AMERICAN SOCIETY FOR MICROBIOLOGY | mBio®

RESEARCH ARTICLE

Sculpting the Bacterial *O*-Glycoproteome: Functional Analyses of Orthologous Oligosaccharyltransferases with Diverse Targeting Specificities

Chris Hadjineophytou,[a] Jan Haug Anonsen,[a]* Tina Svingerud,[a] Tatum D. Mortimer,[b] Yonatan H. Grad,[b,c] Nichollas E. Scott,[d] Michael Koomey[a,e]

[a]Department of Biosciences, Section for Genetics and Evolutionary Biology, University of Oslo, Oslo, Norway
[b]Department of Immunology and Infectious Diseases, Harvard T. H. Chan School of Public Health, Boston, Massachusetts, USA
[c]Division of Infectious Diseases, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA
[d]Department of Microbiology and Immunology, University of Melbourne at the Peter Doherty Institute for Infection and Immunity, Melbourne, Australia
[e]Department of Biosciences, Centre for Ecological and Evolutionary Synthesis, University of Oslo, Oslo, Norway

ABSTRACT Protein glycosylation systems are widely recognized in bacteria, including members of the genus *Neisseria*. In most bacterial species, the molecular mechanisms and evolutionary contexts underpinning target protein selection and the glycan repertoire remain poorly understood. Broad-spectrum *O*-linked protein glycosylation occurs in all human-associated species groups within the genus *Neisseria*, but knowledge of their individual glycoprotein repertoires is limited. Interestingly, PilE, the pilin subunit of the type IV pilus (Tfp) colonization factor, is glycosylated in *Neisseria gonorrhoeae* and *Neisseria meningitidis* but not in the deeply branching species *N. elongata* subsp. *glycolytica*. To examine this in more detail, we assessed PilE glycosylation status across the genus and found that PilEs of commensal clade species are not modified by the gonococcal PglO oligosaccharyltransferase. Experiments using PglO oligosaccharyltransferases from across the genus expressed in *N. gonorrhoeae* showed that although all were capable of broad-spectrum protein glycosylation, those from a deep-branching group of commensals were unable to support resident PilE glycosylation. Further glycoproteomic analyses of these strains using immunoblotting and mass spectrometry revealed other proteins differentially targeted by otherwise remarkably similar oligosaccharyltransferases. Finally, we generated *pglO* allelic chimeras that begin to localize PglO protein domains associated with unique substrate targeting activities. These findings reveal previously unappreciated differences within the protein glycosylation systems of highly related bacterial species. We propose that the natural diversity manifest in the neisserial protein substrates and oligosaccharyltransferases has significant potential to inform the structure-function relationships operating in these and related bacterial protein glycosylation systems.

IMPORTANCE Although general protein glycosylation systems have been well recognized in prokaryotes, the processes governing their distribution, function, and evolution remain poorly understood. Here, we have begun to address these gaps in knowledge by comparative analyses of broad-spectrum *O*-linked protein glycosylation manifest in species within the genus *Neisseria* that strictly colonize humans. Using *N. gonorrhoeae* as a well-defined model organism in conjunction with comparative genomics, intraspecies gene complementation, and glycoprotein phenotyping, we discovered clear differences in both glycosylation susceptibilities and enzymatic targeting activities of otherwise largely conserved proteins. These findings reveal previously unappreciated differences within the protein glycosylation systems of highly related bacterial species. We propose that the natural diversity manifest within *Neisseria* species has significant potential to elucidate the structure-function

Editor Matthew R. Chapman, University of Michigan—Ann Arbor

Address correspondence to Michael Koomey, johnk@ibv.uio.no.

*Present address: Jan Haug Anonsen, Norwegian Research Centre AS, Randaberg, Norway.

The authors declare no conflict of interest.

Received 5 January 2022
Accepted 25 March 2022
Published 26 April 2022

relationships operating in these and related systems and to inform novel approaches to applied glycoengineering strategies.

KEYWORDS *Neisseria*, evolution, glycoproteins, oligosaccharides, pili

Glycosylation is an important element of biological systems across all domains of life through coordinated posttranslational modification of large sets of proteins. Although significant progress has been made in defining the basic enzymatic pathways of many such systems, the processes by which particular proteins are targeted and specific sites within proteins are selected remain incompletely understood. The latter situations persist despite the increasing availability of high-resolution structures for glycan transferases and the identification of domains and motifs inherent to protein target substrates. These gaps in knowledge limit efforts to understand the molecular and evolutionary processes that ultimately define and shape glycoproteomes.

In Gram-negative organisms, the process of broad-spectrum (or general) protein glycosylation is primarily localized to the periplasm and catalyzed by oligosaccharyltransferases (OTases) that are members of the GT-C superfamily (1). These OTases share structural features of multiple, transmembrane helices and the utilization of lipid-linked glycan donors. These include the asparagine-targeting OTases exemplified by *Campylobacter jejuni* PglB (2), as well as a family of OTases that target serine sites and have been documented within *Neis*seria (3, 4), *Acinetobacter* (5), *Ralstonia* (6), *Francisella* (7), *Burkholderia* (8), *Vibrio* (8), and *Mycobacterium* (9) species. The latter class of transferases are termed *O*-OTases based on their *O*-linked glycosylating activities and recognize serine attachment sites within low-complexity regions (LCRs) rich in alanine and proline. *Campylobacter* PglBs and *O*-OTases are also members of the shape, elongation, division, and sporulation (SEDS) protein family of transferases involved in cell wall biogenesis and remodeling and modification of surface glycoconjugates (2). While molecular structures have been solved for orthologous *N*-OTases from bacteria, archaea, and eukaryotes and SEDS proteins (9–13), little is known regarding the structure-function relationships of *O*-OTases.

A unique subset of *O*-OTases found in isolates of *Pseudomonas aeruginosa* and *Acinetobacter* species solely glycosylate subunits of type IV pili (Tfp) or pilin-like molecules (14–16). In the species *Neisseria gonorrhoeae*, *Neisseria meningitidis*, *Ralstonia solanacearum*, and *Francisella tularensis*, Tfp pilins are among the protein targets of the broad-spectrum *O*-OTases (3, 6, 7, 17). In contrast, in *Neisseria elongata* subsp. *glycolytica* (here referred to as *N. elongata*), the broad-spectrum *O*-OTase does not glycosylate its Tfp subunit protein (18). Therefore, a variety of strategies have emerged regarding the adaptability of *O*-OTases to target specific sets of proteins in general and Tfp pilin proteins in particular.

Members of the genus *Neisseria* colonizing mucosal sites of humans include the important pathogens *N. gonorrhoeae* and *N. meningitidis*, as well as additional species groups that are commensal inhabitants. Given their genetic relatedness, common ancestry, and host restriction, these species groups encompass a unique model system in which to study the evolution and ecology of host-microbe interactions. Based on these observations, we set out to examine species-level relationships between PilE glycosylation and *O*-OTase targeting activities when expressed in gonococci as a defined reference background. The findings strongly suggest that glycoprotein repertoires across the genus are shaped by distinct targeting activities of *O*-OTases together with intrinsic structural features of substrate proteins. We propose that standing genetic variation within the genus *Neisseria* provides unique resources to understand the structure-function relationships of protein targeting glycosyltransferases and their protein substrates.

RESULTS

Identification of candidate Tfp subunit genes and validation by complementation in *N. gonorrhoeae*. While studies have identified putative *pilE* genes in other neisserial species groups (19, 20), only those from *N. gonorrhoeae*, *N. meningitidis*, and *N.*

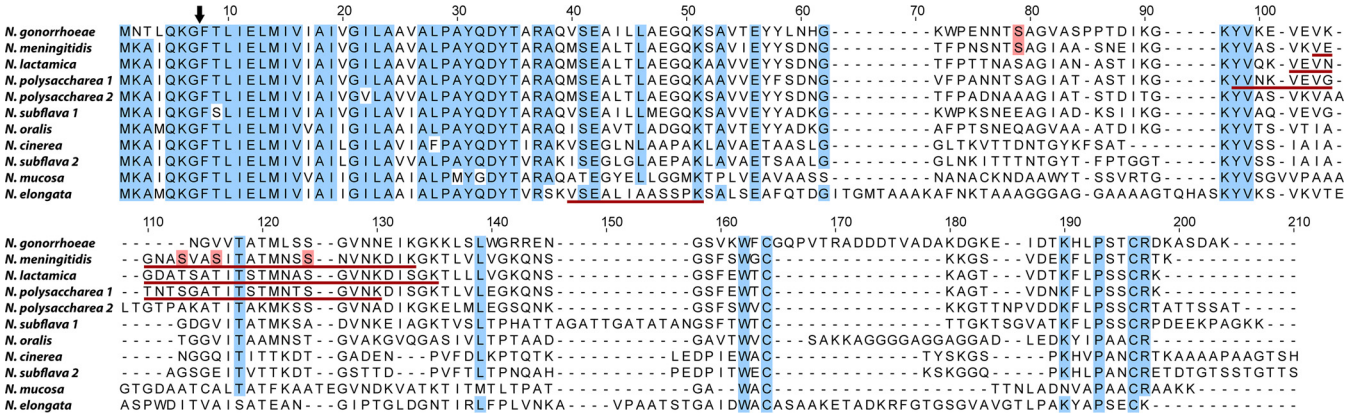

FIG 1 Structural alignment and features of defined and candidate neisserial PilE pilin subunits. The cleavage site for processing to mature pilin is marked with an arrow at the conserved $Gly_{-1}Phe_{+1}$ junction residues. Residues present in at least 10 of the 11 PilE orthologs are highlighted in blue, and established sites of glycan attachment are highlighted in red. Glycopeptides identified in this study are underlined in red (see also Fig. S3). The strains used to generate these data are listed in Table S1A.

elongata have been conclusively validated as such. We first aimed to identify orthologous *pilE* candidates within the genomes of other species available in the PubMLST database using BLASTp analyses and BLASTn with PilE sequences from *N. gonorrhoeae* and *N. elongata* as queries (Fig. 1; see Fig. S1A in the supplemental material). Candidate orthologues from each type strain were expressed in *N. gonorrhoeae* and assessed for their ability to complement Tfp-associated phenotypes. These included those from two type strains identified as isolates of *Neisseria polysaccharea*, but for which further analyses had defined as being members of two unique clusters (designated here as *N. polysaccharea* 1 and 2) and two strains of *Neisseria subflava*, one of which was previously identified as *Neisseria flavescens* (designated here as *N. subflava* 1 and 2, respectively) (21, 22). Immunoblotting using an antiserum recognizing a conserved epitope within Tfp pilins confirmed expression of each of the candidate PilEs (Fig. 2, top panel). Competence for natural transformation was assessed as Tfp expres-

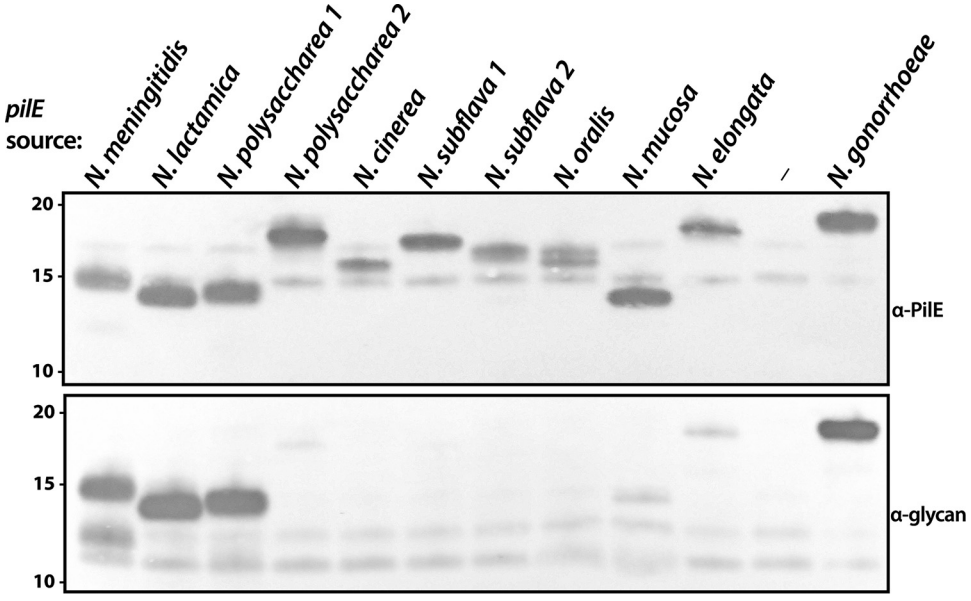

FIG 2 Neisserial Tfp pilins vary in glycosylation susceptibility when expressed in *N. gonorrhoeae*. Immunoblots were performed using equal amounts of bacterial whole-cell lysates. (Top panel) Samples probed with a polyclonal antiserum raised against purified *N. gonorrhoeae* pili from strain MS11. (Bottom panel) The same samples as in the top panel probed with the glycan-specific monoclonal antibody. The strains used are (from left to right) CH24, CH16, CH12, CH78, CH8, CH20, CH28, CH32, CH36, CH4, KS122, and CH114.

sion is essential for this phenotype. All alleles tested supported high levels of transformability, being over 1,000-fold above the *pilE* negative-control background (Fig. 1B). Although there were differences in the abundance of Tfp-like appendages seen by transmission electron microscopy (TEM) (Fig. S1C), peptides corresponding to each of the open reading frames (ORFs) were detected in enriched shear extract fractions by mass spectrometry (unpublished data).

***Neisseria* Tfp pilins vary in glycosylation susceptibility when expressed in *N. gonorrhoeae*.** Having confirmed that the genes identified encode Tfp pilins, we examined if their products were modified by the *N. gonorrhoeae* *O*-glycosylation system. Using immunoblotting with the monoclonal antibody (MAb) npg1, which recognizes a glycan-associated epitope, we found that the PilEs from the *N. meningitidis*, *N. lactamica*, and *N. polysaccharea* 1 isolates were glycosylated, while those from the remaining species showed either no reactivity or weak reactivity in the cases of PilEs from *N. polysaccharea* 2, *N. mucosa*, and *N. elongata* (Fig. 2, bottom panel). To estimate the relative glycosylation levels for the latter proteins, we performed immunoblotting using serial dilutions of whole cells from a *pglA* null background as standards. By comparing their immunoblot signal intensity to that of the *N. gonorrhoeae* PilE standard at various dilutions, glycosylation levels were estimated to be approximately 5% for *N. mucosa* and 10 to 15% for *N. elongata* (see Fig. S2 in the supplemental material).

Mass spectrometry (MS) analyses of PilE from enriched shear extracts identified glycopeptides for all detectably glycosylated pilins, save for those from *N. mucosa* and *N. polysaccharea* 2 (see Fig. S3 in the supplemental material). The glycopeptides found for *N. meningitidis*, *N. lactamica*, and *N. polysaccharea* PilEs included conserved serine attachment sites defined earlier for *N. meningitidis* PilE, for which there was clear evidence for multiple modified residues (23). Despite its low level of glycosylation, a glycopeptide for *N. elongata* PilE was identified, and the residues in this peptide resemble LCR attachment sites identified in *N. gonorrhoeae* and other broad-spectrum systems (Fig. 1; Fig. S3). PilE glycosylation was also assessed by immunogold labeling/transmission electron microscopy of Tfp, where the results were concordant with those from immunoblotting and MS studies (Fig. S1C).

To examine the degree of glycosylation macroheterogeneity occurring in gonococci, the relative mobilities of pilins in backgrounds expressing tri-, di-, and monosaccharide glycoforms along with a *pgl* null control were assessed (see Fig. S4 in the supplemental material). As seen by the stepwise, retarded migration associated with glycan presence and increasing glycan mass, PilEs from *N. meningitidis*, *N. lactamica*, and *N. polysaccharea* 1 were completely glycosylated. In contrast, migration variation was absent for the other PilEs, save for that from *N. elongata*, where a low level with altered migration was found that corresponded with its reduced glycosylation status detected by glycan MAb reactivity.

Phylogeny and diversity of *pglO* across the genus *Neisseria*. The ability of *N. elongata* PilE to be glycosylated (albeit inefficiently) in *N. gonorrhoeae* but not in *N. elongata* (18) suggested that factors independent of PilE structure might impact glycosylation propensity and that this situation might relate to varying *O*-OTase targeting activity. To test this directly, we first identified candidate *O*-OTase genes in neisserial genomes in PubMLST by BLASTp and BLASTn analyses using *N. gonorrhoeae* PglO/*pglO* and *N. meningitidis* PglL/*pglL* as queries. (Note that despite the distinct nomenclature, *pglO* and *pglL* are orthologous [4, 24], so we use the terminology PglO/*pglO* here.) Single, quality hits were identified in the genomes of all human-associated neisserial isolates, and analyses of these revealed patterns of phylogenetic relationships congruent with those established using other common gene sets, including protein glycosylation (*pgl*) genes (25–27) (Fig. 3). Interestingly, closer examination revealed two allelic isoforms in *N. gonorrhoeae* defined by the variable presence of a single nucleotide polymorphism (SNP) in the genus-wide consensus stop codon that results in an ORF that is C-terminally extended by 38 amino acids (see Fig. S5A in the supplemental material). This SNP variant appears to have been acquired multiple times across lineages and is present in the gonococcal background used in studies here (see Fig. S6 in the

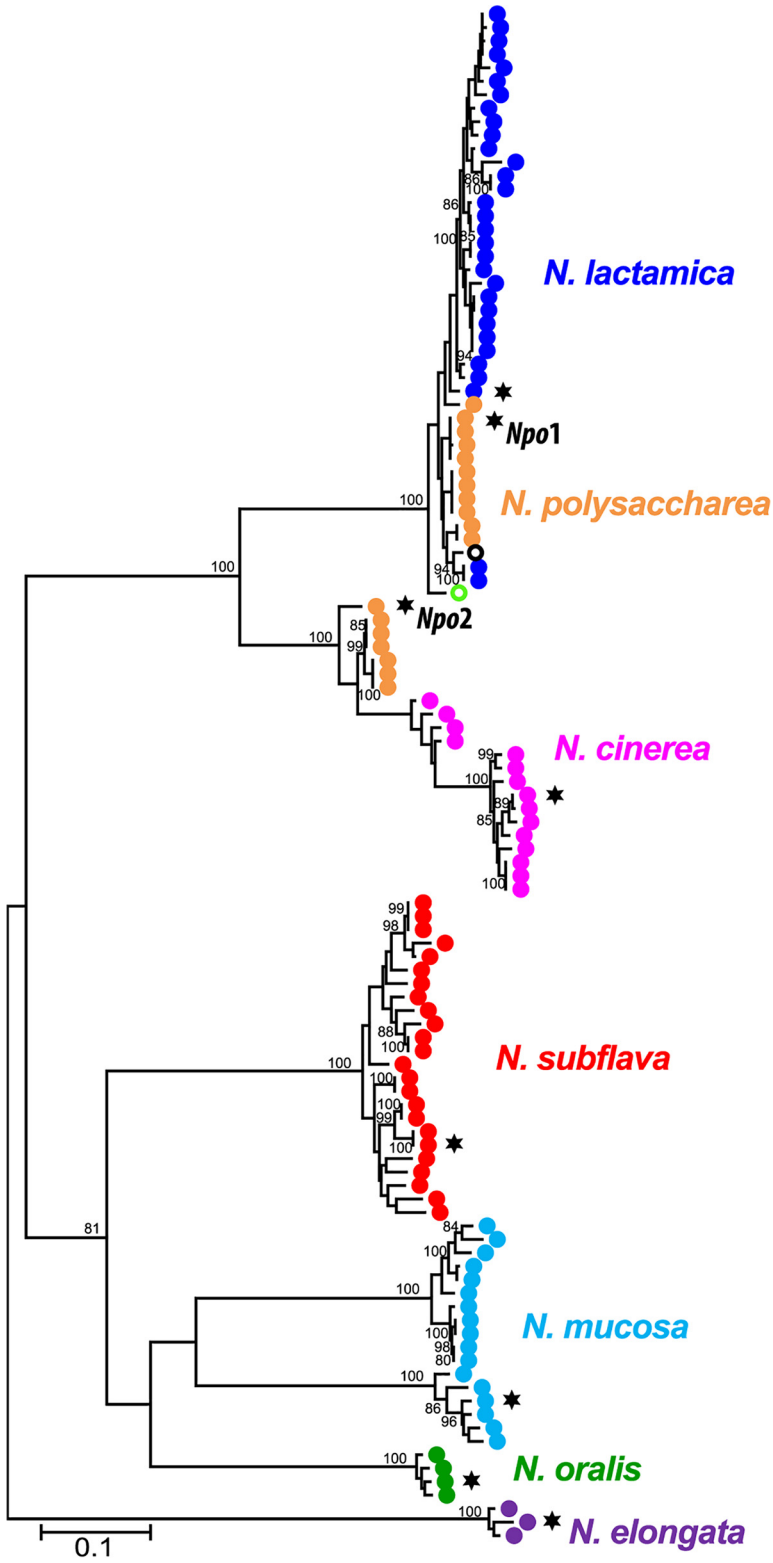

FIG 3 Maximum likelihood phylogenetic tree of neisserial *pglO* alleles. The tree was based on a MUSCLE alignment (66) and was constructed in MEGA X (67) using the Tamura-Nei model (68). Species-representative alleles from *N. gonorrhoeae* MS11 and *N. meningitidis* FAM18 were included for reference (green and black open circles, respectively). Alleles of *pglO* expressed in *N. gonorrhoeae* for the glycosylation complementation experiments are marked with a star. The tree was inferred using 114 sequences at 2,052 sites and replicated 500 times. Bootstrap values with <80% confidence are excluded from the final figure. The strains used can be found in Table S1C.

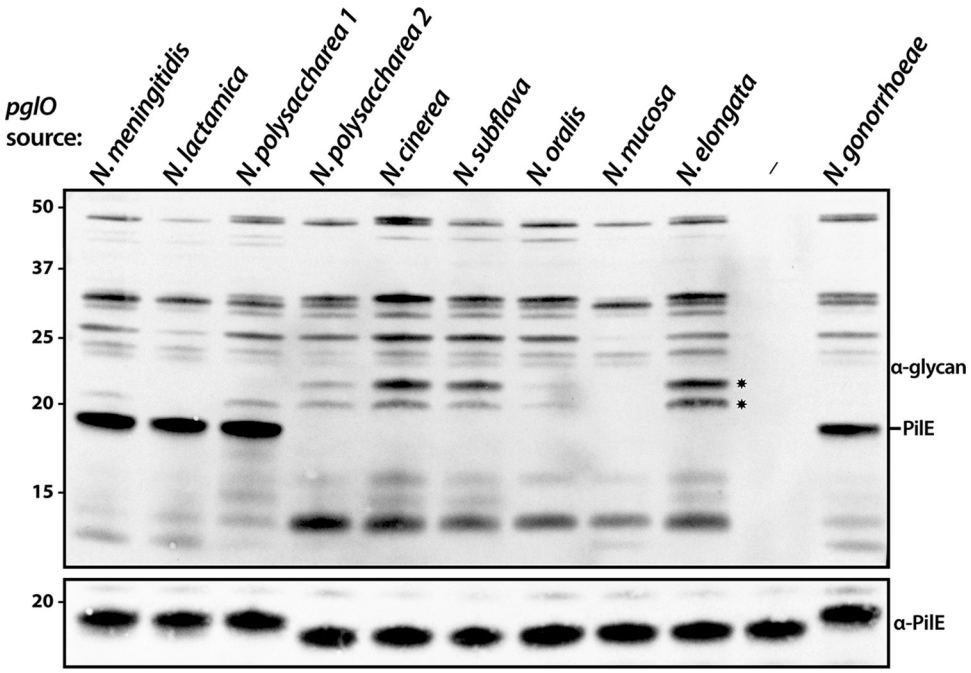

FIG 4 Oligosaccharyltransferases from *Neisseria* species support broad-spectrum glycosylation when expressed in *N. gonorrhoeae*. (Top panel) Samples probed with a glycan-specific monoclonal antibody. (Bottom panel) The same samples as in the top panel probed with a polyclonal antiserum raised against purified *N. gonorrhoeae* PilE. The strains used are (from left to right) CH66, CH65, CH63, CH64, CH59, CH60, CH62, CH61, CH48, CH47, and KS127. Two asterisks denote glycoproteins chosen for further analyses (Fig. 5).

supplemental material). We corrected the SNP to regenerate the stop codon in the strain used here but found no gross differences in either PilE glycosylation or glycoprotein profile determined by immunoblotting using a glycan-specific MAb (Fig. S5B).

Oligosaccharyltransferases from *Neisseria* species support broad-spectrum glycosylation but vary in PilE targeting proficiency. To examine the functionality of PglO from diverse species groups, alleles from the same backgrounds from which *pilE* alleles were derived were introduced into the *N. gonorrhoeae* strain using an allelic replacement protocol. As indicated by the large number of reactive glycoproteins detected via immunoblotting using a glycan-specific MAb, all *pglO* alleles supported broad-spectrum protein glycosylation (Fig. 4). In addition to the expected targeting of endogenous PilE by PglO from both *N. gonorrhoeae* and *N. meningitidis*, pilin glycosylation was also seen for the *N. lactamica* and *N. polysaccharea* 1 PglOs, but not from the remaining species (Fig. 4).

Identification of additional glycoproteins differentially modified by *N. gonorrhoeae* and *N. elongata O*-OTases. We next sought to identify other resident proteins differentially glycosylated by *O*-OTases when expressed in *N. gonorrheae*. We focused on two candidates with relative mobilities of 22 kDa and 23 kDa that were strongly reactive with the glycan MAb in the *N. elongata* and other *O*-OTase backgrounds (marked with asterisks in Fig. 4). Based on previous knowledge of *N. gonorrhoeae* glycoproteins, we identified these as the lipoproteins Ngo0983 (known as Lip) and Ngo0994 (known as Laz [for lipid-linked azurin]) by virtue of their immunoblot reactivity with antibodies raised against recombinant Laz protein (Fig. 5A). Earlier studies showed that although each of these was a glycoprotein, Lip and Laz are poorly glycosylated in *N. gonorrhoeae* (as shown by their failure to show glycosylation-dependent shifts in relative mobilities) (17, 28). Notably, Lip and Laz lipoproteins share limited sequence identities, other than highly conserved amino-terminal domains rich in pentapeptide (AAEAP) repeats that also encompass serine residues implicated as sites of glycan occupancy (Fig. 5B).

We also employed an unbiased, MS-based methodology based on high-field asymmetric waveform ion mobility spectrometry (FAIMS) fractionation (29) to compare the

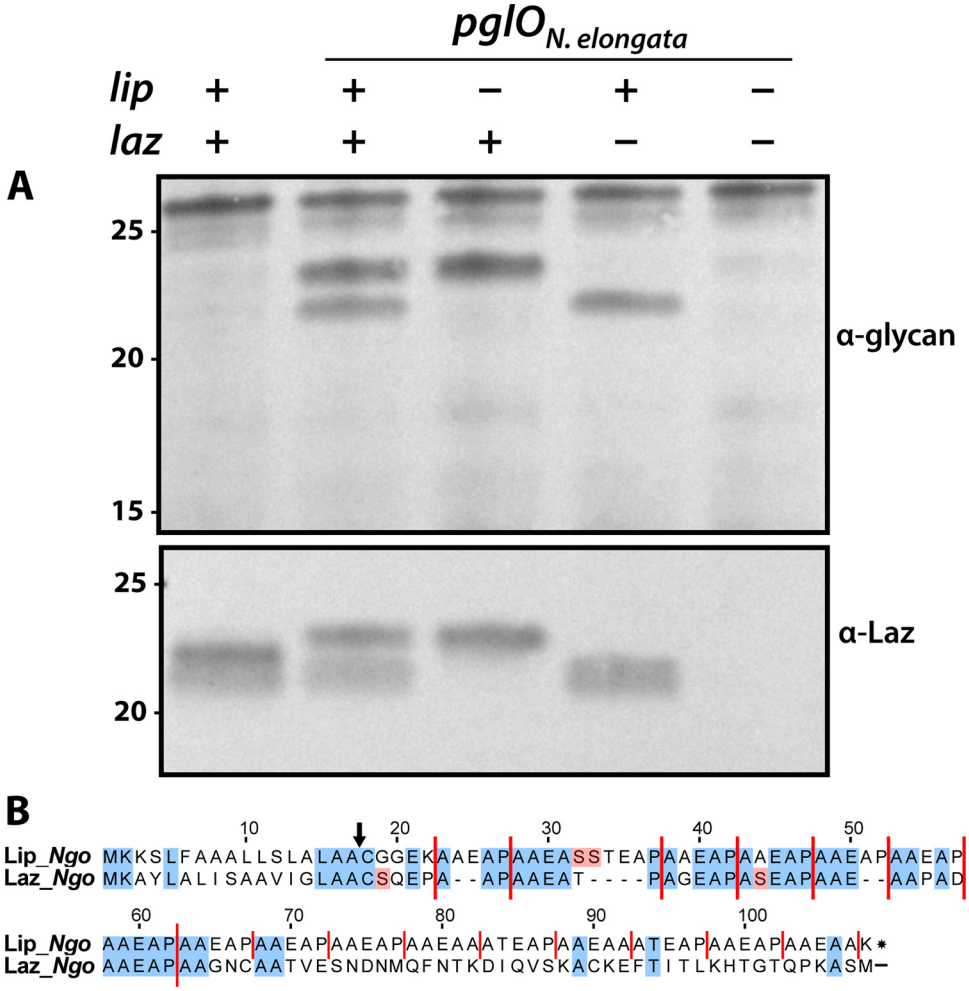

FIG 5 The structurally related lipoproteins Lip and Laz are differentially modified by *N. elongata* versus *N. gonorrhoeae* *O*-OTases. Shown are immunoblots of *N. gonorrhoeae* strains expressing either the endogenous *pglO* allele or the *N. elongata pglO* allele along with ORF-inactivating mutations in *lip* and/or *laz*. Samples were probed with either a glycan-recognizing MAb (top panel) or a polyclonal antiserum raised against recombinant Laz purified from *E. coli* (bottom panel). The latter antibodies react with both Lip and Laz. The strains used are (from left to right) KS127, CH48, CH156, CH155, and CH157. (B) Alignment of Lip and the amino terminus of Laz from *N. gonorrhoeae*. Conserved residues are highlighted in blue, and serine residues (potential sites of glycan attachment) are highlighted in red. Vertical red lines define AEAAP pentapeptide repeat units (and degenerate forms thereof). The vertical black arrow shows the site of N-terminal proteolytic cleavage of the diacylglyceryl-prolipoprotein.

whole proteomes and glycoproteomes of *N. gonorrhoeae* expressing either the endogenous *O*-OTase or that from *N. elongata*. Here, alterations in the abundances of numerous glycopeptides were noted, including those derived from the glycoproteins Ag473 (Ngo1043), the secretin PilQ, the multidrug efflux transporter component MtrC, and the minor pilin subunit PilV (Fig. 6A). Changes in glycopeptide detection here could reflect alterations in the abundance of the proteins themselves from which they are derived. In the cases of PilQ and PilV, however, proteomic analysis showed that these differences could not be accounted for by such alterations in overall protein abundance (Fig. 6B and C). Further analyses nonetheless revealed that there were clear differences in the relative abundances of a large number of other proteins between samples derived from the two PglO backgrounds (see Fig. S7 in the supplemental material).

Chimeric *O*-OTases reveal domains influencing PilE targeting specificity. To begin to assess structural features of *O*-OTases associated with distinct substrate protein targets, we sought features that might explain their abilities to differentially glycosylate PilE. Consensus models from membrane protein structure algorithms predicted that all

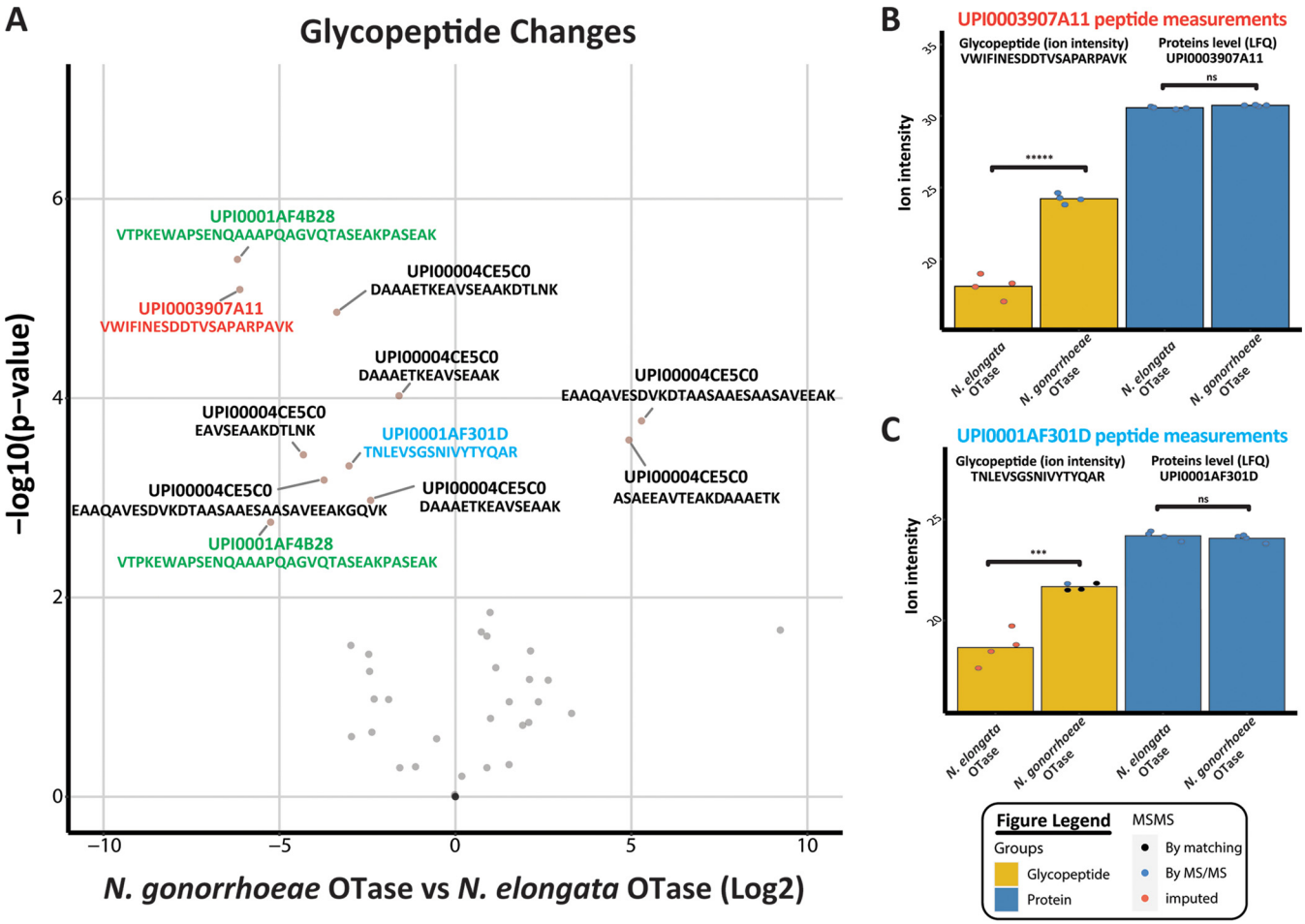

FIG 6 Glycoproteomic/proteomic analysis supports the alteration of glycosylated substrates independent of changes in protein abundance. (A) A volcano plot of quantified glycopeptides reveals alterations in the abundances of glycopeptides within samples expressing different OTases. Examination of peptides identified for the proteins PilQ (UPI0003907A11 [red]) and PilV (UPI0001AF301D [blue]) reveals glycopeptides of these proteins are only observed within strains expressing the *N. gonorrhoeae* OTase. Other glycopeptides showing significant differences in the two backgrounds are derived from MtrC (UPI0001AF4B28 [green]) and Ag473 (UPI00004CE5C0 [black]). Further analyses of the nonglycosylated peptides derived from PilQ and PilV reveal similar levels of protein abundances in both backgrounds (panels B and C, respectively).

neisserial PglOs have similar topologies, with 13 transmembrane domains and three large extrahelical loops oriented toward the periplasm (see Fig. S8A in the supplemental material). These patterns mirrored those of the *N*-linked PglB OTase from *Campylobacter lari* (11) and other members of SEDS family proteins (2). Comparative sequence alignment of neisserial PglOs revealed multiple regions of interspecies diversity, but no obvious patterns were correlated with their observed activities on PilE (Fig. S8B, top panel). Therefore, we set out to identify domains responsible for PilE targeting by creating hybrids between the two most closely related *O*-OTases that differed in PilE targeting activities. To this end, we created chimeras by fusing ORF-encoding sequences from the *O*-OTases of *N. meningitidis* and *Neisseria cinerea* that share 69.9% identity and 75.1% similarity (Fig. S8B, bottom panel). Sites of fusion were targeted to highly conserved stretches of amino acids (Fig. 7: Fig. S8B, top panel). The chimeras were then expressed in *N. gonorrhoeae* via allelic replacement and their targeting activities assessed by immunoblotting using glycan-recognizing antibodies.

All chimeras glycosylated multiple gonococcal proteins, although each was associated with distinct patterns of protein targeting (Fig. 7B and C). PilE glycosylation was only observed for hybrids carrying C-terminal segments of *N. meningitidis* PglO. This activity was retained in hybrid 7, which carried the last 317 residues of *N. meningitidis* PglO, but it was absent from hybrids 9 and 11, which carried the last 247 and 144 residues,

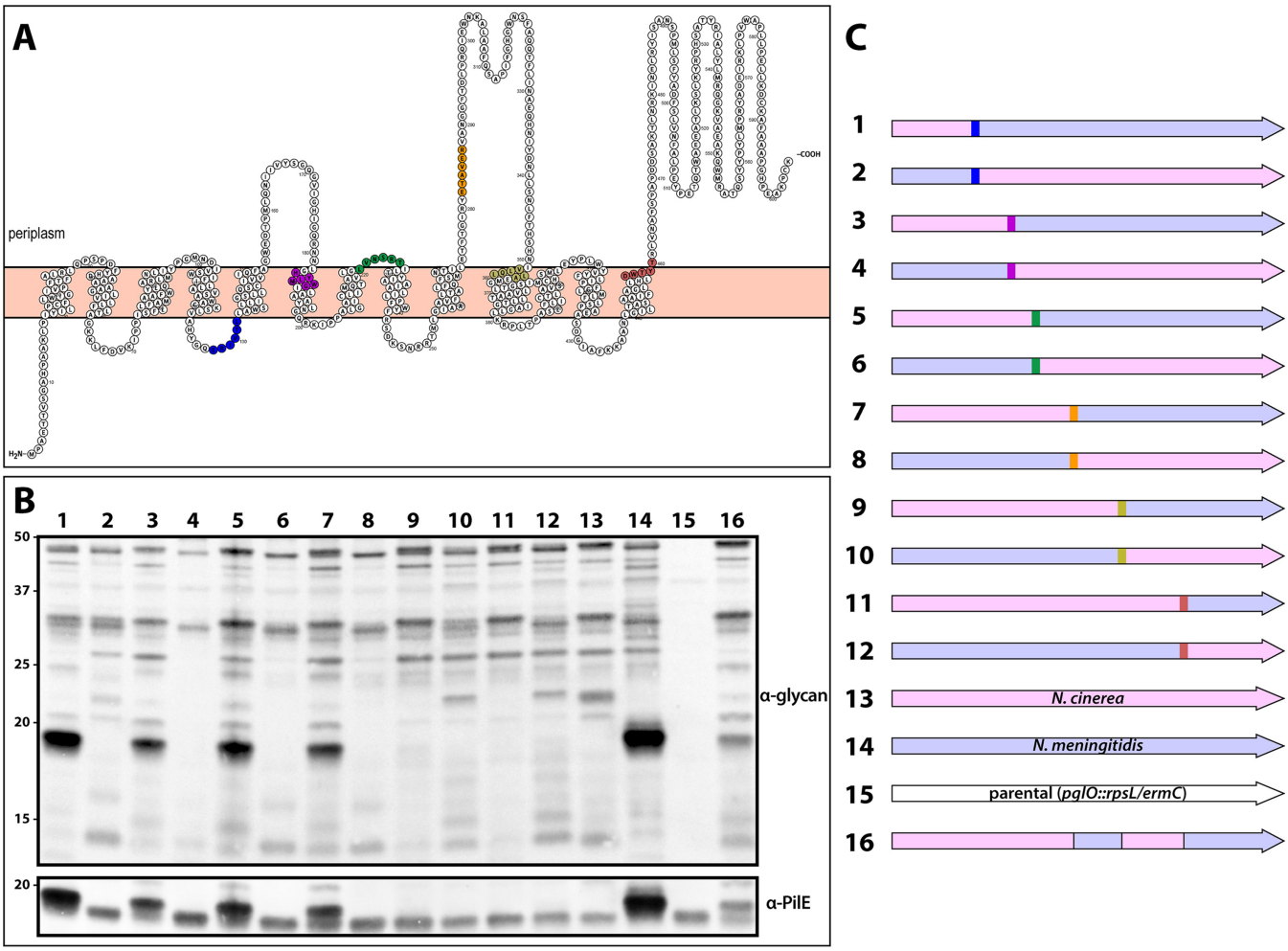

FIG 7 Chimeric *O*-OTases reveal domains influencing PilE targeting specificity. (A) Predicted transmembrane topologies of *N. meningitidis* PglO based on Phobius (69) and visualized using Protter (70). Color-coded stretches indicate conserved amino acid residues shared between *N. meningitidis* and *N. cinerea* alleles that were used as fusion sites. (B) Immunoblot of whole-cell lysates of *N. gonorrhoeae* strains expressing *pglO* chimeric constructs. (Top panel) Membrane probed with a glycan-specific MAb. (Bottom panel) The same membrane shown in top panel reprobed with the PilE-recognizing polyclonal antiserum. (C) Cartoon illustration of the PglO chimeras color coded so as to delineate fusion sites as shown in panel A. The numbering here corresponds to the lane numbers in panel B. The strains used are (from left to right in panel B) CH91 to -100, CH105, CH106, CH59, CH66, CH47, and CH152.

respectively. Based on these results, we hypothesized that the PilE targeting specificity might minimally require the two major, putative periplasmic domains of *N. meningitidis* PglO. Indeed, a hybrid consisting of an *N. cinerea* PglO backbone carrying solely these two segments from *N. meningitidis* PglO was capable of glycosylating PilE (Fig. 7).

DISCUSSION

Broad-spectrum protein glycosylation systems are found in all branches of life. Despite progress in a number of systems, there remain significant gaps in knowledge as to what factors shape glycoproteome content and how protein-targeting transferases select their substrates. Addressing these shortcomings seems crucial to understanding the biological significance of protein glycosylation at both the molecular and evolutionary levels. As shown in this study, the *O*-linked glycosylation systems expressed by human-associated species groups of the genus *Neisseria* have high potential to address both of these questions. We are unaware of other studies examining OTase structure-function diversity at this level. In fact, with the notable exception of dedicated *O*-OTases, the general consensus from limited studies of *O*-OTases in *Burkholderia* and *Acinetobacter* isolates and *N*-OTases in *Campylobacter* species and related *Deltaproteobacteria*/*Epsilonproteobacteria*

members seems to be that highly structurally related OTases have similar targeting activities, respectively (30–33).

The restricted distribution of the PilE glycosylation phenotype (and underlying genotypes) within neisserial species groups raises questions as to its biological, evolutionary, and ecological significance. In gonococci, PilE glycosylation can influence cell-cell interactions occurring within microcolonies that likely stem from subtle changes in pilus extrusion-retraction dynamics and biophysical properties of pili themselves (34). Glycan modification has also been suggested to act as a cloaking device to mask conserved PilE protein epitopes or redirect the humoral response toward the glycan, which in and of itself is antigenically variable. Such scenarios have been particularly invoked in the case of the structurally invariant class II pilins expressed by some meningococcal isolates (23). In addition, disrupted glycosylation of some PilE variants in *N. meningitidis* was associated with a dramatic decrease in the levels of Tfp (23). Pilin subunit glycosylation in *P. aeruginosa* has also been implicated in having an impact on Tfp expression levels and organelle dynamics, as well as blocking predation by bacteriophages that use Tfp as primary receptors (35, 36). Filamentous phages using Tfp as primary receptors have been described in gonococci and meningococci, although there are no reports that PilE glycosylation influences phage susceptibility (37, 38). Thus, while the basis for the variable distribution of PilE glycosylation remains to be determined, we speculate that it involves differences in the roles Tfp serve in niches occupied by distinct neisserial species groups.

A point of interest with regard to PilE glycosylation is its association with particular glycoforms and *pgl* glycosyltransferase genes. Specifically, distribution of PilE glycosylation coincides remarkably with that of the *pglA* and *pglE* glycosyltransferase genes that incorporate specifically galactose sugars into di- and trisaccharide glycoforms (26). Terminal galactose on PilE-associated glycans has been proposed to promote adherence to human cervical mucosa via an interaction with the I-domain of complement receptor 3 (39). Another connection between PilE glycosylation and glycan expression entails the correlation with phase-variable glycosyltransferase alleles whose on-off expression underlies glycan antigenic variation (40, 41, 71). Such phase-variable alleles of *pglA*, *pglG*, and *pglH* are uniquely restricted to species isolates defined here as capable of the PilE glycosylation phenotype (27). As abundant surface-exposed glycoproteins and critical colonization factors in *N. gonorrhoeae* and *N. meningitidis* (and likely in *N. polysaccharea* and *N. lactamica*), PilE glycosylation and glycan antigenic variation may have coevolved as synergistic mechanisms to evade immune surveillance of Tfp, as suggested by others (23).

The distinct substrate targeting specificities and relatively high degrees of sequence identity exhibited by neisserial PglOs appear to provide a unique set of reagents to define structure-function relationships of *O*-OTases. While the potential of the neisserial system in this regard is exemplified by the findings using PglO chimeras, it is also of interest to define the structural features of protein targets that influence glycosylation susceptibility and productive interaction. That prospect is supported here by the finding that Lip and Laz as two proteins coordinately and differentially glycosylated by PglO from *N. elongata* share only a short segment of amino acid identity restricted to their N termini (Fig. 5). It should be possible then to apply the polypeptide chimera approach used in the PglO analyses to assess and define key features of protein substrates as well. Further analyses of the glycopeptide profiles and glycoproteomes associated with distinct PglO isoforms should also provide valuable information on the details of substrate-OTase interactions. That said, we cannot formally rule out that the PglO targeting specificities seen here may be influenced by the efficacies with which exogenous *O*-OTases use gonococcal glycan donors.

In conclusion, neisserial *O*-OTases and their distinct targeting activities provide a unique and facile set of naturally occurring reagents to address fundamental questions about the molecular mechanisms by which proteins are targeted for glycosylation as well as the evolution of protein glycosylation systems.

MATERIALS AND METHODS

Bacterial strains and culture conditions. *Neisseria* strains were grown for 16 to 18 h (at 37°C with 5% CO_2) on GC medium (Difco), supplemented with Kellogg's supplement (42). The gonococcal strains used were derivatives of strain MS11 carrying an isopropyl-β-D-thiogalactopyranoside (IPTG)-inducible *pilE* allele (43). Gonococcal transformants were selected on antibiotics at the following concentrations: streptomycin, 750 μg/mL; kanamycin, 50 μg/mL; and chloramphenicol, 10 μg/mL. Genetic constructs and mutations in glycosylation genes ($pglE_{on}$, *pglA*, and *pglC*) were introduced into various strains using transformation as previously described (44). *Escherichia coli* TOP10 cells (Invitrogen) were used for plasmid propagation. The strains, plasmids, and primers used in this study are listed in Table S1A and B in the supplemental material.

Genome analyses and bioinformatics. Candidate *pilE* and *pglO* genes were identified using BLASTp and BLASTn queries with default parameters against all human-specific *Neisseria* spp. deposited in the Bacterial Isolate Genome Sequence Database (BIGSdb) (http://pubmlst.org/neisseria/) (45). The list of strains utilized is in Table S1C.

Expression of exogenous *pilE* alleles in *Neisseria*. The *pilE* alleles from selected *Neisseria* species were ectopically expressed in an intergenic region of *N. gonorrhoeae* as translational fusions to the native *N. gonorrhoeae pilE* at their conserved $Gly_{-1}Phe_{+1}$ junction (see Fig. S9 in the supplemental material) as previously described (46, 47). This was achieved by PCR amplifying the two overlapping fragments, followed by PCR-based splicing by overhang extension (SOE) using the flanking primers. The hybrid constructs were digested with restriction enzymes PacI and FseI and ligated within the intergenic chromosomal site between the *lctP* and *aspC* genes of *N. gonorrhoeae* in plasmid pGCC6 (48). The plasmids were then linearized and transformed into gonococcal backgrounds in which the endogenous *pilE* allele is placed under an IPTG-inducible promoter (KS101-derived strains). *N. gonorrhoeae* strains expressing exogenous *pilE* alleles were selected on chloramphenicol-containing media.

Quantitative transformation assays. Transformation assays were performed as previously described (43). Briefly, *gyrB*-containing DNA was PCR amplified from plasmid pSY6 using primers CHP80 and CHP81 (47). Then, 0.1 μg of purified DNA was mixed with a bacterial suspension in 0.5 mL of transformation medium (CO_2-saturated GC medium supplemented with IsoVitaleX and 7 mM $MgCl_2$) and incubated at 37°C for 20 min in the presence of 5% CO_2. This was followed by a 5-fold dilution with transformation medium and transfer in a 37°C shaking incubator for 3 h. Finally, cells were appropriately diluted and plated onto GC plates with and without nalidixic acid. Transformation frequencies were calculated by dividing the number of nalidixic acid-resistant CFU by the total CFU.

Allelic exchange of the *pglO* locus in *N. gonorrhoeae*. Various neisserial *pglO* alleles were introduced into *N. gonorrhoeae* via a two-step mutagenesis strategy that allows markerless gene replacement (Fig. S9) by exploiting the fact that the *N. gonorrhoeae* MS11 strain is naturally resistant to streptomycin. This method utilizes a two-gene cassette that contains a selectable marker (*ermC*) and counterselectable marker (*rpsL*) (49). First, the immediate flanking regions of *pglO* from *N. gonorrhoeae* strain were PCR amplified using primers CHP105 and CHP60 for the upstream fragment and CHP191 and CHP56 for the downstream fragment. Then, the *rpsL_ermC* locus was PCR amplified from plasmid pFLOB4300 (49) with primers CHP66 and CHP67, which include sequence homology to the flanking regions of the *pglO* allele. The three fragments were PCR purified and spliced in a PCR using the flanking primers CHP105 and CHP56. The product of this reaction was transformed into strain KS127 and selected on erythromycin to create the intermediate strain CH47. Strain CH47 was then used in a second transformation to replace the *rpsL_ermC* locus with *pglO* allele ORFs (precisely from the start codon to the stop codon) from different *Neisseria* species, creating strains CH59 to CH67. This was achieved by PCR amplifying the *pglO* ORFs from various species (primers listed in Table S1B) and splicing them in a PCR with the *N. gonorrhoeae pglO* flanking products described above. These *pglO* constructs were cloned into the pCRII-TOPO vector to create pCH59-67, sequenced, and transformed into the intermediate strain CH47. Transformants were selected on streptomycin and screened for erythromycin sensitivity.

For chimeric *O*-OTase constructs, we first identified short stretches of conserved amino acid residues that are shared by both *N. meningitidis* and *N. cinerea pglO* alleles. These conserved regions were used in primer design to create homologous overlapping fragments between the two alleles. For templates, we used the pCRII-TOPO plasmids pCH66 and pCH59 (described above) to generate two overlapping fragments that were independently PCR amplified, spliced in a new PCR with flanking primers, and cloned in a plasmid. Finally, the constructs were transformed into strain CH47. The primers used were CHP105 and CH56 (described above) in conjunction with the hybrid-specific primers listed in Table S1B.

Construction strategies for mutants. The *lip* and *laz* genes from *N. gonorrhoeae* were inactivated via the insertion of a kanamycin cassette within their ORFs, resulting in the deletion of 83% and 94% of the *lip* and *laz* genes, respectively. First, the kanamycin gene was amplified from pKAN with primers CHP172 and CHP173. To delete *laz*, primer pairs CHP215 and CHP216 and CHP217 and CHP218 were used to amplify the flanking regions of *laz*, which were then spliced in a PCR-based SOE reaction with the kanamycin cassette. The PCR product was TOPO cloned, transformed into CH48, and selected on kanamycin to create strain CH155. Similarly, the flanking regions of *lip* were PCR amplified with primer pairs CHP219 and CHP220 and CHP221 and CHP222 and then spliced with the kanamycin cassette. Finally, the PCR product was TOPO cloned, transformed into CH48 and CH155, and selected on kanamycin to create strains CH156 and CH157.

We used site-directed mutagenesis to introduce a stop codon to the extended *pglO* ORF of *N. gonorrhoeae*. This was achieved by mutating the Arg605 residue of *pglO* to a stop codon (CGA-TGA). Two PCR fragments were generated using primer pairs CHP105 and CHP213 and CHP212 and CHP56, which

carried the SNP in their overlapping homologous regions. The two fragments were spliced in a new PCR using flanking primers CHP105 and CHP56, purified, and cloned into a TOPO vector. Finally, the plasmid was linearized and transformed into CH47, creating strain CH101 (Fig. S9).

SDS-PAGE, immunoblotting, and pilus purification. Procedures for SDS-PAGE and immunoblotting have been previously described (17). Briefly, glycoproteins were detected by immunoblotting of whole-cell lysates using rabbit antibodies and alkaline phosphatase-coupled goat anti-rabbit secondary antibodies (Sigma). Rabbit monoclonal antibodies npg1 and npg3 were used at a 1:10,000 dilution to detect di-*N*-acetylbacillosamine (diNAcBac) or diNAcBac-Gal-Gal-carrying glycoproteins, respectively (50). To detect PilE subunits, the rabbit polyclonal antibody 903 was used at a 1:2,500 dilution (51). Rabbit antiserum to the Laz protein was a gift from James Moir (52). Pili were purified by shear extraction in ethanolamine buffer and subsequent ammonium sulfate concentration as previously described (53).

LC-MS2 analysis of trypsin generated pilin peptides. Procedures for in-gel digestion was performed as previously described using trypsin (Sigma) (54) with the following adjustments. Destaining of gel pieces was done overnight with Milli-Q water. Generated peptides were Zip-tipped as described by the manufacturer (Merck), and the dried peptides were resuspended in a mixture of 2.5% acetonitrile (Acn) and 0.1% formic acid (FA) prior to liquid chromatography-tandem mass spectrometry (LC-MS2) analysis.

LC-MS2 was performed on a system consisting of a Dionex Ultimate 3000 RSLCnano-LC system (Sunnyvale CA, USA) connected to a hybrid quadrupole Orbitrap (QExactive) mass spectrometer (ThermoElectron, Bremen, Germany) equipped with a nanoelectrospray ion source. For LC separation, an Acclaim PepMap 100 column (C_{18}, 3 μm, 100 Å) (Dionex, Sunnyvale CA, USA) capillary of 25-cm bed length was used with a flow rate of 300 nL/min. Two solvents, A (2.5% Acn, 0.1% FA) and B (aqueous 90% Acn in 0.1% FA), were used to elute the peptides from the nano column. Peptide separation were achieved on a gradient from 3% to 55% (solvent B) over 45 min, before a final wash sequence of 90% solvent B for 2 min. The mass spectrometer was operated in the data-dependent mode to automatically switch between MS and MS2 acquisition. Survey full-scan MS spectra (from *m/z* 200 to 2,000) were acquired with the resolution $R = 70{,}000$ at *m/z* 200, with an automatic gain control (AGC) target of 1e6. The maximum allowed ion accumulation times were 100 ms. The sequential isolation of up to the seven most intense ions, depending on signal intensity (intensity threshold of 1e4), was considered for fragmentation using higher-energy collisional induced dissociation (HCD) at a target value of 100,000 charges and a resolution $R = 17{,}500$ with normalized collision energy (NCE) of 28. Target ions already selected for MS/MS were dynamically excluded for 30 s. The isolation window was $m/z = 1.5$ without offset. The maximum allowed ion accumulation for the MS/MS spectrum was 180 ms. For accurate mass measurements, the lock mass option was enabled in MS mode for internal recalibration during the analysis.

For the data analysis, the generated MS2 spectra were manually investigated using Qual browser 2.2 (Thermo Scientific) and glycopeptide spectra extracted based on the presence of reporter ions for di-*N*-acetylbacillosamine (diNAcBAc) (at *m/z* 229.118 and *m/z* 211.1079), diNAcBAc-Gal (at *m/z* 391.170), and diNAcBAc-AcGal (at *m/z* 433.181). Selected spectra were verified using ProteinProspector (http://prospector.ucsf.edu/prospector/mshome.htm), operated by the UCSF mass spectrometry facility.

Processing of proteome samples. Precipitated proteome samples were prepared using S-trap columns (Protifi, USA) according to the manufacturer's instructions. Briefly samples were resuspended in 5% SDS by boiling, and then protein amounts were quantified using a bicinchoninic acid (BCA) assay (Thermo Fisher Scientific). One hundred micrograms of each sample was then reduced with 10 mM dithiothreitol (DTT) at 95°C for 10 min, allowed to cool to room temperature, and then alkylated with 40 mM chloroacetamide for 30 min in the dark. Samples were then acidified with phosphoric acid to a final concentration of 1.2% and mixed with 7 volumes of 90% methanol–100 mM triethylammonium bicarbonate (TEAB) (pH 7.1) before being applied to S-trap minicolumns. Samples were washed four times with 90% methanol–100 mM TEAB (pH 7.1) to remove SDS, and then 8 μg of trypsin–Lys-c (Promega, USA) in 100 mM TEAB (pH 8.5) was spun through the S-trap columns. Samples were digested for 4 h at 47°C and then collected from the S-traps by washing with 100 mM TEAB (pH 8.5), followed by 0.2% FA, followed by 0.2% FA–50% Acn. Peptide washes were pooled, dried, and then resuspended in buffer A* (0.1% trifluoroacetic acid [TFA], 2% Acn) before being cleaned up with homemade high-capacity StageTips composed of 1 mg Empore C_{18} material (3M) and 5 mg of Oligo R3 reverse-phase resin (Thermo Fisher Scientific, USA) as previously described (55, 56). Columns were wet with buffer B (0.1% FA, 80% Acn) and conditioned with buffer A* prior to use. Resuspended samples were loaded onto conditioned columns and washed with 10 bed volumes of buffer A*, and bound peptides were eluted with buffer B before being dried and then stored at −20°C.

LC-MS analysis of proteome samples. StageTip cleaned up samples were resuspended in buffer A* and separated using a two-column chromatography setup composed of a PepMap100 C_{18} 20-mm by 75-μm trap and a PepMap C_{18} 500-mm by 75-μm analytical column (Thermo Fisher Scientific) coupled to a Orbitrap Fusion Lumos Tribrid mass spectrometer equipped with a FAIMS Pro interface (Thermo Fisher Scientific). Gradients of 145 min were run for each sample, altering the buffer composition from 2% buffer B to 28% B over 126 min, then from 28% B to 40% B over 9 min, then from 40% B to 80% B over 3 min, the composition was held at 80% B for 2 min, and then the composition was dropped to 2% B over 2 min and held at 2% B for another 3 min. The Lumos mass spectrometer was operated in a stepped FAIMS data-dependent mode at three different FAIMS compensation voltages (CVs) of −25, −45, and −65 as previously described (57). For each FAIMS CV, a single Orbitrap MS scan (*m/z* 350 to 2,000, maximal injection time of 50 ms, and AGC set to a maximum of 4×10^5 ions with a resolution of 60,000) was acquired every 1.35 s, followed by Orbitrap MS/MS HCD scans of precursors (NCE of 30%, maximal injection time of 80 ms, and AGC set to a maximum of 1.25×10^5 ions with a resolution of

30,000). HCD scans containing the oxonium ions (*m/z* 204.0867, 138.0545, 366.1396, 229.1189, or 211.1082) triggered three additional product-dependent MS/MS scans (58) of potential glycopeptides: a Orbitrap EThcD scan (NCE of 15%, maximal injection time of 250 ms, AGC set to a maximum of 2×10^5 ions with a resolution of 30,000, and using the extended mass range setting to improve the detection of high-mass glycopeptide fragment ions) (59); a ion trap CID scan (NCE of 35%, maximal injection time of 40 ms, and AGC set to a maximum of 5×10^4 ions) and a stepped collision energy HCD scan (NCE of 35% with 5% stepping, maximal injection time of 250 ms, and AGC set to a maximum of 2×10^5 ions with a resolution of 30,000).

Proteomic analyses. Data files were separated into individual FAIMS fractions using the FAIMS MzXML Generator (29) and processed with MaxQuant (v1.6.17.0) (60). Searches were performed against the *N. gonorrhoeae* MS11 proteome (UniProt accession no. UP000016457 [2,047 proteins]) and the *N. elongata pglO* sequence (UniProt accession no. D4DS59). Searches were undertaken using "Trypsin" enzyme specificity with carbamidomethylation of cysteine as a fixed modification. Oxidation of methionine and the glycan diNAcBAc-Hex2 (chemical composition: $C_{22}H_{36}O_{14}N_2$ [552.2166 Da]) were included as variable modifications, and a maximum of 2 missed cleavages was allowed. To enhance the identification of peptides between samples, the "Match between Runs" option was enabled with a precursor match window set to 2 min and an alignment window of 20 min with the label free quantitation (LFQ) option enabled (61). The resulting outputs were processed within the Perseus (v1.6.0.7) analysis environment (62) to remove reverse matches and common protein contaminants prior to further analysis. For LFQ comparisons, biological replicates were grouped, and missing values were then imputed based on the observed total peptide intensities, with a range of 0.3σ and a downshift of 2.5σ using Perseus. Student's *t* tests were undertaken at the protein and modified peptide levels to compare between groups, and the resulting data were exported to be visualized using ggplot2 (63) within R.

Data availability. The MS data and search results have been deposited into the PRIDE ProteomeXchange Consortium repository (64, 65) and can be accessed with the identifier PXD024737.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

FIG S1, TIF file, 18.6 MB.

FIG S2, TIF file, 3.4 MB.

FIG S3, TIF file, 0.5 MB.

FIG S4, TIF file, 2.7 MB.

FIG S5, TIF file, 0.7 MB.

FIG S6, TIF file, 0.8 MB.

FIG S7, TIF file, 2.3 MB.

FIG S8, TIF file, 1.7 MB.

FIG S9, TIF file, 0.3 MB.

TABLE S1, DOCX file, 0.1 MB.

ACKNOWLEDGMENTS

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. N.E.S. is supported by an Australian Research Council Future Fellowship (FT200100270) and an ARC Discovery Project Grant (DP210100362).

REFERENCES

1. Bai L, Li H. 2021. Protein *N*-glycosylation and *O*-mannosylation are catalyzed by two evolutionarily related GT-C glycosyltransferases. Curr Opin Struct Biol 68:66–73. https://doi.org/10.1016/j.sbi.2020.12.009.
2. Lizak C, Gerber S, Numao S, Aebi M, Locher KP. 2011. X-ray structure of a bacterial oligosaccharyltransferase. Nature 474:350–355. https://doi.org/10.1038/nature10151.
3. Power PM, Seib KL, Jennings MP. 2006. Pilin glycosylation in *Neisseria meningitidis* occurs by a similar pathway to *wzy*-dependent O-antigen biosynthesis in *Escherichia coli*. Biochem Biophys Res Commun 347:904–908. https://doi.org/10.1016/j.bbrc.2006.06.182.
4. 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 65:607–624. https://doi.org/10.1111/j.1365-2958.2007.05806.x.
5. Iwashkiw JA, Seper A, Weber BS, Scott NE, Vinogradov E, Stratilo C, Reiz B, Cordwell SJ, Whittal R, Schild S, Feldman MF. 2012. Identification of a general *O*-linked protein glycosylation system in *Acinetobacter baumannii* and its role in virulence and biofilm formation. PLoS Pathog 8:e1002758. https://doi.org/10.1371/journal.ppat.1002758.
6. Elhenawy W, Scott NE, Tondo ML, Orellano EG, Foster LJ, Feldman MF. 2016. Protein *O*-linked glycosylation in the plant pathogen *Ralstonia solanacearum*. Glycobiology 26:301–311. https://doi.org/10.1093/glycob/cwv098.
7. Egge-Jacobsen W, Salomonsson EN, Aas FE, Forslund AL, Winther-Larsen HC, Maier J, Macellaro A, Kuoppa K, Oyston PC, Titball RW, Thomas RM,

Forsberg A, Prior JL, Koomey M. 2011. *O*-linked glycosylation of the PilA pilin protein of *Francisella tularensis*: identification of the endogenous protein-targeting oligosaccharyltransferase and characterization of the native oligosaccharide. J Bacteriol 193:5487–5497. https://doi.org/10.1128/JB.00383-11.
8. Gebhart C, Ielmini MV, Reiz B, Price NL, Aas FE, Koomey M, Feldman MF. 2012. Characterization of exogenous bacterial oligosaccharyltransferases in *Escherichia coli* reveals the potential for *O*-linked protein glycosylation in *Vibrio cholerae* and *Burkholderia thailandensis*. Glycobiology 22:962–974. https://doi.org/10.1093/glycob/cws059.
9. Liu CF, Tonini L, Malaga W, Beau M, Stella A, Bouyssie D, Jackson MC, Nigou J, Puzo G, Guilhot C, Burlet-Schiltz O, Riviere M. 2013. Bacterial protein-*O*-mannosylating enzyme is crucial for virulence of *Mycobacterium tuberculosis*. Proc Natl Acad Sci U S A 110:6560–6565. https://doi.org/10.1073/pnas.1219704110.
10. Bai L, Wang T, Zhao G, Kovach A, Li H. 2018. The atomic structure of a eukaryotic oligosaccharyltransferase complex. Nature 555:328–333. https://doi.org/10.1038/nature25755.
11. Matsumoto S, Shimada A, Nyirenda J, Igura M, Kawano Y, Kohda D. 2013. Crystal structures of an archaeal oligosaccharyltransferase provide insights into the catalytic cycle of *N*-linked protein glycosylation. Proc Natl Acad Sci U S A 110:17868–17873. https://doi.org/10.1073/pnas.1309777110.
12. Sjodt M, Brock K, Dobihal G, Rohs PDA, Green AG, Hopf TA, Meeske AJ, Srisuknimit V, Kahne D, Walker S, Marks DS, Bernhardt TG, Rudner DZ, Kruse AC. 2018. Structure of the peptidoglycan polymerase RodA resolved by evolutionary coupling analysis. Nature 556:118–121. https://doi.org/10.1038/nature25985.
13. Wild R, Kowal J, Eyring J, Ngwa EM, Aebi M, Locher KP. 2018. Structure of the yeast oligosaccharyltransferase complex gives insight into eukaryotic *N*-glycosylation. Science 359:545–550. https://doi.org/10.1126/science.aar5140.
14. Castric P. 1995. *pilO*, a gene required for glycosylation of *Pseudomonas aeruginosa* 1244 pilin. Microbiology (Reading) 141:1247–1254. https://doi.org/10.1099/13500872-141-5-1247.
15. Harding CM, Nasr MA, Kinsella RL, Scott NE, Foster LJ, Weber BS, Fiester SE, Actis LA, Tracy EN, Munson RS, Jr, Feldman MF. 2015. *Acinetobacter* strains carry two functional oligosaccharyltransferases, one devoted exclusively to type IV pilin, and the other one dedicated to *O*-glycosylation of multiple proteins. Mol Microbiol 96:1023–1041. https://doi.org/10.1111/mmi.12986.
16. Kus JV, Kelly J, Tessier L, Harvey H, Cvitkovitch DG, Burrows LL. 2008. Modification of *Pseudomonas aeruginosa* Pa5196 type IV pilins at multiple sites with D-Araf by a novel GT-C family arabinosyltransferase, TfpW. J Bacteriol 190:7464–7478. https://doi.org/10.1128/JB.01075-08.
17. 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.
18. Anonsen JH, Vik A, Borud B, Viburiene R, Aas FE, Kidd SW, 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.
19. Aho EL, Keating AM, McGillivray SM. 2000. A comparative analysis of pilin genes from pathogenic and nonpathogenic *Neisseria* species. Microb Pathog 28:81–88. https://doi.org/10.1006/mpat.1999.0325.
20. Wormann ME, Horien CL, Johnson E, Liu G, Aho E, Tang CM, Exley RM. 2016. *Neisseria cinerea* isolates can adhere to human epithelial cells by type IV pilus-independent mechanisms. Microbiology (Reading) 162:487–502. https://doi.org/10.1099/mic.0.000248.
21. Bennett JS, Jolley KA, Sparling PF, Saunders NJ, Hart CA, Feavers IM, Maiden MC. 2007. Species status of *Neisseria gonorrhoeae*: evolutionary and epidemiological inferences from multilocus sequence typing. BMC Biol 5:35. https://doi.org/10.1186/1741-7007-5-35.
22. Diallo K, Trotter C, Timbine Y, Tamboura B, Sow SO, Issaka B, Dano ID, Collard JM, Dieng M, Diallo A, Mihret A, Ali OA, Aseffa A, Quaye SL, Bugri A, Osei I, Gamougam K, Mbainadji L, Daugla DM, Gadzama G, Sambo ZB, Omotara BA, Bennett JS, Rebbetts LS, Watkins ER, Nascimento M, Woukeu A, Manigart O, Borrow R, Stuart JM, Greenwood BM, Maiden MCJ. 2016. Pharyngeal carriage of *Neisseria* species in the African meningitis belt. J Infect 72:667–677. https://doi.org/10.1016/j.jinf.2016.03.010.
23. Gault J, Ferber M, Machata S, Imhaus AF, Malosse C, Charles-Orszag A, Millien C, Bouvier G, Bardiaux B, Pehau-Arnaudet G, Klinge K, Podglajen I, Ploy MC, Seifert HS, Nilges M, Chamot-Rooke J, Dumenil G. 2015. *Neisseria meningitidis* type IV pili composed of sequence invariable pilins are masked by multisite glycosylation. PLoS Pathog 11:e1005162. https://doi.org/10.1371/journal.ppat.1005162.
24. Ku SC, Schulz BL, Power PM, Jennings MP. 2009. The pilin *O*-glycosylation pathway of pathogenic *Neisseria* is a general system that glycosylates AniA, an outer membrane nitrite reductase. Biochem Biophys Res Commun 378:84–89. https://doi.org/10.1016/j.bbrc.2008.11.025.
25. Bennett JS, Jolley KA, Earle SG, Corton C, Bentley SD, Parkhill J, Maiden MC. 2012. A genomic approach to bacterial taxonomy: an examination and proposed reclassification of species within the genus *Neisseria*. Microbiology (Reading) 158:1570–1580. https://doi.org/10.1099/mic.0.056077-0.
26. Hadjineophytou C, Anonsen JH, Wang N, Ma KC, Viburiene R, Vik A, Harrison OB, Maiden MCJ, Grad YH, Koomey M. 2019. Genetic determinants of genus-level glycan diversity in a bacterial protein glycosylation system. PLoS Genet 15:e1008532. https://doi.org/10.1371/journal.pgen.1008532.
27. Wang N, Anonsen JH, Hadjineophytou C, Reinar WB, Borud B, Vik A, Koomey M. 2021. Allelic polymorphisms in a glycosyltransferase gene shape glycan repertoire in the *O*-linked protein glycosylation system of *Neisseria*. Glycobiology 31:477–491. https://doi.org/10.1093/glycob/cwaa073.
28. 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.
29. Hebert AS, Prasad S, Belford MW, Bailey DJ, McAlister GC, Abbatiello SE, Huguet R, Wouters ER, Dunyach JJ, Brademan DR, Westphall MS, Coon JJ. 2018. Comprehensive single-shot proteomics with FAIMS on a hybrid Orbitrap mass spectrometer. Anal Chem 90:9529–9537. https://doi.org/10.1021/acs.analchem.8b02233.
30. Hayes AJ, Lewis JM, Davies MR, Scott NE. 2021. Burkholderia PglL enzymes are serine preferring oligosaccharyltransferases which target conserved proteins across the *Burkholderia* genus. Commun Biol 4:1045. https://doi.org/10.1038/s42003-021-02588-y.
31. Ielmini MV, Feldman MF. 2011. *Desulfovibrio desulfuricans* PglB homolog possesses oligosaccharyltransferase activity with relaxed glycan specificity and distinct protein acceptor sequence requirements. Glycobiology 21:734–742. https://doi.org/10.1093/glycob/cwq192.
32. Mills DC, Jervis AJ, Abouelhadid S, Yates LE, Cuccui J, Linton D, Wren BW. 2016. Functional analysis of *N*-linking oligosaccharyl transferase enzymes encoded by deep-sea vent proteobacteria. Glycobiology 26:398–409. https://doi.org/10.1093/glycob/cwv111.
33. 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.
34. Zollner R, Cronenberg T, Kouzel N, Welker A, Koomey M, Maier B. 2019. Type IV pilin post-translational modifications modulate material properties of bacterial colonies. Biophys J 116:938–947. https://doi.org/10.1016/j.bpj.2019.01.020.
35. Harvey H, Bondy-Denomy J, Marquis H, Sztanko KM, Davidson AR, Burrows LL. 2018. *Pseudomonas aeruginosa* defends against phages through type IV pilus glycosylation. Nat Microbiol 3:47–52. https://doi.org/10.1038/s41564-017-0061-y.
36. Harvey H, Kus JV, Tessier L, Kelly J, Burrows LL. 2011. *Pseudomonas aeruginosa* D-arabinofuranose biosynthetic pathway and its role in type IV pilus assembly. J Biol Chem 286:28128–28137. https://doi.org/10.1074/jbc.M111.255794.
37. Bille E, Meyer J, Jamet A, Euphrasie D, Barnier JP, Brissac T, Larsen A, Pelissier P, Nassif X. 2017. A virulence-associated filamentous bacteriophage of *Neisseria meningitidis* increases host-cell colonisation. PLoS Pathog 13:e1006495. https://doi.org/10.1371/journal.ppat.1006495.
38. Piekarowicz A, Kłyż A, Majchrzak M, Szczêsna E, Piechucki M, Kwiatek A, Maugel TK, Stein DC. 2014. *Neisseria gonorrhoeae* filamentous phage NgoPhi6 is capable of infecting a variety of Gram-negative bacteria. J Virol 88:1002–1010. https://doi.org/10.1128/JVI.02707-13.
39. Jennings MP, Jen FE, Roddam LF, Apicella MA, Edwards JL. 2011. *Neisseria gonorrhoeae* pilin glycan contributes to CR3 activation during challenge of primary cervical epithelial cells. Cell Microbiol 13:885–896. https://doi.org/10.1111/j.1462-5822.2011.01586.x.
40. Banerjee A, Wang R, Supernavage SL, Ghosh SK, Parker J, Ganesh NF, Wang PG, Gulati S, Rice PA. 2002. Implications of phase variation of a gene (*pgtA*) encoding a pilin galactosyl transferase in gonococcal pathogenesis. J Exp Med 196:147–162. https://doi.org/10.1084/jem.20012022.
41. 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.
42. Kellogg DS, Jr, Cohen IR, Norins LC, Schroeter AL, Reising G. 1968. *Neisseria gonorrhoeae*. II. Colonial variation and pathogenicity during 35 months *in vitro*. J Bacteriol 96:596–605. https://doi.org/10.1128/jb.96.3.596-605.1968.
43. Wolfgang M, van Putten JP, Hayes SF, Dorward D, Koomey M. 2000. Components and dynamics of fiber formation define a ubiquitous biogenesis pathway for bacterial pili. EMBO J 19:6408–6418. https://doi.org/10.1093/emboj/19.23.6408.
44. Tonjum T, Freitag NE, Namork E, Koomey M. 1995. Identification and characterization of *pilG*, a highly conserved pilus-assembly gene in pathogenic *Neisseria*. Mol Microbiol 16:451–464. https://doi.org/10.1111/j.1365-2958.1995.tb02410.x.
45. Jolley KA, Bray JE, Maiden MCJ. 2018. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. Wellcome Open Res 3:124. https://doi.org/10.12688/wellcomeopenres.14826.1.
46. Salomonsson E, Forsberg A, Roos N, Holz C, Maier B, Koomey M, Winther-Larsen HC. 2009. Functional analyses of pilin-like proteins from *Francisella tularensis*: complementation of type IV pilus phenotypes in *Neisseria gonorrhoeae*. Microbiology (Reading) 155:2546–2559. https://doi.org/10.1099/mic.0.028183-0.
47. Aas FE, Wolfgang M, Frye S, Dunham S, Lovold C, Koomey M. 2002. Competence for natural transformation in *Neisseria gonorrhoeae*: components of DNA binding and uptake linked to type IV pilus expression. Mol Microbiol 46:749–760. https://doi.org/10.1046/j.1365-2958.2002.03193.x.
48. Mehr IJ, Long CD, Serkin CD, Seifert HS. 2000. A homologue of the recombination-dependent growth gene, *rdgC*, is involved in gonococcal pilin antigenic variation. Genetics 154:523–532. https://doi.org/10.1093/genetics/154.2.523.
49. Johnston DM, Cannon JG. 1999. Construction of mutant strains of *Neisseria gonorrhoeae* lacking new antibiotic resistance markers using a two gene cassette with positive and negative selection. Gene 236:179–184. https://doi.org/10.1016/s0378-1119(99)00238-3.
50. Borud 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 192:2816–2829. https://doi.org/10.1128/JB.00101-10.
51. Winther-Larsen HC, Wolfgang MC, van Putten JP, Roos N, Aas FE, Egge-Jacobsen WM, Maier B, Koomey M. 2007. *Pseudomonas aeruginosa* type IV pilus expression in *Neisseria gonorrhoeae*: effects of pilin subunit composition on function and organelle dynamics. J Bacteriol 189:6676–6685. https://doi.org/10.1128/JB.00407-07.
52. Deeudom M, Huston W, Moir JW. 2015. Lipid-modified azurin of *Neisseria meningitidis* is a copper protein localized on the outer membrane surface and not regulated by FNR. Antonie Van Leeuwenhoek 107:1107–1116. https://doi.org/10.1007/s10482-015-0400-z.
53. Wolfgang M, Lauer P, Park HS, Brossay L, Hebert J, Koomey M. 1998. PilT mutations lead to simultaneous defects in competence for natural transformation and twitching motility in piliated *Neisseria gonorrhoeae*. Mol Microbiol 29:321–330. https://doi.org/10.1046/j.1365-2958.1998.00935.x.
54. Anonsen JH, Egge-Jacobsen W, Aas FE, Borud B, Koomey M, Vik A. 2012. Novel protein substrates of the phospho-form modification system in *Neisseria gonorrhoeae* and their connection to *O*-linked protein glycosylation. Infect Immun 80:22–30. https://doi.org/10.1128/IAI.05920-11.
55. Ishihama Y, Rappsilber J, Mann M. 2006. Modular stop and go extraction tips with stacked disks for parallel and multidimensional peptide fractionation in proteomics. J Proteome Res 5:988–994. https://doi.org/10.1021/pr050385q.
56. Rappsilber J, Mann M, Ishihama Y. 2007. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat Protoc 2:1896–1906. https://doi.org/10.1038/nprot.2007.261.
57. Ahmad Izaham AR, Ang CS, Nie S, Bird LE, Williamson NA, Scott NE. 2021. What are we missing by using hydrophilic enrichment? Improving bacterial glycoproteome coverage using total proteome and FAIMS analyses. J Proteome Res 20:599–612. https://doi.org/10.1021/acs.jproteome.0c00565.
58. Saba J, Dutta S, Hemenway E, Viner R. 2012. Increasing the productivity of glycopeptides analysis by using higher-energy collision dissociation-accurate mass-product-dependent electron transfer dissociation. Int J Proteomics 2012:560391. https://doi.org/10.1155/2012/560391.
59. Pandey A, Cleary DW, Laver JR, Gorringe A, Deasy AM, Dale AP, Morris PD, Didelot X, Maiden MCJ, Read RC. 2018. Microevolution of *Neisseria lactamica* during nasopharyngeal colonisation induced by controlled human infection. Nat Commun 9:4753. https://doi.org/10.1038/s41467-018-07235-5.
60. Cox J, Mann M. 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol 26:1367–1372. https://doi.org/10.1038/nbt.1511.
61. Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, Mann M. 2014. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Mol Cell Proteomics 13:2513–2526. https://doi.org/10.1074/mcp.M113.031591.
62. Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, Mann M, Cox J. 2016. The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat Methods 13:731–740. https://doi.org/10.1038/nmeth.3901.
63. Wickham H. 2016. ggplot2: elegant graphics for data analysis. Springer-Verlag, New York, NY.
64. Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, Inuganti A, Griss J, Mayer G, Eisenacher M, Perez E, Uszkoreit J, Pfeuffer J, Sachsenberg T, Yilmaz S, Tiwary S, Cox J, Audain E, Walzer M, Jarnuczak AF, Ternent T, Brazma A, Vizcaino JA. 2019. The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res 47:D442–D450. https://doi.org/10.1093/nar/gky1106.
65. Vizcaino JA, Csordas A, Del-Toro N, Dianes JA, Griss J, Lavidas I, Mayer G, Perez-Riverol Y, Reisinger F, Ternent T, Xu QW, Wang R, Hermjakob H. 2016. 2016 update of the PRIDE database and its related tools. Nucleic Acids Res 44:11033. https://doi.org/10.1093/nar/gkw880.
66. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797. https://doi.org/10.1093/nar/gkh340.
67. Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870–1874. https://doi.org/10.1093/molbev/msw054.
68. Tamura K, Nei M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol 10:512–526. https://doi.org/10.1093/oxfordjournals.molbev.a040023.
69. Kall L, Krogh A, Sonnhammer EL. 2007. Advantages of combined transmembrane topology and signal peptide prediction—the Phobius web server. Nucleic Acids Res 35:W429–32. https://doi.org/10.1093/nar/gkm256.
70. Omasits U, Ahrens CH, Muller S, Wollscheid B. 2014. Protter: interactive protein feature visualization and integration with experimental proteomic data. Bioinformatics 30:884–886. https://doi.org/10.1093/bioinformatics/btt607.
71. 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 49:833–847. https://doi.org/10.1046/j.1365-2958.2003.03602.x.