Characterization of exogenous bacterial oligosaccharyltransferases in *Escherichia coli* reveals the potential for O-linked protein glycosylation in *Vibrio cholerae* and *Burkholderia thailandensis*

Carol Gebhart², Maria Veronica Ielmini², Bela Reiz³, Nancy L Price², Finn Erik Aas⁴, Michael Koomey⁴, and Mario F Feldman¹,²

¹Alberta Glycomics Centre, Department of Biological Sciences, University of Alberta, Edmonton, CW405 Biological Sciences Building, AB, Canada T6G 2E9; ²Mass Spectrometry Facility, Department of Chemistry, University of Alberta, Edmonton, AB, Canada T6G 2G2; and ³Department of Molecular Biosciences, Centre for Molecular Biology and Neuroscience, University of Oslo, 0316 Oslo, Norway

Received on January 12, 2012; revised on February 27, 2012; accepted on February 28, 2012

Bacterial protein glycosylation systems from varying species have been functionally reconstituted in *Escherichia coli*. Both N- and O-linked glycosylation pathways, in which the glycans are first assembled onto lipid carriers and subsequently transferred to acceptor proteins by an oligosaccharyltransferase (OTase), have been documented in bacteria. The identification and characterization of novel OTases with different properties may provide new tools for engineering glycoproteins of biotechnological interest. In the case of OTases involved in O-glycosylation (O-OTases), there is very low sequence homology between those from different bacterial species. The Wzy_C signature domain common to these enzymes is also present in *Waal* ligases; enzymes involved in lipopolysaccharide biosynthesis. Therefore, the identification of O-OTases using solely bioinformatic methods is problematic. The hypothetical proteins BTH_I0650 from *Burkholderia thailandensis* E264 and VC0393 from *Vibrio cholerae* N16961 contain the Wzy_C domain. In this work, we demonstrate that both proteins have O-OTase activity and renamed them PglL<sub>Bt</sub> and PglL<sub>VC</sub>, respectively, similar to the *Neisseria meningitidis* counterpart (PglL<sub>Nm</sub>). In *E. coli*, PglL<sub>Bt</sub> and PglL<sub>VC</sub> display relaxed glycan and protein specificity. However, effective glycosylation depends upon a specific combination of the protein acceptor, glycan and O-OTase analyzed. This knowledge has important implications in the design of glycoconjugates and provides novel tools for use in glycoengineering applications. The codification of enzymatically active O-OTase in the genomes of members of the *Vibrio* and *Burkholderia* genera suggests the presence of still unknown O-glycoproteins in these organisms, which might have a role in bacterial physiology or pathogenesis.

Keywords: *Burkholderia* / glycoengineering / oligosaccharyltransferase / protein glycosylation / *Vibrio cholerae*

Introduction

The number of prokaryotic protein glycosylation systems identified, both N- and O-linked, is increasing rapidly (Szymanski et al. 2003; Fletcher et al. 2009; Vik et al. 2009; Egge-Jacobsen et al. 2011). Glycoproteins can be generated via glycosyltransferases in the cytoplasm or by an en bloc mechanism through the activity of oligosaccharyltransferases (OTases); enzymes that transfer a preassembled glycan from a lipid carrier to select amino acids in their target proteins in the periplasm (Nanthaft and Szymanski 2010). In bacteria, both N- and O-glycosylation can be carried out en bloc via an N- or an O-OTase, respectively (Hug and Feldman 2010). The best characterized bacterial OTase is PglB, an N-OTase homolog to the eukaryotic Stt3, which is the key component of the OTase complex responsible for N-glycosylation in the lumen of the endoplasmic reticulum. In *Campylobacter jejuni*, PglB is responsible for the N-glycosylation of more than 60 periplasmic and membrane proteins (Szymanski and Wren 2005; Nothaft et al. 2010). N-OTases seem to be common in many epsilonids and in a few delta proteobacteria, particularly in the genus *Desulfovibrio*. N-OTases from different bacterial species appear to have particular sequence recognition sequences and diverse sugar specificities (Jervis et al. 2010; Ielmini and Feldman 2011; Schwarz et al. 2011).

Two bacterial O-OTases have been well-characterized. PilO is responsible for O-glycosylation of a single protein, the type IV pilin, in several *Pseudomonas aeruginosa* strains (Castric 1995). The O-OTases PglL from *Neisseria meningitidis* and PglO from *N. gonorrhoeae* are able to glycosylate multiple proteins (Faridmoayer et al. 2007; Ku et al. 2009; Vik et al. 2009).
2009; Borud et al. 2010). PgIO and PgIL primary sequences differ only at a few sites, and for simplicity, this manuscript will use PgIL\textsubscript{Nm} to refer to both enzymes. O-OTases share a conserved Pfam \texttt{Wzy}\_C domain and a predicted similar membrane topology with the O-antigen ligase, WaaL, an enzyme involved in the attachment of O-antigen to lipid A during lipopolysaccharide (LPS) biosynthesis (Ruan et al. 2012). Due to these shared commonalities and the availability of only a few characterized O-OTases, current bioinformatic methods are unable to differentiate whether a particular gene encodes an O-antigen ligase or an O-OTase. Thus, the identification of O-glycosylation activity requires experimental demonstration.

The activities of both PilO and PgIL have been reconstituted in \textit{Escherichia coli}, where they exhibited very relaxed glycogen specificity (DiGiandomenico et al. 2002; Faridmoayer et al. 2007, 2008). No requirements for specific stereochemistry or substitutions in the reducing sugar have been determined for glycogen recognition to date. However, there is a difference in the length of the glycogen that can be transferred; although PgIL has no apparent length limit, PilO is only able to transfer short oligosaccharides (Faridmoayer et al. 2007). To date, there is also no determined consensus sequence for the acceptor protein, other than that it must encompass a serine or threonine residue. Particularly, PilO appears to only require a terminal serine residue (Qutyan et al. 2010). It was proposed that PgIL\textsubscript{Nm} recognizes “low-complexity regions (LCRs)” around the glycosylation site that includes an abundance of alanine, valine and proline residues, similar to what is found in the O-glycosylation sites of eukaryotes (Vik et al. 2009). However, such a sequence is neither necessary nor sufficient for glycosylation. Although some requirements and properties have been described, the mechanism governing bacterial protein O-glycosylation remains largely unknown. Studying more of these enzymes from various bacterial systems will provide insight into the molecular basis of the O-glycosylation mechanism and may increase the already enormous potential of these enzymes for glycoengineering novel glycoproteins, which have promising applications in the vaccine design and as diagnostic tools (Feldman 2009; Ielmini and Feldman 2011). Interestingly, both proteins are well conserved within \textit{Vibrio} and \textit{Burkholderia} species (Supplementary data, Figures S2 and S3). Therefore, we chose to focus further characterization on VC0393 and BTH\_I0650.

VC0393 and BTH\_I0650 have O-OTase activity

To determine if VC0393 and BTH\_I0650 have O-OTase activity, the genes were cloned and the proteins expressed in \textit{E. coli} with C-terminal polyhistidine tags. The VC0393 ORF consists of 597 residues and has a predicted mass of 67.4 kDa, whereas BTH\_I0650 has 595 residues and a predicted molecular weight of 65 kDa. Expression of the proteins was analyzed by immunoblotting using anti-His-tag antibodies. VC0393 expression at 30°C resulted in significantly greater amounts of the protein than at 37°C, whereas the opposite was found for BTH\_I0650 (Supplementary data, Figure S4A and B).

To test the functionality of an O-OTase in \textit{E. coli}, the enzyme has to be co-expressed with an appropriate protein acceptor and a glycan donor (Feldman et al. 2005). Flagellin glycosylation has been reported in \textit{B. pseudomallei} (Scott et al. 2011). However, there is no evidence of an \textit{en bloc} glycosylation system in \textit{V. cholerae} or \textit{Burkholderia} spp. reported to date. Consequently, no information about possible native protein acceptors or glycan donors was available. It was previously shown that the lipid-linked heptasaccharide from \textit{C. jejuni} (CjLLO) can serve as a successful glycan donor for several \textit{N-} and O-OTases, including PgIL\textsubscript{Nm} and PilO (Faridmoayer et al. 2007; Ielmini and Feldman 2011). We speculated that it could also act as a substrate for the putative \textit{V. cholerae} and \textit{B. thailandensis} O-Tases analyzed. As potential protein acceptors, we chose three well-characterized O-linked glycoproteins; the type IV pilin subunit protein from \textit{N. meningitidis} MC58 (Power et al. 2006), the Laz (lipid-azurin) protein from \textit{N. gonorrhoeae} (Vik et al. 2009) and the disulphide isomerase DsbA1 protein from \textit{N. meningitidis} MC58 (Lafaye et al. 2009). The corresponding genes were cloned and expressed in \textit{E. coli} as C-terminal His-tagged proteins.

To test the OTase activity of the \textit{V. cholerae} and \textit{B. thailandensis} ORFs, plasmids expressing one of the putative OTases and one of the three protein acceptors (pilin, Laz or DsbA1) were co-transformed into \textit{E. coli} strain CLM24 carrying a plasmid that contained the \textit{C. jejuni} pgI genes required for the O-glycosylation of the pilin subunit. The pilin subunit-glycogen fusion proteins were isolated from cell lysates using nickel chloride and analyzed by SDS-PAGE and Coomassie staining. The resulting gels were imaged using a Genesis II Imaging System (Bio-Rad Laboratories, Hercules, CA, USA) and analyzed using Quantity One software (Bio-Rad Laboratories, Hercules, CA, USA). The results showed that both VC0393 and BTH\_I0650 were capable of transferring short oligosaccharides to the pilin subunit, indicating their activity as O-OTases.

Results

Identification of potential \textit{B. thailandensis} and \textit{V. cholerae} OTases

O-OTases characterized to date have overall low sequence similarity. However, they share two main common features that can be used for their identification. These enzymes have the Pfam \texttt{Wzy}\_C domain (PF04932) and are membrane proteins with at least 12 transmembrane domains (Power et al. 2006). Using bioinformatic techniques that included BLASTP and BLASTX searches (Altschul et al. 1997), hidden Markov modeling (Eddy 2009) and transmembrane profiling software, we identified open reading frames (ORFs) that could encode for WaaL ligases or O-OTases in both \textit{V. cholerae} O1 El Tor N16961 (VC0237 and VC0393) and \textit{B. thailandensis} E264 (BTH\_I0988 and BTH\_I0650). VC037 has been functionally characterized as a true WaaL ligase, responsible for O-antigen transfer during LPS synthesis in \textit{V. cholerae} (Schild et al. 2005). Despite overall low sequence identities between themselves, VC0393 and BTH\_I0650 could be distinguished from VC0237 and BTH\_I0988 by virtue of a relatively extended C-terminus encompassing a DUF3366 domain that is also present in PgIL\textsubscript{Nm} (Supplementary data, Figure S1).

Interestingly, both proteins are well conserved within \textit{Vibrio} and \textit{Burkholderia} species (Supplementary data, Figures S2 and S3). Therefore, we chose to focus further characterization on VC0393 and BTH\_I0650.
synthesis of the CjLLO. The CLM24 strain is a convenient system for glycosylation studies, because the absence of WaaL results in the accumulation of the Und-PP-linked glycans that serve for OTase substrates (Feldman et al. 2005). In our experiments, we included the previously characterized OTase PgLLNm from *N. meningitidis* MC58 for comparison purposes. Whole-cell lysates were separated on sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) and analyzed by western blot using an infrared fluorescent imaging system. This system allows the simultaneous detection of two different molecules (protein and glycan) in separate fluorescent channels: red and green. Co-localization of both signals is visualized in yellow, which allows the detection of modifications to proteins in the same membrane. Western blot analysis was performed using the polyclonal anti-glycan HR6 antibody for detection of the *C. jejuni* heptasaccharide and monoclonal anti-His antibody to detect the His-tagged recombinant proteins.

All the three acceptor proteins tested in the presence of VC0393 exhibited a slower migrating band compared with each unmodified protein control that was reactive toward both antibodies (Figure 1). The electrophoretic shift noticed was compatible with the addition of the CjLLO to each protein. This suggested that VC0393 was able to transfer the CjLLO to the three acceptor proteins tested. In a parallel experiment, BTH_I0650 transferred the oligosaccharide only to Laz and

---

**Fig. 1.** O-OTases BTH_I0650 and VC0393 transfer a heptasaccharide to specific protein acceptors. Whole-cell lysates of *E. coli* CLM24 cells expressing pilin (A), Laz (B) or DsbA1 (C) as acceptor proteins together with CjLLO and different O-OTases were separated by SDS–PAGE and analyzed by western blotting. His-tagged proteins were detected using the monoclonal anti-His antibody (green). The CjLLO glycan was detected using the polyclonal HR6 antibody (red). The overlapping signals are shown in yellow (merge). The glycosylation of pilin by PgLLNm has been shown elsewhere (Faridmoayer et al. 2007). Some bands corresponding to unspecific binding of the antibodies or *E. coli* protein glycosylated with the heptasaccharide can be seen in panels B and C.
DsbA1, but not to pilin. These results indicate that indeed VC0393 and BTH I0650 have OTase activity and were renamed PgLLNm and PgLNBt, respectively. Surprisingly, PgLLNm only recognized as substrates pilin and DsbA1. This was unexpected since Laz is an acceptor of the native oligosaccharides produced in N. gonorrhoeae. DsbA1 was the only protein recognized by the three OTases, PgLLNm, PgLNBt and PgLNVc (Figure 1C, lower panel). More than one band reacting with both the anti-glycan and the anti-His antibodies were detected for this acceptor protein. These bands likely reflect the glycosylation of two different forms of DsbA1, possibly resulting from proteolytic degradation or incomplete processing of the lipoprotein.

Mapping the glycosylation sites of DsbA1 glycopeptides

Western blot analysis indicated that DsbA1 was glycosylated by all three OTases with the glycan tested. Given the potential use of this protein for biotechnological applications, we wanted to further confirm the glycosylation results observed by western blot and to determine the precise glycosylation site(s) on this protein by mass spectrometric (MS) analysis. DsbA1 glycosylated with the C. jejuni heptasaccharide by PgLNBt was purified using a Triton X-114 phase separation method optimized for the purification of lipoproteins (Supplementary data, Figure S5), followed by affinity nickel chromatography. The purified glycosylated DsbA1 was subjected to SDS-PAGE and visualized by Coomassie staining. The same sample was also analyzed by western blot revealing four bands corresponding to unglycosylated and three glycosylated forms of DsbA1 (Figure 2A). These bands were processed by proteolytic in-gel digestion with trypsin and the peptide mixture was analyzed by MS. Previous work on N. gonorrhoeae DsbA1 homolog demonstrated that this protein is glycosylated by PgLNBt in an LCR at the N-terminal part of the protein (Vik et al. 2009). The in silico trypptic digestion of DsbA1 predicts the generation of a relatively long peptide, VQTVSPADSAPAASAAAPAGLIEEEQNYTVLNNPIQP-QQAAGK, of m/z 4059.0782 Da (Figure 2B), which contains the LCR indicated in bold. Two major ions at m/z 4059.18 and 4081.22 Da were present in the matrix-assisted laser desorption ionization (MALDI)-MS spectrum of the unglycosylated DsbA1 (upper panel) corresponding to the [M + H]+ and [M + Na]+ ions, respectively. The same ions were found in the spectrum of the glycosylated forms of DsbA1 (lower panel) together with two new ions at m/z 5439.98 and 5462.04 Da. The difference between these set of ions is ~1381 Da, which could correspond to the mass of a (HexNAc)3Hex moiety, one of the glycosylation forms of the CJLO synthesized in E. coli, containing N-acetylglucosamine (HexNAc) instead of the 2,4-di-N-acetylglucosamine (diNAcBac) produced in the original host (Wacker et al. 2002; Reid et al. 2009). The identity of the attached glycan was confirmed by β-elimination of the O-linked glycans followed by permethylation and MALDI-MS and MS/MS analysis (Figure 3). The high-resolution MALDI-MS spectrum showed a peak at 1743.88888 Da corresponding to an elemental composition of C_{74}H_{126}N_{33}O_{36} + Na with a mass error of ~0.05 ppm, which corresponds to the above-mentioned glycan moiety. A sequence of the glycan moiety has also been confirmed by MS/MS analysis. Similar results were obtained when glycosylated samples of DsbA1 by PgLNVc were analyzed (data not shown). Together, these data provided strong evidence that both OTases recognize and glycosylate a site located in the same LCR in the N-terminal part of the protein, in agreement with the findings previously reported for PgLNNm.

Glycan donor specificity of PgLNBt and PgLNVc

To test if PgLNVc and PgLNBt also present relaxed glycan specificity, we compared their ability to transfer mono- and polysaccharides to the protein acceptors. Monosaccharides are not transferred by the N-OTase PgLNB (Schwarz et al.) We tested if the undecaprenyl pyrophosphate attached to the monosaccharide diNAcBac could serve as a glycan donor. This sugar was chosen because glycosylation can be easily detected by western blot using a specific monoclonal antibody (mAb) that recognizes a diNAcBac-associated epitope. Moreover, diNAcBac is found as the reducing sugar in several glycans present in N-glycoproteins from C. jejuni and O-glycoproteins from Neisseria spp. (Young et al. 2002; Power and Jennings 2003). Whole-cell lysates of the E. coli strain CLM24 expressing one OTase, one of the protein acceptors and carrying the N. gonorrhoeae pglFBCD genes, sufficient for the synthesis of lipid-linked diNAcBac, were separated by SDS-PAGE and analyzed by western blot. The addition of a monosaccharide does not significantly change the electrophoretic mobility of the glycoproteins relative to the unglycosylated control. However, glycosylation can be detected as an overlapping signal, visualized in yellow, between the protein and glycan signals. As is shown in Figure 4, PgLNNm and PgLNVc were able to transfer diNAcBac to pilin and DsbA1, but not to Laz, whereas PgLNBt transferred the monosaccharide only to Laz and DsbA1. It was surprising to find that although the three OTases were able to efficiently transfer diNAcBac to DsbA1, the transfer of this monosaccharide to Laz was inefficient. Both PgLNVc and PgLNBt glycosylated Laz with the CjLLO; therefore, Laz contains all the information required for recognition by this OTase. Furthermore, these results indicate that certain combinations of sugars and proteins result in product-specific glycosylation. Interestingly, pilin is not a substrate of PgLNBt but it is recognized by PgLNVc and PgLNNm, whereas Laz is a substrate of PgLNVc and PgLNBt. This indicates that these OTases do not recognize the same structural determinants in the acceptor proteins.

PgLNBt and PgLNVc can transfer polysaccharides

The ability of PgLNBt and PgLNVc to glycosylate DsbA1 with both mono- and oligosaccharides prompted us to test if these enzymes could also transfer polysaccharides to the acceptor proteins, a property that may be relevant for glycoengineering purposes. The E. coli K12 strains naturally expresses the O16 O-antigen, which has been shown to be a substrate for PgLNNm (Faridmoayer et al. 2008). However, all the common lab strains, including CLM24, have an insertion in the rhamnosyltransferase (WbbL) gene, which is essential for O16-antigen biosynthesis. Plasmid pMF19, encoding WbbL, restores the synthesis of this polysaccharide (Feldman et al. 1999). Escherichia coli CLM24 cells expressing DsbA1 were
transformed with pMF19 and either pCDG10 (pggL<sub>Bt</sub>) or pCDG4 (pggL<sub>Vc</sub>). Glycosylation of DsbA1 with the <i>E. coli</i> O16 polysaccharide was analyzed using a polyclonal anti-His and polyclonal anti-HR6 antibody directed against the CjLLO (right panel). (B) The four bands shown in (A) (labeled 1–4) were cut and subjected to in-gel digestion with trypsin. The structure of the peptide and glycan is shown, with potential glycosylation sites underlined and italicized. The LCR is indicated by bold letters. MALDI-MS spectra of singly charged ion at <i>m/z</i> 5440<sup>+</sup> corresponds to the 41-amino acid residue peptide (4059 <i>m/z</i>) modified with the CjLLO (1381 <i>m/z</i>) composed of HexNAc residues. In the figure, squares represent N-acetylgalactosamine residues, and circles represent glucose residues.

Fig. 2. MALDI-MS analysis of glycosylated DsbA1 with the <i>C. jejuni</i> heptasaccharide. (A) Purified glycosylated DsbA1 was separated on SDS–PAGE and analyzed by Coomassie stain (left panel) or immunodetected with monoclonal anti-His and polyclonal anti-HR6 antibody directed against the CjLLO (right panel). (B) The four bands shown in (A) (labeled 1–4) were cut and subjected to in-gel digestion with trypsin. The structure of the peptide and glycan is shown, with potential glycosylation sites underlined and italicized. The LCR is indicated by bold letters. MALDI-MS spectra of singly charged ion at <i>m/z</i> 5440<sup>+</sup> corresponds to the 41-amino acid residue peptide (4059 <i>m/z</i>) modified with the CjLLO (1381 <i>m/z</i>) composed of HexNAc residues. In the figure, squares represent N-acetylgalactosamine residues, and circles represent glucose residues.

transformed with pMF19 and either pCDG10 (pggL<sub>Bt</sub>) or pCDG4 (pggL<sub>Vc</sub>). Glycosylation of DsbA1 with the <i>E. coli</i> O16 polysaccharide was analyzed using a polyclonal anti-O16 antibody. The typical pattern of bands corresponding to protein attached to O-antigen was detected with both the anti-His antibodies and anti-O16 antibodies, indicating that PgL<sub>Bt</sub> was able to glycosylate DsbA1 with fully polymerized <i>E. coli</i> O16 antigen (Figure 5). A band of a molecular mass compatible in size with PgL<sub>Vc</sub> was visualized with anti-His antibody only in cells expressing PgL<sub>Vc</sub> in the absence of O-antigen (Figure 5, middle panel). Interestingly, PgL<sub>Nm</sub> and PgL<sub>Vc</sub> were unable to transfer the polysaccharide to DsbA1. This was unexpected, because PgL<sub>Nm</sub> is able to glycosylate pilin with the same polysaccharide.

*PgL<sub>Vc</sub> can glycosylate itself*

Carbohydrate-specific staining methods failed to show evidence of PgL<sub>Vc</sub> activity in <i>V. cholerae</i> (data not shown). To identify the possible endogenous glycosylation substrates, we
transformed *V. cholerae* cells with plasmid pCDG4 (*pglLVc*) and pEF33 (contains the enzymes to synthesize lipid-linked diNAcBac). Only one band matching the expected electrophoretic mobility of PglLVC was detected with the anti-glycan antibody in the presence of both the OTase and the glycan (Figure 6A). This band was also detected with the anti-His antibodies, suggesting that PglLVC was glycosylated. To confirm this hypothesis, PglLVC was purified by affinity nickel chromatography. The purified protein was subjected to SDS-PAGE and visualized by Coomassie staining. The band was in-gel digested with trypsin and the peptide mixture was analyzed by LC-ESI-QTOF-MS/MS (Figure 6B).

The MS/MS results identified a triply charged peak detected at *m/z* 749.63 Da having the peptide sequence of...
VNSATSDAEVPSAATPHH with two diNacBac residues attached to it. This peptide is located at the C-terminal end of PglL\textsubscript{vc} and also contains two of the histidine residues derived from the His-tag. This result demonstrated that PglL\textsubscript{vc} contains the molecular determinants necessary for glycosylation and that in the presence of a suitable glycan donor, it can self-glycosylate. This analysis does not rule out the possibility that other proteins, present in minor amounts, are also glycosylated, but not detected in these conditions.

**Discussion**

Most of the bacterial OTases described so far exhibit relaxed glycan specificity. This feature makes bacterial OTases useful tools for glycoengineering of novel glycoconjugates with promising applications in vaccinology and diagnostics. The identification and study of additional OTases may result in novel and more efficient enzymes, which could have distinct biotechnological applications. It has been proposed that bacterial glycoproteins could constitute a novel generation of conjugate vaccines (Feldman et al. 2005; Feldman 2009). Recently, promising applications of bacterial engineered glycoproteins for the diagnostics of bacterial infections have been demonstrated (Iwashkiw et al. 2012). The first OTases characterized were *C. jejuni* PglB N-OTase and multiple related \( \varepsilon \)-proteobacteria homologs (Szymanski et al. 1999; Wacker et al. 2002, 2006; Kowarik et al. 2006; Jervis et al. 2010; Schwarz et al. 2011). Although initially it was assumed that all bacterial PglB homologs will recognize the same protein acceptor sequence DxNxS/T, it was established that PglB from the \( \delta \)-proteobacterium *Desulfovibrio desulfuricans* recognizes a different sequence in its target protein (Ielmini and Feldman 2011). O-Glycosylation seems to be more widespread among bacteria. However, the activities of only three
O-OTases have been conclusively demonstrated. These are *P. aeruginosa* PilO, *Neisseria* PglL and *Francisella tularensis* PglA (Castric 1995; Faridmoayer et al. 2007; Egge-Jacobsen et al. 2011). The acceptor sequences and glycan specificities of PilO and PglL are different. PilO does not transfer polysaccharides and recognizes a sequence located at the C-terminal of the protein (Qutyan et al. 2010). Although PglL can transfer polysaccharides of virtually any glycan structure to the pilin protein, the acceptor requirements have not been conclusively elucidated (Faridmoayer et al. 2008). Several acceptor proteins, but intriguingly not the pilin protein, are glycosylated at LCRs, rich in S, A and P (Vik et al. 2009). For some of the remaining O-glycosylation systems described, like the ones present in *Bacteroides fragilis* and *Tannerella forsythia*, the acceptor sequence is known to be D-(S/T)-(A/I/L/M/T/V), but the O-OTases involved remain elusive (Fletcher et al. 2009; Posch et al. 2011). In other cases, like *P. aeruginosa* TfpW, the OTase activity remains to be experimentally demonstrated (Kus et al. 2008).

In this work, through their active expression in *E. coli*, we have shown that the genomes of *B. thailandensis* and *V. cholerae* encode O-OTases. Both novel OTases are present in bacteria where no glycosylation systems have been reported previously, and therefore their native substrates were unknown. However, the OTase activity could be analyzed using proteins that are known substrates of PglLNm, which demonstrates that the three enzymes are able to recognize and glycosylate the same motifs. This is supported by the MS results that map the glycosylation site on the acceptor protein DsbA1 used by PglLVc and PglLBt at the same region previously described for PglLNm.

Pilin, Laz and DsbA1 were glycosylated by PglLVc with the CjLLO, which demonstrates that these three protein substrates carry the molecular determinants recognized by this OTase. PglLVc was also capable of transferring a monosaccharide, diNAcBac, to two of these proteins, pilin and DsbA1, indicating that the enzyme can utilize the lipid-linked diNAcBac donor molecule. However, PglLVc failed to glycosylate Laz, with only diNAcBac, which means that although Laz and diNAcBac can be substrates in the presence of a different sugar or acceptor protein, they cannot be effectively linked by this enzyme. Similarly, PglLNm was able to use DsbA1 as a
substrate with diNAcBac and the CjLLO, but was unable to transfer the \textit{E. coli} O16 polysaccharide, a known substrate for this enzyme when pilin is the acceptor protein. Therefore, it appears that O-OTases require a specific combination of protein and glycan structures for a productive interaction resulting in glycosylation. It is tempting to speculate that one of the substrates binds first to the enzyme causing a conformational change, which in turn may impose certain steric constraints that will allow the second substrate to productively reach the catalytic site. This finding has relevance for glycoengineering purposes, as some glycan structures of interest may be efficiently attached to certain carriers, but not to others.

The presence of active O-OTases in the genomes of both \textit{B. thailandensis} E264 and \textit{V. cholerae} O1 El Tor N16961 demonstrates that these strains have the potential to express glycosylation systems. It remains unclear if such activities are indeed functional in the native backgrounds. It is interesting to note that in \textit{V. cholerae}, pglL\textsubscript{Vc} maps near mshA, the gene for the major component of one of the type IV pilin that this bacterium expresses, whereas in \textit{B. thailandensis} and other \textit{Burkholderia} species, pglL\textsubscript{Bt} and orthologous genes map immediately adjacent to pilA, encoding a likely type IV pilus protein. We have attempted to glycosylate MshA and \textit{B. thailandensis} pilin in \textit{E. coli} using strategies employed to glycosylate the \textit{Neisseria} proteins, but these efforts were unsuccessful. However, introduction of an inducible plasmid expressing PglL\textsubscript{Vc} and lipid-linked diNAcBac biosynthetic genes under a constitutive promoter into \textit{V. cholerae} showed

---

**Fig. 6.** PglL\textsubscript{Vc} glycosylates itself. (A) Whole-cell lysates of \textit{V. cholerae} expressing PglL\textsubscript{Vc} and diNAcBac were separated by SDS–PAGE and analyzed by western blotting. His-tagged proteins were detected using the polyclonal anti-His antibody and the diNAcBac glycan was detected using the polyclonal antibody UOS-2. (B) Analysis of the deconvoluted spectrum of a triply charged peak detected at m/z 749.63 Da positively identified the peptide sequence of VNSATSDEVP-SAATPHH with two diNAcBac residues attached to it. The spectrum shows an almost complete series of y-type, some b-type peptide fragment ions along with strong peaks for the peptide only at m/z 1790.7 Da and peptide with one diNAcBac attached at m/z 2018.9 Da. The difference between the peak at 2018.9 Da and the molecular ion at m/z 2247.0 Da demonstrates the attachment of a second diNAcBac. No peaks within the spectrum were found to show the exact location of the attached glycan residues. The presence of diBacNAc residues is also confirmed by the presence of the fragment ions detected at m/z 229.1, 211.1 and 169.1 Da, as reported previously (Chamot-Rooke et al. 2007).
that in the presence of a suitable lipid-linked sugar, PgL_{Vc} can self-glycosylate. Interestingly, the peptide glycosylated in PgL_{Vc}, as most of the acceptor sequences shown for PgL, contains multiple S, A, and P residues. The high degree of conservation of the O-OTases in different *Vibrio* and *Burkholderia* species may be taken as an indication of the existence of a strong evolutionary pressure to retain the O-OTases in the genomes of these strains. These results prompt us to speculate that O-glycosylation in these strains is a highly regulated process, which may occur only under certain environmental conditions. In this context, it is important to recall that *Vibrio* and *Burkholderia* species also express O-poly saccharides that are dependent on WaaL ligases and pathways that parallel to those of protein O-glycosylation. Such regulation might be required to coordinate these processes so as to mitigate potential biosynthetic and metabolic cross-talk or incompatibilities. However, strains of *F. tularensis* subsp. *tularensis* simultaneously express WaaL-dependent, O-poly saccharide modified LPS and O-OTase-dependent protein glycosylation systems (Egge-Jacobsen et al. 2011). Future work will be required to determine if and when protein O-glycosylation occurs in these species.

The number of bacterial species known to carry O-linked protein glycosylation is rapidly increasing. It is now clear that many bacterial pathogens are able to O-glycosylate proteins. The involvement of O-glycosylation in bacterial pathogenesis remains unknown. It will be interesting to test if glycosylation plays a role in pathogenesis of *V. cholerae* or *Burkholderia* spp., since glycoproteins could also play a role in biofilm formation. The discovery of PgL_{Vc} and PgL_{Bt} enlarges the repertoire of enzymes available for glycoengineering, which irrespective of the role played by glycosylation in bacterial physiology and/or pathogenesis, and may contribute to the design of novel and improved diagnostics tools and conjugate vaccines against bacterial infections.

**Materials and methods**

**Bacterial strains, plasmids and growth conditions**

The bacterial strains and plasmids used in this study are listed in Table I. *Escherichia coli* and *V. cholerae* strains were grown overnight on Luria–Bertani broth at 37 and 30°C, respectively. Trimethoprim (Tp, 100 μg/mL), tetracycline (Tet, 20 μg/mL), spectinomycin (Sp, 80 μg/mL), chloramphenicol (Cm, 20 μg/mL), kanamycin (Km, 50 μg/mL), streptomycin (Sm, 200 μg/mL) and ampicillin (Amp, 100 μg/mL) were added to the medium as required. Plasmid Construction.

Genomic DNA was isolated using the DNeasy Blood & Tissue kit (Qiagen Inc., Canada). Plasmid DNA was isolated with either the QIAprep Spin Miniprep Kit (Qiagen Inc., Canada) or the QIAGEN Plasmid Midi Kit (Qiagen Inc., Canada). The PglFBCD1 gene encoding the O-OTase of *V. cholerae* O1 El Tor N16961 was amplified by polymerase chain reaction (PCR) with oligonucleotides PgL{F,D}1 (5′-ATA GAA TTC ATG AAC TGC ATG GCA ACT TTG CTG TTA AGT

![DNA](https://academic.oup.com/glycob/article-abstract/22/7/962/1988040/)

### Table I. Bacterial strains and plasmids used in this study

<table>
<thead>
<tr>
<th><em>Escherichia coli</em></th>
<th>Description</th>
<th>Reference or source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top10</td>
<td>merA, Δmer-hus/RMS-merBC, ΔlacX74, deoR, recA1, araD139 Δ(ara-leu7)697, galK, rpsL, endA1, nupG</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>CLM24</td>
<td>Constructed from <em>E. coli</em> W3110 (In(rrnD-rrnE)1) pβ-1. <em>waaL</em> mutant</td>
<td>Feldman et al. (2005)</td>
</tr>
<tr>
<td>V. cholerae</td>
<td>O1 El Tor N16961 Clinical strain isolated in 1971 in Bangladesh, hemolysin production, Sm′</td>
<td>Dr. Stefan Pukatzki, University of Alberta, Canada</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pSPORT1</td>
<td>Cloning vector, Amp′</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>pMLBAD</td>
<td>Cloning and expression vector, arabinose inducible, Ty′</td>
<td>Lefebre and Valvano (2002)</td>
</tr>
<tr>
<td>pEXT20</td>
<td>Cloning and expression vector, IPTG inducible, Amp′</td>
<td>Dykhoo et al. (1996)</td>
</tr>
<tr>
<td>pEXT21</td>
<td>Cloning and expression vector, IPTG inducible, Sp′</td>
<td>Dykhoo et al. (1996)</td>
</tr>
<tr>
<td>pACYC184</td>
<td>Cloning and expression vector, IPTG inducible, derived from pACT3</td>
<td>Chang and Cohen (1978)</td>
</tr>
<tr>
<td>pACYPeglB&lt;sub&gt;mut&lt;/sub&gt;</td>
<td>pACYC184-based plasmid encoding the <em>C. jejuni</em> protein glycosylation locus cluster containing mutations W458A and D459A in PglB, Cm′</td>
<td>Wacker et al. (2002)</td>
</tr>
<tr>
<td>pHVC32</td>
<td>pVK102-based cosmid encoding the O7 antigen cluster from <em>E. coli</em>, Tet′</td>
<td>Valvano and Cosa (1989)</td>
</tr>
<tr>
<td>pMF19</td>
<td>0.9-kb PCR ampiclon containing the wbbL (rhamnosyltransferase) gene cloned into pEXT21, Sp′</td>
<td>Feldman et al. (1999)</td>
</tr>
<tr>
<td>pAMF10</td>
<td>C-10× His-tagged PgL&lt;sub&gt;Nac&lt;/sub&gt; cloned into pEXT20, Amp′</td>
<td>Faridmoayer et al. (2008)</td>
</tr>
<tr>
<td>pAMF16</td>
<td>C-10× His-tagged Pilin cloned into pMLBAD, Ty′</td>
<td>Faridmoayer et al. (2008)</td>
</tr>
<tr>
<td>pEF186</td>
<td>pJB658-based plasmid containing BTH&lt;sub&gt;10650&lt;/sub&gt; from <em>B. thailandensis</em> E264, toluate inducible, Amp′</td>
<td>Egge-Jacobsen et al. (2011)</td>
</tr>
<tr>
<td>pEF33</td>
<td>pACYC184-based plasmid containing the pglFBCD locus from <em>N. gonorrhoeae</em>, Cm′</td>
<td>Egge-Jacobsen et al. (2011)</td>
</tr>
<tr>
<td>pEF110</td>
<td>pET24a-based plasmid containing NG00994 (Laz) from <em>N. gonorrhoeae</em> MS11, Km′</td>
<td>This study</td>
</tr>
<tr>
<td>pMRR1</td>
<td>pBCSK-based plasmid containing NG00994 from pEF110, Cm′</td>
<td>This study</td>
</tr>
<tr>
<td>pCDG2</td>
<td>C-6× His-tagged VC0393 cloned in pSPORT1, Amp′</td>
<td>This study</td>
</tr>
<tr>
<td>pCDG4</td>
<td>C-6× His-tagged VC0393 cloned into pEXT20, Amp′</td>
<td>This study</td>
</tr>
<tr>
<td>pCDG10</td>
<td>C-6× His-tagged BTH&lt;sub&gt;10650&lt;/sub&gt; cloned into pEXT20, Amp′</td>
<td>This study</td>
</tr>
<tr>
<td>pLAZ1</td>
<td>pEXT21-based plasmid containing Laz from pMRM1, Sp′</td>
<td>This study</td>
</tr>
<tr>
<td>pAMF22</td>
<td>C-10× His-tagged DsbA1 from <em>N. meningitidis</em> MC58 cloned into pMLBAD, Ty′</td>
<td>Faridmoayer A. and Feldman MF. (unpublished)</td>
</tr>
</tbody>
</table>
the desired protein acceptor, O-OTase and glycan biosynthetic coli strain CLM24 transformed with three plasmids containing pLaz1 (Table I).

All protein glycosylation reactions were carried out in the in vivo protein glycosylation section. Cultures were induced with IPTG (0.1 M) and arabinose (0.2%). Bacterial cells were harvested during stationary phase at 5000 × g for 15 min at 4°C and were washed with phosphate-buffered saline (PBS). Resulting bacterial pellets resuspended in a ratio of 1 g of pellet to 10 mL of buffer #1 (1 μg/mL DNase I, complete EDTA-free protease inhibitor cocktail, 2% Triton X-114 in PBS). Lysis was allowed to occur during incubation on ice for 4 h. Cellular debris was removed by centrifugation at 10,000 × g for 10 min at 4°C. The supernatant was incubated at 37°C for a minimal time to induce phase separation. The aqueous phase was removed after centrifugation at 10,000 × g for 10 min at 30°C. The remaining detergent layer containing the glycosylated lipoproteins was diluted to the original solution volume with buffer #1 without detergent. Ice incubation and aqueous phase removal were repeated twice, although the ice incubation was reduced to 1 h. To the detergent phase, 10 mM imidazole and 300 mM NaCl were added prior to the addition of an appropriate volume of 50% nickel-nitritriacetic acid (Ni-NTA) resin equilibrated in buffer #2 (2% Triton X-114, and 30 mM NaCl in PBS). This solution was allowed to mix overnight at 4°C. The Ni-NTA resin was packed and then washed three times with buffer #2 containing 50 mM imidazole. Desired His-tagged lipoproteins were eluted twice with buffer #2 containing 300 mM imidazole and were concentrated by TCA precipitation.

In vivo protein glycosylation

All protein glycosylation reactions were carried out in the E. coli strain CLM24 transformed with three plasmids containing the desired protein acceptor, O-OTase and glycan biosynthetic machinery. Cells containing pGL12-Bp were cultured at 37°C, whereas those containing pGL12-vc or pGL12-nn were cultured at 30°C. Cells were grown to an optical density at 600 nm of 0.4–0.6 before induction with 0.1 mM IPTG and/or 0.2% arabinose, when appropriate. Those cultures requiring arabinose induction had an additional dose added after a minimum of 4 h of induced growth. Whole-cell lysates were harvested at a stationary phase. Glycosylation efficiency was determined by western blot analysis.

SDS–PAGE, Coomassie and western blot analysis

Glycoproteins obtained from whole-cell lysates were separated with 12% SDS–PAGE gels, using the standard protocols. All samples were loaded at a concentration of 0.1 OD. Separated glycoproteins were either transferred to nitrocellulose for detection with western blotting or immediately visualized with Coomassie blue staining. Primary and secondary antibodies used in this study are listed in Supplementary data, Table S2. Membranes were visualized using the Odyssey® Infrared Imaging System (Li-Cor Biosciences, USA), which allows for direct IR fluorescence detection at 685 and 785 nm simultaneously. Hence, the same membrane can be blotted with polyclonal antibody and mAb at once, allowing for simple confirmation of overlapping signals.

Purification of glycosylated DsbA1

Purification of the 6× His-tagged glycosylated lipoprotein, DsbA1, was adapted from previously published protocols (Hasebe et al. 2004; Thakran et al. 2008). Briefly, 2 L of E. coli CLM24 cells expressing pAMF22, pCDG10 or pCDG4 and pACYCplBmut were cultured for glycosylation as described above in in vivo protein glycosylation section. Cultures were induced with IPTG (0.1 M) and arabinose (0.2%). Bacterial cells were harvested during stationary phase at 5000 × g for 15 min at 4°C and were washed with phosphate-buffered saline (PBS). Resulting bacterial pellets resuspended in a ratio of 1 g of pellet to 10 mL of buffer #1 (1 μg/mL DNase I, complete EDTA-free protease inhibitor cocktail, 2% Triton X-114 in PBS). Lysis was allowed to occur during incubation on ice for 4 h. Cellular debris was removed by centrifugation at 10,000 × g for 10 min at 4°C. The supernatant was incubated at 37°C for a minimal time to induce phase separation. The aqueous phase was removed after centrifugation at 10,000 × g for 10 min at 30°C. The remaining detergent layer containing the glycosylated lipoproteins was diluted to the original solution volume with buffer #1 without detergent. Ice incubation and aqueous phase removal were repeated twice, although the ice incubation was reduced to 1 h. To the detergent phase, 10 mM imidazole and 300 mM NaCl were added prior to the addition of an appropriate volume of 50% nickel-nitritriacetic acid (Ni-NTA) resin equilibrated in buffer #2 (2% Triton X-114, and 30 mM NaCl in PBS). This solution was allowed to mix overnight at 4°C. The Ni-NTA resin was packed and then washed three times with buffer #2 containing 50 mM imidazole. Desired His-tagged lipoproteins were eluted twice with buffer #2 containing 300 mM imidazole and were concentrated by TCA precipitation.

In-gel protein digestion

In-gel enzymatic digestion for Coomassie blue stained glycoylated and unglycosylated DsbA1 bands was done as published previously (Shevchenko et al. 1996), with the following modifications. Briefly, purified, glycosylated DsbA proteins were run on a 12% SDS–PAGE gel and stained with Coomassie blue. Cysteine bonds were reduced by treatment with 10 mM DTT in 50 mM ammonium bicarbonate for 1 h at 37°C. Thiol groups were alkylated with 50 mM iodoacetamide in 50 mM ammonium bicarbonate at room temperature for 1 h. Following subsequent washing, peptides were digested with the protease, trypsin (Promega, USA; 0.02 mg/mL of 50 mM ammonium bicarbonate) for at least 16 h at 37°C. Peptide fragments were eluted from the gel slices using acetoni trile (100%) and lyophilized. Salts were removed from the peptides using ZipTipC18 (Millipore, USA) according to the manufacturer’s instructions. Desalted peptides were dried by Speed-Vac, dissolved in 0.1% formic acid and used for subsequent MS analyses. Alternatively, dried samples were subjected to β-elimination and glycan permethylation for detailed glycan MS analyses, as described below.

β-Elimination and glycan permethylation

The β-elimination reactions were done according to a previously published protocol (Goetz et al. 2009). Briefly, 1 μL of 5 μg/μL ammonia–borane complex in 28% ammonium hydroxide was added to ~1 μg of lyophilized glycoprotein solution prepared as above and incubated for 24 h at 60°C. Cooled solutions were neutralized with 1 M hydrochloric acid. Dried samples were repeatedly resuspended in 100 μL of methanol (100%) and lyophilized for three times. Released

C Gebhart et al.
glycans were permethylated according to the published protocols (Kang et al. 2008; Alley et al. 2010). Briefly, NaOH beads were resuspended in acetoneitrile and loaded onto SpinColumnTM reactors (Harvard Apparatus, USA). Beads were washed with dimethylformamide (DMF). Dried glycans were solubilized in 100 μL of 70:25:5% DMF:iodomethane:water solution. This solution was loaded onto the NaOH beads in the spin column and incubated for 15 min at room temperature. To ensure that all glycans had bound to the beads, an additional 25 μL of iodomethane was added and allowed to incubate for an additional 15 min. Spin columns were washed twice with 100 μL of acetoneitrile. Glycan samples were recovered by liquid/liquid extraction and added to a glass vial containing 300 μL of dichloromethane. To this, 1 mL of 0.5 M sodium chloride was added to separate the organic and aqueous phases. The aqueous phase was removed and phase separation was repeated two additional times. Samples were washed with HPLC grade water three times and then lyophilized. Dried samples were resuspended in a 75:25 organic and aqueous phases. The aqueous phase was removed and phase separation was repeated two additional times. Samples were washed with HPLC grade water three times and then lyophilized. Dried samples were resuspended in a 75:25 solution of methanol to water prior to being spotted on the MALDI plates needed for MS analysis.

MALDI-MS and MS/MS analysis

Desalted samples were spotted on an MTP AnchorChipTM 600/384 target plate and air dried. A 0.5-μL aliquot of a 10-mg/mL Desalted samples were spotted on an MTP AnchorChipTM 600/384 target plate and air dried. A 0.5-μL aliquot of a 10-mg/mL solution of 2.5-di-hydroxybenzoic acid (Burker Daltonics Ltd., Canada) prepared in 70% acetoneitrile and 30% Optima LC-MS Grade H2O (Fisher, Canada) was spotted on the MALDI plate for MALDI analysis. The MALDI-TOF MS and MS/MS mass spectra were obtained on an UltraFlieXtreme MALDI TOF/TOF mass spectrometer (Bruker Daltonics, Germany). The analyzer was used in the reflectron mode of operation. Calibration was done externally using a mixture of standard peptides. The peptide ion peak picking and mass assignments were done automatically using the peak picking software of the system. The high-resolution MALDI-MS spectrum of the permethylated glycan was obtained using a Bruker 9.4T Apex-Qe FTICR (Bruker Daltonics, USA). Calibration was done externally using a sodium cationized polyethylene glycol.

Supplementary data

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

Funding

This work was supported by grants from Alberta Innovates Health Solutions, the Natural Sciences and Engineering Research Council of Canada (NSERC), the Alberta Glycomics Centre to MFF, the Research Council of Norway (166931 and 183814), the Functional Genomics initiative (FUGE) directed through the Consortium of Advanced Microbial Sciences and Technologies (CAMST) and by funds and infrastructure from the Department of Molecular Biosciences, the Centre for Molecular Biology and Neurosciences and the Mathematics and Natural Sciences Faculty at the University of Oslo to MK. MFF is an Alberta Heritage Foundation for Medical Research (AHFMR) scholar and a Canadian Institutes of Health Research (CIHR) New Investigator.

Abbreviations

Amp, ampicillin; CjLLO, lipid-linked heptasaccharide from C. jejuni; Cm, chloramphenicol; diNAcBac, 2,4-di-N-acetylbacillosamine; DMF, dimethylformamide; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; HexNAc, N-acetylglucosamine; HPLC, high-performance liquid chromatography; IPTG, isopropyl-β-D-thio-galactoside; Km, kanamycin; LC-ESI-QTOF-MS, liquid chromatography-electrospray ionization-quadrupole-time-of-flight mass spectrometer; LPS, lipopolysaccharide; LCR, low-complexity region; mAb, monoclonal antibody; MALDI, matrix-assisted laser desorption ionization; MS, mass spectrometry; Ni-NTA, nickel-nitrilotriacetic acid; OD, optical density; ORF, open reading frame; OTase, oligosaccharyltransferase; PBS, phosphate-buffered saline; PCR, polymerase chain reaction; SDS–PAGE, sodium dodeyl sulfate–polyacrylamide gel electrophoresis; Sm, streptomycin; Sp, spectinomycin; Tet, tetracycline; TCA, trichloroacetic acid; Tp, trimethoprim.

References
