Relaxed acceptor site specificity of bacterial oligosaccharyltransferase in vivo

Flavio Schwarz, Christian Lizak, Yao-Yun Fan, Susanna Fleurkens, Michael Kowarik, and Markus Aebi

Department of Biology, Institute of Microbiology, ETH Zürich, Wolfgang-Pauli-Str. 10, CH-8093 Zürich, Switzerland

Received on June 24, 2010; revised on August 20, 2010; accepted on August 25, 2010

A number of proteobacteria carry the genetic information to perform N-linked glycosylation, but only the protein glycosylation (pgl) pathway of Campylobacter jejuni has been studied to date. Here, we report that the pgl gene cluster of Campylobacter lari encodes for a functional glycosylation machinery that can be reconstituted in Escherichia coli. We determined that the N-glycan produced in this system consisted of a linear hexasaccharide. We found that the oligosaccharyltransferase (OST) of C. lari conserved a predominant specificity for the primary sequence D/E–X$_{1}$–N–X$_{1}$–S/T (where X$_{1}$ and X$_{1}$ can be any amino acid but proline). At the same time, we observed that this enzyme exhibited a relaxed specificity toward the acceptor site and modified asparagine residues of a protein at sequences DANSG and NNNST. Moreover, C. lari pgl glycosylated a native E. coli protein. Bacterial N-glycosylation appears as a useful tool to establish a molecular description of how single-subunit OSTs perform selection of glycosyl acceptor sites.

Keywords: Campylobacter lari, N-glycosylation, oligosaccharyltransferase

Introduction

N-linked glycosylation is a ubiquitous protein modification that involves the decoration of selected asparagine residues within the sequences N–X$_{1}$–S/T (X$_{1}$≠P) with an oligosaccharide. This process is organized in a topologically equivalent way in eukaryotes and prokaryotes (Yurist-Doutch et al. 2008) and involves (I) the assembly of the glycan donor on a lipid anchor at the cytoplasmic side of the ER or of the plasma membrane, (II) reorientation to the opposite face of the membrane and (III) subsequent attachment to the acceptor protein (Burda and Aebi 1999). The defining event of N-glycosylation is the formation of the N-glycosidic bond between the oligosaccharide and the asparagine side chain of the acceptor protein. This reaction is catalyzed by the oligosaccharyltransferase (OST), a single-subunit enzyme in bacteria, archaean and protozoa, but a multimeric protein complex in other eukaryotes (Kelleher and Gilmore 2006). The Stt3 protein represents the central component of the OST complex, and it is believed to be the catalytic subunit (Feldman et al. 2005; Igura et al. 2008). However, limited structure/function information on the Stt3 protein limits our understanding of how the poorly reactive amido group of the asparagine side chain is activated to form the N-glycosidic bond to the oligosaccharide (Igura et al. 2008; Maita et al. 2009).

An OST identifies acceptor sequences in polypeptides and modifies only some of the potential sites. Amino acids around the consensus sequence impact the probability of glycosylation of the particular site (Petrescu et al. 2004). According to the current model, different strategies evolved to extend the substrate range of N-linked glycosylation. Campylobacter jejuni OST (called PglB) transfers a heptasaccharide to the acceptor protein, stringently within the sequon D/E–X$_{1}$–N–X$_{1}$–S/T (X$_{1}$, X$_{1}$≠P) (Kowarik, Young et al. 2006). In protozoa, substrate diversity was extended through duplication of the STT3 loci with subsequent diversification of OST specificity. In fact, Leishmania major and Trypanosoma brucei bear different OST isoforms with distinct donor and acceptor specificities (Nasab et al. 2008; Izquierdo et al. 2009). That is, protozoa are able to assemble different glycans on distinct sequons, thus increasing the complexity of their glycoproteome. In other eukaryotes, the core Stt3 is supplemented with a variable number of additional functions that assist and refine the glycosylation process (Kelleher and Gilmore 2006). Along this line, ribophorin I and Ost3/6p are connected to the glycosylation status of membrane proteins and the site occupancy of specific sequons, respectively (Wilson and High 2007; Schulz et al. 2009). These auxiliary OST subunits are believed to preserve distinct protein substrates in an accessible, unfolded state, leading to an increased number of sites accessible to the Stt3 protein.

Here, we attempted to mine the diversity of natural bacterial OSTs in order to extend our understanding of bacterial N-glycosylation. In particular, we investigated the role of the acidic amino acid invariantly present in the consensus sequence of C. jejuni glycoproteins and analyzed whether this
requirement was conserved by N-glycosylation machineries of other bacterial species. We found that the protein glycosylation (pgl) locus of *Campylobacter lari* encodes for a functional glycosylation system, able to synthesize a linear hexaasaccharide and to transfer it to proteins. We observed that CIPgIPB modified variants of the *C. jejuni* glycoprotein AcrA that lack the canonical consensus sequence D/E-X,₁₋₁-N-X,₁₋₁-S/T. We proved that CIPgIPB transfers the glycan to sites that lack an acidic residue at the −2 position of the AcrA protein, albeit with low efficiency. These data suggested that the negatively charged amino acid of the glycosylation site was not absolutely required for catalysis by bacterial OST, and it raised the question whether the amino acid at the −2 position was involved in the recognition of the acceptor site. Furthermore, we observed that the nonconsensus sequence DANSGT was glycosylated by CIPgIPB.

**Results**

*C. lari* pgl glycosylates proteins with a linear hexaasaccharide

The *C. jejuni* pgl cluster is the first and the only bacterial glycosylation system characterized to date. However, in the last few years, genome sequencing unveiled a broad distribution of pgl loci among ε- and δ-proteobacteria (Szymanski and Wren 2005). Some of the pgl loci share a genetic organization analogous to the one of *C. jejuni*, whereas others present insertions of a variable number of open reading frames (ORFs) between the pgl genes. Interestingly, *Helicobacter canadensis* and *Helicobacter pullorum* carry two pgl II paralogs in distinct locations of the chromosome (Langdon et al. 2009; Jervis et al. 2010). While pgl II homologs exhibit a variable degree of similarity in their sequences and there is not yet conclusive evidence that they encode for functional OSTs, they all display the conserved motif WWDXG and are predicted to arrange in a comparable topology. Moreover, analysis of the pgl I loci revealed the presence of functions involved in the bacillosamine biosynthesis, suggesting that this monosaccharide might constitute a common trait of *Campylobacter*-N-glycosylation.

We studied the *C. lari* glycosylation system since we identified significant elements of difference among an overall similarity to the *C. jejuni* pathway. The pgl cluster extends for about 12 kb on the *C. lari* chromosome and contains 10 ORFs (Figure 1A). The gene (pglI) encoding for the glycosyltransferase that attaches the branching glucose to the *C. jejuni* glycan was not found, whereas the other functions of the *C. jejuni* pgl locus were identified based on sequence similarity of the ORFs. CIPgIPB presents the WWDXG motif and shares 56% identity to CIPgIPB. We isolated the *C. lari* pgl locus and expressed it in combination with the *C. jejuni* AcrA protein in *Escherichia coli* cells. Immunoblot analysis indicated glycosylation of AcrA (Figure 1B). Purified glycoproteins were detected using serum specific for the *C. jejuni* N-glycan (hR6) (Figure 1C). We have previously showed that this serum recognizes a linear stretch of GalNAc residues (Schwarz et al. 2010). Treatment of glycosylated AcrA with an exo-α-N-acetyl-galactosaminidase (α-GalNAc-asel) resulted in a mobility shift after sodium dodecyl sulfate (SDS)–polyacrylamide gel electrophoresis (PAGE) and in the gradual reduction of reactivity to hR6 over time. This suggested that the *C. lari* glycan contained a linear chain of GalNAc residues, linked in α conformation. To determine the structure of the N-glycan of *C. lari*, AcrA was purified from *E. coli* cells carrying either the *C. jejuni* or the *C. lari* pgl cluster, incubated with proteinase K, permethylated and subjected to mass spectrometric (MS) analysis (Figure 1D). Comparison of the two spectra revealed a difference in mass of 204 Da, matching with a permethylated hexose. The MS/MS spectrum of the ion at m/z = 2008.97 determined fragmentation behavior compatible with a linear Bac(HexNAc)₅ (Figure 1E). Altogether, these data proved that the *C. lari* N-glycan consisted of GalNAc-α-GalNAc-α-GalNAc-α-GalNAc-α-GalNAc-1,3-Bac.

*C. jejuni* and *C. lari* pgl exhibit different acceptor specificities

We aimed to dissect the *C. lari* PglB specificity toward the acceptor protein and attempted a direct comparison to its *C. jejuni* counterpart. AcrA presents five N-X-S/T sequences (Figure 2A); two of them exhibit an acidic amino acid at the −2 position to the asparagine and are glycosylated by CIPgIPB (Kowarik, Young et al. 2006). Expression of AcrA in *E. coli* bearing the *C. jejuni* or the *C. lari* pgl cluster led to a comparable pattern, as judged by immunoblot, corresponding to AcrA glycosylated at two sites (Figure 2B, lanes 5 and 6). The synthesis of glycosylation-incompetent AcrA N²⁴–⁴⁷ in the presence of the *C. jejuni* pgl resulted in an unmodified acceptor (Figure 2B, lane 3), but expression of AcrA N²⁴–⁴⁷ in combination with the *C. lari* pgl gave a pattern indicative of glycosylation (Figure 2B, lane 4). The slowest migrating protein, barely visible in the anti-AcrA immunoblot, produced a signal with hR6, implying double glycosylation. This suggested that the activity of CIPgIPB did not strictly depend on the 5-mer consensus sequence as defined in *C. jejuni*, but it exhibited a different or broader substrate specificity. Moreover, the hR6 serum specifically reacted with an...
additional periplasmic protein of *E. coli* cells bearing the *C. lari* pgl cluster (Figure 2B, indicated with a star). In order to identify this protein, we separated glycosylated proteins from periplasmic extracts by soy bean agglutinin (SBA) lectin affinity chromatography (Supplementary data, Figure S1A) and analyzed the resulting specimens by MS. We identified the spermidine–putrescine-binding protein PotD of *E. coli*, previously reported not to be a substrate of the *C. jejuni* pgl machinery (Kowarik, Young et al. 2006). Immunoblot analysis confirmed that endogenous PotD was modified by the *C. lari* machinery (Figure 2C). Lastly, we confirmed that the PotD protein was modified at the sites DDN_{26}NT and ESN_{62}ET by expression of histidine-tagged PotD variants presenting point mutations in these sites (Supplementary data, Figure S1B).

Altogether, these data suggested that *C. lari* and *C. jejuni* pgl possess different yet overlapping acceptor specificities.

**C. lari PglB glycosylates AcrA lacking the D/E−X−1−N−X−1−S/T site**

To investigate whether bacterial OSTs generally exhibit substrate specificities different from that of *C. jejuni*, we designed two versions of the AcrA acceptor protein. In the *C. jejuni* system, the AcrA N273Q mutant conserves a single glycosylation site, while the double mutation D121A/N273Q prevents glycosylation. We analyzed the glycosylation profile of these AcrA variants in *E. coli* SCM3 bearing a plasmid to direct synthesis of the *C. jejuni* glycan and a plasmid encoding for the PglB from *C. jejuni*, *Campylobacter coli* or *C. lari* (Figure 3).
Both C. coli and C. lari PglB transferred the C. jejuni glycan to AcrA N273Q, albeit with very low efficiency in the case of C. coli (Figure 3, lanes 5 and 7). Strikingly, we could observe glycosylation of the AcrA D121A/N273Q variant only by C. lari PglB (Figure 3, lane 8). We speculated that the detected glycosylated product could be either due to glycosylation at a distinct site or due to a mixed population of AcrA glycoforms, each weakly glycosylated at a different site.

Therefore, we undertook a comprehensive analysis of AcrA glycosylation by the C. lari system. We first generated mutant forms of AcrA containing N-to-L replacements at each of the N–X–S/T sites and studied their glycosylation profile (Figure 4A). Disruption of the sites at N117 (N1L), N145 (N2L) and N274 (N5L) gave a phenotype similar to the wild-type AcrA (Figure 4A, lanes 3, 5 and 7). The N123L (N3L) mutation clearly affected glycosylation, implying that the site N123 was used by CIPgLB (Figure 4A, lane 4). Interestingly, mutation of the N273L (N5L) reduced the ratio between doubly and mono-glycosylated AcrA (Figure 4A, lane 6). This suggested that the inactivation of the N4 site led to inefficient modification of an alternative site. At this point, we speculated that the mutation of N4 directed glycosylation of the adjacent site N5.

We further constructed a series of AcrA variants presenting multiple point mutations. These proteins were expressed in E. coli carrying the C. lari pgl locus, purified via nickel affinity chromatography and analyzed by SDS–PAGE. To convincingly demonstrate glycosylation, we incubated the AcrA preparations with α-GalNAc-ase. Purified fractions from cells expressing the triple mutants N117L/N123L/N273L (N1,2,4L) and N123L/N145L/N273L (N2,3,4L) AcrAs were resolved as a major protein with a mass of about 37 kDa and two larger proteins as visualized by Coomassie staining (Figure 4B, lanes 1 and 3). These proteins were sensitive to glycosidase treatment, suggesting glycosylation (lanes 2 and 4). The hR6 immunoblot supported this observation. Comconitent mutation of the N123, N273 and N274 sites (N2,4,5L) resulted in a single hR6-reactive protein (Figure 4B, lane 5), hR6 reactivity disappeared after glycosidase treatment (lane 6). This led us to conclude that N5 was glycosylated by CIPgLB. Surprisingly, simultaneous mutation of four (N1,2,4,5L) or all five (N1,2,3,4,5L) asparagine residues within the N–X–S/T sites gave a phenotype equivalent to the triple mutant N2,4,5L (Figure 4B, lanes 7–10). After careful inspection of the AcrA sequence, we suspected that the site 257-DANSGT262 might represent an alternative, suboptimal glycosylation site. Indeed, inactivation of this site by mutation of the asparagine residue (N6) prevented glycosylation by CIPgLB (Figure 4B, lane 11).

In order to conclusively prove that the N5 and N6 sites were glycosylated, we performed the MS analysis. To overcome the issue of limiting ionization efficiency of glycopeptides, we took advantage of the newly generated data on the C. lari N-glycan and developed a procedure to produce useful N-glycopeptides. We incubated nickel affinity-purified AcrA variants with trypsin and then, after heat inactivation of the protease, added α-GalNAc-ase. In this way, we produced a mixture of peptides and corresponding bacillosamine-tagged glycopeptides that were subjected to the liquid chromatography-electrospray ionization mass spectrometry (LC-ESI-MS) analysis. The fragmentation spectrum of the triply charged precursor ion at m/z = 923.8 matched with the unmodified peptide AVFDQNQSTLLPGAFATISETGEFQK (Supplementary data, Figure S2A), whereas the fragmentation profile of the triply charged precursor ion at m/z = 999.8 was consistent with the presence of bacillosamine at N5 of the same peptide (Figure 4C). Similarly, the MS/MS profile of the doubly charged precursor ion at m/z = 871.5 matched with the peptide LYFIDSVIDAN6SGTVK (Supplementary data, Figure S2B), and ions generated by the fragmentation of the species at m/z = 985.5 (doubly charged) were in agreement with bacillosamine-tagged N6 of the same peptide (Figure 4D). Consequently, glycosylation at the sites DQNQRST, DNLNST and DANSGT illustrated that precise positioning of the D/E and S/T residues was not essential on specific sites of AcrA.

A D/E residue at –2 position is not absolutely required for CIPgLB glycosylation

Next, we examined the glycosylation profile of AcrA N2L presenting a D-to-A substitution in the 271-DNN4N5A 276 site (Figure 5A). Remarkably, CIPgLB transferred a glycan to the ANN4N5ST site, implying that aspartic acid was not essential for catalysis (lanes 1 and 2 and 7 and 8). MS/MS analysis of the precursor ion at the mass corresponding to the glycosylated peptide (m/z = 2940) did not allow us to clearly assign the glycan at N4 or N5 (Supplementary data, Figure S3).

We extended the study to glycosylation of the DNN4N5AT and DNN4N5SA sites. We observed a gradual increase in the glycosylation efficiency, with ANN4N5ST site being modified the least (Figure 5A, lane 1) and DNN4N5SA the most (Figure 5A, lane 5). This indicated that the aspartic acid had an active role in recruiting the substrate to PglB. Although we did not determine the position of the glycan in the DNN4N5AT site, we reasoned that this was a similar situation to DNNQ5ST and DNLN5ST. Accordingly, the S275A mutation provided a suboptimal substrate, whereas T276A resulted in an optimal acceptor site sequence for CIPgLB.

CIPgLB does not glycosylate a eukaryotic protein at the native glycosylation site

We ruled out the hypothesis that the acidic amino acid in the –2 position was strictly required for catalysis by CIPgLB. In
Fig. 4. C. lari pgl glycosylates AcrA at two noncanonical sites. (A) Periplasmic extracts from E. coli expressing single-site AcrA variants were analyzed by SDS–PAGE and immunoblot and probed with anti-AcrA serum (top) or hr6 (bottom). Numbers at the right of the gel frame indicate the N-glycans on AcrA. (B) Nickel affinity-purified AcrA variants were separated by SDS–PAGE, stained by Coomassie (top) or transferred to nitrocellulose membranes and probed with hr6 serum (bottom). AcrAs were incubated with (+) or without (−) α-GalNAc-ase. (C) LC-ESI-MS/MS analysis of AcrA D121A/N273Q, purified from E. coli cells bearing the C. lari pgl cluster. Spectrum from fragmentation of the triply charged precursor ion at m/z = 999.8 corresponds to the glycopeptide AVFDNQ(Bac)STLLPGAATITSEGFIQK. (D) LC-ESI-MS/MS analysis of AcrA N123L D271A. Spectrum from fragmentation of the doubly charged ion at m/z = 985.5 matches with the glycopeptides LYFIDSVIDAN(Bac)SGTVK.
During N-linked glycosylation, only a subset of potential acceptor sites is chosen for glycosylation. In yeast and higher eukaryotes, site selection is aided by the concerted activity of the OST subunits, which present specific polypeptides to the St3 protein. A similar advantage has been achieved in some protozoa by duplication of the STT3 gene: OST isoforms featuring various degrees of similarity direct glycosylation of distinct subsets of polypeptides. A limited number of proteobacteria possess a protein N-glycosylation system, probably acquired by the horizontal transfer from archaea. Although duplication of the gene encoding for the OST has been observed in rare cases, most bacteria carry a single copy of pglB on the genome. In this respect, N-glycosylating bacteria are likely to mimic an early evolutionary stage. Therefore, bacterial systems appear as ideal candidates to establish an accurate molecular description of how a single-subunit OST performs site selection. This will in turn shed light on the catalytic mechanism of N-glycosylation. Lastly, access to bacterial glycosylation systems represents an important means to explore and exploit the diversity of this glycosylation.

C/PglB modifies asparagine residues within the sequence D/E-X1-N-X2-S/T (where X1 and X2 are any amino acids but proline) when located in flexible, exposed structural elements (Kowarik, Numao et al. 2006; Kowarik, Young et al. 2006). The presence of an acidic group at the −2 position to the asparagine seems to be a common prerequisite of C/PglB and C/PglB for efficient glycosylation. However, weak glycosylation of AcrA at the NNNST site by C/PglB illustrates that the acidic amino acid is dispensable for catalysis in vivo and corroborates the hypothesis that the formation of the N-glycosidic bond proceeds via a reaction scheme conserved in bacteria and eukaryotes. Similar conclusions have been drawn from the observation that tripeptides of sequence NLT are modified in vitro by membrane fractions containing C/PglB, in the presence of synthetic lipid-linked oligosaccharides (Glover et al. 2005). Both in vivo and in vitro, the presence of the aspartic acid at the −2 position greatly enhances the efficiency of the reaction. Although it is tempting to invoke involvement of the negatively charged amino acid in the recruitment of acceptor substrate to PglB, the exact role of the −2 position remains elusive. In fact, peptides with the semi-consensus sequence DFNVX (X = cysteine, homoserine or diaminopropionio acid) are neither substrates of C/PglB nor behave as inhibitors in vitro (Chen et al. 2007). Similarly, DFQRT peptides do not seem to reduce glycosylation of AcrA in vitro (data not shown). This evidence supports a more complex role for the −2 position. It might be that the aspartic acid, which is a poor α-helix former (Pace and Scholtz 1998), enhances the flexibility required by PglB. Thus, the D121A mutation on AcrA might preclude glycosylation at N2 by a double effect—reduction of flexibility plus a drop in recognition by PglB. In contrast, the N3/N5 sites are predicted to fold in a flexible conformation (Kowarik, Young et al. 2006), and their local structure might be unaffected by the mutation of D271. However, while this hypothesis supports site occupancy of AcrA, it does not explain why IgG1-Ct12 is not glycosylated.

Although we cannot formally exclude the possibility that heterologous expression of pgl genes in E. coli might affect
their specificities or that Pgl protein levels might differ, selective modification of the PotD protein by CPglB demonstrates that bacterial glycosylation machineries feature different glycosylation phenotypes. However, the evidence that native *E. coli* proteins are rarely glycosylated suggests that *Campylobacter* acceptor proteins co-evolved with *pgl* systems and are thus tuned to accept N-glycans.

Glycosylation of the sequence DANSG is unusual. The vast majority of all N-glycans is found within the sequence N-X-S/T (where X can be any amino acid but proline) (Marshall 1974). Atypical N-glycosylation sites bearing the NXC sequence have been reported for eukaryotic proteins including CD69 (Vance et al. 1997), immunoglobulin μ chain (Kehry et al. 1979), human leptin receptor (Haniu et al. 1998) and von Willebrand factor (Titani et al. 1986). Additionally, a recent global LC-MS/MS analysis of murine glycopeptides (Kehry et al. 1979), human leptin receptor (Haniu et al. 1998) and von Willebrand factor (Titani et al. 1986). Additionally, a recent global LC-MS/MS analysis of murine glycopeptides has extended the spectrum of nonconsensus recognition motifs to include N-G-X and N-X-V sequences (Zielinska et al. 2010). Structural studies have indicated that the amide bond is involved in the formation of the non-canonical NSG sequence of the CH1 domain of human IgG1 and partially glycosylated (Valliere-Douglass et al. 2009). This finding seems similar to what we report on the NSG site of AcrA. Both observations demonstrate that bacterial and eukaryotic O-linked glycosylation sites have a similar basis for the recognition of glycosyl acceptor sites. Moreover, occurrence of this particular glycosylation both in bacteria and in eukaryotes implies that this activity is intrinsic in the PglB/Stt3 protein, and it is not due to the absence of potential acceptor substrates. In a follow-up paper, Valliere-Douglass et al. (2010) identified additional noncanonical glycosylation sites in IgG1 and proposed that the common feature of these unusual low-efficiency glycosylation sites is the presence of a "reverse-conensus" motif (serine/threonine residue at the −2 position relative to the modified asparagine residue). Therefore, the DANSG site of AcrA might not belong to this class of atypical glycosylation sequences.

Further investigation will be needed to unravel the biological significance of this relaxed glycosylation, whether this has to be regarded as a limitation in the fidelity of OST or as a strategy to extend the number of N-glycosylations.

Materials and methods
Constitution of plasmids
All the strains and plasmids used in this study are listed in Supplementary data, Table Sl. Oligonucleotides were purchased from Microsynth (Balgrach, Switzerland). Unless otherwise stated, *E. coli* DH5α was chosen as host for cloning. Restriction enzymes were purchased from NEB or Fermentas. T4 DNA ligase was from NEB. Plasmid pFLA29 was constructed by annealing of oligos 5'-TGTAGTTATTACGATGTTAATAACGATGATTTATCGTT and 5'-GCTCAGCAGCAGCTAGGCTAGCAGCTAGCAGCAG-3' and 5'-GCTCAGCAGCAGCTAGGCTAGCAGCTAGCAGCAG-3'. AATTAAAC-3' and subsequent ligation into pACYC184, previously cut with *BfiI* and *Bpu1102I*. Cloning of the *C. lari* pgl cluster was performed by three overlapping polymerase chain reaction (PCR) fragments and homologous recombination in yeast (Oldenburg et al. 1997). Integration of the PCR fragments into the Smal-linearized YEp352 vector was achieved by homologous recombination in SS328 yeast strain during transformation. Transformants were selected on synthetic minimal medium lacking uracil, and recombinant plasmids were isolated from yeast cells (Sherman 2002). After re-transformation in *E. coli*, plasmids were confirmed by restriction digestion. A 12 kb fragment was released from YEp352(pgl4) by digestion with *PacI* and Bpu1102I and ligated into the corresponding sites of pFLA29, thus producing pACYC(pgl4). Plasmid pFLA26 was constructed by subcloning the Psil/StpAcrA fragment from pMIK43 into pFLA19. The pglB ORFs were amplified by PCR using *C. coli* gDNA and *C. lari* gDNA. Fragments containing CcpglB were cut with Xbal/HindIII and ligated into pMLBAD, previously cut with the same enzymes. Fragments containing CpglB were cut with EcoRI and *PstI* and ligated into pMLBAD, previously cut with the same enzymes. All pglB ORFs are in frame with an HA tag at the C-terminus. The *potD* gene of *E. coli* MC1067 was amplified and cloned into pEC415 previously cut with *NdeI* and *EcoRI*. All point mutations were inserted by site-directed mutagenesis with the QuickChange Kit from Stratagene. All constructs were confirmed by restriction analysis and sequencing of relevant fragments (Syngene AG, Switzerland).

**Protein expression, purification and analysis**
*E. coli* SCM3 or SCM6 harboring pACYC(pgl) or pACYC(pgl4) and a plasmid coding for an acceptor protein were grown overnight from single colonies at 37°C in volumes of 10 mL of Luria-Bertani (LB) medium. Ampicillin (100 mg/L), trimethoprim (100 mg/L) and chloramphenicol (25 mg/L) were added to the medium, as needed. Preparation of periplasmic extracts was carried out by lysozyme treatment, consisting of an incubation in 30 mM Tris–HCl pH 8.5, 20% (w/v) sucrose, 1 mM EDTA, 1 mg/mL lysozyme (Sigma), for 1 h on ice. Spheroplasts were removed by centrifugation. For protein purification, cultures were scaled up to 0.5 L. For AcrA purification, periplasmic extracts were supplemented with imidazole to reach a final concentration of 10 mM, sterile-filtered (0.22 μm) and purified by nickel affinity chromatography using Ni-NTA agarose (Qiagen). For the separation of glycosylated proteins, periplasmic extracts were dialyzed twice against phosphate-buffered saline (PBS) pH 7.4 and loaded onto an SBA-agarose (Vector laboratories Inc., CA) column. After washing with PBS, bound proteins were eluted with PBS containing 0.5 M galactose. SDS–PAGE was performed according to Lämmli; immunodetection was performed with polyclonal anti-AcrA serum (Wacker et al. 2002), polyclonal hR6 serum (S. Amber and M. Aebi, in preparation), polyclonal anti-PotD serum (Furuchi et al. 1991) and monoclonal anti-His antibody (Qiagen). For the removal of the GalNAc residues from N-glycan, 3 μg of glycoproteins were incubated for 3 h (unless differently specified) with 40 U of exo-α-N-acetyl-galactosaminidase (NEB) in the presence of 4 μg of bovine serum albumin in 50 mM Tris buffer pH 7.5.
For the characterization of the N-glycan of C. lari, 50 µg of glycoproteins was incubated with proteinase K (Sigma), overnight at 37°C, at a 1:50 (w/w) enzyme:substrate ratio. Digestion products were passed through a carbon column (ENVI-Carb SPE, Supelco), washed with 0.1% formic acid (FA) and eluted with 25% acetonitrile (ACN), 0.1% FA. Eluent was permethylated and spotted on the plate with 2,5-dihydroxybenzoic acid (DHB) matrix (10 mg/mL in 50% ACN and 0.1% FA). MALDI-MS and MS/MS were acquired in the positive ion reflector mode by MALDI-TOF/TOF MS with 4800 MALDI TOF/TOF Analyzer instrument (Applied Biosystems, CA). For glycophosphate identification, 0.1 mg of glycoproteins was incubated with trypsin (Promega), overnight at 37°C. After heat inactivation, tryptic digestion products were further treated with exo-α-N-acetyl-galactosaminidase from 8 h to overnight. Sep-Pak C18 was used to desalt and remove enzymes. Peptides were eluted with 70% ACN and 0.1% FA and analyzed with an LTQ-FT-ICR-MS instrument (Thermo Scientific, MA). Samples were injected into an Eksigent-nano-HPLC system (Eksigent Technologies, CA) by an autosampler and separated on a self-made reverse-phase tip column (75 µm × 80 mm) packed with C18 material (AQ, 3 µm, 200 A, Bischoff GmbH, Germany). The column was equilibrated with solvent A (A: 3% ACN and 0.2% FA, in water). Peptides were eluted using the following gradient: 0–50 min, 0–60% B; 50–53 min, 60–97% B; 53–60 min, 97% B (B: 80% ACN and 0.2% FA, in water) at a flow rate of 0.2 µL/min. High accuracy mass spectra were acquired at an LTQ-ICR-FT in the mass range of 300–2000 m/z and a target value of 5 × 10^5 ions. Up to four data-dependent MS/MS spectra of the most intense ions with charge state 2+ or higher were recorded in parallel at the ion trap using collision-induced dissociation.

**Supplementary data**

Supplementary data for this article is available online at http://glycob.oxfordjournals.org/.

**Funding**

This work was supported by the Swiss National Science Foundation (grant 3100AQ-105541 to M.A.) and the ETH Zurich.

**Acknowledgments**

We thank the members of the Aebl laboratory for fruitful discussions. We are grateful to Manuela Mally and Karin Ilg for critically reading the manuscript. We thank Miguel Valvano and Cristina Marolda for the gift of E. coli SCM6 and Brendan Wren and Herbert Hächler for the genomic DNA of C. lari and C. coli. F.S., C.L. and Y.-Y.F. are students of the Life Sciences Zurich Graduate School.

**Conflict of interest statement**

None declared.

---

**References**


Furuchi T, Kashivagi K, Kobayashi H, Igarashi K. 1991. Characteristics of α-60% B; 50 m/z; 53–60 min, 97% B (B: 80% ACN and 0.1% FA, in water) at a flow rate of 0.2 µL/min. High accuracy mass spectra were acquired at an LTQ-ICR-FT in the mass range of 300–2000 m/z and a target value of 5 × 10^5 ions. Up to four data-dependent MS/MS spectra of the most intense ions with charge state 2+ or higher were recorded in parallel at the ion trap using collision-induced dissociation.

**Supplementary data**

Supplementary data for this article is available online at http://glycob.oxfordjournals.org/.

**Funding**

This work was supported by the Swiss National Science Foundation (grant 3100AQ-105541 to M.A.) and the ETH Zurich.

**Acknowledgments**

We thank the members of the Aebl laboratory for fruitful discussions. We are grateful to Manuela Mally and Karin Ilg for critically reading the manuscript. We thank Miguel Valvano and Cristina Marolda for the gift of E. coli SCM6 and Brendan Wren and Herbert Hächler for the genomic DNA of C. lari and C. coli. F.S., C.L. and Y.-Y.F. are students of the Life Sciences Zurich Graduate School.

**Conflict of interest statement**

None declared.

---

**Abbreviations**

ACN, acetonitrile; Bac, bacillolysamine; CjPgIB, Campylobacter jejuni PgIB; CIPgIB, Campylobacter lari PgIB; DHB, 2,5-dihydroxybenzoic acid; FA, formic acid; GalNAc, N-acetyl-galactosamine; IgG, immunoglobulin G; LB, Luria-Bertani; LC-ESI, liquid chromatography-electrospray ionization mass spectrometry; MALDI-MS, matrix-assisted laser desorption/ionization mass spectrometry; MS, mass spectrometry; ORF, open reading frame; OST, oligosaccharyltransferase; PBS, phosphate-buffered saline; PCR, polymerase chain reaction; pgI, protein glycosylation; SBA, soy bean agglutinin; SDS–PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

**References**


