Glycobiology, 2017, vol. 27, no. 9, 888–899
doi: 10.1093/glycob/cwx032
Advance Access Publication Date: 23 May 2017
Original Article

OXFORD

Microbial Biology

Structural and genetic analyses of glycan *O*-acetylation in a bacterial protein glycosylation system: evidence for differential effects on glycan chain length

Jan Haug Anonsen[2,3,4,7], Bente Børud[5,7], Åshild Vik[2], Raimonda Viburiene[2], and Michael Koomey[1,2,4,6]

[2]Department of Biosciences, [3]IBV Mass Spectrometry and Proteomics Unit, and [4]Center for Integrative Microbial Evolution (CIME), University of Oslo, 0371 Oslo, Norway, [5]Department of Molecular Biology, Norwegian Institute of Public Health, 0403 Oslo, Norway, and [6]Centre for Ecological and Evolutionary Synthesis, University of Oslo, 0371 Oslo, Norway

[1]To whom correspondence should be addressed: Tel: +47-22854091; Fax: +47-22854726; e-mail: j.m.koomey@ibv.uio.no

[7]These authors contributed equally.

Received 28 February 2017; Revised 8 April 2017; Editorial decision 10 April 2017; Accepted 20 April 2017

Abstract

O-acetylation is a common modification of bacterial glycoconjugates. By modifying oligosaccharide structure and chemistry, *O*-acetylation has important consequences for biotic and abiotic recognition events and thus bacterial fitness in general. Previous studies of the broad-spectrum *O*-linked protein glycosylation in pathogenic *Neisseria* species (including *N. gonorrhoeae* and *N. meningitidis*) have revealed *O*-acetylation of some of their diverse glycoforms and identified the committed acetylase, PglI. Herein, we extend these observations by using mass spectrometry to examine a complete set of all glycan variants identified to date. Regardless of composition, all glycoforms and all sugars in the oligosaccharide are subject to acetylation in a PglI-dependent fashion with the only exception of di-*N*-acetyl-bacillosamine. Moreover, multiple sugars in a single oligosaccharide could be simultaneously modified. Interestingly, *O*-acetylation status was found to be correlated with altered chain lengths of oligosaccharides expressed in otherwise isogenic backgrounds. Models for how this unprecedented phenomenon might arise are discussed with some having potentially important implications for the membrane topology of glycan *O*-acetylation. Together, the findings provide better insight into how *O*-acetylation can both directly and indirectly govern glycoform structure and diversity.

Key words: *O*-acetylation, oligosaccharide, *O*-linked protein glycosylation, top-down mass spectrometry

Introduction

Bacterial extracytoplasmic glycoconjugates such as peptidoglycan (PG), capsular polysaccharides (CPS), exopolysaccharides and lipopolysaccharide/lipoologosaccharides (LPS/LOS) play important and influential roles in the fitness of microbes within diverse environments. As in many other biological systems, glycan diversity in these extracytoplasmic glycoconjugates is generated by the use of a vast array of unique monosaccharides linked in linear and branched chains with varying stereochemistry and configurations. These structures can be further elaborated by enzyme-mediated

chemical modifications including methylation, sulfation, phosphorylation, acylation and epimerization as well as by the covalent addition of amino acids (Varki et al. 2009). One chemical modification that is particularly widespread among bacterial glycoconjugates is *O*-acetylation. This modification occurs at free hydroxyl groups on sugar residues and is mediated by *O*-acetyltransferase enzymes (Clarke et al. 2000). Some bacterial systems also make use of esterases that *O*-deacetylate glycoforms leading to dynamic levels of modification. The latter phenomenon is exemplified by the PG remodeling documented in many species (Moynihan and Clarke 2010).

Two species of pathogenic *Neisseria* (*Neisseria gonorrhoeae* (Ng) and *Neisseria meningitidis* (Nm)) have proven to be interesting models to study glycoconjugate *O*-acetylation as they have multiple, dedicated *O*-acetylation systems targeting PG, LOS, and in the case of Nm, CPS (Antignac et al. 2003; Gudlavalleti et al. 2004; Kahler et al. 2006). Both species also express highly related, broad-spectrum *O*-linked protein glycosylation (*pgl*) systems in which some glycoforms are subject to *O*-acetylation (Aas et al. 2007). Glycan biosynthesis, modification and transfer to protein have been examined at both the genetic and biochemical levels in these species (Aas et al. 2007; Vik et al. 2009; Børud et al. 2010, 2011). The *pgl* core locus products act in the synthesis of undecaprenyl diphosphate (UndPP) linked monosaccharides (PglB, PglC and PglD) and its flipping into the periplasm (PglF). PglB acts a bifunctional protein with an acetyltransferase domain and a phospho-glycosyltransferase domain responsible for synthesis of UndPP-*N*,*N*′-diacetylbacillosamine (diNAcBac). Interestingly, nearly one half of Nm isolates express a variant allele of *pglB*, designated *pglB2*, with an ATP grasp domain and a phospho-glycosyltransferase domain responsible for synthesis of UndPP-glyceramido-acetamido trideoxyhexose (GATDH) (Kahler et al. 2001; Power et al. 2003, Chamot-Rooke et al. 2007). PglA is a galactosyltransferase that extends both UndPP-monosaccharides (diNAcBac and GATDH-based), whereas PglE is a galactosyltransferase that further elaborates the PglA-generated UndPP-linked disaccharide to a trisaccharide (illustrated in Figure 1).

In addition to that involving *pglB/B2*, another significant polymorphism at the core *pgl* locus involves the variable presence of the *pglH* gene that lies between the *pglF* and *pglB/B2* genes. It encodes a glucosyltransferase that acts on UndPP-monosaccharides (both diNAcBac and GATDH-based) to generate glucose (Glc)-containing

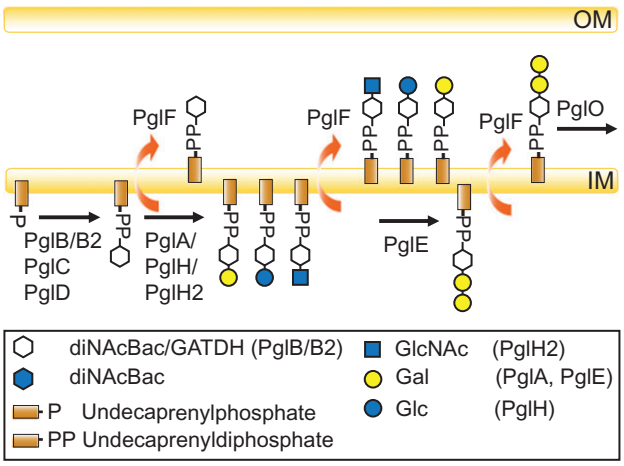

Fig. 1. Current model of the broad-spectrum *O*-linked glycosylation systems in *Neisseria*. OM, outer membrane; IM, inner membrane. This figure is available in black and white inprint and in color at *Glycobiology* online.

disaccharides (Kahler et al. 2001; Power et al. 2003; Børud et al. 2011). An allelic variant of *pglH*, termed *pglH2*, results in incorporation of an *N*-acetylglucosamine (GlcNAc) instead of Glc. In strains simultaneously expressing PglA and PglH, mixtures of Gal- and Glc-containing disaccharides are observed (Børud et al. 2014). In contrast to the situation with *pglA*, the *pglH*-derived UndPP-disaccharide is recalcitrant to further elaboration by PglE. Some strains of Ng and Nm display intrastrain glycan length variation (microheterogeneity) that results from phase-variable, slipped-strand mispairing events within the *pglA*, *pglE* and *pglH/H2* genes (Power et al. 2003; Aas et al. 2007; Børud et al. 2011). In addition, some strains display glycan microheterogeneity due to hypomorphic *pglA* alleles as well as recoding within *pglA* that are associated with shifting glycosyltransferase activity (Johannessen et al. 2012). In summary, individual strains of Ng and Nm have the ability to express arrays of multiple glycans sequentially and simultaneously due to phase variation, recoding and hypomorphic glycosyltransferase alleles. Interstrain glycan differences result from the same processes in addition to differences in glycosyltransferase gene content and status.

Bacterial glycoconjugate *O*-acetylation requires a member of a broad family of acetyltransferases using acetyl-CoA as a donor precursor. In the case of neisserial *pgl*-associated glycans, this process is mediated by PglI that it structurally related to *O*-acetyltransferases modifying LPS in Gram negatives (Slauch et al. 1996; Warren et al. 2004, Aas et al. 2007). Shared structural features within these family members include a conserved N-terminal domain with multiple membrane spanning domains (Pfam_PF01757). Comparative MS-based analyses of *pgl* glycans from wt and *pglI* mutants demonstrated the role of PglI in *O*-acetylation (Aas et al. 2007; Børud et al. 2011). Interestingly, *pglI* is predicted to be subject to phase-variable expression in Nm but not in Ng. The functional implications of *pgl* oligosaccharide *O*-acetylation remain undefined. However, *O*-acetylation has been shown to alter *pgl* glycan antigenicity as shown by both immunoblotting and immunogold labeling with glycoform specific mAbs. Those results indicated that *O*-acetylation masks epitopes present on sugar moieties (Børud et al. 2010).

There remain some discrepancies in the literature as to the role of *pglI*. Warren et al. originally concluded that PglI was involved in the biosynthesis of the basal diNAcBac residue (Warren et al. 2004). In contrast, MS-based studies found no role for PglI in diNAcBac synthesis but showed that modifications were limited to distal sugars in di- and trisaccharide glycoforms (Aas et al. 2007; Børud et al. 2011, 2014).

Here, we examined the effects of *O*-acetylation on *pgl* glycan structure in further detail and report a unique association between *O*-acetylation status and glycoform chain length as well as microheterogeneity. These findings have potentially important implications for the mechanisms by which *O*-acetylases affect oligosaccharide structure and expression.

Results

Acetylation of diNAcBac-based glycans

diNAcBac-based disaccharides

To characterize the extent of *O*-acetylation complexity and the effect of *O*-acetylation on neisserial glycan variation, we carried out top-down electrospray ionization mass spectrometry (ESI-MS) of glycosylated PilE, the major component of the type IV pili, from various glycan expressing backgrounds. The deconvoluted mass spectrum derived by top-down ESI MS of PilE protein from the wild type Ng N400 strain (KS100) revealed a broad distribution of

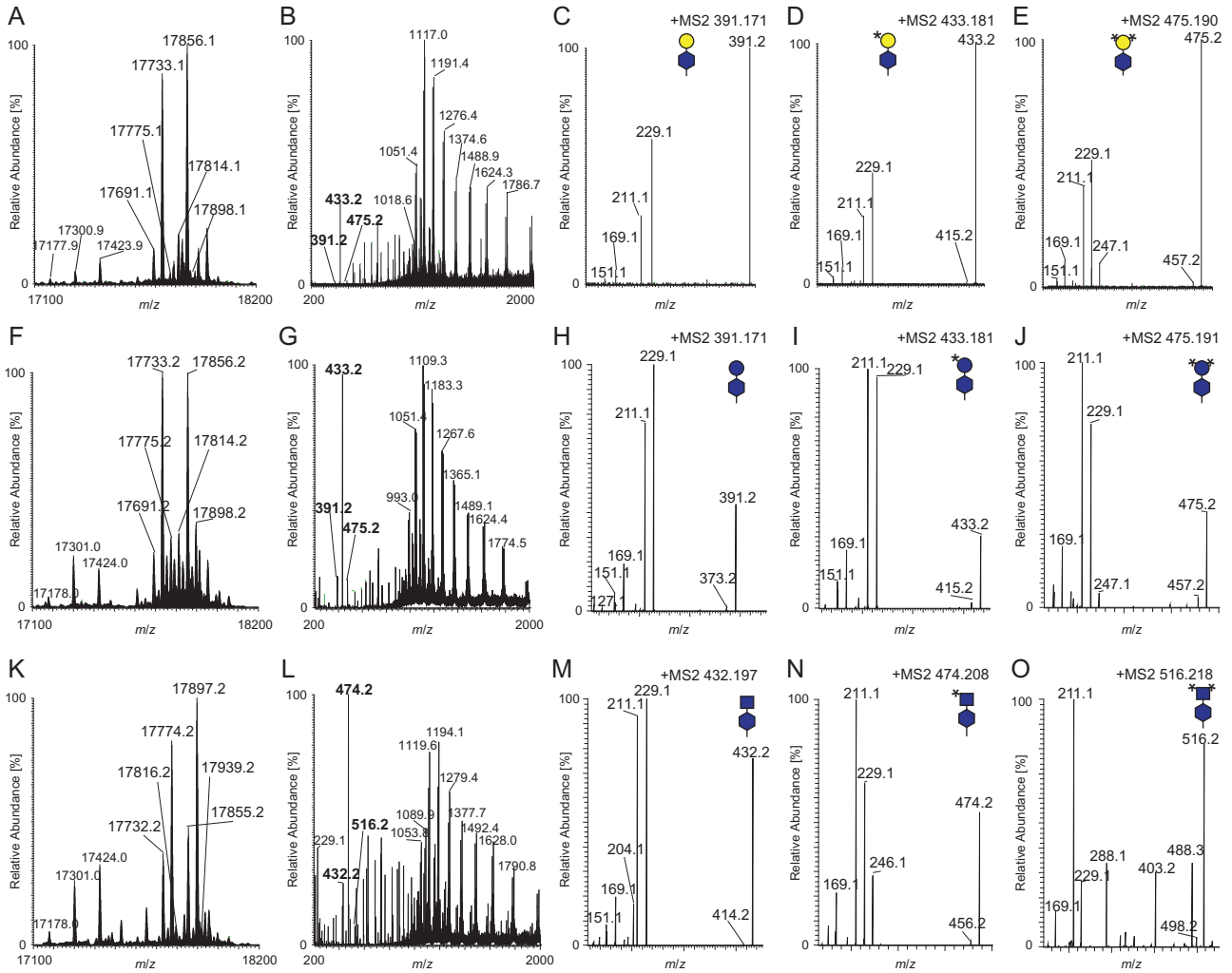

Fig. 2. Acetylation of the diNAcBac-based disaccharides. (**A**) Deconvoluted mass spectrum of intact PilE by ESI-MS analyses from the KS100 strain expressing diNAcBac-Gal-based glycans. The two major peaks at *m/z* 17,733.1 and at *m/z* 17,856.1 represent PilE modified with the mono-acetylated diNAcBac-Gal glycan in addition to one or two PE moieties respectively. The peaks at *m/z* 17,691.1 and at *m/z* 17,814.1 represent PilE modified with the unacetylated diNAcBAc-Gal glycan in addition to one PE and two PE moieties, respectively. The peaks at *m/z* 17,775.1 and at *m/z* 17,898.1 represent PilE modified with di-acetylated diNAcBac-Gal in addition to one PE and two PE moieties, respectively. (**B**) ESI mass spectrum of PilE from strain KS100. Glycan reporter ions detected at *m/z* 391.2 (diNAcBAc-Gal) at *m/z* 433.2 (diNAcBAc-AcGal) and at *m/z* 475.2 (diNAcBAc-diAcGal) in bold. (**C**) HCD MS2 spectrum of the diNAcBac-Gal glycan at *m/z* 391.171 generated by in-source fragmentation of PilE. The figure depicts the unacetylated diNAcBac-Gal glycan fragmented. (**D**) HCD MS2 spectrum of the diNAcBac-AcGal glycan at *m/z* 433.181 generated by in-source fragmentation of PilE. The figure depicts the mono-acetylated diNAcBac-Gal glycan fragmented. (**E**) HCD MS2 spectrum of the diNAcBac-diAcGal glycan at *m/z* 475.190 generated by in-source fragmentation of PilE. The figure depicts the di-acetylated diNAcBac-Gal glycan fragmented. (**F**) Deconvoluted mass spectrum of intact PilE by ESI-MS analyses from the KS352 strain expressing diNAcBac-Glc-based glycans. The two major peaks at *m/z* 17,733.2 and at *m/z* 17,856.2 represent PilE modified with the mono-acetylated diNAcBac-Glc glycan in addition to one or two PE moieties, respectively. The peaks at *m/z* 17,691.2 and at *m/z* 17,814.2 represent PilE modified with the unacetylated diNAcBAc-Glc glycan in addition to one PE and two PE moieties, respectively. The peaks at *m/z* 17,775.2 and at *m/z* 17,898.2 represent PilE modified with di-acetylated diNAcBac-Glc in addition to one PE and two PE moieties, respectively. (**G**) ESI mass spectrum of PilE from strain KS352. Glycan reporter ions detected at *m/z* 391.2 (diNAcBAc-Glc) at *m/z* 433.2 (diNAcBAc-AcGlc) and at *m/z* 475.2 (diNAcBAc-diAcGlc) in bold. (**H**) HCD MS2 spectrum of the diNAcBac-Glc glycan at *m/z* 391.171 generated by in-source fragmentation of PilE. The figure depicts the unacetylated diNAcBac-Glc glycan fragmented. (**I**) HCD MS2 spectrum of the diNAcBac-AcGlc glycan at *m/z* 433.181 generated by in-source fragmentation of PilE. The figure depicts the mono-acetylated diNAcBac-Glc glycan fragmented. (**J**) HCD MS2 spectrum of the diNAcBac-diAcGlc glycan at *m/z* 475.190 generated by in-source fragmentation of PilE. The figures depict the di-acetylated diNAcBac-Glc glycan fragmented. (**K**) Deconvoluted mass spectrum of intact PilE by ESI-MS analyses from the KS985 strain expressing diNAcBac-GlcNAc-based glycans.The two major peaks at *m/z* 17,774.2 and at *m/z* 17,897.2 represent PilE modified with the mono-acetylated diNAcBac-GlcNAc glycan in addition to one or two PE moieties, respectively. The peaks at *m/z* 17,732.2 and at *m/z* 17,855.2 represent PilE modified with the unacetylated diNAcBAc-GlcNAc glycan in addition to one PE and two PE moieties, respectively. The peaks at *m/z* 17,816.2 and at *m/z* 17,939.2 represent PilE modified with di-acetylated diNAcBac-GlcNAc in addition to one PE and two PE moieties, respectively. (**L**) ESI mass spectrum of PilE from strain KS985. glycan reporter ions detected at *m/z* 432.2 (diNAcBAc-GlcNAc) at *m/z* 474.2 (diNAcBAc-AcGlcNAc) and at *m/z* 516.2 (diNAcBAc-diAcGlcNAc) in bold. (**M**) HCD MS2 spectrum of the diNAcBac-GlcNAc glycan at *m/z* 432.197 generated by in-source fragmentation of PilE. The figure depicts the unacetylated diNAcBac-GlcNAc glycan fragmented. (**N**) HCD MS2 spectrum of the diNAcBac-AcGlcNAc glycan at *m/z* 474.208 generated by in-source fragmentation of PilE. The figure depicts the mono-acetylated diNAcBac-GlcNAc glycan fragmented. (**O**) HCD MS2 spectrum of the diNAcBac-diAcGlcNAc glycan at *m/z* 516.218 generated by in-source fragmentation of PilE. The figure depicts the unacetylated diNAcBac-GlcNAc glycan fragmented. A complete list of all glycan reporter ions with *m/z*-values and corresponding molecular weight values of all PilE species are found in Tables II and III. Symbols representing glycan structures are explained in the legend of Figure 1. This figure is available in black and white in print and in color at *Glycobiology* online.

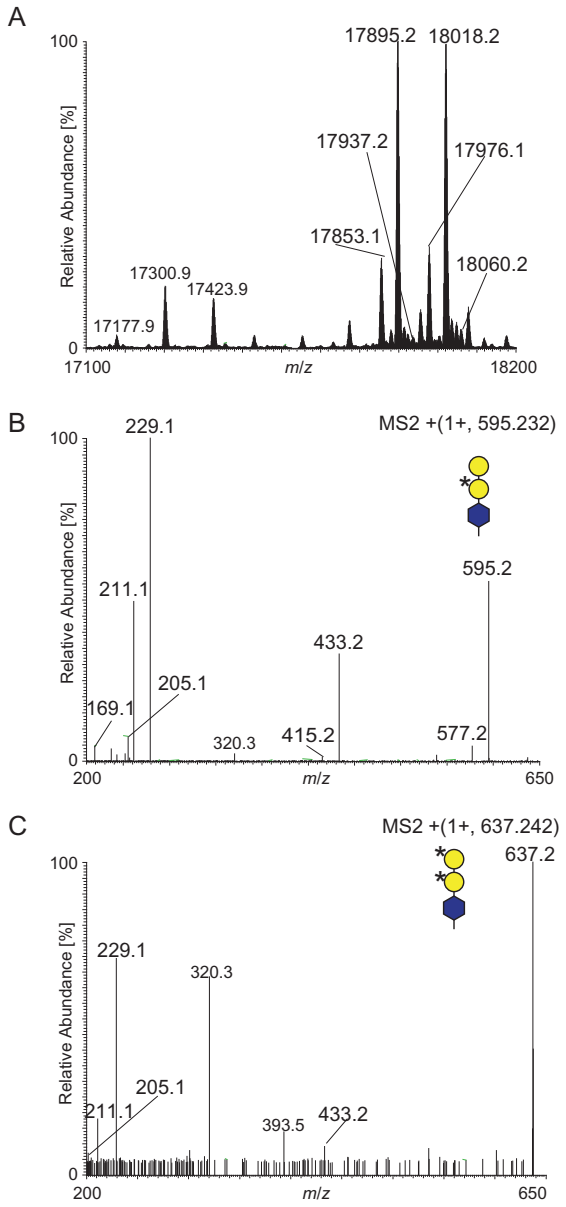

Fig. 3. Acetylation of the diNAcBac-Gal-Gal trisaccharide. (**A**) Deconvoluted mass spectrum of intact PilE by ESI-MS analyses from the KS142 strain expressing diNAcBac-Gal-Gal-based trisaccharides. The two major peaks at *m/z* 17,895.2 and at *m/z* 18,018.2 represent PilE modified with the acetylated diNAcBac-Gal-Gal glycan (594.2 Da) in addition to one or two PE moieties, respectively. The peaks at *m/z* 17,853.1 an at *m/z* 17,976.1 represent PilE modified with the unacetylated diNAcBAc-Gal-Gal glycan (552.2 Da) in addition to one PE and two PE moieties, respectively. The low intensity peaks at *m/z* 17,937.2 and at *m/z* 18,060.2 represent PilE modified with di-acetylated diNAcBac-Gal-Gal (636.2 Da) in addition to one PE and two PE moieties, respectively. (**B**) HCD MS2 spectrum of the acetylated diNAcBac-Gal-Gal glycan at *m/z* 595.232 generated by in-source fragmentation of PilE. The figure illustrates the structure of the mono-acetylated diNAcBac-Gal-Gal glycan fragmented. (**C**) HCD MS2 spectrum of the di-acetylated diNAcBac-Gal-Gal glycan at *m/z* 637.242 generated by in-source fragmentation of PilE. The figure illustrates the structure of the di-acetylated diNAcBac-Gal glycan fragmented. Ions at *m/z* 320.3 and at *m/z* 393.5 were background ions present in all MS and MS2 spectra and not connected to glycan fragmentation. A complete list of all glycan reporter ions with *m/z*-values and corresponding molecular weight values of all PilE species are found in Tables II and III. Symbols representing glycan structures are explained in the legend of Figure 1. This figure is available in black and white in print and in color at *Glycobiology* online.

modified forms (Figure 2A). As previously shown (Aas et al. 2007), the mature form of the PilE protein (*processed, methylated and with a single disulfide bond*) appear in the spectrum at *m/z* 17,177.9. The species appearing at *m/z* 17,300.9 and *m/z* 17,423.9 correspond to protein modified with one or two phosphoethanolamine (PE) moieties (mass addition of 123.0 Da), respectively (Hegge et al. 2004). Addition of the diNAcBac-Gal disaccharide glycan (mass addition of 390.2 Da) to the PE modified protein forms leads to the species appearing at *m/z* 17,691.1 and *m/z* 17,814.1 whereas the most abundant protein species present in the spectrum, at *m/z* 17,733.1 and at *m/z* 17,856.1, represent the PilE protein modified with a mono-acetylated (mass addition of 42.0 Da) diNAcBac-Gal glycan (total mass addition of 432.2 Da) (Aas et al. 2007) and either one or two PE moieties, respectively. Two low abundant protein species, at *m/z* 17,775.1 and *m/z* 17,898.1, corresponding to a mass addition of 474.2 Da to the PE modified protein forms, equaling to two mass additions of 42.0 Da to the diNAcBac-Gal glycan, were also detected. No ions at *m/z* 17,775.1 or at *m/z* 17,898.1 were detected in the glycosylation null mutant strain (KS104), nor were mass additions of 84.0 or 42.0 Da to the unmodified, one PE or two PE PilE protein species (Supplementary data, Fig. 1a) detected demonstrating that the second 42 Da mass addition was associated with the glycan modification. Consequently, we used in-source fragmentation of the purified PilE protein (Figure 2B) and targeted the reporter ion appearing at *m/z* 475.2, corresponding to the $[H]^+$ reporter ion of the novel 474.2 Da modification, for CID/HCD fragmentation (MS2) coupled with high resolution fragment ion detection analysis. The MS2 spectrum of the ion at *m/z* 475.2 (Figure 2E) showed the same general fragmentation pattern as previously described for diNAcBac-Gal at *m/z* 391.2 (Figure 2C) and diNAcBac-AcGal at *m/z* 433.2 (Figure 2D) (Aas et al. 2007). In addition to the loss of water (18.0 Da), commonly experienced in glycan MS fragmentation, fragmentation of the *m/z* 475.2 ion showed a complete loss of the modified distal Gal to the reporter ions associated with the complete reducing end diNAcBac at *m/z* 229.1 (diNAcBac fragment ions, Supplementary data, Fig. 1e or Aas et al. 2007), demonstrating that the new PilE mass addition was a diNAcBac-based disaccharide. Moreover, no ion at *m/z* 475.2 was detected in the *pglA* mutant strain (KS141) (Supplementary data, Fig. 1c), demonstrating that the new mass addition was located on the distal Gal sugar of the protein attached glycan. The high mass accuracy afforded by the Orbitrap showed that the mass addition (42.009 Da) to the distal Gal was nearly exactly the theoretical mass of an acetylation (42.011 Da) (i.e., the mass difference between the precursor glycan ions at *m/z* 475.190 and *m/z* 433.181 seen in Figure 2F and D, respectively). Moreover, no species was detected at *m/z* 475.2 and no mass additions of neither 42.0 nor 84.0 Da to the glycan were detected in an acetyltransferase lacking (*pglI* null) background (KS144) (Supplementary data, Fig. 2a and b). Thus, the O-acetyltransferase PglI is capable of attaching at least two acetyl groups to the *Neisseria* diNAcBac-Gal glycan.

Two additional diNAcBac-based disaccharides have previously been described in *Neisseria*, diNAcBac-Glc and diNAcBac-GlcNAc (Børud et al. 2011, 2014). Our earlier work has shown a single acetylation of the diNAcBac-Glc glycan (Børud et al. 2011) while no acetylation was reported for the diNAcBac-GlcNAc glycan (as this work was performed in *pglI* mutant backgrounds) (Børud et al. 2014). To investigate if the di-acetylation is a general feature on diNAcBac-based disaccharides, we applied high mass resolution top-down ESI MS in combination with in-source fragmentation and targeted CID/HCD fragmentation of different PilE proteins modified with either diNAcBac-Glc or the diNAcBac-GlcNAc-based disaccharides in backgrounds expressing PglI. The

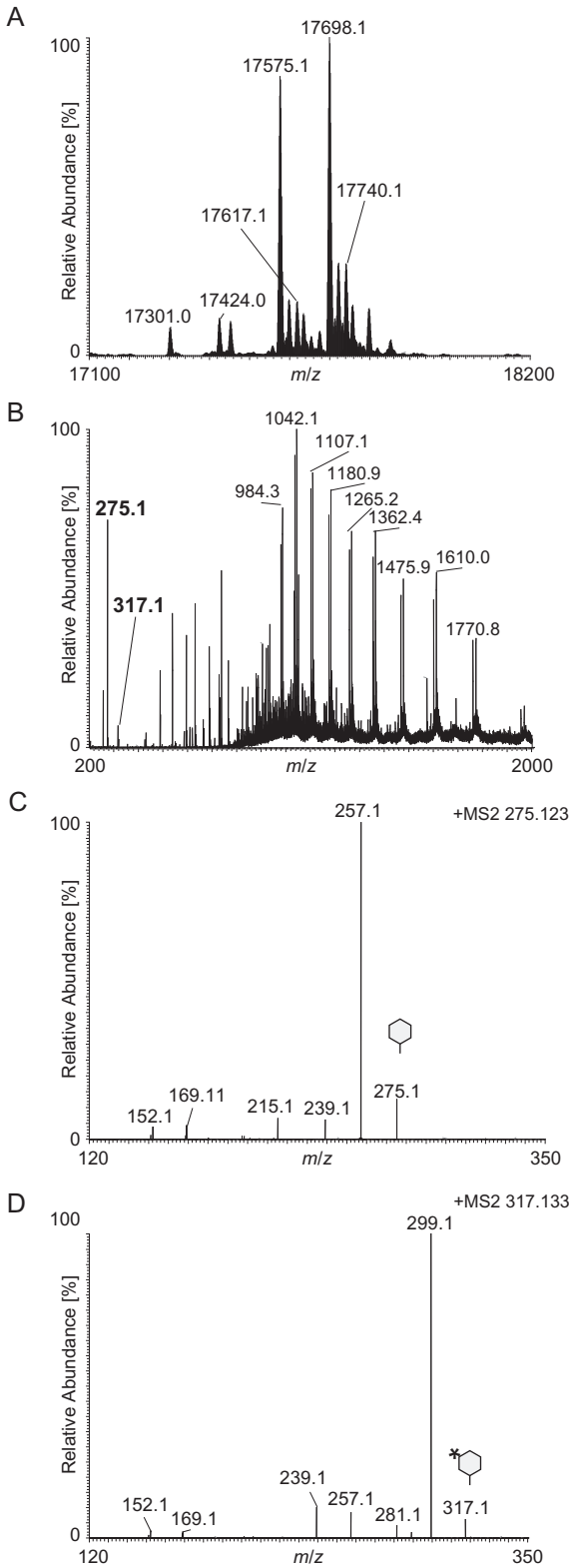

Fig. 4. Acetylation of the GATDH. (**A**) Deconvoluted mass spectrum of intact PilE from strain KS309 by ESI-MS analyses. The two major peaks at *m/z* 17,575.1 and at *m/z* 17,698.1 represent PilE modified with the GATDH glycan (274.1 Da) in addition to one or two PE moieties, respectively. (**B**) ESI mass spectrum of PilE from strain KS309. Glycan reporter ions detected at *m/z* 275.1 (GATDH) and t *m/z* 317.133 (AcGATDH) in bold. (**C**) HCD MS2 spectrum of the GATDH glycan at *m/z* 275.123 generated by in-source fragmentation of PilE. The white diamond figure illustrates the nonacetylated GATDH glycan fragmented. (**D**) HCD MS2 spectrum of the acetylated GATDH glycan at *m/z* 317.133 generated by in-source fragmentation of PilE. The white diamond figure illustrates the mono-acetylated GATDH glycan fragmented. A complete list of all glycan reporter ions with *m/z*-values and corresponding molecular weight values of all pilin/PilE species are found in Tables II and III. Symbols representing glycan structures are explained in the legend of Figure 1.

deconvoluted MS spectrum of PilE purified from the strain expressing diNAcBac-Glc-based disaccharide (KS352) showed protein species consistent with no acetylation (at *m/z* 17,691.2 and at *m/z* 17,814.2), mono-acetylation (at *m/z* 17,733.2 and at *m/z* 17,856.2) and di-acetylation of the disaccharide (at *m/z* 17,775.2 and at *m/z* 17,898.2) (Figure 2F). In-source fragmentation revealed reporter ions consistent with no- (*m/z* 391.2), mono- (*m/z* 433.2) and di-acetylation (*m/z* 475.2) of the diNAcBac-Glc disaccharide (Figure 2G). The MS2 spectrum of the ion at *m/z* 475.2 (Figure 2J) showed a similar fragmentation pattern as the corresponding ions generated from the diNAcBac-Gal-based glycan (Figure 2E), with the appearance of a di-acetylated Glc ion at *m/z* 247.1. Moreover, the MS2 spectrum resembled the previously reported diNAcBac-Glc (Figure 2H) and diNAcBac-AcGlc (Figure 2I) with an addition 42 Da located at the distal Glc sugar. No ion at *m/z* 475.2 or protein forms consistent with acetylation of diNAcBac-Glc was detected in the *pglI* mutant background (KS421) (Supplementary data, Figs. 2c and 2d). The diNAcBac-Glc disaccharide may therefore be doubly acetylated at the distal Glc sugar. For the strain producing a diNAcBac-GlcNAc-based disaccharide (KS985), the same pattern of acetylation was observed as seen for the diNAcBac-Gal and the diNAcBac-Glc glycan. Species consistent with both single (at *m/z* 17,774.2 and at *m/z* 17,897.2) and double (at *m/z* 17,816.2 and at *m/z* 17,939.2) acetylation were detected (Figure 2K). The MS2 spectrum of the ion at *m/z* 474.2 (Figure 2N) (42.011 Da larger than the diNAcBac-GlcNAc precursor ion at *m/z* 432.197 (Figure 2M)) showed a fragment pattern consistent with a single acetylation of the distal GlcNAc sugar, with the appearance of the acetylated GlcNAc ion at *m/z* 246.1. Whereas the MS2 spectrum of the ion at *m/z* 516.2 (Figure 2O) (42.010 Da larger than the ion at *m/z* 474.208) showed a fragmentation pattern consistent with a double acetylation of the distal GlcNAc sugar. No ions indicative of mass additions of 42.0 or 84.0 to the diNAcBac-GlcNAc glycan in the *pglI* background (KS966) were detected (Supplementary data, Figs. 2e and 2f). Thus, all known diNAcBac-based disaccharides, can be doubly acetylated by the *O*-acetyltransferase PglI.

Acetylation of diNAcBac-based trisaccharides

The only glycosyltransferase in pathogenic *Neisseria* known to be capable of generating a trisaccharide glycoform is PglE. Although the number of possible acetylation sites increases on a trisaccharide, only a single PglI-dependent *O*-acetylation has been found for both trisaccharide glycoforms diNAcBac-Gal-Gal and GATDH-Gal-Gal (Aas et al. 2007; Børud et al. 2011). Therefore, to investigate the number of *O*-acetylations present on the diNAcBac-Gal-Gal trisaccharide, we performed top-down MS of PilE from the KS142 strain (expressing an acetylated diNAcBac-Gal-Gal glycan). As shown in Figure 3A, the deconvoluted mass spectrum of PilE from the KS142 strain with the most abundant protein species at *m/z* 17,895.2 (1 PE) and at *m/z* 18,018.2 (2 PE), representing the mono-acetylated

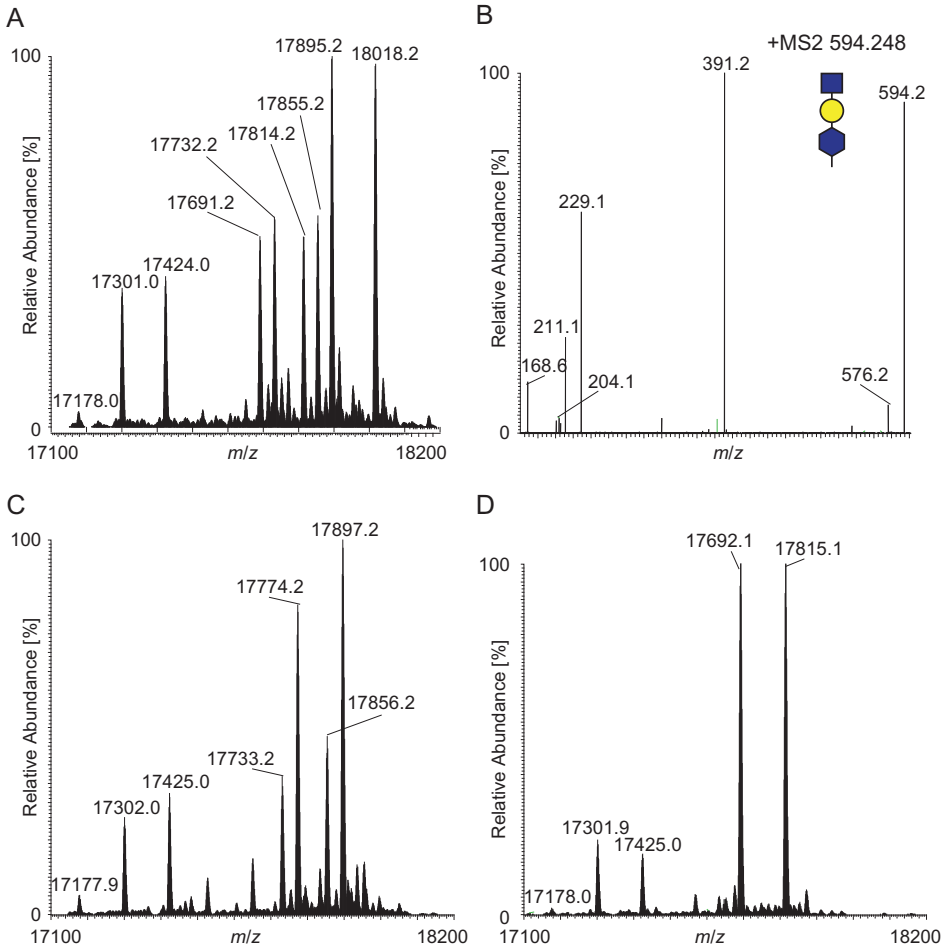

Fig. 5. Acetylation determines glycan length. (**A**) Deconvoluted mass spectrum of intact PilE by ESI-MS analyses from strain KS983. The two rightmost peaks in the mass spectrum at *m/z* 17,895.2 and at *m/z* 18,018.2 represent PilE modified with the new trisaccharide diNAcBac-Gal-GlcNAc and 1 or 2 PE, respectively. The peaks at *m/z* 17,691.2 and at *m/z* 17,814.2 represent PilE modified with the diNAcBac-Gal disaccharide and 1 or 2 PE moieties, respectively. The peaks at *m/z* 17,732.2 and at *m/z* 17,855.2 represent PilE modified with the diNAcBac-GlcNAc disaccharide and 1 or 2 PE moieties, respectively. (**B**) HCD MS2 spectrum of the diNAcBac-Gal-GlcNAc glycan at *m/z* 594.248 generated by in-source fragmentation of PilE. The figure illustrates the structure of the diNAcBAc-Gal-GlcNAc glycan fragmented. (**C**) Deconvoluted mass spectrum of intact PilE by ESI-MS analyses from strain KS391.The two major peaks in the mass spectrum at *m/z* 17,774.2 (1 PE) and at *m/z* 17,897.2 (2 PE) represent PilE modified with mono-acetylated diNAcBac-GlcNAc. The peaks at *m/z* 17,733.2 and at *m/z* 17,856.2 represent both the mono-acetylated diNAcBac-AcGal (at *m/z* 17,733.1 and at *m/z* 17,856.1) and the unacetylated diNAcBac-GlcNAc (at *m/z* 17,732.1 and at *m/z* 17,855.1) glycoforms. (**D**) Deconvoluted mass spectrum of intact PilE by ESI-MS analyses from strain RV710.The two major peaks in the mass spectrum at *m/z* 17,692.1 and at *m/z* 17,815.1 represent a mix of PilE modified with diNAcBac-Gal and diNAcBac-Glc glycoforms and 1 and 2 PE moieties respectively. A complete list of all glycan reporter ions with *m/z*-values and corresponding molecular weight values of all PilE species are found in Tables II and III. Symbols representing glycan structures are explained in the legend of Figure 1. This figure is available in black and white in print and in color at *Glycobiology* online.

trisaccharide glycan modified protein species. The diNAcBac-Gal-Gal protein forms could be seen at *m/z* 17,853.1 and at *m/z* 17,976.1 whereas only small peaks at *m/z* 17,937.2 and at *m/z* 18,060.2, consistent with di-acetylated trisaccharide glycan modified protein species were detected. No peak representing a triply acetylated trisaccharide modified protein species was detected in the deconvoluted mass spectrum. The MS2 spectrum of the Ac-diNAcBac-Gal-Gal ion at *m/z* 595.2 from in-source fragmentation of PilE (Figure 3B) was consistent with previous data (Aas et al. 2007) showing only a loss of H_2O (18.0 Da) or Gal (162.0 Da) from the 595.2 precursor ion to the *m/z* 433.2 mono-acetylated disaccharide ion. No ion representing diNAcBac-Gal (at *m/z* 391.2) was detected, demonstrating that for the mono-acetylated trisaccharide glycoform, the acetylation was primarily located at the first Gal (diNAcBac-AcGal-Gal). The low abundance of protein species corresponding to di-acetylated trisaccharide, yielding a reporter ion at *m/z* 637.2 from in-source fragmentation of PilE (Supplementary data, Fig. 3) when selectively fragmented, generated an MS2 spectrum close to the noise level (Figure 3C). The loss of 204.0 Da from the ion at *m/z* 637.2 to the diNAcBac-AcGal ion at *m/z* 433.2 indicated that for di-acetylated trisaccharide, the second acetylation may be located at the distal Gal residue (diNAcBac-AcGal-AcGal). However, the low abundance of glycan related fragment ions obscures any other combination of location for di-acetylation on the trisaccharide. No ion corresponding to a third acetylation of the diNAcBac-Gal-Gal was detected and only ions related to nonacetylated trisaccharide was detected in the KS304 *pglI* mutant strain (Supplementary data, Fig. 4a–d). The di-acetylation of the diNAcBac-Gal-Gal trisaccharide is therefore dependent upon PglI and, since the distal Gal only gets acetylated when two acetylations are present, seems to be hierarchical.

Acetylation of GATDH-based sugar

Acetylation of GATDH

In addition to diNAcBac-based glycans, a second reducing end sugar, GATDH, has been described in Nm (Chamot-Rooke et al. 2007). Using an Ng strain expressing *pglB2*, previous work reported both GATDH-Gal and GADTH-Glc to be mono-acetylated (Børud et al. 2011) while GATDH-GlcNAc was only characterized in a *pglI* mutant background (Børud et al. 2014). To investigate the acetylation status of GATDH-based glycans, we performed top-down ESI MS analysis and in-source fragmentation of PilE from strains expressing PglI and GATDH-based glycans and targeted glycan reporter ions for HCD/CID MS2. In Figure 4A, the deconvoluted PilE mass spectrum from top-down ESI MS of PilE from the strain expressing only the GATDH-monosaccharide (KS309) showed two major protein species representing PilE with a single GATDH modification and either one PE (at *m/z* 17,575.1) or two PE (at *m/z* 17,698.1) modifications. Two peaks 42.0 Da larger than the major protein species, consistent with acetylated GADTH, were seen at *m/z* 17,617.1 (1 PE) and at *m/z* 17,740.1 (2 PE) (Figure 4A). To further investigate the acetylation status of GATDH, we targeted the reporter ion for GATDH, at *m/z* 275.123, and the reporter ion consistent with an acetylated GATDH, at *m/z* 317.133 (Figure 4B) for MS2 fragmentation. The MS2 spectrum of GATDH shown in Figure 4C was consistent with previously reported fragmentation spectrum of GATDH (Chamot-Rooke et al. 2007). The MS2 spectrum of the ion at *m/z* 317.1 (Figure 4D) showed a fragmentation pattern that, except for the 42.0 Da mass addition, followed the GATDH fragmentation pattern. Moreover, the two MS2 spectra converged at the *m/z* 257.1 ion (*m/z* 215.1 + 42.0 Da) and followed the same fragmentation pathway demonstrating that the core of the two molecules are the same. No ion consistent with two 42.0 Da mass additions to the GATDH glycan was detected and no ion at *m/z* 317.1 was detected in the acetylation negative *pglI* background (N60) (Supplementary data, Fig. 5). Thus, PglI is capable of attaching a single acetate moiety to the glyceracetamido group of GATDH.

Acetylation of GATDH-based di- and trisaccharide

When investigating the GATDH-based disaccharides, nonacetylated, acetylated and di-acetylated forms of GATDH-Gal (KS308) (Supplementary data, Fig. 6a-d), GATDH-Glc (KS361) (Supplementary data, Fig. 6 e-h) and GATDH-GlcNAc (KS968) (Supplementary data, Fig. 6 i-l) were detected. MS2 spectra of mono-acetylated GATDH-Gal (Supplementary data, Fig. 6b) and GATDH-Glc (Supplementary data, Fig. 6f) (ions at *m/z* 479.2) as well as mono-acetylated GATDH-GlcNAc (Supplementary data, Fig. 6J) (ion at *m/z* 520.2) showed fragment ions both at *m/z* 275.1 (GATDH) and at *m/z* 317.1 (AcGATDH). These results demonstrated that for the mono-acetylated GATDH-based disaccharides, there was a mix of glycoforms with acetylation either at the basal (GATDH) residue or on the distal (disaccharide) position. MS2 spectra of di-acetylated GATDH-Gal (Supplementary data, Fig. 6c) and GATDH-Glc (Supplementary data, Fig. 6g) (ions at *m/z* 521.2) also showed ions at *m/z* 275.1 and at *m/z* 317.1 as well as at *m/z* 247 (2xAcHex) demonstrating that the di-acetylated glycoforms of GATDH-Gal and GATDH-Glc were a mix of AcGATDH-AcGal (AcGATDH-AcGlc) and the GATDH-diAcGal (GATDH-diAcGlc) glycoforms. However, no ion at *m/z* 275.1 was detected in MS2 spectra of di-acetylated GATDH-GlcNAc (Supplementary data, Fig. 6k) (ion at *m/z* 562.2) indicating that only a single di-acetylated glycoform, AcGATDH-AcGlcNAc (strain KS968) was present. For the GATDH-based trisaccharide (GATDH-Gal-Gal, strain KS310), the deconvoluted mass spectra (Supplementary data, Fig. 6p) showed abundant species

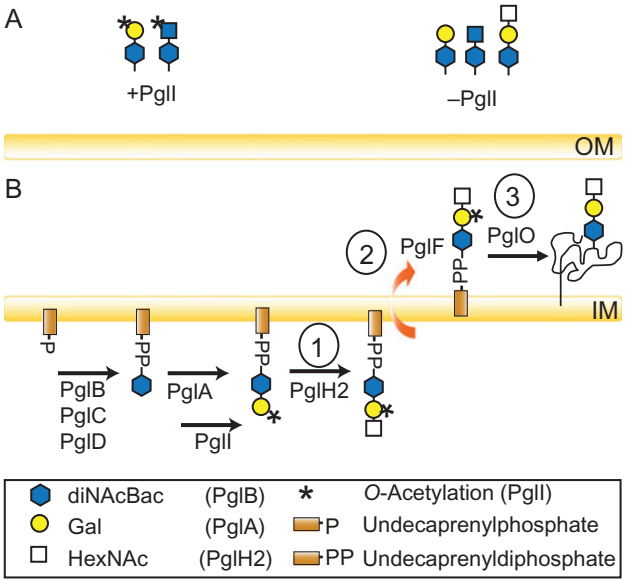

Fig. 6. Expression of *pgl*-related glycoforms is influenced by *O*-acetylation status. (**A**) Repertoires of glycans expressed by strains simultaneously expressing PglA and PglH2 detected in the presence (+) and absence (–) of *O*-acetylation (mediated by PglI). (**B**) Models for how *O*-acetylation of the GlcNAc residue might preclude expression of the diNAcBac-Gal-HexNAc trisaccharide include (1) sterically hindering incorporation of the distal HexNAc by PglH2, (2) precluding membrane translocation/flipping into the periplasm (mediated by PglF) and (3) precluding the ability to be used as a donor substrate by the PglO oligosaccharide transferase. See text for more details. This figure is available in black and white in print and in color at *Glycobiology* online.

consistent with a nonacetylated trisaccharide (mass addition of 598.2 Da, at *m/z* 17,777.1, at *m/z* 17,900.1 (1 PE) and at *m/z* 18,023.1 (2 PE)) as well as mono-acetylated trisaccharide (mass addition of 640.2 Da, at *m/z* 17,819.1, at *m/z* 17,942.1 (1 PE) and at *m/z* 18,065.1 (2 PE)). However, no mass consistent with a di-acetylated protein species could be unambiguously identified. In-source fragmentation (Supplementary data, Fig. 5m) and targeted MS2 of the ion at *m/z* 641.2 (Supplementary data, Fig. 5n) showed fragmentation patterns with ions at *m/z* 275.1 (GATDH), *m/z* 317.1(AcGATDH), *m/z* 437.2 (GATDH-Gal) and at *m/z* 479.2 (GATDH-AcGal). Both the presence and intensity of these ions demonstrated that all three residues of the GATDH-Gal-Gal glycan may be singly acetylated in the mono-acetylated glycoform. Targeted MS2 of the ion at *m/z* 683.2 (Supplementary data, Fig. 5o), consistent with a di-acetylated trisaccharide glycoform, produced a fragment pattern indicating that multiple overlapping precursor ions had been selected for fragmentation thereby preventing a clear results. However, in the small mass area, reporter ions consistent with AcGATDH (at *m/z* 317.1, at *m/z* 299.1, at *m/z* 257.2, and at *m/z* 215.1) show that there is a di-acetylated GATDH-Gal-Gal glycoform although at very low levels. No tri-acetylated glycoforms were detected for any of the GATDH-based di- or trisaccharides. Moreover, no glycoforms consistent with a 42.0 Da mass addition were detected in *O*-acetyltransferase negative (*pglI*) backgrounds of GATDH-based glycoform expressing strains (KS901, KS420, KS970, KS414) (Supplementary data, Fig. 7). PglI is therefore capable of modifying GATDH-based disaccharides with mono- and di-acetylation at the disaccharide position or a combination of mono-acetylation of the disaccharide position with a single acetylation of GATDH. Moreover, PglI is capable of acetylating the GATDH-Gal-Gal glycan with a single acetyl group on all three sugar positions.

Table I. *N. gonorrhoeae* strains used in this study

Strain	Relevant genotype	Glycan	References
KS100	*recA6*	diNAcBac-Gal diNAcBac-AcGal diNAcBac-diAcGal	Tonjum et al. (1995) and Figure 2
KS104	*pglC::kan*	None	Hegge et al. (2004) and Supplementary data, Figure S1
KS141	*pglA::kan*	diNAcBac	Aas et al. (2007) and Supplementary data, Figure S1
KS144	*pglI::kan*	diNAcBac-Gal	Aas et al. (2007) and Supplementary data, Figure S2
KS352	pgl_{ST640} *pglA::kan*	diNAcBac-Glc diNAcBac-AcGlc diNAcBac-diAcGlc	Børud et al. (2011) and Supplementary data, Figure 2
KS421	pgl_{Z2491} *pglA::erm pglI::kan*	diNAcBac-Glc	Børud et al. (2011) and Supplementary data, Figure S2
KS985	*lct::*$pglH2_{SK\text{-}93\text{-}1035}$ *pglA::kan*	diNAcBac-GlcNAc diNAcBac-AcGlcNAc diNAcBac-diAcGlcNAc	Børud et al. (2014) and Figure 2
KS966	*pglA::erm pglI::kan lct::*$pglH2_{SK\text{-}93\text{-}1035}$	diNAcBac-GlcNAc	Børud et al. (2014) and Supplementary data, Figure S2
KS142	$pglE_{on}$	diNAcBAc-Gal-Gal diNAcBac-AcGal-Gal diNAcBac-Gal-AcGal	Aas et al. (2007), Figure 3 and Supplementary data, Figure S3
KS304	$pglE_{on}$ *pglI::kan*	diNAcBAc-Gal-Gal	Børud et al. (2010) and Supplementary data, Figure S4
KS309	$pglB2_{8013}$ *pglA::kan*	GATDH AcGATDH	Børud et al. (2010) and Figure 4
N60	*pgl::8013 pglI::kan pglA::erm*	GATDH	Supplementary data, Figure S5
KS308	$pglB2_{8013}$	GATDH-Gal AcGATDH-Gal GATDH-AcGal AcGATDH-AcGal GATDH-diAcGal	Børud et al. (2010) and Supplementary data, Figure S6
KS361	pgl_{FAM18} *pglA::kan*	GATDH-Glc AcGATDH-Glc GATDH-AcGlc AcGATDH-AcGlc GATDH-diAcGlc	Børud et al. (2011) and Supplementary data, Figure S6
KS968	$pglB2_{8013}$ *pglA::kan lct::*$pglH2_{SK\text{-}93\text{-}1035}$	GATDH-GlcNAc AcGATDH-GlcNAc GATDH-AcGlcNAc AcGATDH-AcGlcNAc GATDH-diAcGlcNAc	Børud et al. (2014) and Supplementary data, Figure S6
KS310	$pglB2_{8013}$ $pglE_{on}$	AcGATDH-Gal-Gal GATDH-AcGal-Gal GATDH-Gal-AcGal AcGATDH-AcGal-Gal	Børud et al. (2010) and Supplementary data, Figure S6
KS901	$pglB2_{8013}$ *pglI::kan*	GATDH-Gal	Supplementary data, Figure S7
KS420	$pglB2_{FAM18}$ *pglA::erm pglI::kan*	GATDH-Glc	Supplementary data, Figure S7
KS970	$pglB2_{8013}$ *pglA::erm pglI::kan lct::*$pglH2_{SK\text{-}93\text{-}1035}$	GATDH-GlcNAc	Supplementary data, Figure S7
KS414	$pglB2_{8013}$ *pglEon pglI::kan*	GATDH-Gal-Gal	Supplementary data, Figure S7
KS983	*lct::*$pglH2_{SK\text{-}93\text{-}1035}$ *pglI::kan*	diNAcBac-GlcNAc diNAcBac-Gal diNAcBac-Gal-GlcNAc	Figure 5 and Supplementary data, Figure S8
KS391	*lct::*$pglH2_{SK\text{-}93\text{-}1035}$	diNAcBac-GlcNAc diNAcBac-AcGlcNAc diNAcBac-diAcGlcNAc diNAcBac-AcGal diNAcBac-diAcGal	Figure 5 and Supplementary data, Figure S8
RV710	*pglE pglI::kan lct::*$pglH2_{A303Q}$	diNAcBac-Gal diNAcBac-GlcNAc diNAcBac-Gal-GlcNAc	Figure 5 and Supplementary data, Figure S8

Acetylation can influence glycan chain length elongation

When investigating the effect of *O*-acetylation on glycans from a strain simultaneously expressing both diNAcBac-Gal (PglA-based) and diNAcBac-GlcNAc (PglH2-based) glycoforms in *pglI* background, a previously undescribed glycoform was found. As seen in Figure 5A, the deconvoluted mass spectrum of PilE from this strain (KS983), showed signals corresponding to the protein modified with nonacetylated diNAcBac-Gal (together with 1 PE at *m/z* 17,691.2 and with 2 PE at *m/z* 17,814.2) and nonacetylated diNAcBac-GlcNAc (with 1 PE at *m/z* 17,732.2 and with 2 PE at *m/z* 17,855.2) glycoforms. Interestingly, two previously undescribed forms were detected at *m/z* 17,895.2 and at *m/z* 18,018.2 that correspond to a

Table II. Mass additions to PilE and reporter ion *m/z*

PTM	Mass addition (Da)	Reporter ion (*m/z*)
Acetylation	42.0106	
Phosphoethanolamine	123.009	142.026
Gal or Glc	162.053	163.06
GlcNAc	203.079	204.086
diNAcBac	228.111	229.118
diNAcBac-Gal	390.164	391.171
diNAcBac-AcGal	432.1746	433.1816
diNAcBac-diAcGal	474.1852	475.1922
diNAcBac-Gal-Gal	552.217	553.224
diNAcBac-AcGal-Gal	594.2276	595.2346
diNAcBac-AcGal-AcGal	636.2382	637.2452
diNAcBac-Glc	390.164	391.171
diNAcBac-AcGlc	432.1746	433.1816
diNAcBac-diAcGlc	474.1852	475.1922
diNAcBac-GlcNAc	431.19	432.197
diNAcBac-AcGlcNAc	473.2006	474.2076
diNAcBac-diAcGlcNAc	515.2112	516.2182
diNAcBac-Gal-GlcNAc	593.243	594.25
GATDH	274.116	275.123
GATDHAc	316.1266	317.1336
GATDH-Gal	436.169	437.176
GATDH-AcGal	478.1796	479.1866
GATDH-diAcGal	520.1902	521.1972
GATDH-Gal-Gal	598.222	599.229
GATDH-AcGal-Gal	640.2326	641.2396
GATDH-AcGal-AcGal	682.2432	683.2502
GATDH-Glc	436.169	437.176
GATDH-AcGlc	478.1796	479.1866
GATDH-diAcGlc	520.1902	521.1972
GATDH-GlcNAc	477.195	478.202
GATDH-AcGlcNAc	519.2056	520.2126
GATDH-diAcGlcNAc	561.2162	562.2232
GATDH-Gal-GlcNAc	639.248	640.255

593.2 Da protein modification in addition to 1 PE and 2 PE modifications, respectively. In-source fragmentation of PilE and targeted MS2 of the ion at *m/z* 594.2 (Supplementary data, Fig. 8a) generated a fragmentation pattern consistent with a previously undescribed trisaccharide (Figure 5B). The 203.0 Da difference between the precursor ion at *m/z* 594.2 and the ion at *m/z* 391.2, together with the presence of the reporter ion at *m/z* 204.1 demonstrated that the distal moiety was HexNAc. The ion at *m/z* 391.2, 162.1 Da larger than the diNAcBAc ion (at *m/z* 229.1), corresponded to diNAcBac-Hex, showing that the second position of the trisaccharide was occupied by Gal. Moreover, the presence of only a single ion representing the second sugar at *m/z* 391.2 (diNAcBac-Gal) and not a combination of disaccharide ions at *m/z* 391.2 and at *m/z* 432.2 (diNAcBac-GlcNAc) indicated that PglA and PglH2 operate strictly sequentially, generating only a diNAcBac-Gal-HexNAc glycoform. In the low mass area, ions characteristic of diNAcBac fragmentation (at *m/z* 229.1, *m/z* 211.1 and at *m/z* 168.6) were detected. This is in contrast to our previous results from a strain expressing pilin from an otherwise isogenic strain containing an intact *pglI* allele (Figure 5C), where no 593.2 Da mass addition to PilE was detected (and thus no ion species at *m/z* 17,895.2 nor at *m/z* 18,010.2). Neither was a glycan reporter ion at *m/z* 594.28 detected by in-source fragmentation of PilE from strain KS391 (Supplementary data, Fig. 8b) demonstrating that the presence of the diNAcBac-Gal-HexNAc trisaccharide was associated with the lack of glycan *O*-acetylation.

To further investigate the basis for the conditional expression of the unique trisaccharide, we exploited the availability of a strain expressing a *pglH2*-derived allele that leads to incorporation of Glc (rather than GlcNAc) at the second position (Børud et al. 2014). This allele differs from that which incorporates GlcNAc by a single amino acid substitution at position 303 (alanine instead of glutamine). The deconvoluted mass spectrum derived from this background (RV710) yielded a single abundant diNAcBac-Hex form corresponding to a mixture of Glc and Gal glycoforms (at *m/z* 17,692.1 together with 1 PE and at *m/z* 17,815.1 together with 2 PE) (Figure 5D). That is, no species consistent with a trisaccharide or a HexNAc containing glycoform were detected. (Figure 5D and Supplementary data, Fig 8c). The trisaccharide glycoform is therefore specifically associated with the GlcNAc-incorporating PglH2 glycosyltransferase.

Discussion

The *O*-acetylation of bacterial surface glycoconjugates such as CPS and *O*-specific polysaccharides of LPS often has profound consequences for glycan structure and immunochemistry. In such instances, differential phenotypes almost invariably reflect the simple presence or absence of *O*-acetate on an otherwise unmodified glycan structure.

In this report, we confirm and extend our understanding of *O*-acetylation in the *pgl* systems of Ng and Nm. In concordance with our previous data, we found no evidence for PglI impacting on diNAcBac structure or biosynthesis. Likewise, clear evidence was seen for *O*-acetylation of the Gal residue (associated with *pglA*) as well as the Glc residue (associated with *pglH*) in both diNAcBac- and GATDH-based disaccharides. For the first time, we demonstrated that the GlcNAc residue (associated with *pglH2*) in both diNAcBac- and GATDH-based disaccharides also undergo *O*-acetylation. Similarly, there is now clear evidence that the Gal, Glc and GlcNAc residues in disaccharides were also susceptible to being modified with two acetate groups. In addition, the MS data here revealed the previously unrecognized *O*-acetylation of the distal Gal residue in both diNAcBac- and GATDH-based trisaccharides. It is noteworthy that the levels of *O*-acetylation of the distal Gal in the trisaccharides, although detectable, were relatively low and that in diNAcBac-based trisaccharide such modifications were only seen on glycoforms in which the central Gal residue was also *O*-acetylated. These findings may be indicative of a hierarchical mode of *O*-acetylation in which modification of the third residue relies on modification of the second residue. Another finding of note is the presence of *O*-acetate on the basal GATDH residue.

Perhaps the most striking findings involve the genetic interactions observed between *pglA*, *pglH2* and *pglI*. Strains expressing both PglA and PglH2 together with PglI expressed mixtures of disaccharide forms carrying *O*-acetylated Gal and *O*-acetylated GlcNAc at the nonreducing ends. In a *pglI* null background, an unanticipated and prominent reporter ion was detected together with those corresponding to the expected nonacetate modified disaccharides. CID/HCD MS2 and further fragmentation analysis of this unique reporter ion at *m/z* 594.2 defined it as corresponding to a diNAcBac-Gal-HexNAc trisaccharide. Thus, PglI expression and glycan *O*-acetylation are associated with altered oligosaccharide chain length.

In determining how *O*-acetylation might exert its impact here, we envision two basic possibilities based on knowledge of the *pgl* biosynthetic pathway (Hartley et al. 2011) (illustrated in Figure 6).

Table III. PilE species with corresponding *m/z*

Glycan modifications to pilin/PilE	*m/z* pilin/PilE wo PE	*m/z* pilin/PilE w 1 PE	*m/z* pilin/PilE w 2 PE	
pilin/PilE	17,178.0	17,301.0	17,424.0	±0.2
diNAcBac	17,406.1	17,529.1	17,652.1	±0.2
diNAcBac-Gal	17,568.2	17,691.2	17,814.2	±0.2
diNAcBac-AcGal	17,610.2	17,733.2	17,856.2	±0.2
diNAcBac-diAcGal	17,652.2	17,775.2	17,898.2	±0.2
diNAcBac-Gal-Gal	17,730.2	17,853.2	17,976.2	±0.2
diNAcBac-AcGal-Gal	17,772.2	17,895.2	18,018.2	±0.2
diNAcBac-AcGal-AcGal	17,814.2	17,937.2	18,060.3	±0.2
diNAcBac-Glc	17,568.2	17,691.2	17,814.2	±0.2
diNAcBac-AcGlc	17,610.2	17,733.2	17,856.2	±0.2
diNAcBac-diAcGlc	17,652.2	17,775.2	17,898.2	±0.2
diNAcBac-GlcNAc	17,609.2	17,732.2	17,855.2	±0.2
diNAcBac-AcGlcNAc	17,651.2	17,774.2	17,897.2	±0.2
diNAcBac-diAcGlcNAc	17,693.2	17,816.2	17,939.2	±0.2
diNAcBac-Gal-GlcNAc	17,771.2	17,894.3	18,017.3	±0.2
GATDH	17,452.1	17,575.1	17,698.1	±0.2
AcGATDH	17,494.1	17,617.1	17,740.1	±0.3
GATDH-Gal	17,614.2	17,737.2	17,860.2	±0.2
GATDH-AcGal	17,656.2	17,779.2	17,902.2	±0.2
AcGATDH-AcGal	17,698.2	17,821.2	17,944.2	±0.2
GATDH-Gal-Gal	17,776.2	17,899.2	18,022.2	±0.2
GATDH-AcGal-Gal	17,818.2	17,941.2	18,064.3	±0.2
AcGATDH-AcGal-Gal	17,860.2	17,983.3	18,106.3	±0.2
GATDH-Glc	17,614.2	17,737.2	17,860.2	±0.2
GATDH-AcGlc	17,656.2	17,779.2	17,902.2	±0.2
AcGATDH-AcGlc	17,698.2	17,821.2	17,944.2	±0.2
GATDH-GlcNAc	17,655.2	17,778.2	17,901.2	±0.2
GATDH-AcGlcNAc	17,697.2	17,820.2	17,943.2	±0.2
AcGATDH-AcGlcNAc	17,739.2	17,862.2	17,985.2	±0.2
GATDH-Gal-GlcNAc	17,817.2	17,940.3	18,063.3	±0.2

In one scenario, the unique UndPP-trisaccharide in its *O*-acetylated form would be synthesized but is incompatible with downstream events. These would be the penultimate and final steps in the pathway entailing flipping of the UndPP-glycoform into the periplasm (Figure 6(2)) and its use as a donor substrate by PglO (Figure 6(3)), the protein targeting oligosaccharyltransferase respectively. However, both neisserial flippases in general and the neisserial PglO oligosaccharyltransferase specifically are very promiscuous in substrate utilization (Aas et al. 2007; Faridmoayer et al. 2008). In fact, both steps clearly accommodate a large array of glycoforms in their *O*-acetylated and nonacetylated forms. A more favored scenario then is that *O*-acetylation of the UndPP-diNAcBac-Gal disaccharide blocks its further elaboration by a glycosyltransferase (Figure 6(1)). Given the dependence of this phenomenon on PglH2 (and not PglH), it seems possible that PglH2 has two activities: one that adds GlcNAc to UndPP-diNAcBac and a second that adds a HexNAc to the UndPP-diNAcBac-Gal. This latter activity would be inhibited by *O*-acetylation of the UndPP-disaccharide acceptor and thus only detected in a *pglI* background.

The latter model has potentially significant implications for *O*-acetylation mechanisms of bacterial glycans destined ultimately to an extracytoplasmic location. In perhaps the best studied system involving PG, *O*-acetylation is generally thought to occur in the periplasm (Moynihan et al. 2014). There, translocation of donor acetate groups from the cytoplasm (acetyl-CoA being the presumed source) to the periplasm is thought to be carried out by an integral membrane protein, PatA or homologs thereof. Subsequent transfer to PG is then carried out by PatB, a peripheral membrane *O*-acetyltransferase. In Gram-positive bacteria, a single composite membrane protein, OatA, is proposed to catalyze both reactions of the process. The situation with regard to LPS *O*-acetylation remains less clear. PglI shows clear structural similarities to other predicted *O*-acetylases with the Acyl_transf_3-Pfam PF01757 domain and multiple membrane spanning elements implicated in LPS modifications. Few if any biochemical or enzymatic parameters have been defined for any of them. Assuming that PglI-mediated acetate addition blocks oligosaccharide chain elongation (as opposed to blocking downstream UndPP oligosaccharide processing) and given the fact that oligosaccharide chain extension takes place in the cytoplasm, it stands to reason that *O*-acetylation also takes place on the cytoplasmic side of the inner membrane in this system. Whether this finding can be extended to LPS-related processes remains to be determined. Nonetheless, the cytoplasmic partitioning of this process would be consistent with the absence of any identifiable acetyl-CoA translocating element within the neisserial *pgl* and LPS systems.

We are aware of only one other example of an association between *O*-acetylation and glycan chain elongation. This occurs in association with cell wall biosynthesis in *N. meningitidis* where de-*O*-acetylation of specific PG muropeptides impacts on glycan chain elongation (Veyrier et al. 2013). In contrast to the situation with *pgl* glycoform acetylation, the absence of O-acetylation (mediated by the action of Ape1, a deacetylase) is associated with glycan chain elongation (truncation) whereas in our case O-acetylation blocks glycan extension.

Together, these results reveal heretofore unrecognized levels of complexity and microheterogeneity of *pgl* glycoform *O*-acetylation. Moreover, the finding that *O*-acetylation impacts on oligosaccharide chain length is noteworthy in that *pglI* expression in Nm (but not in

Ng) is subject to phase-variable expression. Thus on–off PglI expression in Nm would result not only in shifts between oligosaccharides with or without acetate modification but also in the basic repertoires of oligosaccharides and the levels of glycoform microheterogeneity. Further studies are required to discern the precise mechanisms by which *O*-acetylation modify oligosaccharide chain length in this system.

Materials and methods

Bacterial strains and culture conditions

All bacterial strains used in this study are Ng, described in Table I, and were grown on conventional GC medium as previously described (Freitag et al. 1995).

Construction of *pgl* mutants

pgl mutations (*pglA, pglEon, pglI, pglH, pglH2, pglB2*) were introduced into various strain backgrounds using transformation of constructs as detailed in Table I and previously described (Aas et al. 2006, 2007; Børud et al. 2010, 2011). Antibiotics were used for selection of transformants at the following concentrations: streptomycin, 750 μg mL^{-1}; erythromycin, 8 μg mL^{-1}; kanamycin, 50 μg mL^{-1}; gradient chloramphenicol, 25 μg ml^{-1}. The *pglE* mutation was made by deletion of the *Hinf*I fragment of *pglE* (549 bp) in the pCRII-*pglE* plasmid, this deletion was spot transformed into the *pglE* gene in the N400 background, and correct *pglE* deletion strains were screened and confirmed by PCR.

Sample preparation and top-down MS analysis of intact protein

PilE purification, sample preparation and top-down ESI-MS on the LTQ Orbitrap were performed as previously described (Vik et al. 2012). Intact protein mass spectra were acquired with a resolution of 60 000 at *m/z* 400. Deconvoluted protein masses are reported as monopronated [M+H+]. Masses of unmodified and modified proteins were determined from calculated theoretical masses, mass differences and previous work. Masses and reporter ions of PilE PTM modifications are reported in Table II. Relevant *m/z* values of modified PilE species are reported in Table III.

Supplementary data

Supplementary data is available at *Glycobiology* online.

Funding

This work was supported by the Research Council of Norway grant [NFR 214442]; the Centre for Integrative Microbial Evolution at the Department of Biosciences and by funds from the University of Oslo Faculty of Mathematics and Natural Sciences.

Conflict of interest statement

None declared.

Abbreviations

CPS, capsular polysaccharides; diNAcBac, *N,N′*-diacetylbacillosamine; ESI MS, electrospray ionization mass spectrometry; Gal, galactose; GATDH, glyceramido-acetamido trideoxyhexose; Glc, glucose; GlcNAc, *N*-acetyl-glucosamine; HexNAc, *N*-acetyl-hexosamine; LPS/LOS, lipopolysaccharide/lipoologosaccharides; mAbs, monoclonal antibodies; MS, mass spectrometry; MS2, CID/HCD fragmentation; Ng, *Neisseria gonorrhoeae*; Nm, *Neisseria meningitidis*; PE, phosphoethanolamine; PG, peptidoglycan; *pgl*, protein glycosylation; UndPP, undecaprenyl diphosphate.

References

Aas FE, Egge-Jacobsen W, Winther-Larsen HC, Lovold C, Hitchen PG, Dell A, Koomey M. 2006. *Neisseria gonorrhoeae* Type Iv Pili undergo multisite, hierarchical modifications with phosphoethanolamine and phosphocholine requiring an enzyme structurally related to lipopolysaccharide phosphoethanolamine transferases. *J Biol Chem*. 281(38):27712–27723.

Aas FE, Vik A, Vedde J, Koomey M, Egge-Jacobsen W. 2007. *Neisseria gonorrhoeae O*-linked pilin glycosylation: functional analyses define both the biosynthetic pathway and glycan structure. *Mol Microbiol*. 65(3):607–624.

Antignac A, Rousselle JC, Namane A, Labigne A, Taha MK, Boneca IG. 2003. Detailed structural analysis of the peptidoglycan of the human pathogen *Neisseria meningitidis*. *J Biol Chem*. 278(34):31521–31528.

Børud B, Aas FE, Vik A, Winther-Larsen HC, Egge-Jacobsen W, Koomey M. 2010. Genetic, structural and antigenic analyses of glycan diversity in the *O*-linked protein glycosylation systems of human neisseria Species. *J Bacteriol*. 192(11):2816–2829.

Børud B, Anonsen JH, Viburiene R, Cohen EH, Samuelsen ABC, Koomey M. 2014. Extended glycan diversity in a bacterial protein glycosylation system linked to allelic polymorphisms and minimal genetic alterations in a glycosyltransferase gene. *Mol Microbiol*. 94(3):688–699.

Børud B, Viburiene R, Hartley MD, Paulsen BS, Egge-Jacobsen W, Imperiali B, Koomey M. 2011. Genetic and molecular analyses reveal an evolutionary trajectory for glycan synthesis in a bacterial protein glycosylation system. *Proc Natl Acad Sci U S A*. 108(23):9643–9648.

Chamot-Rooke J, Rousseau B, Lanternier F, Mikaty G, Mairey E, Malosse C, Bouchoux G, Pelicic V, Camoin L, Nassif X, Duménil G. 2007. Alternative *Neisseria Spp*. Type Iv pilin glycosylation with a glyceramido acetamido trideoxyhexose residue. *Proc Natl Acad Sci U S A*. 104(37):14783–14788.

Clarke AJ, Starting H, Blackburn N. (2000). Pathways for the O-acetylation of bacterial cell wall polymers. In: Doyle RJ, editor, *Glycomicrobiology*. New York: Plenum Publishing Co. Ltd. pp. 187–212.

Faridmoayer A, Fentabil MA, Haurat MF, Yi W, Woodward R, Wang PG, Feldman MF. 2008. Extreme substrate promiscuity of the neisseria oligosaccharyl transferase involved in protein O-glycosylation. *J Biol Chem*. 283(50):34596–34604.

Freitag NE, Seifert HS, Koomey M. 1995. Characterization of the *pilF-pilD* pilus-assembly locus of *Neisseria gonorrhoeae*. *Mol Microbiol*. 16(3):575–586.

Gudlavalleti SK, Datta AK, Tzeng YL, Noble C, Carlson RW, Stephens DS. 2004. The *Neisseria meningitidis* serogroup a capsular polysaccharide O-3 and O-4 acetyltransferase. *J Biol Chem*. 279(41):42765–42773.

Hartley MD, Morrison MJ, Aas FE, Børud B, Koomey M, Imperiali B. 2011. Biochemical caracterization of the *O*-linked glycosylation pathway in *Neisseria gonorrhoeae* responsible for biosynthesis of protein glycans containing *N,N′*-diacetylbacillosamine. *Biochem*. 50(22):4936–4948.

Hegge FT, Hitchen PG, Aas FE, Kristiansen H, Lovold C, Egge-Jacobsen W, Panico M, Leong WY, Bull V, Virji M, et al. 2004. Unique modifications with phosphocholine and phosphoethanolamine define alternate antigenic forms of Neisseria gonorrhoeae type IV Pili. *Proc Natl Acad Sci USA*. 101(29):10798–10803.

Johannessen C, Koomey M, Børud B. 2012. Hypomorphic glycosyltransferase alleles and recoding at contingency loci influence glycan microheterogeneity in the protein glycosylation system of *Neisseria* Species. *J Bacteriol*. 194(18):5034–5043.

Kahler CM, Lyons-Schindler S, Choudhury B, Glushka J, Carlson RW, Stephens DS. 2006. *O*-Acetylation of the terminal *N*-acetylglucosamine of the lipooligosaccharide inner core in *Neisseria meningitidis*. Influence on inner core structure and assembly. *J Biol Chem*. 281(29):19939–19948.

Kahler CM, Martin LE, Tzeng YL, Miller YK, Sharkey K, Stephens DS, Davies JK. 2001. Polymorphisms in pilin glycosylation locus of *Neisseria meningitidis* expressing class II pili. *Infect Immun*. 69(6):3597–3604.

Moynihan PJ, Clarke AJ. 2010. O-acetylation of peptidoglycan in gram-negative bacteria: Identification and characterization of peptidoglycan O-acetyltransferase in *Neisseria gonorrhoeae*. *J Biol Chem*. 285(17): 13264–13273.

Moynihan PJ, Sychantha D, Clarke AJ. 2014. Chemical biology of peptidoglycan acetylation and deacetylation. *Bioorg chem*. 54:44–50.

Power PM, Roddam LF, Rutter K, Fitzpatrick SZ, Srikhanta YN, Jennings MP. 2003. Genetic characterization of pilin glycosylation and phase variation in *Neisseria meningitidis*. *Mol Microbiol*. 49(3):833–847.

Slauch JM, Lee AA, Mahan MJ, Mekalanos JJ. 1996. Molecular characterization of the oafa locus responsible for acetylation of *Salmonella typhimurium* O-antigen: Oafa Is a member of a family of integral membrane transacylases. *J Bacteriol*. 178(20):5904–5909.

Tonjum T, Freitag NE, Namork E, Koomey M. 1995. Identification and characterization of *pilG*, a highly conserved pilus-assembly gene in pathogenic *Neisseria*. *Mol Micro*. 16(3):451–464.

Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, Etzler ME. 2009. *Essentials of Glycobiology*. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press, Essentials of Glycobiology. Varki A, Cummings RD, Esko JD_et al..

Veyrier FJ, Williams AH, Mesnage S, Schmitt C, Taha MK, Boneca IG. 2013. De-O-acetylation of peptidoglycan regulates glycan chain extension and affects in vivo survival of *Neisseria meningitidis*. *Mol Microbiol*. 87(5): 1100–1112.

Vik A, Aas FE, Anonsen JH, Bilsborough S, Schneider A, Egge-Jacobsen W, Koomey M. 2009. Broad spectrum *O*-linked protein glycosylation in the human pathogen *Neisseria gonorrhoeae*. *Proc Natl Acad Sci USA*. 106 (11):4447–4452.

Vik A, Aspholm M, Anonsen JH, Børud B, Roos N, Koomey M. 2012. Insights into type Iv pilus biogenesis and dynamics from genetic analysis of a C-terminally tagged pilin: A role for O-linked glycosylation. *Mol Microbiol*. 85(6):1166–1178.

Warren MJ, Roddam LF, Power PM, Terry TD, Jennings MP. 2004. Analysis of the role of *pglI* in pilin glycosylation of *Neisseria meningitidis*. *FEMS Immunol Med Microbiol*. 41(1):43–50.