Jervis, AJ; Wood, AG; Cain, JA; Butler, JA; Frost, H; Lord, E; Landon, R; Cordwell, SJ; Wren, BW; Linton, D (2018) Functional analysis of the Helicobacter pullorum N-linked protein glycosylation system. Glycobiology. ISSN 0959-6658 DOI: https://doi.org/10.1093/glycob/cwx110

Downloaded from: http://researchonline.lshtm.ac.uk/4646138/

DOI: 10.1093/glycob/cwx110

Usage Guidelines

Please refer to usage guidelines at http://researchonline.lshtm.ac.uk/policies.html or alternatively contact researchonline@lshtm.ac.uk.

Available under license: http://creativecommons.org/licenses/by/2.5/
Microbial Biology

Functional analysis of the *Helicobacter pullorum* N-linked protein glycosylation system

Adrian J Jervis²,†, Alison G Wood³,†, Joel A Cain⁴, Jonathan A Butler⁵, Helen Frost⁶, Elizabeth Lord⁶, Rebecca Langdon⁷, Stuart J Cordwell⁴, Brendan W Wren⁷, and Dennis Linton⁶,¹

²Manchester Institute of Biotechnology, SYNBIOCHEM, University of Manchester, Manchester, UK, ³Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield S10 2TN, UK, ⁴School of Molecular Bioscience and Charles Perkins Centre, The University of Sydney, ⁵School of Healthcare Science, Manchester Metropolitan University, Manchester M1 5GD, UK, ⁶Faculty of Biology, Medicine and Health, Michael Smith Building, University of Manchester, Manchester M13 9PT, UK, and ⁷Pathogen Molecular Biology Unit, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK

¹To whom correspondence should be addressed: Tel: +44-161-275-1541; e-mail: james.d.linton@manchester.ac.uk

†Both authors contributed equally to this work.

Received 19 October 2017; Revised 20 December 2017; Editorial decision 22 December 2017; Accepted 10 January 2018

Abstract

N-linked protein glycosylation systems operate in species from all three domains of life. The model bacterial N-linked glycosylation system from *Campylobacter jejuni* is encoded by *pgl* genes present at a single chromosomal locus. This gene cluster includes the *pglB* oligosaccharyltransferase responsible for transfer of glycan from lipid carrier to protein. Although all genomes from species of the *Campylobacter* genus contain a *pgl* locus, among the related *Helicobacter* genus only three evolutionarily related species (*H. pullorum*, *H. canadensis* and *H. winghamensis*) potentially encode N-linked protein glycosylation systems. *Helicobacter* putative *pgl* genes are scattered in five chromosomal loci and include two putative oligosaccharyltransferase-encoding *pglB* genes per genome. We have previously demonstrated the in vitro N-linked glycosylation activity of *H. pullorum* resulting in transfer of a pentasaccharide to a peptide at asparagine within the sequon (D/E)NXS/T. In this study, we identified the first *H. pullorum* N-linked glycoprotein, termed HgpA. Production of histidine-tagged HgpA in the background of insertional knockout mutants of *H. pullorum* *pgl/wbp* genes followed by analysis of HgpA glycan structures demonstrated the role of individual gene products in the PglB1-dependent N-linked protein glycosylation pathway. Glycopeptide purification by zwitterionic-hydrophilic interaction liquid chromatography coupled with tandem mass spectrometry identified six glycosites from five *H. pullorum* proteins, which was consistent with proteins reactive with a polyclonal antiserum generated against glycosylated HgpA. This study demonstrates functioning of a *H. pullorum* N-linked general protein glycosylation system.

**Key words:** bacteria, glycoprotein, glycosylation, Helicobacter, N-linked

© The Author(s) 2018. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.
Introduction

In all three domains of life subsets of proteins are modified by the covalent attachment of sugars to asparagine residues within a conserved consensus sequon of N-X-S/T (Calo et al. 2010; Nothaft and Szymanski 2010; Larkin and Imperiali 2011; Schwarz and Aebl 2011; Eichler 2013). Among bacterial species, two such distinct N-linked glycosylation systems have been identified. The first involves attachment of monosaccharide to asparagine by a cytoplasmic N-glycosyltransferase and has been characterized in Haemophilus influenzae (Grass et al. 2003, 2008, 2010) and Actinobacillus pleuropneumoniae (Choi et al. 2010; Kawai et al. 2011; Schwarz et al. 2011a; Naegeli et al. 2014; Cacciu et al. 2017), with further N-glycosyltransferase orthologues identified in a number of other species including pathogenic Yersinia spp., and enterotoxigenic Escherichia coli (Grass et al. 2010). The second type of bacterial N-linked protein glycosylation system was discovered in Campylobacter jejuni (Szymanski et al. 1999) and subsequently other related species from the Epsilon subdivision of the Proteobacteria (Nothaft and Szymanski 2010). In this system cytoplasmic assembly of an oligosaccharide on an isoprenoid lipid, is followed by transfer across the inner membrane and attachment onto proteins in the periplasm mediated by an integral membrane oligosaccharyltransferase (OTase). The prototypical OTase-dependent N-linked protein glycosylation system of C. jejuni has been intensively studied. More than 60 extracytoplasmic proteins are known to be glycosylated within an extended N-X-S/T sequon containing an acidic residue (D/E) at the −2 position (Kowarik et al. 2006; Wacker et al. 2006; Chen et al. 2007; Gerber et al. 2013) although examples of non-classical occupied sequons (without the D/E at the −2 position or S/T at the +2 position) have also been demonstrated (Scott et al. 2014). A single locus contains genes required for biosynthesis, transport and linkage of a conserved heptasaccharide to protein. Five cytoplasmic glycosyltransferases (PgL, PgL, PglH and PglI) assemble the heptasaccharide on the lipid carrier (Glover et al. 2005a; Linton et al. 2005; Glover et al. 2006). This is transported across the inner membrane and into the periplasm by the "flipase" PglK (Alaimo et al. 2006) and transferred onto protein by the OTase, PglB (Wacker et al. 2002; Glover et al. 2005b; Lizak et al. 2011). Three further proteins (PgLd, PgL and PglF) are required for biosynthesis of the reducing end sugar, a diacetamidotrixylosehexose known as di-N-acetyl bacillosamine (diNACb), from N-acetyl glucosamine (Oliver et al. 2006). The C. jejuni pgl gene locus when expressed in E. coli results in N-linked protein glycosylation (Wacker et al. 2002) and can be used to glycosylate an array of target proteins with diverse glycan (Feldman et al. 2005; Iwashkiw et al. 2012; Jervis et al. 2012). The specificity of the C. jejuni PglB for lipid-linked oligosaccharides (LLOs) with an acetamido group on the C-2 carbon of the reducing end sugar and accessibility of the target sequon on the surface of the folded target protein, are the major limitations of this approach to generating N-linked glycoproteins of choice.

Characterization of further bacterial PglBxs has led to identification of OTases from many Campylobacter species (Jervis et al. 2012; Nothaft et al. 2012), Desulfovibrio desulfuricans (Ielmini and Feldman 2011) and deep sea vent dwelling organisms Nitratireductor terarcus, Sulfitobacterium lithotrophicum and Deferrribacter desulfuricans with some displaying differing sequon recognition and glycan promiscuity (Mills et al. 2016). In addition to pglB genes, pgl gene-containing genetic loci are present in the Campylobacter species genomes sequenced to data (Nothaft and Szymanski 2010), and the structures of many of the corresponding N-linked glycans were characterized (Jervis et al. 2012; Nothaft et al. 2012). Although the genera Campylobacter and Helicobacter are closely related, orthologues of the Campylobacter pgl genes are absent in genomes of most Helicobacter species, including Helicobacter pylori. However, pgl gene orthologues are present in a single evolutionarily related group of three Helicobacter species: Helicobacter pullorum, Helicobacter canadensis and Helicobacter winghamensis (Jervis et al. 2010). In contrast to Campylobacter species, the Helicobacter pgl genes are scattered around five loci (Figure 1). A further significant deviation from the C. jejuni model is the presence in these Helicobacter species of not one but two pglB genes potentially encoding distinct putative N-linked OTases. In our previous work we have demonstrated that a H. pullorum membrane extract is capable of in vitro N-linked peptide glycosylation with a linear pentasaccharide glycan consisting of HexNAc-216-217-217-HexNAc where 216 and 217 represent the mass differences between species generated by glycan fragmentation and correspond to residues with MH+ values of 217 and 218 Daltons, respectively. Peptide N-glycosylation was H. pullorum pglB1 dependent and required an acidic residue at the −2 position of the sequon as for C. jejuni (Jervis et al. 2010). In this more in-depth study, we have demonstrated the in vivo functioning of a H. pullorum PglB1-dependent N-linked general protein glycosylation pathway.

Results

H. pullorum N-linked protein glycosylation loci

Orthologues of C. jejuni pglABCDEFHJK genes that encode the well-characterized N-linked protein glycosylation system are present

Fig. 1. Schematic representation of the fragmented H. pullorum pgl gene loci. Individual coding sequences are represented by horizontal arrows and gene designations are based on significant levels of sequence similarity to C. jejuni pgl genes (pglABCDEFHJK) or Pseudomonas aeruginosa wbp genes (wbpOS). The putative gene labeled pglIL encodes an as yet uncharacterized glycosyltransferase. The genes encoding putative N-linked OTases or PglB proteins are shaded black, those encoding putative sugar biosynthesis enzymes are unshaded, those encoding glycosyltransferases are lightly shaded, and the putative transporter or flipase encoding pglK is striped. Vertical arrowheads indicate genes that were disrupted through insertion of a kanamycin resistance cassette with the same transcriptional polarity as the mutated gene and characterized herein.
in *H. pullorum* (Figure 1) as well as the closely related species *H. canadensis* and *H. wilinghamensis* but not in other *Helicobacter* species. These *Helicobacter* species also possess a putative glycosyltransferase-encoding gene located between *pglI* and *pglJ* that is absent in *C. jejuni* and is here designated *pggL* (Figure 1). Two further genes co-located with *Helicobacter* *pgl* genes also lack *C. jejuni* orthologues. Their predicted products have significant levels of sequence similarity to *WbpOS* enzymes involved in sugar biosynthesis (King et al. 2010) and we have thus named them *wbpO* and *wbpS* (Figure 1). In contrast to *C. jejuni* and the majority of *Campylobacter* species where *pgl* genes are located in a single locus, in *Helicobacter* species these genes are present in five distinct loci. A notable feature of the *Helicobacter pgl* gene loci is the presence of two orthologues (*pglB1* and *pglB2*) of the single *Campylobacter* pgl *B* gene encoding the OTase (Figure 1). Amino acid sequence alignment of the *Helicobacter PglB* with the structurally and mechanistically characterized PglB of *Campylobacter lari* and the well-characterized *C. jejuni* PglB showed a high degree of conservation of the known essential residues for oligosaccharyltransferase activity. Catalytically active residues D454, R145, D152, D154, E316 and R372 and the 455WWD457 motif required for peptide binding (Lizal et al. 2011; Gerber et al. 2013) are all absolutely conserved in the *Helicobacter PglB1* and PglB2 enzymes (Figure S1). This strongly indicates both PglB1 and PglB2 possess oligosaccharyltransferase or related activity with potentially two distinct N-linked protein glycosylation systems operating in *H. pullorum*. To investigate N-linked protein glycosylation in this species we first determined whether both *pglB1* and *pglB2* genes were expressed. Specific intragenic primers were designed for both genes (Table S1) and RT-PCR used to detect corresponding transcripts. The generation of RT-PCR products of the predicted sizes (Figure S2) indicates both genes were transcribed during in vitro growth. To demonstrate N-linked OTase activity of *H. pullorum* PglBs we expressed both genes in the background of a *C. jejuni* *pglB* insertion knockout mutant (*pglB*::*aphA*). In this N-linked, mutant glycoproteins are not produced and we propose that the presence of lipid-linked heptasaccharide and numerous sequon-containing target proteins provides a sensitive and convenient assay for detecting related N-linked OTase activities. The *pglB* genes were recombined onto the *C. jejuni* *pglB*::*aphA* chromosome within pseudogene Cj0223 (see Methods), and complementation in this way with the *C. jejuni* *pglB* gene fully restored glycosylation as detected by reactivity of numerous proteins with the N-linked heptasaccharide specific antisera hR6 (Figure 2). Complementation with the *H. pullorum* *pglB1* gene also restored hR6 immunoreactivity though relatively few proteins were glycosylated (Figure 2). In contrast, the *pglB2* gene did not restore detectable levels of hR6 immunoreactivity (Figure 2). These data confirm that *pglB1* encodes an N-linked OTase able to transfer the *C. jejuni* heptasaccharide glycan onto protein whilst the activity of PglB2 remains elusive.

**Identification of *H. pullorum* N-linked glycoproteins**

To directly demonstrate activity of the *H. pullorum* N-linked protein glycosylation system we sought to identify corresponding N-linked glycoproteins. Initial efforts to identify a lectin that interacts with such glycoproteins were unsuccessful (data not shown). We therefore undertook an unbiased approach based on glycopeptide enrichment and site-specific identification using MS/MS. Whole cell protein lysates were digested with trypsin and glycopeptides enriched using ZIC-HILIC prior to identification by CID MS/MS to provide N-glycan structural information and HCD MSMS to identify the peptide backbone (Scott et al. 2011). We identified 62 glycopeptides modified with the HexNAc-216-217-HexNAc pentasaccharide (Figure S3, Table SIV), which represent six confirmed sites of N-glycosylation from five *H. pullorum* proteins (Table I). These proteins are predominantly of unknown function, however all are predicted periplasmic or membrane-associated proteins. Given that many *C. jejuni* N-linked glycoproteins have now been identified (Scott et al. 2011), we searched the five confirmed *H. pullorum* glycoprotein sequences against the *C. jejuni* NCTC 11168 genome. *C. jejuni* contained an orthologue for all five *H. pullorum* glycoproteins (Table I), with sequence identities of between 24.0% and 38.0% (data not shown). Examination of the literature confirmed that four of the five orthologues are known *C. jejuni* glycoproteins, with only Cj1259 as an unknown N-glycoprotein. Cj1259 is the major outer membrane protein (PorA or MOMP) in *C. jejuni* and the NCTC 11168 sequence contains no N-linked sequons. MOMP has however, recently been identified as a unique O-glycoprotein modified with a four residue glycan at a single threonine residue (Mahdavi et al. 2014). We additionally noted that *H. pullorum* glycoprotein Hp00510 is an orthologue of the *C. jejuni* N-linked glycoprotein A or CgpA (Wacker et al. 2002). We therefore named this protein HgpA (*Helicobacter* glycoprotein A). HgpA is a predicted periplasmic protein with a single N-linked glycosylation sequon of ENNDT and is annotated as HPMG_01281 in the *H. pullorum* MIT 98-548970 genome sequence. To further investigate N-linked glycosylation of HgpA, the corresponding gene was cloned and expressed in *E. coli* from plasmid pQElhgpA (Table SII and Methods) along with the *C. jejuni pgl* locus on a second plasmid (Wacker et al. 2002). Western blotting with *C. jejuni* N-linked heptasaccharide glycan-specific antisera hR6 demonstrated that HgpA was glycosylated in a *C. jejuni* PglB-dependent manner (Figure 3). The HgpA protein was similarly glycosylated in *E. coli* by *H. pullorum* PglB1, but not PglB2 (Figure 3). Glycosylation of HgpA by both *C. jejuni* PglB and *H. pullorum* PglB1 was dependent on presence of...
Table I. Identification of glycopeptides from *H. pullorum* using ZIC-HILIC and LC-MS/MS

<table>
<thead>
<tr>
<th>Gene number (Gp. No.)</th>
<th>Protein identification</th>
<th>Protein mass (Da)</th>
<th>Precursor mass (Charge)</th>
<th>MASCOT score</th>
<th>Peptide sequencea</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hp00296c (Cj0114)</td>
<td>Tetraicosopeptide Repeat Protein</td>
<td>35,661</td>
<td>1888.8647 (3+)</td>
<td>69.87</td>
<td>KDTIKEDSVENN6GSPNA177</td>
</tr>
<tr>
<td>Hp00314c (Cj1565c)</td>
<td>Flagellar Functional Protein</td>
<td>93,678</td>
<td>2371.1319 (3+)</td>
<td>27.25</td>
<td>PNNANANIATIESAENNQKESKQ196c</td>
</tr>
<tr>
<td>Hp00510c (Cj670c)</td>
<td>Hypothetical Protein (HgpA)</td>
<td>28,021</td>
<td>1896.9534 (3+)</td>
<td>56.58</td>
<td>HAHYYQMLLQNPKDEAEKKEIQALDDTLNYEDDA465Tk</td>
</tr>
<tr>
<td>Hp00561c (Cj633c)</td>
<td>Hypothetical Protein</td>
<td>38,588</td>
<td>4941.5575 (6+)</td>
<td>37.78</td>
<td>TQQSTIPQITPPTISQESKPTKPTQIPS KPQ4CQK111</td>
</tr>
<tr>
<td>Hp01062c (Cj1259)</td>
<td>Putative Membrane Protein</td>
<td>56,457</td>
<td>1447.6162 (3+)</td>
<td>63.45</td>
<td>DRDVDNGDYNQS4k</td>
</tr>
</tbody>
</table>

aMASCOT score represents the best score provided for the identified glycosylation sequon.
bSequon is shown in italics, with glycosylation site in bold and underlined.

Full sequon was not identified and the underlined Asn is predicted only.

**Identification of *H. pullorum* genes involved in HgpA N-linked glycosylation**

To identify genes involved in HgpA N-linked glycosylation, we constructed insertional knockout mutants in seven *H. pullorum* pgl/wbp genes from five loci (Figure 1). These mutants were constructed in the *H. pullorum* Hp47 genetic background that produces HgpAhis10 (see above and Table III) and the relative electrophoretic mobility of the N-linked glycoprotein investigated by SDS-PAGE of whole-cell lysates followed by Western blotting with an anti-His antiserum. HgpAhis10 electrophoretic mobility increased when derived from a pglB1 insertional knockout mutant compared to that derived from the wild-type strain indicating modification via a pglB1 dependent pathway (Figure 4). Similar analyses of *H. pullorum* strains with disrupted glycosyltransferase-encoding pgc, pgla, pglh and pglj genes resulted in increased but varying mobility of the corresponding HgpAhis10 proteins in SDS-PAGE (Figure 4) indicating their involvement in assembly of the N-linked pentasaccharide. We also investigated wbpO and wbpS (Figure 1), products of which have significant levels of amino acid sequence identity (32% and 63% respectively) to *Pseudomonas aeruginosa* WbpO and WbpS involved in biosynthesis of 2-acetamido-2-deoxy-β-D-galacturonic acid (GalNAcAN). In *P. aeruginosa*, WbpO converts UDP-GalNAc to UDP-2-acetamido-2-deoxy-β-D-glucaric acid (UDP-GlcNAc), and the activated form of this sugar is then converted to UDP-GalNAc by the isomerase WbpP (King et al. 2010). It was further proposed that WbpS amidotransferase activity is responsible for production of UDP-GalNAcAN (King et al. 2010). The role of *P. aeruginosa* wbpOS gene products in biosynthesis of these sugars indicated that the *H. pullorum* orthologues may be involved in biosynthesis of similar sugars and this is consistent with presence of HexNACAN/HexNAcA in the *H. pullorum* N-linked pentasaccharide. Indeed insertional knockout mutagenesis of wbpO and wbpS resulted in increased mobility of the HgpA protein (Figure 4).

To further investigate HgpA produced in these pgl/wbp genetic backgrounds, a chromosomal hgpA his10 gene was introduced into these backgrounds (see Methods). The HgpAhis10 protein was purified and intact mass values were determined by electrospray ionization MS (Table II). The predicted mass of unmodified HgpAhis10 is 26,071 Da and the observed masses of HgpAhis10 proteins derived from the wild type and pglB1:aphA mutant were 27,127 and 26,070 Da, respectively. This difference in electrophoretic mobility and mass indicates pglB1-dependent modification of HgpAhis10 with the asparagine (underlined) within the ENNDT N-linked glycosylation sequon identified above (Figure 3).

Further characterization of HgpA N-linked glycosylation

A C-terminal deca-histidine-tagged version of hgpA (hgpAhis) was recombined onto the *H. pullorum* NCTC 12824 chromosome via plasmid pHPC2hpAhis (see Methods) to produce strain Hp47 (Table SIII). In order to verify that HgpAhis10 was glycosylated, it was purified by nickel affinity chromatography (see Methods). ZIC-HILIC enrichment and MS/MS of tryptic peptides from HgpAhis10 identified a glycopeptide with mass (MH+) of 4208.96 Da. CID MS/MS confirmed the presence of the pentasaccharide glycan (Figure S4a) while HCD MS/MS identified the peptide containing the previously identified sequon ENNDT (Figure S4b). HCD MS/MS also shows intense singly charged oxonium ions at 204.08 (HexNacA), 217.08 and 218.07 m/z. Elemental composition analysis suggests peaks with monoisotopic masses of 217.08190 and 218.06591 likely correspond to C6H11N2O3 and C7H12N2O4 respectively (Figure S4c). These compositions are consistent with those determined for the uronamide HexNAcAN and the urionate HexNAcA. Overall these data suggest a *H. pullorum* N-linked pentasaccharide with the structure HexNac-HexNAcAN-HexNAcA-HexNAcA-HexNAcA.
consistent values in intact mass analysis experiments likely due to demonstrating intermediate mobility of HgpAhis from different mutants. Following SDS-PAGE of whole-cell lysates and western blotting, HgpAhis was detected with anti-his antibody. A single band of increased mobility (relative to that produced in the wild type strain) but varying mobility was observed in individual mutants. For comparison, extracts from wild type and pglB1::aphA genetic backgrounds were mixed (labeled pglB1 + wt) to demonstrate intermediate mobility of HgpAhis from different mutants.

Table II. Intact mass values of HgpA determined by LC-ESI–MS in Helicobacter pylori pgl and wbp gene insertion knockout mutants

<table>
<thead>
<tr>
<th>H. pylori genetic background</th>
<th>HgpAhis intact mass (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCTC 12824 23S::eryhgpAhis6 (Hp31)</td>
<td>27,127</td>
</tr>
<tr>
<td>NCTC 12824 23S::eryhgpAhis6 pglB1::aphA (Hp67)</td>
<td>26,070</td>
</tr>
<tr>
<td>NCTC 12824 23S::eryhgpAhis6 pglA::aphA (Hp25)</td>
<td>26,297</td>
</tr>
<tr>
<td>NCTC 12824 23S::eryhgpAhis6 pglC::aphA (Hp26)</td>
<td>26,069</td>
</tr>
<tr>
<td>NCTC 12824 23S::eryhgpAhis6 pglH::aphA (Hp27)</td>
<td>26,923</td>
</tr>
<tr>
<td>NCTC 12824 23S::eryhgpAhis6 pglJ::aphA (Hp29)</td>
<td>26,488</td>
</tr>
<tr>
<td>NCTC 12824 23S::eryhgpAhis6 wbpS::aphA (Hp68)</td>
<td>26,490</td>
</tr>
</tbody>
</table>

| NCTC 12824 23S::eryhgpAhis6 (Hp31) | 27,127 |
| NCTC 12824 23S::eryhgpAhis6 pglB1::aphA (Hp67) | 26,070 |
| NCTC 12824 23S::eryhgpAhis6 pglA::aphA (Hp25) | 26,297 |
| NCTC 12824 23S::eryhgpAhis6 pglC::aphA (Hp26) | 26,069 |
| NCTC 12824 23S::eryhgpAhis6 pglH::aphA (Hp27) | 26,923 |
| NCTC 12824 23S::eryhgpAhis6 pglJ::aphA (Hp29) | 26,488 |
| NCTC 12824 23S::eryhgpAhis6 wbpS::aphA (Hp68) | 26,490 |

Table II. Intact mass values of HgpA determined by LC-ESI–MS in Helicobacter pylori pgl and wbp gene insertion knockout mutants

<table>
<thead>
<tr>
<th>H. pylori genetic background</th>
<th>HgpAhis intact mass (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCTC 12824 23S::eryhgpAhis6 (Hp31)</td>
<td>27,127</td>
</tr>
<tr>
<td>NCTC 12824 23S::eryhgpAhis6 pglB1::aphA (Hp67)</td>
<td>26,070</td>
</tr>
<tr>
<td>NCTC 12824 23S::eryhgpAhis6 pglA::aphA (Hp25)</td>
<td>26,297</td>
</tr>
<tr>
<td>NCTC 12824 23S::eryhgpAhis6 pglC::aphA (Hp26)</td>
<td>26,069</td>
</tr>
<tr>
<td>NCTC 12824 23S::eryhgpAhis6 pglH::aphA (Hp27)</td>
<td>26,923</td>
</tr>
<tr>
<td>NCTC 12824 23S::eryhgpAhis6 pglJ::aphA (Hp29)</td>
<td>26,488</td>
</tr>
<tr>
<td>NCTC 12824 23S::eryhgpAhis6 wbpS::aphA (Hp68)</td>
<td>26,490</td>
</tr>
</tbody>
</table>

To identify further N-linked glycoproteins, a polyclonal antiserum was raised against purified HgpAhis10 from H. pylori (see Methods). When a wild type H. pylori whole-cell lysate was probed with this antiserum a number of immunoreactive bands were observed (Figure 6, arrowheads). Major bands at 35–40 kDa are consistent with the predicted masses of proteins Hp00296c and Hp00561c identified by MS/MS (Table I), with additional minor bands above 50 kDa and approximately 90 kDa consistent with the identification of Hp01062c and Hp00314c, respectively. The majority of these bands (Figure 6; shaded arrowheads) were no longer detected in a pglB1 insertion knockout mutant indicating these represented glycoproteins glycosylated in a pglB1-dependent manner. Their immunoreactivity was also dependent on pglABC and wbpOS genes demonstrating their role in this general protein glycosylation pathway. In pglJ and particularly pglH knockout backgrounds, bands retained some immunoreactivity suggesting that the antiserum recognizes the shortened glycans likely present on proteins in these backgrounds. Immune-reactivity of an approximately 28 kDa band was unaffected by pgl and wbp gene mutations, however the mobility of this band was increased in these mutants (Figure 6). This band likely represents HgpA detected by antibodies against both the N-linked glycan and the protein itself and this is

Structural characterization of N-linked glycans produced in mutant backgrounds

To further characterize the variety of HgpA modifications produced in pgl/wbp mutant backgrounds, purified HgpAhis9 proteins were digested with trypsin and peptides analyzed by MALDI-TOF MS. A number of peaks were observed corresponding to the predicted masses of HgpAhis-derived tryptic peptides confirming protein identity (data not shown). However, a peak at m/z 2429 was present in the spectrum for HgpAhis derived from the wild type H. pylori but not from the pglB1 knockout. This m/z value corresponds to the mass of the predicted tryptic peptide containing the N-linked glycosylation sequon (EIPTPNENNDTK – 1371.4 Da) plus the mass of the H. pylori pentasaccharide N-linked glycan (1056 Da). This peak was selected for fragmentation by MALDI-LIFT-TOF/TOF MS and the resultant spectrum (Figure 5) is consistent with the previously observed HexNAc-216-217-217-HexNAc pentasaccharide observed on an in vitro generated glycopeptide (Jervis et al. 2010) and present on glycoproteins identified above. Insertional knockout mutagenesis of the H. pylori pglH gene results in production of an N-linked glycopeptide with a tetrasaccharide structure, lacking the non-reducing end residue (Figure 5). Similarly, inactivation of the pglJ gene resulted in production of an N-linked disaccharide glycan consisting of the first two sugars of the pentasaccharide (Figure 5). Inactivation of the pglA gene resulted in production of a monosaccharide N-linked glycan, although the spectrum produced indicated presence of a diNACAc instead of a HexNAc residue (Figure 5). In a pglC insertional knockout mutant there was no detectable N-linked glycan suggesting its role as the initiating transferase as seen in C. jejuni (Linton et al. 2005; Glover et al. 2006). These data indicate that biosynthesis of the H. pylori N-linked pentasaccharide glycan involves sequential action of the PglC, PglA, PglJ and PglH glycosyltransferases and are entirely consistent with the intact mass analysis for each of the mutants described above. The MALDI-LIFT-TOF/TOF MS spectra generated from HgpAhis produced in both the wbpO and wbpS mutants indicated presence of a disaccharide N-linked glycan. In the wbpO mutant the disaccharide consisted of a diNACAc reducing end sugar and a HexNAc residue, whilst in the wbpS mutant the disaccharide was composed of a reducing end HexNAc and a HexNAcAN residue, again consistent with intact mass data presented above (Figure 5).
consistent with a predicted size of 26 kDa for glycosylated HgpA. The increased HgpA mobility in these backgrounds is presumably due to reduction in size of glycan structures, an interpretation consistent with data obtained above. These and previous data demonstrate the functioning of an *H. pullorum* PglB1-dependent general protein glycosylation system.

**Discussion**

It is well established that *C. jejuni* encodes a general N-linked protein glycosylation pathway and similar N-linked glycan structures have been identified in other *Campylobacter* species (Jervis et al. 2012; Nothhaft et al. 2012). Less well characterized are the putative N-linked protein glycosylation systems present in a small number of species.
from the related Helicobacter genus. The Helicobacter N-linked protein glycosylation systems are notable for the presence of two distinct OTase encoding pglB genes (Figure 1), with the H. pullorum PglB1 protein more similar to the C. jejuni PglB (31% amino acid sequence identity) than to H. pullorum PglB2 (23% identity). Through a variety of approaches in E. coli (Jervis et al. 2010), C. jejuni and H. pullorum we have demonstrated N-linked oligosaccharyltransferase activity of H. pullorum PglB1 but not PglB2. Furthermore evidence was provided for a H. pullorum PglB1 directed general protein glycosylation pathway (Figure 6). MS/MS analysis of enriched glycopeptides confirmed the H. pullorum N-glycan structure and identified six sites of glycosylation within five proteins. Intriguingly, only five of six glycopeptides contained the anticipated N-linked bacterial sequon D/E-X-N-X-S/T, with a second glycopeptide in Hp00296c containing a nonclassical sequon with a lysine at the +2 position (ENNQK; Table I). Recent work demonstrated that C. jejuni PglB is able to glycosylate nonclassical sequons with three examples identified in the NCTC 11168O strain (Scott et al. 2014). These included sequons lacking the D/E at position −2, or S/T at position +2, but never both. A similar phenomenon has been described for Campylobacter lari PglB (Schwarz et al. 2011b). For Hp00296c, we noted that the occupied nonclassical sequon is close to the identified “classical” sequon (N171 versus N190; Table I), suggesting that hierarchical site occupancy as observed for the C. jejuni lipoprotein JlpA (Scott et al. 2009) may occur, however this is yet to be determined. Knockout of several genes from the H. pullorum pgl gene loci that encode putative glycosyltransferases and enzymes involved in sugar biosynthesis altered the structure of the PglB1-dependent N-linked glycan (Figures 4 and 5). Based on these data we propose a model for H. pullorum N-linked protein glycosylation (Figure 7). The proposed activities of H. pullorum PglCAJH are broadly consistent with those of their similarly named counterparts in C. jejuni (Glover et al. 2005a; Linton et al. 2005). The H. pullorum wbpS mutant produced an HgpA-linked disaccharide of HexNAc–HexNAcA (Figure 5). Again this is consistent with the predicted role for WbpS in the biosynthesis of UDP-HexNAcAN from UDP-HexNAcA in P. aeruginosa (King et al. 2010). Thus if the H. pullorum wbpS knockout mutant is

Fig. 6. Immunoreactivity of whole-cell lysates from H. pullorum pgl and wbp gene mutants. Whole cell lysates from H. pullorum NCTC 12824 and corresponding pglBCAJH and wbpOS insertional knockout mutants were separated by SDS-PAGE, Western blotted and probed with a polyclonal antibody raised against purified HgpA glycoprotein. Shaded arrowheads indicate position of bands corresponding to putative glycoproteins whilst the unshaded arrowhead indicates the band corresponding to HgpA itself.

Fig. 7. Model of H. pullorum N-linked protein glycosylation pathways. Direct evidence for role of PglCAJHB1 and WbpOS is provided in this study. The proposed roles of PglDEFK and PglB2 are based on established function of C. jejuni orthologues. The H. pullorum PglB1-dependent N-linked pentasaccharide glycan is assembled through sequential action of glycosyltransferases PglCAJH, flipped into the periplasm by PglK and transferred to proteins as indicated in the periplasm by PglB1. A second proposed glycan, with a reducing end 228 kDa residue synthesized by PglDEF activity, is similarly assembled and flipped into the periplasm where it is transferred to as yet unidentified protein(s) via PglB2 activity.
unable to synthesize HexNACAN, the structurally related and biosynthetic precursor HexNACA is transferred to the reducing end HexNAC in its place by the PglJ transferase presumably unable to further extend the glycans structure. In the background of pglAlabpA and wbpO:abpA mutants, HgpA was modified with glycans including a reducing end diNacBac residue absent in the pentasaccharide N-glycan produced in the wild type background (Figure 5), but consistent with the presence in H. pullorum of orthologs of the C. jejuni pglDEF genes encoding enzymes required for diNacBac biosynthesis (Figure 1). Surprisingly the H. pullorum pglDEF genes are not required for biosynthesis of the PglB1-dependent H. pullorum N-linked pentasaccharide (Jervis et al. 2010) and we hypothesize that they are involved in biosynthesis of an as yet uncharacterized glycans potentially transferred by the H. pullorum PglB2 dependent pathway. Disruption of the PglB1 pathway at the earlier stages of pentasaccharide assembly may result in cross-talk between the two systems. Due to its similarity to P. aeruginosa WbpO it is plausible that H. pullorum WbpO converts UDP-HexNAc to UDP-HexNACA. Indeed disruption of wbpO resulted in substitution of the HexNACAN of the N-linked glycans with a HexNAA residue albeit with a diNacBac reducing end. Despite this, we found no evidence of enriched glycopeptides with Asn-linked glycans containing a diNacBac reducing end sugar in the H. pullorum wild-type strain. Glycopeptide data analysis was performed by manual interpretation of all MS/MS scans containing the HexNA oxonium ion 204.086 m/z, which would identify glycopeptides with truncated HexNA-containing glycans or those containing diNacBac–HexNAC disaccharides. Searches were also performed to extract MS/ MS scans containing diNacBac-associated ions, however bacillosamine does not produce intense oxonium ions during vibrational fragmentation approaches. Database searches using diNacBac, diNacBac + HexNAC and other combinations of possible glycans masses also failed to identify additional glycopeptides (data not shown). This suggests that, in accordance with the production of diNacBac glycopeptides in pglAlabpO mutants, such glycans are only produced under specific conditions that influence pglAlabpO expression and these are yet to be determined. We have shown the involvement of a number of pgl genes in the HpglB1-dependent glycosylation system but further genes associated with these pgl loci remain uncharacterized. These include pgll and pgll encoding putative glycosyltransferases, which we were unable to disrupt and the putative diNacBac synthesis genes pglDEF that are not involved in the HpglB1-dependent system (Jervis et al. 2010). This suggests the N-linked protein glycosylation system in H. pullorum may be more complex than in the C. jejuni model. The parasites Leishmania major and Trypanosoma brucei possess four and three single subunit N-linked OTases, respectively, with different glycan and acceptor protein specificities and varying growth phase dependent expression patterns resulting in production of distinct subsets of the N-linked glycoproteome (Nasab et al. 2008; Izquierdo et al. 2009). Archaeal species with multiple OTases have not yet been studied experimentally though Haloferax volcanii has been shown to produce two structurally distinct S-layer protein N-linked glycans (Kaminski et al. 2013a) and Archaeoglobus fulgidus encodes three OTases and produces two structurally distinct LLOs (Taguchi et al. 2016). Genetic analysis shows that most OTase gene duplication events are relatively ancient (Kaminski et al. 2013b). Our combined data enable proposal of a model for N-linked protein glycosylation in H. pullorum (Figure 7) with experimental evidence for biosynthesis of pentasaccharide transferred onto proteins in the periplasm by PglB1 and a more speculative proposal for PglB2 function that will require experimental verification.

Materials and methods

Bacterial strains

All E. coli strains were grown in Luria–Bertani (LB) broth or on LB agar plates. C. jejuni NCTC 11168 and H. pullorum NCTC 12824 strains were grown on Columbia agar containing 5% defibrinated horse blood (TCS Biosciences) at 42°C in a modified atmosphere (85% N2, 10% CO2 and 5% O2) generated with a VA500 workstation (Don Whitley Ltd.). Antibiotics were used at the following concentrations: kanamycin 50 μg/mL, chloramphenicol 17 μg/mL, tetracycline 10 μg/mL, ampicillin 100 μg/mL and erythromycin 300 μg/mL. Primers, plasmids and strains are described in Tables S1, S2 and S3, respectively.

Reverse transcriptase PCR

Total RNA was extracted from H. pullorum harvested from 48 h blood agar plates using the Qiagen RNeasy kit with an additional in-solution DNase I digestion step. Reverse transcriptase PCR was performed using the Qiagen OneStep RT PCR kit according to manufacturer’s instructions.

Integration of pglB genes onto the C. jejuni chromosome

Various pglB1 genes were integrated onto the C. jejuni 11168 pglB:: aphA chromosome using a modification of a previous method (Gerber et al. 2013). A chlamydomycin resistance cassette (van Vliet et al. 1998) was PCR amplified using primers Cm-F (restriction sites SpeI and XhoI) and Cm-R (NcoI, Nhel and SpeI), digested with SpeI and cloned into the SpeI site of a previously constructed vector consisting of pUC18 backbone with a region of the C. jejuni 11168 pseudogene Cj0223 cloned into the Smal site (Hitchen et al. 2010). The resultant plasmid was named pCJC1.

Complete C. jejuni pglB, H. pullorum pglB1 and H. pullorum pglB2 genes were PCR amplified with primer pairs PglB1cmp-F/PglB1cmp-R, HppglB1cmp-F/HppglB1cmp-R and HppglB2cmp-F/HppglB2cmp-R, respectively, to include approximately 70 bp upstream of the first start codon. Primers PglB1cmp-F, HppglB1cmp-F and HppglB2cmp-F and HppglB2cmp-R site cloned into Ncol site at the 5’ end and the corresponding reverse primers encoded a deca-histidine tag at the 3’ end along with either an Nhel (PglB1cmp-R) or SpeI (HppglB1cmp-R and HppglB2cmp-R) site. Products were digested with Ncol and either Nhel or SpeI as appropriate and cloned immediately downstream of, and in the same transcriptional orientation as, the chlamydomycin resistance cassette of pCJC1. The resulting plasmids pCJC1pglBCj, pCJC1pglHp1 and pCJC1HppgB2 (Table SII) were electroporated into C. jejuni 11168 pglB::aphA cells, chlamydomycin-resistant colonies selected and the anticipated double crossover integration events verified by PCR.

Expression of hgpA in E. coli

The predicted H. pullorum NCTC 12824 hgpA coding sequence was PCR amplified with primers hgpA-F and hgpA-R to introduce SplI and BglIII restriction sites at the 5’ and 3’ end, respectively. Following digestion with SplI and BglIII, PCR products were ligated into similarly digested vector pQE70 to generate pQEhgpA encoding a C-terminal hexa-his tagged protein. A variant, pQEhgpAN46Q, was created by site directed mutagenesis using primers hgpAN46SDM-F and hgpAN46SDM-R. Plasmids pQEhgpA and pQEhgpAN46Q were transformed into E. coli Novablide (Strategene) cells harboring
pACYCpGlb::aphA (24) along with plasmids expressing either C. jejuni pgB (pMAF10), H. pullorum pgBl (pMLHp1) or H. pullorum pgB2 (pMLHp2) (Jervis et al. 2010). Transformants were grown overnight at 37°C, diluted 1 in 100 in LB broth supplemented with appropriate antibiotics, grown to an optical density at 600 nm (A600) of 0.6 and expression of hgpAhis and pgBl genes induced with 1 mM IPTG and 0.2% (w/v) arabinose. Following overnight incubation at 37°C cultures were harvested, normalized by A600, resuspended in SDS-PAGE loading buffer and incubated at 95°C for 5 min. Whole-cell lysates were separated by SDS-PAGE, transferred to nitrocellulose membrane, and probed with hR6 antiserum (gift from Markus Aebi) followed by an IRDye 800CW goat anti-rabbit IgG secondary antibody (LI-COR) and mouse anti-penta-His (Qiagen) followed by goat anti-mouse IgG secondary antibody (LI-COR). Blots were imaged using a LI-COR Odyssey Infrared Imaging System.

Introduction of oligo-histidine tagged hgpA onto the H. pullorum chromosome
A 1.8 kb internal fragment of the H. pullorum 23S rRNA gene (rrl) was amplified using primers Hp23s-F and Hp23s-R, ligated into pGEM T-easy to create plasmid pHPG and digested with HindIII to excise 276 bp of rrl. An erythromycin resistance cassette (ermC) lacking a transcriptional terminator was amplified using primers Ery-F (containing a HindIII site) and Ery-R (HindIII and BamHI sites), and ligated into HindIII digested pHPC in the same transcriptional orientation as flanking rrl fragments to create (pHPG1). A derivative of this plasmid, termed pHPC2, was constructed by cloning the promoter region of the C. jejuni porA gene into the pHPC BamHI site located immediately downstream of the erythromycin resistance cassette with the promoter in the same transcriptional orientation as rrl gene and erythromycin cassette. The promoter region was PCR amplified using primers porAP-F (containing XhoI site) and porAP-R (Ndel and XhoI sites). Two H. pullorum hgpA expression systems were constructed based on either pHPC1 or pHPC2. In the first the hgpA ORF was PCR amplified using primers hgpABamHI-F and hgpABamHI-R to incorporate a C-terminal hexa-his tag and ligated into the BamHI site at the 3’ end of the erythromycin resistance cassette in pHPC1 to create pHPC1hgpAhis. In the second, the hgpA ORF was PCR amplified with primers hgpANdel-F and hgpANdel-R to include a C-terminal deca-his tag and cloned into the Ndel site of the plasmid pHPC2 downstream of the porA promoter region to create plasmid pHPC2hgpAhis.

Plasmids pHPC1hgpAhis and pHPC2hgpAhis were electroporated into H. pullorum cells (van Vliet et al. 1998; Jervis et al. 2010) and erythromycin resistant colonies screened by PCR for the predicted double crossover events within the chromosomal rrl gene.

Purification of HgpAhis from H. pullorum
Approximately 1 g of H. pullorum NCTC 12824 rrl:ermC hgpAhis (Hp31) cells were resuspended in 3 mL of Binding Buffer (50 mM Na2HPO4, 300 mM NaCl, 30 mM imidazole, pH 8.0) containing protease inhibitors phenylmethylsulfonylfluoride (0.1 mM) and benzamide (1 mM). Cells were lysed in a French press (Thermo Scientific, UK), centrifuged at 8000 x g for 20 min and the supernatant incubated with 150 µL of Ni-NTA Magnetic Agarose Beads (Qiagen, UK) for 1 h at room temperature with mixing. Beads were washed three times with 1 mL of Binding Buffer and bound protein eluted in 50 µL of 50 mM Na2HPO4, 300 mM NaCl, 500 mM imidazole, pH 8.0 at room temperature for 5 min.

Intact mass analysis by ESI–MS
Purified proteins were dialyzed against 25 mM Tris-HCl, 25 mM NaCl (pH 8.0) and analyzed by LC-ESI–MS using a Dionex PepSwift RP column (200 µm x 50 mm) connected to a Micromass LCT ESI-MS. Spectra were deconvoluted using the MaxEnt 1 software (Micromass).

Glycan analysis by MALDI-MS
Coomassie stained SDS-PAGE bands were excised, lyophilized and digested with trypsin (E.C.3.4.21.4, Promega) overnight. Peptides were extracted from gel pieces using a C18 ZipTip (Millipore, UK) according to the manufacturer’s protocol and eluted in 10 µL of 50% acetonitrile, 0.1% formic acid. MALDI-TOF MS and MALDI-LIFT-TOF/TOF MS spectra were acquired by laser-induced dissociation (LID) using a Bruker UltrafleXX II mass spectrometer in the positive-ion reflection mode with a matrix of 20 mg/mL 2,5-dihydroxybenzoic acid (DHB) (30% acetonitrile, 0.1% TFA). Data were analysed with FlexAnalysis 3.0 software (Bruker Daltonics).

Enrichment of glycopeptides using zwitterionic-hydrophilic interaction liquid chromatography (ZIC-HILIC) and identification by reversed phase LC–MS/MS
Identification of glycopeptides from H. pullorum NCTC 12824 WT was conducted as previously described (17, 33). Lysates were suspended in 6 M urea, 2 M thiourea, 40 mM NH4HCO3, and reduced and alkylated with 20 mM dithiothreitol and 40 mM iodoacetamide, respectively, each for 1 h at room temperature. Samples were diluted 1:10 with 40 mM NH4HCO3 and digested with porcine sequencing grade trypsin (Promega, Madison WI; 1:100) overnight at 37°C. Peptides were acidified with 2% (v/v) formic acid and 0.1% (v/v) TFA, then desalted by hydrophilic lipophilic-balance solid phase extraction (HLB-SPE) (Waters, Milford MA). ZIC-HILIC enrichments were carried out according to Scott et al. (2011). Fractions were resuspended in 0.1% formic acid and loaded directly onto a 20 cm, 75 µm inner diameter, 360 µm outer diameter Repulsor Gold C18 AQ 1.9 µm (Dr. Maisch, Ammerbuch-Entringen, Germany) reversed phase (RP) column using a trappel EASY-nLC II system (Proxence, Odense Denmark) coupled to an LTQ-Orbitrap Velos Pro mass spectrometer (Thermo Scientific, San Jose, CA). Peptides were loaded in 95% buffer A (0.1% FA) and eluted at 250 nL/min using a linear gradient of buffer B (80% ACN, 0.1% FA) from 5% to 40% over 120 min. The column was washed with 90% buffer B for 10 min before being returned to 95% buffer A. The LTQ-Orbitrap Velos Pro was operated using Xcalibur v2.2 (Thermo Scientific) with a capillary temperature of 200°C in a data-dependent mode automatically switching between MS and higher energy collisional dissociation (HCD)/collision-induced dissociation (CID) MS/MS. For each MS scan, the three most abundant precursor ions were selected for HCD (normalized collision energy 45) and CID (normalized collision energy 35). Data processing was carried out as previously described (Scott et al. 2014). Briefly, HCD scans from .raw files were processed in Proteome Discoverer v1.4.1.14 (Thermo Scientific) and searched using SEQUEST against an in-house, translated H. pullorum NCTC 12824 database. MS/MS scans that did not result in identifications were exported as .mgf files. The “mgf graph” feature within the MSMS module of GPMAW 10.0 (Lighthouse Data, Odense, Denmark) was used to highlight all scan events containing the diagnostic HexNAc oxonium ion 204.086 mz, in addition to the oxonium ions for the HexNAcA (218.0665 mz).
and HexNAcAN (217.0824 m/z) species that are constituents of the *H. pullorum* N-glycan. MASCOT v2.2 searches were conducted against the *H. pullorum* NCTC 12824 database with parent ion mass accuracy of 20 ppm and product ion accuracy of 0.02 Da, no protease specificity, instrument set to MALDI-QIT-TOF, as well as the fixed modification carbamidomethyl (C) and variable modifications oxidation (M) and deamidation (N). All spectra were searched with the decoy option enabled, and no matches were detected (FDR 0%). HCD and CID scans from matched spectra (MASCOT scores >20) were manually inspected to ensure all major peaks were matched, and to validate attachment and composition of the N-glycan. Isotopic distribution analysis was performed with the MS Isotope module of Protein Prospector (http://prospector.ucsf.edu/prospector).

Construction of *H. pullorum* insertional knockout mutants

*H. pullorum* genes were inactivated by insertion of the *aphA* gene via double crossover recombination events with appropriately constructed suicide vectors introduced into cells by electroporation (van Vliet et al. 1998). To create suicide vectors, PCR products of approximately 2 kb were generated that incorporated regions of target genes. These were ligated into pGEM-T Easy, and the *aphA* cassette lacking a transcriptional terminator cloned into BamHI or HindIII sites within the central region of cloned PCR products. If cassettes were used to validate attachment and composition of the N-glycan.

**pglA:** Primers pglA-F and pglA-R were used to amplify a 1.6 kb fragment encompassing the complete pglA gene with a naturally occurring central HindIII restriction site.

**pglC:** A 729 bp fragment consisting of the first 111 bp at the 5′ end of pglC plus upstream region was amplified using primers pglC-UF and pglC-UR to include a 3′ HindIII site. A 689 bp fragment consisting of 340 bp of the 3′ end of pglC and downstream region was amplified using primers pglC-DF and pglC-DR to include a 5′ HindIII site. Overlap PCR was performed using primers pglC-UF and pglC-DR.

**pglH:** A 912 bp fragment consisting of the first 561 bp at the 5′ end of pglH plus upstream region was amplified using primers pglH-UF and pglH-UR to include a 3′ HindIII site. A 929 bp fragment consisting of 527 bp of the 3′ end of pglH and downstream region was amplified using primers pglH-DF and pglH-DR to include a 5′ HindIII site. Overlap PCR was performed using primers pglH-UF and pglH-DR.

**pglI:** An 823 bp fragment consisting of the first 236 bp at the 5′ end of pglI plus upstream region was amplified using primers pglI-UF and pglI-UR to include a 3′ HindIII site. An 857 bp fragment consisting of 651 bp of the 3′ end of pglI and downstream region was amplified using primers pglI-DF and pglI-DR to include a 5′ HindIII site. Overlap PCR was performed using primers pglI-UF and pglI-DR.

**pglJ:** An 879 bp fragment consisting of the first 204 bp at the 5′ end of pglJ plus upstream region was amplified using primers pglJ-UF and pglJ-UR to include a 3′ HindIII site. An 808 bp fragment consisting of 196 bp of the 3′ end of pglJ and downstream region was amplified using primers pglJ-DF and pglJ-DR to include a 5′ HindIII site. Overlap PCR was performed using primers pglJ-UF and pglJ-DR.

**pglL:** An 857 bp fragment consisting of the first 180 bp at the 5′ end of pglL plus upstream region was amplified using primers pglL-UF and pglL-UR to include a 3′ HindIII site. An 885 bp fragment at the 3′ end of pglL was amplified using primers pglLD-F and pglLD-R to include a 5′ HindIII site. Overlap PCR was performed using primers pglLU-F and pglLD-R.

**wbpO:** Primers wbpO-F and wbpO-R were used to amplify a 1.9 kb fragment including the complete wbpO CDS with a naturally occurring central HindIII restriction site.

**wbpS:** Primers wbpS-F and wbpS-R were used to amplify a 1.2 kb internal fragment of wbpS. A BamHI restriction site was created using site directed mutagenesis with primers wbpSBamHI-F and wbpSBamHI-R.

Supplementary data

Supplementary data is available at GLYCOBIOLOGY online.

Funding

This work was supported by the UK Biotechnology and Biological Science Research Council [Grant numbers BB/F009321/1, BB/H017542/1]; the Wellcome Trust [Grant 102979/13/Z]; and the Australian Research Council [DP110103573 to S.J.C.]. A.W. was supported by a BBSCDTP studentship. J.A.C. was supported by an Australian Postgraduate Award.

Acknowledgements

The authors thank Dr. David Knight and staff of the Biomolecular Analysis core research facility, Faculty of Biology, Medicine and Health, University of Manchester for assistance with mass spectrometry. Glycopeptide analysis on the LTQ-Orbitrap Velos Pro was facilitated by access to the Mass Spectrometry Core Facility (MSCF) at the University of Sydney. The anti-glycan bR6 antiserum was kindly provided by Professor Markus Aebi and the erythromycin resistance cassette by Dr. Arnoud van Vliet.

Conflict of interest statement

None declared.

References


