et al.** 1991. Structure of the O-antigen of *Francisella tularensis* strain 15. Carbohydr. Res. **214**:289–297.
49. **Wacker M, et al.** 2006. Substrate specificity of bacterial oligosaccharyltransferase suggests a common transfer mechanism for the bacterial and eukaryotic systems. Proc. Natl. Acad. Sci. U. S. A. **103**:7088–7093.
50. **Wacker M, et al.** 2002. N-linked glycosylation in *Campylobacter jejuni* and its functional transfer into E. coli. Science **298**:1790–1793.
51. **Young NM, et al.** 2002. Structure of the N-linked glycan present on multiple glycoproteins in the gram-negative bacterium, *Campylobacter jejuni*. J. Biol. Chem. **277**:42530–42539.
52. **Zhao X, et al.** 2000. WbpO, a UDP-*N*-acetyl-D-galactosamine dehydrogenase from *Pseudomonas aeruginosa* serotype O6. J. Biol. Chem. **275**:33252–33259.