Disrupted Synthesis of a Di-*N*-acetylated Sugar Perturbs Mature Glycoform Structure and Microheterogeneity in the O-Linked Protein Glycosylation System of *Neisseria elongata* subsp. *glycolytica*

Nelson Wang,^a Jan Haug Anonsen,^{a,b*} Raimonda Viburiene,^a Joseph S. Lam,^c Åshild Vik,^{a*} Michael Koomey^a

^aDepartment of Biological Sciences, Center for Integrative Microbial Evolution, University of Oslo, Oslo, Norway

^bDepartment of Biological Sciences, Proteomics and Mass Spectrometry Unit, University of Oslo, Oslo, Norway

^cDepartment of Cellular and Molecular Biology, University of Guelph, Guelph, Ontario, Canada

ABSTRACT The genus *Neisseria* includes three major species of importance to human health and disease (*Neisseria gonorrhoeae*, *Neisseria meningitidis*, and *Neisseria lactamica*) that express broad-spectrum O-linked protein glycosylation (Pgl) systems. The potential for related Pgl systems in other species in the genus, however, remains to be determined. Using a strain of *Neisseria elongata* subsp. *glycolytica*, a unique tetrasaccharide glycoform consisting of di-*N*-acetylglucosamine and glucose as the first two sugars followed by a rare sugar whose mass spectrometric fragmentation profile was most consistent with di-*N*-acetyl hexuronic acid and a *N*-acetylhexosamine at the nonreducing end has been identified. Based on established mechanisms for UDP-di-*N*-acetyl hexuronic acid biosynthesis found in other microbes, we searched for genes encoding related pathway components in the *N. elongata* subsp. *glycolytica* genome. Here, we detail the identification of such genes and the ensuing glycosylation phenotypes engendered by their inactivation. While the findings extend the conservative nature of microbial UDP-di-*N*-acetyl hexuronic acid biosynthesis, mutant glycosylation phenotypes reveal unique, relaxed specificities of the glycosyltransferases and oligosaccharyltransferases to incorporate pathway intermediate UDP-sugars into mature glycoforms.

IMPORTANCE Broad-spectrum protein glycosylation (Pgl) systems are well recognized in bacteria and archaea. Knowledge of how these systems relate structurally, biochemically, and evolutionarily to one another and to others associated with microbial surface glycoconjugate expression is still incomplete. Here, we detail reverse genetic efforts toward characterization of protein glycosylation mutants of *N. elongata* subsp. *glycolytica* that define the biosynthesis of a conserved but relatively rare UDP-sugar precursor. The results show both a significant degree of intra- and trans-kingdom conservation in the utilization of UDP-di-*N*-acetyl-glucuronic acid and singular properties related to the relaxed specificities of the *N. elongata* subsp. *glycolytica* system

KEYWORDS *Neisseria*, O-linked glycosylation, commensal, mass spectrometry, microheterogeneity, *pgl*, posttranslational modification

Glycoconjugates varying in function and structure are common on surfaces of bacterial cells. Given their prevalence and significance, it is interesting to understand the forces driving glycan diversification. Defining the selective pressures is difficult because carbohydrates are secondary gene products and the biosynthetic pathways and the glycans themselves are not always completely characterized. Systems

Citation Wang N, Anonsen JH, Viburiene R, Lam JS, Vik Å, Koomey M. 2019. Disrupted synthesis of a di-*N*-acetylated sugar perturbs mature glycoform structure and microheterogeneity in the O-linked protein glycosylation system of *Neisseria elongata* subsp. *glycolytica*. *J Bacteriol* 201:e00522-18. <https://doi.org/10.1128/JB.00522-18>.

Editor Yves V. Brun, Indiana University Bloomington

Copyright © 2018 American Society for Microbiology. All Rights Reserved.

Address correspondence to Michael Koomey, j.m.koomey@ibv.uio.

* Present address: Jan Haug Anonsen, International Research Institute of Stavanger, Randaberg, Norway; Åshild Vik, Research Council of Norway, Oslo, Norway.

Received 28 August 2018

Accepted 4 October 2018

Accepted manuscript posted online 15 October 2018

Published 7 December 2018

N. gonorrhoeae
FA1090

N. elongata subsp. *glycolytica*
ATCC29315

FIG 1 Comparison of *pgl* gene content and synteny between *N. gonorrhoeae* strain FA1090 and *N. elongata* subsp. *glycolytica* strain ATCC 29315. The core *pgl* gene locus of *N. elongata* subsp. *glycolytica* contains *pglG*, *pglH*, a split, two-ORF version of *pglB*, *pglC*, *pglJ*, and *pglK*. *pglD*, *pglO*, and the linked *pglN* and *pglM* alleles are located separately outside the core locus. Gonococcal genomes lack orthologues of *pglJ*, *-K*, *-N*, or *-M* but do have *pglA* and *pglE* (which are absent in *N. elongata* subsp. *glycolytica* strain ATCC 29315).

involving lipopolysaccharides, capsules, and cell wall-associated moieties show robust inter- or intrastrain glycan variation, suggesting that such diversification is adaptive. However, strict correlations between glycosylation genotypes and phenotypes have been difficult to establish. In the same vein, reconciling the glycosylation gene repertoire with particular bacterial populations is challenging. Thus, there remain significant knowledge gaps relating to both the short- and the long-term evolutionary trajectories in these systems.

Bacterial protein glycosylation systems in which multiple, diverse proteins are targeted for modification are continually being identified. In particular, studies have strived to elucidate how these systems are evolving at the species and genus levels. In the case of *N*-linked protein glycosylation exhibited by species within the genus *Campylobacter*, there is conservation of a basic oligosaccharide glycoform within species of the thermotolerant taxa (1). In contrast, an unexpected diversity of *N*-glycans varying in length and structure was seen in nonthermotolerant *Campylobacter* species (1). In *Acinetobacter*, multiple distinct glycoforms associated with general *O*-linked protein glycosylation have been found (2), and a study of *O*-linked protein glycosylation in *Bacteroidetes* revealed that while core glycan structures are conserved across the phylum, terminal glycans are clearly divergent (3).

O-linked protein glycosylation (Pgl) systems expressed by the important human bacterial pathogens *Neisseria gonorrhoeae* and *Neisseria meningitidis* have been well characterized at both the genetic and biochemical levels (Fig. 1). From these studies, a consensus model for neisserial *pgl*-dependent protein glycosylation has been identified (4–8). A NAD⁺-dependent dehydratase (PglC) and aminotransferase (PglD) convert UDP-GlcNAc to UDP-2-acetamido-4-amino-2,4,6-trideoxy- α -D-glucose (8). The bifunctional enzyme (PglB) then catalyzes the amino acetylation of the above-mentioned precursor to form UDP-di-*N*-acetylbacillosamine (diNAcBac) and the subsequent transfer of the phosphosugar to the lipid carrier undecaprenyl phosphate (Und-P) (8). PglB2, encoded by *pglB2* alleles found in some *N. meningitidis* strains, contains an altered C-terminal domain involved in the transfer of a glycerol moiety (instead of the acetyl group) to produce 4-glyceramido-2-acetamido-2,4,6-trideoxy- α -D-hexose (GATDH) (9). Subsequent elaboration of these undecaprenyl diphosphate (Und-PP) monosaccharide sugars occurs via two parallel pathways using distinct glycosyltransferases. One involves PglH or its allelic variant-encoded PglH2, which attaches a Glc or GlcNAc, respectively, to the Und-PP-monosaccharides. The second pathway utilizes the PglA

FIG 2 Proposed biosynthetic pathway of UDP-GlcNAc(3NAc)A in *N. elongata* subsp. *glycolytica*. The *N. elongata* subsp. *glycolytica* strain ATCC 29315 genome contains four ORFs whose corresponding products could function in UDP-GlcNAc(3NAc)A biosynthesis. The experimentally defined *P. aeruginosa* WbpABED pathway components are shown in gray (15). Identity percentages are shown in parentheses. *N. elongata* subsp. *glycolytica* Pgl homologue designations: Nelon10580, PglJ; Nelon10585, PglK; Nelon05890, PglM; and Nelon05885, PglN.

and PglE glycosyltransferases to add successive Gal units to produce a trisaccharide (4, 10). The PglH/PglH2-generated Und-PP-disaccharides are recalcitrant to further extension by PglE (5, 6). However, strains simultaneously expressing both PglA and PglH can generate both glycoforms, and all di- and trisaccharide forms can be further modified via O-acetylation mediated by PglI (4, 5, 11).

O-linked protein glycosylation is not a classical virulence determinant as it is not differentially distributed among disease-associated and carriage strains of *N. meningitidis* and is found in nonpathogenic species such as *Neisseria lactamica* and *Neisseria elongata* subsp. *glycolytica* (7). Despite this, the glycoform structure appears to be under positive, diversifying selection in both *N. gonorrhoeae* and *N. meningitidis*, and major shifts in *pgl* genotypes have been documented within otherwise stable clonal complexes due to horizontal gene transfer (12, 13).

We sought to gain a broader understanding of *pgl* biology by examining a strain of *N. elongata* subsp. *glycolytica* and identified an unrecognized tetrasaccharide glycoform consisting of di-*N*-acetylbaucosamine, glucose, di-*N*-acetylhexuronic acid, and *N*-acetylhexosamine [diNAcBac-Glc-HexNAc(3NAc)A-HexNAc] (14). In that background, the *pglBCD-pglH* pathway generates a disaccharide, diNAcBac-Glc, and is extended by the product of *pglG*, a gene found in the *pgl* loci of most pathogenic *Neisseria* spp., to form diNAcBac-Glc-HexNAc(3NAc)A (6, 7) (Fig. 1). The glycosyltransferase responsible for the addition of the terminal HexNAc in the *N. elongata* subsp. *glycolytica* glycoform has yet to be identified.

Glycan serotyping utilizing previously established antibodies and mass spectrometric (MS) data identified the glycan components of the tetrasaccharide with the exception of that defined as diNAcHexA, for which only MS-based fragmentation data were available (14). Conserved pathways for HexNAc(3NAc)A biosynthesis have been defined based on data collected from mutagenesis and biochemical experiments in *Pseudomonas aeruginosa* and *Bordetella pertussis*, where it is found in the lipopolysaccharides (LPS) of some strains (15), and in archaeal *Methanococcus* species, where it is a component of the *N*-linked protein glycosylation glycan (16). Studies in *P. aeruginosa* PAO1 have shown that GlcNAc(3NAc)A is synthesized by four enzymes in the Wbp pathway, sequentially using WbpA, WbpB, WbpE, and WbpD, starting from the precursor UDP-GlcNAc (Fig. 2) (15). WbpA oxidizes the C-6 hydroxyl group to form UDP-GlcNAcA, followed by the action of the dehydrogenase WbpB and aminotransferase WbpE to yield UDP-GlcNAc(3NH₂)A. WbpB and WbpE act as a dehydrogenase/aminotransferase pair converting UDP-GlcNAcA to UDP-GlcNAc(3NH₂)A in a coupled reaction via a NAD(+) recycling pathway that uses α -ketoglutarate as an oxidant. Next, the acetyltransferase WbpD catalyzes acetylation of UDP-GlcNAc(3NH₂)A to yield UDP-

FIG 3 Cross-complementation of *P. aeruginosa* PAO1 *wbp* null mutants by the respective *N. elongata* subsp. *glycolytica* *pgl* genes. Immunoblotting of LPS preparations was performed using MAb MF15-4, which is specific against O-specific antigen (OSA) of *P. aeruginosa*.

GlcNAc(3NAc)A (Fig. 2). To confirm the identity of the third sugar moiety of the *N. elongata* subsp. *glycolytica* oligosaccharide, we identified genes orthologous to those required for HexNAc(3NAc)A synthesis in these other microbes. Here, we report the isolation of mutants with defined mutations in the *N. elongata* subsp. *glycolytica* genes and the ensuing effects on *pgl* glycoform expression.

RESULTS

Identification of genes involved in UDP-HexNAc(3NAc)A biosynthesis. Using the *P. aeruginosa* PAO1 Wbp proteins in queries of the annotated *N. elongata* subsp. *glycolytica* genome sequence, significantly related open reading frames (ORFs) were found for the first four enzymes in the pathway (Fig. 2). NELON_10580 shared 76% sequence identity to WbpA, while NELON_10585 shared 20% sequence identity to WbpB. The two associated genes were tandemly arrayed just 3' of *pglC* (encoding an aminotransferase essential for the synthesis of UDP-*N,N'*-diacetyl bacillosamine). NELON_05890 and NELON_05885 shared 75% and 76% similarities to WbpE and WbpD, respectively, and their genes were also oriented tandemly at a locus far removed from the so-called core *pgl* locus (Fig. 1). Using UniProtKB, all four *N. elongata* subsp. *glycolytica* ORFs were annotated, and their products were predicted to have enzymatic activity consistent with the Wbp enzymes. We were unsuccessful in identifying an *N. elongata* subsp. *glycolytica* orthologue for the WbpI epimerase. Amino acid alignments of the predicted products from *N. elongata* subsp. *glycolytica* with their counterparts in both *P. aeruginosa* PAO1 and *B. pertussis* are shown in Fig. S1 in the supplemental material.

Cross-genus complementation restores OSA biosynthesis in *P. aeruginosa*. To verify that the four identified *N. elongata* subsp. *glycolytica* gene-encoded enzymes can support the synthesis of UDP-GlcNAc(3NAc)A, we attempted to complement a panel of defined *P. aeruginosa* PAO1 mutants bearing corresponding null mutations in the O-antigen pathway. We introduced plasmids expressing each of the *N. elongata* subsp. *glycolytica* ORFs into the corresponding *P. aeruginosa* PAO1 *wbpA*, *wbpB*, *wbpE*, and *wbpD* mutant strains and investigated the status of O-specific antigen (OSA) expression. In *P. aeruginosa*, smooth LPS containing OSA runs on an SDS-polyacrylamide gel as a ladder that represents individual LPS species with differing numbers of O-antigen repeats (15). The LPS forms can be specifically detected by their reactivity in immunoblotting utilizing monoclonal antibody (MAb) MF15-4. Previous work has shown that restitution of OSA O-antigen expression requires plasmid-based transcomplementation with the individual corresponding gene (15). We prepared LPS from the *wbp* mutants and strains complemented with the putative *N. elongata* subsp. *glycolytica* orthologues and carried out immunoblotting analyses. LPS of PAO1 mutants individually lacking *wbpA*, *wbpB*, *wbpE*, or *wbpD* failed to react with MAb MF15-4, whereas LPS purified from strains complemented with the putatively orthologous *N. elongata* subsp. *glycolytica* *pgl* genes showed reactivity with the MAb (Fig. 3). These results demonstrate that the *N. elongata* subsp. *glycolytica* genes are orthologous to their *P. aeruginosa* PAO1

counterparts, and we have designated them, according to *pgl* nomenclature, *pglJ* (NELON_10580), *pglK* (NELON_10585), *pglM* (NELON_05890), and *pglN* (NELON_05885).

Effects of *pgl* null mutations on *N. elongata* subsp. *glycolytica* protein glycosylation. We next examined the effects of loss-of function-mutations in the newly identified *N. elongata* subsp. *glycolytica* *pgl* genes on glycoform expression and protein glycosylation. Accordingly, we constructed a panel of *N. elongata* subsp. *glycolytica* mutants in which each of the four genes was individually deleted using a previously described strategy (14). These mutations were created in an *N. elongata* subsp. *glycolytica* background expressing a *nirK* allele in which its ORF was tagged with a C-terminal hexahistidine (6×His) extension. NirK was then affinity purified from all four mutant backgrounds, subjected to proteolytic digestion, and examined by liquid chromatography-tandem mass spectrometry (LC-MS²) analysis. As previously described (14), the MS² spectrum of the wild-type *N. elongata* subsp. *glycolytica* tetrasaccharide displays a series of glycan-specific oxonium reporter ions: diNAcBac at *m/z* 211.110, diNAcBac-Hex at *m/z* 391.170, diNAcBac-Hex-diNAcHexA at *m/z* 649.254, and HexNAc at *m/z* 204.086.

The LC-MS² extracted-ion chromatograms (XIC) of chymotrypsin-generated NirK-His from a *pglJ* background revealed overlapping peaks representing diNAcBac and diNAcBac-Hex glycan reporter ions (Fig. 4A). This indicated that the glycopeptides detected were modified solely with a diNAcBac-Hex disaccharide and that the *pglJ* mutant phenocopied a *pglG* null mutant (14). Confirmation that the peaks representing the glycan reporter ions were derived from glycopeptides can be seen in a representative MS² spectrum selected from the XIC because of the overlapping reporter ion peaks. In the chromatogram of the chymotrypsin-generated glycopeptide 371KGTGA AAGAASGASGASAPAAPASSASGSSNPY403 (Fig. 4B), two glycan reporter ions representing diNAcBac (at *m/z* 211.1) and diNAcBac-Hex (at *m/z* 391.2) (i.e., a disaccharide) were detected in the low-mass area.

Based on the data from prior studies and the complementation data presented here, we anticipated that subsequent disruption of the GlcNAc(3NAc)A biosynthetic pathway in *N. elongata* subsp. *glycolytica* mutants might lead to incorporation of oligosaccharide moieties with distinct, signature reporter ions for each blocked step, i.e., diNAcBac-Hex-HexNAcA (*m/z* 608.229), diNAcBac-Hex-HexNAc(3keto)A (*m/z* 607.245), and diNAcBac-Hex-HexNAc(3NH₂)A (*m/z* 607.222), each together with a terminal HexNAc (*m/z* 204.1). The LC-MS² chromatogram and representative spectrum of the corresponding glycopeptide from the *pglK* background displayed reporter ions representing a tetrasaccharide with HexNAcA (at *m/z* 218.1) rather than the diNAcHexA at the third position, at *m/z* 608.2 (together with diNAcBac [*m/z* 211.1], diNAcBac-Hex [*m/z* 391.2], and a HexNAc [*m/z* 204.1]) (Fig. 4C and D). In addition, this sample yielded evidence for microheterogeneity as seen by the simultaneous presence of a population of peptides bearing just the disaccharide diNAcBac-Hex (*m/z* 391.2) across the XIC of the specific glycan reporter ions. Interestingly, the same tetrasaccharide reporter ions and signs of disaccharide-associated microheterogeneity seen in the *pglK* background were detected in both the chromatogram and spectrum of the NirK-His glycopeptide from the *pglM* background (Fig. 4E and F). Identical findings for the latter two mutants could be attributed to the likelihood that *N. elongata* subsp. *glycolytica* PglK and PglM act as a dehydrogenase/aminotransferase pair that converts UDP-GlcNAcA to UDP-GlcNAc(3NH₂)A in an obligatory, coupled reaction involving NAD(+) recycling, as previously detailed for WbpB and WbpE in *P. aeruginosa* (17).

For the glycopeptide from the *pglN* background, reporter ions consistent with the presence of the two distinct glycoforms seen in the *pglK* and *pglM* samples were evident. In addition, signals indicative of HexNAc(NH₂)A-containing tetrasaccharide were seen in the chromatogram (Fig. 4G and H). In the MS² spectrum of the triply charged precursor ion at *m/z* 1172.907 (Fig. 4H), dominant reporter ions at *m/z* 607.3 and at *m/z* 810.4 corresponding to diNAcBac-Hex-HexNAc(NH₂)A and diNAcBac-Hex-HexNAc(NH₂)A-HexNAc, respectively, were detected. However, it is likely that the MS spectrum (Fig. 4H) contains the fragmentation pattern of both a diNAcBac-Hex-

FIG 4 Effect of *pgl* gene mutations on glycan structure. Shown are LC-MS² chromatograms of proteolytically derived peptides from affinity-purified NirK from *pgl* mutants. Total ion chromatogram (TIC) intensity values represent the amount of peptides entering the mass spectrometer. The selected ion (Continued on next page)

HexNAc(NH₂)A-HexNAc and the diNAcBac-Hex-HexNAcA-HexNAc seen in both the *pglM* and *pglK* backgrounds. This is based on the finding that the ¹³C isotope of each of *m/z* 607.3 and *m/z* 810.4 (i.e., *m/z* 608.3 and *m/z* 811.4, respectively) is more abundant than predicted. These observations suggest the presence of diNAcBac-Hex-HexNAcA and diNAcBac-Hex-HexNAcA-HexNAc glycoforms along with the presence of the reporter ion at *m/z* 218.1 corresponding to HexNAcA.

Together, these results show that PglG has relaxed specificity as evidenced by its ability to incorporate altered HexNAcA-derived sugars resulting from disruptions in the *pglJKMN* biosynthetic pathway. Thus, *pglG* and *pglJ* seem to fulfil the minimal requirements for further elaboration of the disaccharide form. Moreover, the unidentified glycosyltransferase responsible for tetrasaccharide formation must also have relaxed specificity to recognize Und-PP-trisaccharide substrates terminating with either HexNAcA, HexNAc(NH₂)A, or HexNAc(NAc)A.

To corroborate the MS findings, we examined the status of *N. elongata* subsp. *glycolytica* NirK in a panel of *pgl*-relevant backgrounds by immunoblotting using an antibody specific for a tetra-His epitope and a rabbit polyclonal serum specific for diNAcBac-Glc disaccharide epitopes (pDAb2). As previously reported (14), detection using the poly-His-recognizing antibody revealed a clear increase in the relative mobility of tagged-NirK in a *pglC* null background (totally defective in glycosylation) versus that seen for the wild-type background. In neither case did NirK-His react with the pDAb2 antibody. In the *pglH* and *pglG* backgrounds, NirK showed slightly decreased mobility relative to that seen in the *pglC* background (the mono- and disaccharide glycoforms are difficult to resolve on an SDS-polyacrylamide gel) but increased mobility compared to that in the wild-type background. However, as expected, only in the *pglG* background was NirK reactive with antiserum pDAB2, which is specific for the diNAcBac-Glc disaccharide (Fig. 5).

NirK from the *pglJ* background phenocopied the reactivity pattern of the *pglG* background, as no discernible differences between the two backgrounds toward either the anti-His or the pDAb2 antibodies were detected (Fig. 5). In contrast, NirK from *pglK*, *pglM*, and *pglN* all showed similar relative mobilities intermediate to those of the wild-type and *pglG* backgrounds using SDS-PAGE. The “ladder-like banding” observed

FIG 4 Legend (Continued)

chromatograms (XIC) are of the four glycan reporter ions characteristic for a tetrasaccharide, as well as the predicted glycan reporter ions originating from oligosaccharides of the *pglJKMN* pathway: diNAcBac at *m/z* 211.108 (1), diNAcBac-Hex at *m/z* 391.170 (2), diNAcBac-Hex-HexNAcA at *m/z* 608.229 (3), diNAcBac-Hex-(Keto)HexNAcA at *m/z* 607.222 (4), diNAcBac-Hex-(amino)HexNAcA at *m/z* 607.245 (5), diNAcBac-Hex-diNAcHexA at *m/z* 649.254 (6), and HexNAc at *m/z* 204.086 (7). All glycan reporter ions were searched with a 10-ppm window. The MS² spectrum demonstrates the presence of glycan reporter ions (marked in boldface type and numbered as described above). All XIC were normalized according to their abundance, either to the most abundant glycan reporter ion or, for the less abundant reporter ion, to the abundance of the overall least abundant reporter ion, the diNAcBac-Hex-(amino)HexNAcA ion at *m/z* 607.245. (A) LC-MS² chromatogram of trypsin-derived peptides from affinity-purified NirK-His from a *pglJ* background (RV727). The total ion chromatogram intensity value was set at 4E8. All selected ion chromatogram values were normalized to the most abundant glycan reporter ion, diNAcBac-Hex at 2.0E5. (B) MS² spectrum of the singly modified chymotrypsin-generated peptide ³⁷¹KGTGAAAGAASGASGASAPAAPASSASGSSNPY⁴⁰³ from NirK-His from panel A carrying a diNAcBac-Hex disaccharide. (C) LC-MS² chromatogram of chymotrypsin-derived peptides from affinity-purified NirK-His from a *pglK* background (RV731). The total ion chromatogram intensity value was set at 4E8. Abundant XIC values were normalized to the most abundant glycan reporter ion, diNAcBac-Hex-HexNAcA at 2.0E5. (D) MS² spectrum of the singly modified chymotrypsin-generated peptide ³⁷¹KGTGAAAGAASGASGASAPAAPASSASGSSNPY⁴⁰³ from NirK-His from panel C carrying a diNAcBac-Hex-HexNAcA-HexNAc tetrasaccharide. (E) LC-MS² chromatogram of chymotrypsin-derived peptides from affinity-purified NirK-His from a *pglM* background (RV735). The XIC intensity value was set at 4E8. Abundant XIC values were normalized to the most abundant glycan reporter ion, diNAcBac-Hex-HexNAcA at 2.0E5. (F) MS² spectrum of the chymotrypsin-generated peptide ³⁷¹KGTGAAAGAASGASGASAPAAPASSASGSSNPY⁴⁰³ from NirK-His from panel E carrying a diNAcBac-Hex-HexNAcA-HexNAc tetrasaccharide. (G) LC-MS² chromatogram of chymotrypsin-derived peptides from affinity-purified NirK-His from a *pglN* background (RV739). The total ion chromatogram intensity value was set at 4E8. The XIC intensity value was set at 4E8. Abundant XIC values were normalized to the most abundant glycan reporter ion, diNAcBac-Hex-HexNAcA at 2.0E5. (H) MS² spectrum of the chymotrypsin-generated peptide ³⁷¹KGTGAAAGAASGASGASAPAAPASSASGSSNPY⁴⁰³ from NirK-His from panel G carrying a diNAcBac-Hex-(amino)HexNAcA-HexNAc tetrasaccharide.

FIG 5 Immunoblotting analyses of *N. elongata* subsp. *glycolytica* *pgI* mutants reveals associated, altered glycoprotein mobility. (Top) Immunoblot analysis of whole-cell lysates from wild-type (wt) (KS992), *pgIC* (KS994), *pgIH* (KS1052), *pgIG* (KS1034), *pgIJ* (RV727), *pgIK* (RV731), *pgIM* (RV735), and *pgIN* (RV739) strains using a tetra-His epitope-recognizing antibody (detecting the NirK-His glycoprotein). (Bottom) Immunoblot analysis of the same samples as those described above with the diNAcBac-Glc epitope-recognizing polyclonal antibody pDAb2.

here was also detected by both the anti-His and pDAb2 antibodies (Fig. 5). In the case of NirK in a *pgIG* background, earlier MS work revealed microheterogeneity as evidenced by the detection of both mono- and disaccharide glycans (14), and similar findings are reported here for the *pgIK*, *pgIM*, and *pgIN* backgrounds (Fig. 4). In this context, it is also important to recall that *N. elongata* subsp. *glycolytica* NirK has at least five glycan attachment sites (14). Hence, the observed banding pattern exhibited in these backgrounds could be indicative of a variable number of sites occupied by glycans per protein as well as proteins carrying differing glycoforms or even individual proteins carrying mixtures of glycoforms.

Other glycoproteins detected by the diNAcBac-Glc disaccharide antibodies in the *pgIK*, *pgIM*, and *pgIN* backgrounds also show signs of altered migration versus those in the other backgrounds (for example, the CcoP glycoprotein band that migrates just above the NirK band). Therefore, the observed electrophoretic migration anomalies are not unique to NirK.

The immunoblotting data from the *pgIJ*, *pgIK*, *pgIM*, and *pgIN* backgrounds were consistent with the expression of NirK bearing differing glycoforms or mixtures of glycoforms per protein. Based on these data alone, however, one cannot formally rule out the potential influence of differences in site occupancy (microheterogeneity) potentially occurring in these backgrounds on SDS-PAGE migration.

DISCUSSION

In this study, we identified four genes involved in the synthesis of the rare sugar 2,3-diacetamido-2,3-dideoxy-D-glucuronic acid, which is a component of the tetrasaccharide expressed by the broad-spectrum O-linked protein glycosylation system of *N. elongata* subsp. *glycolytica*. These findings reveal an extension of a highly conserved pathway for the production of HexNAc(3NAc)A sugar and its derivatives associated with LPS structures in some strains of *P. aeruginosa* (17) and *B. pertussis* (15) and protein N-linked glycans in the archaeal species *Methanococcus voltae* (16) and *Methanococcus maripaludis* (18). Although we lack nuclear magnetic resonance (NMR) data to define glycan stereochemistry, our results, together with the apparent absence of any ORF within the *N. elongata* subsp. *glycolytica* genome sharing significant relatedness to established UDP-GlcNAc(3NAc)A epimerases [converting the glycan to a UDP-ManNAc(3NAc)A], suggest that the likely end product of the *N. elongata* subsp. *glycolytica* hexuronic acid pathway is a glucuronic acid-based sugar. HexNAc(3NAc)A- and

diNAcBac-containing glycoforms have also been identified in the general O-linked protein glycosylation systems of some *Acinetobacter* species (2). Although *Neisseria* falls within the betaproteobacteria and *Acinetobacter* within the gammaproteobacteria, it seems likely that there may be some common ancestry behind their shared glycans and *pgl* gene repertoires.

The pathway for GlcNAc(3NAc)A sugar synthesis based on biochemistry experiments is well defined in other bacterial biosynthesis systems (15, 17, 18), all of which entail stepwise assembly of a glycan onto polyprenyl-phosphates. In contrast, the *in vivo* steps at which defects in particular genes and their products associated with GlcNAc(3NAc)A synthesis might disrupt downstream oligosaccharide synthesis are unclear. Formally, mutations causing defects in synthesis of intermediates through to UDP-GlcNAc(3NAc)A from UDP-GlcNAcA could result in failure of relevant glycosyltransferases to recognize these precursors, failure of subsequently acting glycosyltransferases to utilize lipid-linked oligosaccharides bearing the aberrant sugar, or failure of the targeting oligosaccharyl-transferase to utilize an altered lipid-linked precursor. In both *P. aeruginosa* and *B. pertussis*, disruption of the genes associated with the UDP-diNAcManA pathway cause lack of insertion of any glycan into OSA LPS and LPS, respectively (15). In the *M. maripaludis* N-linked system, glycoproteins from mutants individually lacking any of the four enzymes required for the synthesis of UDP-GlcNAc(3NAc)A from UDP-GlcNAc carry only a truncated monosaccharide (18). The corresponding *N. elongata* subsp. *glycolytica* *pgl* system mutants showed clear evidence of usage of the noncanonical UDP-sugar. While the *pglJ* dehydrogenase null mutant-derived glycoprotein bore only the diNAcBaC-Glc disaccharide, the glycoproteins from mutants lacking PglK, PglM, or PglN carried mixtures of di-, tri-, and tetrasaccharide glycoforms. In addition, the oxonium ion fragmentations patterns for the tri- and tetrasaccharide glycoforms in LC-MS² analyses were consistent with the presence of the glucuronic acid-derived intermediate predicted from a disrupted biosynthetic pathway. It follows then that the PglG glycosyl-transferase has relaxed specificity for the UDP-sugar donor that minimally requires UDP-glucuronic acid. Moreover, the unidentified glycosyltransferase responsible for the addition of the fourth sugar should also have relaxed specificity for the Und-PP-linked substrate, minimally requiring a terminally situated glucuronic acid.

In conclusion, we have identified a conserved enzymatic pathway for the biosynthesis of di-N-acetylated glucuronic acid in a deeply branching species in the genus *Neisseria*. A current working model of the O-linked glycosylation pathway as expressed in *N. elongata* subsp. *glycolytica* is presented in Fig. 6. The findings presented here show that both PglG and PglJ are required for the extension of the growing oligosaccharide chain before the full-length tetrasaccharide glycan can be transferred to its target glycoproteins. We find it interesting that this and related moieties are so widely but relatively sparsely distributed as an oligosaccharide constituent across bacterial O-linked protein glycosylation, LPS biosynthetic, and capsular polysaccharide systems as well as in N-linked protein glycosylation in archaeal species. However, we currently have little knowledge of the structure-function relationships accounting for the broad but restricted distribution of this modified sugar form. Furthermore, hexuronic acid-related sugars are uncommon as components of oligosaccharides in neisserial species, having been identified only in the meningococcal capsular polysaccharides of serogroup I and K variants as N-acetylmannosaminuronic acid units (19). That said, there is relatively little information on the distribution of O-linked protein glycosylation and potential associated glycoforms in commensal neisserial species. The findings and methods developed here should help address these concerns.

MATERIALS AND METHODS

Bacterial strains and culture conditions. The strains used are described in Table S1 in the supplemental material and were derived from *N. elongata* subsp. *glycolytica* strain ATCC 29315. *N. elongata* subsp. *glycolytica* strains were cultured on conventional GC plates (Difco) as previously described (14). Mutants were obtained by transforming *N. elongata* subsp. *glycolytica* with crude cell lysates, PCR-amplified DNA, or purified plasmid DNA. Transformation of *N. elongata* subsp. *glycolytica* was performed following a protocol described previously for *N. gonorrhoeae* (20), and mutants were selected

FIG 6 Current working model of the broad-spectrum O-linked glycosylation pathway expressed in *N. elongata* subsp. *glycolytica*. The findings discussed in this paper show that PglJ, PglK, PglM, and PglN are all required for the biosynthesis of UDP-GlcNAc(3NAc)A. The PglG glycosyltransferase uses this as a donor to extend the disaccharide chain. A yet-to-be-identified glycosyltransferase (PglI?) transfers the final HexNAc onto the terminal position to generate the tetrasaccharide, which then gets flipped by PglF into the periplasm. The PglO oligosaccharyltransferase transfers the glycan onto glycoproteins.

on plates containing appropriate antibiotics as follows: kanamycin (50 μ g/ml), tetracycline (20 μ g/ml), and chloramphenicol (30 μ g/ml). The antibiotics used for selection in *Escherichia coli* were kanamycin (50 μ g/ml) and tetracycline (15 μ g/ml), and those for *P. aeruginosa* strains were gentamicin (200 μ g/ml) and tetracycline (100 μ g/ml). Plasmid transformations in *P. aeruginosa* were done using an adapted electroporation protocol detailed previously (21), while for *E. coli* plasmids were chemically transformed into TOP10 cells (Invitrogen).

Plasmid and strain construction. Information about *N. elongata* subsp. *glycolytica*-related plasmids and strains used in this study is in the supplemental material.

Construction of cross-complementation plasmids. For complementation of the defined *wbpA*, *wbpB*, *wbpD*, and *wbpE* *P. aeruginosa* deletion mutants (15), *N. elongata* subsp. *glycolytica* genes were cloned into the pUCP26 shuttle vector and used for *P. aeruginosa* complementation analyses. An EcoRI restriction site and the optimized Shine-Dalgarno sequence AGGAGGACAAGCT were included in the forward primer (which contained the start codon of the respective gene). Reverse primer pairs included a short sequence downstream of the stop codon and a HindIII restriction site. Primer sequences are listed in Table S2 in the supplemental material. PCR products were digested accordingly and ligated into an EcoRI/HindIII-digested pUCP26 vector. The resulting plasmids were chemically transformed into TOP10 *E. coli* cells (Invitrogen) and isolated using Qiagen's small-scale miniprep kit, and DNA was sequenced at GATC Sanger sequencing (GATC Biotech, Germany).

Cross-complementation studies and analysis of LPS. Plasmids containing *N. elongata* subsp. *glycolytica* genes were transformed into the corresponding *P. aeruginosa* knockout strains. The unmodified pUCP26 vector was used as a negative control. LPS samples from the resulting strains were then used in complementation analyses by immunoblotting. A small-scale LPS preparation of *P. aeruginosa* strains was prepared as described previously (22). In brief, cells were harvested from overnight liquid cultures, and equal numbers of cells (as measured by optical density [OD]) were resuspended in lysis buffer (22). Samples were then boiled for 30 min, treated with proteinase K, and incubated at 55°C overnight. Equal volumes of LPS samples were run on a 12% Criterion Bis-Tris protein gel (Bio-Rad) and immunoblotting was conducted with the primary monoclonal antibody MAb MF 15-4 (specific for B-band O-antigen) (23).

Affinity purification of glycoproteins and in-gel digestion. Purification of His-tagged proteins and in-gel digestion of Coomassie blue-stained gel slices containing purified proteins with chymotrypsin (Sigma) were performed as previously described (24, 25).

Reverse-phase LC-MS² analysis of tryptic peptides. Reverse-phase nanoflow LC-MS and MS² analysis (nano-LC-MS²) of chymotrypsin-derived peptides was performed as previously described (14).

Immunoblotting of whole-cell lysates. Whole-cell lysates were prepared from overnight cultures grown for 16 to 18 h. Cells were scraped from plates, resuspended in water, and diluted to a final OD of 1, when an equal volume of sample loading buffer was added. Lysates were heated at 80°C for 15 min and saved at -20°C. Samples were run on a 12% Criterion Bis-Tris protein gel (Bio-Rad), and immunoblotting was done with primary antibodies pDAB2 (diNAcBac-Glc epitope-recognizing polyclonal antibody) (5) or tetra-His-recognizing antibodies (Qiagen) as previously described (14).

Genome databases. Genome analysis was done by using deposited sequences found in the Bacterial Isolate Genome Sequence Database (BIGSdb) (26) and from the National Center for Biotechnology Information (NCBI) database servers. *N. elongata* subsp. *glycolytica* accession numbers were found using the Universal Protein Resource (UniProt) (27).

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/JB.00522-18>.

SUPPLEMENTAL FILE 1, PDF file, 0.3 MB.

ACKNOWLEDGMENTS

This research was supported in part by Research Council of Norway grants 166931, 183613, and 183814, by the Center for Integrative Microbial Evolution at the Department of Biosciences, University of Oslo, and by funds from the University of Oslo, Faculty of Mathematics and Natural Sciences. This publication made use of the *Neisseria* Multi Locus Sequence Typing website (<http://pubmlst.org/neisseria/>) developed by Keith Jolley and sited at the University of Oxford. The development of this site has been funded by the Wellcome Trust and European Union. This study also made use of the Meningitis Research Foundation Meningococcus Genome Library (<http://www.meningitis.org/research/genome>) developed by Public Health England, the Wellcome Trust Sanger Institute, and the University of Oxford as a collaboration. That project is funded by the Meningitis Research Foundation.

REFERENCES

- Nothaft H, Scott NE, Vinogradov E, Liu X, Hu R, Beadle B, Fodor C, Miller WG, Li J, Cordwell SJ, Szymanski CM. 2012. Diversity in the protein *N*-glycosylation pathways within the *Campylobacter* genus. *Mol Cell Proteomics* 11:1203–1219. <https://doi.org/10.1074/mcp.M112.021519>.
- Scott NE, Kinsella RL, Edwards AV, Larsen MR, Dutta S, Saba J, Foster LJ, Feldman MF. 2014. Diversity within the *O*-linked protein glycosylation systems of *Acinetobacter* species. *Mol Cell Proteomics* 13:2354–2370. <https://doi.org/10.1074/mcp.M114.038315>.
- Coyne MJ, Fletcher CM, Chatzidakis-Livanis M, Posch G, Schaffer C, Comstock LE. 2013. Phylum-wide general protein *O*-glycosylation system of the *Bacteroidetes*. *Mol Microbiol* 88:772–783. <https://doi.org/10.1111/mmi.12220>.
- Aas FE, Vik Å, 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: 607–624. <https://doi.org/10.1111/j.1365-2958.2007.05806.x>.
- Borud 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* 108:9643–9648. <https://doi.org/10.1073/pnas.1103321108>.
- Borud B, Anonsen JH, Viburiene R, Cohen EH, Samuelsen AB, 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* 94:688–699. <https://doi.org/10.1111/mmi.12789>.
- Borud B, Aas FE, Vik Å, 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* 192:2816–2829. <https://doi.org/10.1128/JB.00101-10>.
- Hartley MD, Morrison MJ, Aas FE, Borud B, Koomey M, Imperiali B. 2011. Biochemical characterization of the *O*-linked glycosylation pathway in *Neisseria gonorrhoeae* responsible for biosynthesis of protein glycans containing *N,N'*-diacetylglucosamine. *Biochemistry* 50:4936–4948. <https://doi.org/10.1021/bi2003372>.
- Chamot-Rooke J, Rousseau B, Lanterrier F, Mikaty G, Mairey E, Malosse C, Bouchoux G, Pelicic V, Camoin L, Nassif X, Dumenil G. 2007. Alternative *Neisseria* spp. type IV pilin glycosylation with a glyceramido acetamido tri-deoxyhexose residue. *Proc Natl Acad Sci U S A* 104: 14783–14788. <https://doi.org/10.1073/pnas.0705335104>.
- Power PM, Roddam LF, Rutter K, Fitzpatrick SZ, Srikantha YN, Jennings MP. 2003. Genetic characterization of pilin glycosylation and phase variation in *Neisseria meningitidis*. *Mol Microbiol* 49:833–847.
- Anonsen JH, Borud B, Vik Å, Viburiene R, Koomey M. 2017. Structural and genetic analyses of glycan *O*-acetylation in a bacterial protein glycosylation system: evidence for differential effects on glycan chain length. *Glycobiology* 27:888–899. <https://doi.org/10.1093/glycob/cwx032>.
- Lamelas A, Harris SR, Roltgen K, Dangy JP, Hauser J, Kingsley RA, Connor TR, Sie A, Hodgson A, Dougan G, Parkhill J, Bentley SD, Pluschke G. 2014. Emergence of a new epidemic *Neisseria meningitidis* serogroup A clone in the African meningitis belt: high-resolution picture of genomic changes that mediate immune evasion. *mBio* 5:e01974-14. <https://doi.org/10.1128/mBio.01974-14>.
- Krauland MG, Dunning Hotopp JC, Riley DR, Daugherty SC, Marsh JW, Messonnier NE, Mayer LW, Tettelin H, Harrison LH. 2012. Whole genome sequencing to investigate the emergence of clonal complex 23 *Neisseria meningitidis* serogroup Y disease in the United States. *PLoS One* 7:e35699. <https://doi.org/10.1371/journal.pone.0035699>.
- Anonsen JH, Vik Å, Borud B, Viburiene R, Aas FE, Kidd SWA, Aspholm M, Koomey M. 2016. Characterization of a unique tetrasaccharide and distinct glycoproteome in the *O*-linked protein glycosylation system of *Neisseria elongata* subsp. *glycolytica*. *J Bacteriol* 198:256–267. <https://doi.org/10.1128/JB.00620-15>.
- Westman EL, Preston A, Field RA, Lam JS. 2008. Biosynthesis of a rare di-*N*-acetylated sugar in the lipopolysaccharides of both *Pseudomonas aeruginosa* and *Bordetella pertussis* occurs via an identical scheme despite different gene clusters. *J Bacteriol* 190:6060–6069. <https://doi.org/10.1128/JB.00579-08>.
- Larkin A, Chang MM, Whitworth GE, Imperiali B. 2013. Biochemical evidence for an alternate pathway in *N*-linked glycoprotein biosynthesis. *Nat Chem Biol* 9:367–373. <https://doi.org/10.1038/nchembio.1249>.
- Larkin A, Imperiali B. 2009. Biosynthesis of UDP-GlcNAc(3NAc)A by WbpB, WbpE, and WbpD: enzymes in the Wbp pathway responsible for *O*-antigen assembly in *Pseudomonas aeruginosa* PAO1. *Biochemistry* 48:5446–5455. <https://doi.org/10.1021/bi900186u>.
- Siu S, Robotham A, Logan SM, Kelly JF, Uchida K, Aizawa S, Jarrell KF. 2015. Evidence that biosynthesis of the second and third sugars of the archaeal tetrasaccharide in the archaeon *Methanococcus maripaludis* occurs by the same pathway used by *Pseudomonas aeruginosa* to make a di-*N*-acetylated sugar. *J Bacteriol* 197:1668–1680. <https://doi.org/10.1128/JB.00040-15>.
- Harrison OB, Claus H, Jiang Y, Bennett JS, Bratcher HB, Jolley KA, Corton C, Care R, Poolman JT, Zollinger WD, Frasch CE, Stephens DS, Feavers I, Frosch M, Parkhill J, Vogel U, Quail MA, Bentley SD, Maiden MC. 2013. Description and nomenclature of *Neisseria meningitidis* capsule locus. *Emerg Infect Dis* 19:566–573. <https://doi.org/10.3201/eid1904.111799>.
- Freitag NE, Seifert HS, Koomey M. 1995. Characterization of the pilF-pilD pilus-assembly locus of *Neisseria gonorrhoeae*. *Mol Microbiol* 16: 575–586. <https://doi.org/10.1111/j.1365-2958.1995.tb02420.x>.
- Hoang TT, Kutchma AJ, Becher A, Schweizer HP. 2000. Integration-proficient plasmids for *Pseudomonas aeruginosa*: site-specific integration and use for engineering of reporter and expression strains. *Plasmid* 43:59–72. <https://doi.org/10.1006/plas.1999.1441>.
- Hitchcock PJ, Brown TM. 1983. Morphological heterogeneity among *Salmonella* lipopolysaccharide chemotypes in silver-stained polyacrylamide gels. *J Bacteriol* 154:269–277.

23. Burrows LL, Pigeon KE, Lam JS. 2000. *Pseudomonas aeruginosa* B-band lipopolysaccharide genes wbpA and wbpI and their *Escherichia coli* homologues wecC and wecB are not functionally interchangeable. FEMS Microbiol Lett 189:135–141. <https://doi.org/10.1111/j.1574-6968.2000.tb09219.x>.
24. 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 U S A 106:4447–4452. <https://doi.org/10.1073/pnas.0809504106>.
25. Anonsen JH, Vik A, Egge-Jacobsen W, Koomey M. 2012. An extended spectrum of target proteins and modification sites in the general O-linked protein glycosylation system in *Neisseria gonorrhoeae*. J Proteome Res 11:5781–5793. <https://doi.org/10.1021/pr300584x>.
26. Jolley KA, Maiden MC. 2010. BIGSdb: scalable analysis of bacterial genome variation at the population level. BMC Bioinformatics 11:595. <https://doi.org/10.1186/1471-2105-11-595>.
27. The UniProt Consortium. 2017. UniProt: the universal protein knowledgebase. Nucleic Acids Res 45:D158–D169. <https://doi.org/10.1093/nar/gkw1099>.