

Glycoproteomic and proteomic analysis of *Burkholderia cenocepacia* reveals glycosylation events within FliF and MotB are dispensable for motility

Jessica M. Lewis,¹ Leila Jebeli,¹ Pauline M. L. Coulon,¹ Catrina E. Lay,¹ Nichollas E. Scott¹

AUTHOR AFFILIATION See affiliation list on p. 18.

ABSTRACT Across the *Burkholderia* genus O-linked protein glycosylation is highly conserved. While the inhibition of glycosylation has been shown to be detrimental for virulence in *Burkholderia cepacia* complex species, such as *Burkholderia cenocepacia*, little is known about how specific glycosylation sites impact protein functionality. Within this study, we sought to improve our understanding of the breadth, dynamics, and requirement for glycosylation across the *B. cenocepacia* O-glycoproteome. Assessing the *B. cenocepacia* glycoproteome across different culture media using complementary glycoproteomic approaches, we increase the known glycoproteome to 141 glycoproteins. Leveraging this repertoire of glycoproteins, we quantitatively assessed the glycoproteome of *B. cenocepacia* using Data-Independent Acquisition (DIA) revealing the *B. cenocepacia* glycoproteome is largely stable across conditions with most glycoproteins constitutively expressed. Examination of how the absence of glycosylation impacts the glycoproteome reveals that the protein abundance of only five glycoproteins (BCAL1086, BCAL2974, BCAL0525, BCAM0505, and BCAL0127) are altered by the loss of glycosylation. Assessing Δ fliF (Δ BCAL0525), Δ motB (Δ BCAL0127), and Δ BCAM0505 strains, we demonstrate the loss of FliF, and to a lesser extent MotB, mirror the proteomic effects observed in the absence of glycosylation in Δ pglL. While both MotB and FliF are essential for motility, we find loss of glycosylation sites in MotB or FliF does not impact motility supporting these sites are dispensable for function. Combined this work broadens our understanding of the *B. cenocepacia* glycoproteome supporting that the loss of glycoproteins in the absence of glycosylation is not an indicator of the requirement for glycosylation for protein function.

IMPORTANCE *Burkholderia cenocepacia* is an opportunistic pathogen of concern within the Cystic Fibrosis community. Despite a greater appreciation of the unique physiology of *B. cenocepacia* gained over the last 20 years a complete understanding of the proteome and especially the O-glycoproteome, is lacking. In this study, we utilize systems biology approaches to expand the known *B. cenocepacia* glycoproteome as well as track the dynamics of glycoproteins across growth phases, culturing media and in response to the loss of glycosylation. We show that the glycoproteome of *B. cenocepacia* is largely stable across conditions and that the loss of glycosylation only impacts five glycoproteins including the motility associated proteins FliF and MotB. Examination of MotB and FliF shows, while these proteins are essential for motility, glycosylation is dispensable. Combined this work supports that *B. cenocepacia* glycosylation can be dispensable for protein function and may influence protein properties beyond stability.

KEYWORDS glycosylation, *Burkholderia cenocepacia*, *Burkholderia*, post-translational modifications, proteomics, PglL, glycoproteomics

Editor Silvia T. Cardona, University of Manitoba, Winnipeg, Canada

Ad Hoc Peer Reviewers Jennifer Geddes-McAlister, University of Guelph, Guelph, Canada; Bart Devreese, Ghent University, Ghent, Belgium

Address correspondence to Nichollas E. Scott, Nichollas.scott@unimelb.edu.au.

Jessica M. Lewis, Leila Jebeli, and Pauline M. L. Coulon contributed equally to this article. Joint first author order was determined by length of time spend working on the project. This project was a true team effort involving multiple team members over the course of three years.

The authors declare no conflict of interest.

See the funding table on p. 19.

Received 10 February 2024

Accepted 16 April 2024

Published 6 May 2024

Copyright © 2024 Lewis et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

The *Burkholderia cepacia* complex (Bcc) is a group of at least 20 phenotypically similar yet genetically distinct species of increasing clinical importance (1–3). Bcc members are associated with chronic and life-threatening infections within sensitive populations, such as individuals with cystic fibrosis (CF) (1, 4), where they can establish persistent infections (4, 5). Bcc colonization of the CF lung is associated with accelerated lung function decline and reduced life expectancy (6, 7), leading to elevated rates of morbidity and mortality compared to other bacterial infections (8–11). While aggressive antibiotic regimes are utilized to control these infections (12), the intrinsic resistance of Bcc members makes the eradication of these infections challenging (13). *Burkholderia cenocepacia*, a predominant member of the Bcc, is of particular concern to the CF community as it is associated with a heightened risk of developing “Cepacia syndrome,” a rapidly necrotizing form of pneumonia that can result in fulminating septicemia, which is nearly uniformly fatal (4, 14). While it is now recognized that *B. cenocepacia* infections are associated with poor clinical prognosis (15) the underlying drivers of these clinical differences are still unclear.

Over the past 20 years, significant strides have been made in our understanding of *B. cenocepacia* CF lung colonization (6, 7), yet how the repertoire of known *B. cenocepacia* virulence factors (16, 17) is coordinated under these conditions remains largely unclear. Within the CF lung, *B. cenocepacia* is observed extracellularly within luminal mucus as well as within macrophages of the mucopurulent material (5, 18). Due to growth under similar conditions, several CF-like media models established for *Pseudomonas aeruginosa*, including artificial sputum medium (ASM) (19, 20) and synthetic cystic fibrosis medium (21), have been used to probe the response of *B. cenocepacia* to extracellular growth in CF-like conditions. Indeed, multiple metabolic (22) and transcriptomic (23–25) studies have now shown that under CF-like conditions, *B. cenocepacia* reveals marked physiological alterations such as changes in motility and flagellin expression (26), as well as shifts in metabolism toward phenylacetic acid degradation (23, 27). While these studies have proven insightful for understanding transcriptional changes, whether these changes reflect proteome-level events remains unclear (28). As post-translational regulation is important for controlling virulence within *B. cenocepacia*, as noted for the iron starvation σ factor OrbS (25), understanding the proteome changes under CF-like conditions may provide valuable insights into the regulation of virulence. This is especially true for post-translational protein modifications, such as glycosylation, which may be influenced by the interplay of glycosylation machinery, sugar precursor levels, and protein substrate availability.

While once thought to be absent in bacterial systems, glycosylation is now known to be commonplace (29, 30) within multiple species including members of the *Campylobacter* (31, 32), *Neisseria* (33–37), and *Burkholderia* (38–40) genera. Within *Burkholderia* species multiple proteins are targeted for glycosylation by a conserved general O-linked glycosylation system encoded by two non-genetically linked components; the O-linked Glycosylation Cluster (*ogc*), responsible for the biosynthesis of the trisaccharide β -Gal-(1,3)- α -GalNAc-(1,3)- β -GalNAc used for O-linked glycosylation (40); and the O-linked oligosaccharyltransferase PglL, required for the *en bloc* transfer of the O-glycan to serine residues of compatible periplasmic proteins (38, 41). Within *B. cenocepacia* the loss of general O-linked glycosylation leads to dramatic changes in virulence (40, 41) with similar impacts on virulence also noted in the non-CF pathogen *Burkholderia pseudomallei* (42). While these observations support the involvement of O-linked glycosylation in *Burkholderia* virulence the specific glycoproteins and glycosylation sites remain unclear. Recent studies have now expanded the known *B. cenocepacia* glycoproteome to more than >100 glycoproteins yet most glycoproteins have no known function (38, 39, 43). Furthermore, no information is currently available on the abundance of these glycoproteins under differing growth conditions including host-like stimuli. Thus, to improve our understanding of the *B. cenocepacia* glycoproteome and define which glycoproteins may be associated with virulence further insights into the dynamics of glycoproteins are needed.

The analysis of bacterial proteomes using liquid chromatography–mass spectrometry (LC-MS) has emerged as the quintessential approach for the quantitation of microbial proteomes and post-translational modifications (44, 45). Within bottom-up proteomics, the most widely implemented form of proteomics, proteins are first digested into peptides prior to analysis (46). Once digested, peptides can be analyzed directly or subjected to enrichment approaches to enhance the detection of specific peptides of interest, such as glycopeptides in the case of glycoproteomics (47). In glycoproteomics, a range of enrichment approaches can be used to isolate glycopeptides (47) with Zwitterionic hydrophilic interaction liquid chromatography (ZIC-HILIC) (38, 39, 41, 48–50) and Field Asymmetric waveform Ion Mobility Spectrometry (FAIMS) (43, 51, 52) being two commonly utilized approaches. These enrichment approaches exploit the hydrophilic nature of glycans, in the case of ZIC-HILIC, or the shape of glycopeptide ions, in the case of FAIMS, to allow the separation of glycosylated peptides from non-glycosylated peptides improving the depth of glycoproteomic analysis. For both enriched and total proteome analyses, the resulting samples are analyzed by LC-MS, where in peptides/glycopeptides are subjected to MS-based fragmentation, and the resulting spectra analyzed using computational tools which match the resulting spectra to a protein database. While traditionally bottom-up proteomics using data-dependent acquisition (DDA) has been used for bacterial proteome analysis (53–55) and the analysis of glycopeptides (38, 39, 41, 43, 48–52), the stochastic nature of this method leads to missing data points across samples compromising quantitative analysis, especially within low abundant proteins (56). To address these issues data-independent acquisition (DIA) has emerged as an alternative to DDA, which leads to a dramatic reduction in missing data points across studies while enabling deep proteome analysis (57, 58). Within DIA, the highly reproducible sampling across the LC gradient enables the reproducible selection and fragmentation of peptides even allowing the detection of peptides not observable on the MS1 level (59, 60), with recent improvements in DIA informatics broadening the accessibility of this approach (57, 61–63). Combined these advancements now allow DIA to provide deep bottom-up proteomic coverage and track even challenging protein classes such as microproteins (<100 amino acid in length) (64) as well as low abundant proteins (59, 60).

Within this study, we have sought to consolidate our understanding of the known *B. cenocepacia* glycoproteome and define the dynamics of glycoproteins under different growth conditions, growth phases, and in response to glycosylation inhibition. Complementing our previous studies (38, 39, 43), we analyzed the glycoproteome of *B. cenocepacia* across multiple growth conditions identifying that at least 141 glycoproteins are subjected to glycosylation within *B. cenocepacia*. Using DIA analysis, we demonstrate most glycoproteins are stably expressed across growth phases and culturing conditions yet a subset of five glycoproteins undergo dramatic reductions in protein abundance in the absence of glycosylation. Of these glycoproteins, we demonstrate the glycoproteins FliF (BCAL0525) and MotB (BCAL0127) are essential for motility yet the known glycosylation sites within these proteins appear dispensable for protein function. Overall, this study expands the known *B. cenocepacia* glycoproteome and demonstrates glycosylation sites, even within proteins that appear impacted by the loss of glycosylation, can be dispensable for protein functionality.

MATERIALS AND METHODS

Bacterial strains and growth conditions

Strains used in this study are listed in Table S10. Strains of *B. cenocepacia* were grown at 37°C in liquid cultures to the desired growth phase in one of three different media: Lysogeny Broth (LB), Tryptic Soy Broth (TSB) or Artificial Sputum Media (ASM). Cultures were inoculated 1/10 with 0.02 OD_{600nm} of overnight LB cultures and allowed to grow for either 8 or 15 h (defined as Log and Stationary phase respectively; Fig. S1) with shaking at 185 rpm. LB and TSB media were prepared according to the manufacturer's

instructions (BD, New Jersey, USA), while the ASM was prepared according to previous reports (65–67). Individual components of the ASM, except for egg yolk emulsion (Oxiod/Thermo Fisher Scientific) and porcine mucin (Sigma), were filter sterilized using a 0.22- μ M filter. Porcine stomach mucin was sterilized by dissolving in Milli-Q water, 0.25 mg/mL, then autoclaving. After sterilization/filtering, individual components of the ASM were combined to a final concentration of 10 mg/mL mucin (Sigma-Aldrich), 1.39 mg/mL salmon sperm DNA (Sigma-Aldrich), 5 mg/mL NaCl, 2.2 mg/mL KCl, 0.22 mg/mL CaCl₂, 5 mg/mL Casamino acids (Bacto/Thermo Fisher Scientific, USA), 10 mg/mL Bovine Serum Albumin (Bovogen Biologicals, Australia), 0.5% egg yolk emulsion, and 0.005% diethylenetriaminepentaacetic acid (Sigma) in NaOH adjusted Milli-Q water (pH 14, Sigma-Aldrich). The pH of the final media was confirmed to be ~7 and the sterility of batches was checked by plating 100 μ L of artificial sputum media onto LB plates and incubating for 24 h at 37°C. *B. cenocepacia* strains were grown to the required growth phase within the media of interest and then collected by centrifugation at 10,000 \times g at 4°C for 10 min and washed three times with ice-cold phosphate-buffered saline (PBS) to remove media contaminants. Washed cells were then snap frozen and then stored at –80°C until processing.

When required, antibiotics were added to a final concentration of 50 μ g/mL trimethoprim for *Escherichia coli* and 100 μ g/mL for *B. cenocepacia*, 20 μ g/mL tetracycline for *E. coli*, and 150 μ g/mL for *B. cenocepacia* and 40 μ g/mL kanamycin for *E. coli*. Ampicillin was used at 100 μ g/mL and polymyxin B at 25 μ g/mL for triparental mating to select against donor and helper *E. coli* strains as previously described (68). Induction of complemented *B. cenocepacia* strains was undertaken by the addition of L-rhamnose monohydrate (Sigma-Aldrich, final concentration 0.1%) to agar plates used for motility assays as well as liquid cultures used for western blotting and proteomic experiments. Antibiotics were purchased from Thermo Fisher Scientific while all other chemicals, unless otherwise stated, were provided by Sigma-Aldrich.

Recombinant DNA methods

Oligonucleotides used in this study are provided in Table S12. pGPI-Scel (69) mutagenesis constructs pGPI-Scel- Δ *flfF* (BCAL0525), pGPI-Scel- Δ BCAM0505, and pGPI-Scel- Δ *motB* (BCAL0127) were generated using Gibson assembly (70) by the insertion of PCR amplified fragments into *Sma*I linearized pGPI-Scel using NEBuilder HiFi DNA master mix according to the manufacturer's instructions (New England Biolabs). The inducible pSCRhaB2-*flfF*-his₁₀ and pSCRhaB2-*motB*-his₁₀ constructs were generated using Gibson assembly (70) by inserting PCR amplified products into *Nde*I and *Xba*I linearized pSCRhaB2 (71) using NEBuilder HiFi DNA master mix. pSCRhaB2 site-directed mutagenesis was undertaken using PCR-based site replacement and *Dpn*I digestion (72). All restriction endonuclease digests, and agarose gel electrophoresis were performed using standard molecular biology techniques with chemically competent *E. coli* pir2 cells transformed using heat shock-based transformation (72). PCR amplifications were carried out using Q5 DNA polymerase (New England Biolabs) according to the manufacturer's recommendations with the addition of 2% DMSO for the amplification of *B. cenocepacia* DNA, due to its high GC content. Genomic DNA isolations were performed using genomic DNA clean-up Kits (Zymo Research), while PCR recoveries and restriction digest purifications were performed using DNA Clean & Concentrator Kits (Zymo Research). Colony and screening PCRs were performed using GoTaq DNA polymerase (Promega). Plasmids were confirmed by either Sanger sequencing undertaken at the Australian Genome Research Facility (Melbourne, Australia) or nanopore plasmid sequencing using Plasmidsaurus (SNPsaurus LLC, Eugene, OR) with a complete list of the plasmids used within this study provided in Table S11.

Construction of unmarked Δ *fliF*, Δ *motB*, and Δ BCAM0505 deletions and the introduction of pSCrhaB2 plasmids into *B. cenocepacia*

Deletions of genes of interest were undertaken using the approach of Flannagan *et al.* for the construction of unmarked, non-polar deletions in *B. cenocepacia* (69). Plasmids were introduced into *B. cenocepacia* K56-2 (73) or *B. cenocepacia* K56-2 Δ *pglL* (55) by triparental mating with the aid of *E. coli* containing pRK2013 (74). Removal of the integrated pGPI-Scel plasmids to generate required deletions was achieved by the introduction of pDAI-Scel-SacB (69, 75) by triparental mating. Mutagenesis was confirmed using screening oligonucleotides which bind outside the region of recombination (Table S12) using GoTaq DNA polymerase-based PCR screening supplemented with 10% DMSO. Complementation (pSCrhaB2) plasmids were introduced into strains of interest by triparental mating with the aid of *E. coli* containing pRK2013 (74) with confirmation of complementation with the correct plasmid undertaken using PCR-based screening and Sanger sequencing.

Preparation of proteomic samples

Frozen whole-cell samples were prepared for analysis using the in-StageTip preparation approach as previously described (76). Cells were resuspended in 4% sodium deoxycholate (SDC), 100 mM Tris pH 8.0 and boiled at 95°C with shaking (2,000 rpm) for 10 min to solubilize the proteome. Samples were allowed to cool for 10 min and then boiled for a further 10 min (95°C, 2,000 rpm) before the protein concentrations were determined by bicinchoninic acid assays (Thermo Fisher Scientific). For Zwitterionic Hydrophilic Interaction Liquid Chromatography (ZIC-HILIC) glycoproteomic enrichments, 1 mg of protein from four biological replicates of each growth media at stationary phase were prepared for glycopeptide enrichment. For FAIMS-based glycopeptide enrichment and quantitative proteomic comparisons, 100 μ g of protein from four biological replicates of each condition at both log and stationary phases were prepared. Samples were reduced/alkylated with the addition of Tris-2-carboxyethyl phosphine hydrochloride and Chloroacetamide (final concentration 10 and 40 mM, respectively), and samples were incubated in the dark for 1 h at 45°C. Following reduction/alkylation samples were digested overnight with Trypsin (1/50 wt/wt Solu-trypsin, Sigma) at 37°C with shaking at 1,000 rpm.

Samples for ZIC-HILIC glycoproteomic enrichment were acidified to a final concentration of 0.1% TFA to precipitate the SDC with the SDC removed by centrifugation at $10,000 \times g$ for 10 min (77, 78). Clarified peptide solutions were loaded on conditioned 50 mg tC18 Sep-Pak columns (Waters Corporation, Milford, MA, USA) and peptides were desalted according to the manufacturer's instructions. tC18 Sep-Pak columns were conditioned with 10 bed volumes of Buffer B (0.1% formic acid, 80% acetonitrile), then equilibrated with 10 bed volumes of Buffer A* (0.1% TFA, 2% acetonitrile) before use. Samples were loaded onto equilibrated columns then columns washed with at least 10 bed volumes of Buffer A* before bound peptides were eluted with Buffer B. Eluted peptides were dried by vacuum centrifugation at room temperature and stored at -20°C . Quantitative proteomic samples were mixed with 1.25 volumes of isopropanol to prevent the precipitation of SDC (79) and samples were acidified with TFA to a final concentration of 1% TFA. In-house-made SDB-RPS StageTips were prepared according to previously described protocols (76, 80, 81) with five frits of SDB-RPS excised using a blunt 16-gauge Hamilton needle and loaded into 200 μ L tips. SDB-RPS StageTips were placed in a Spin96 tip holder (80) to enable batch-based spinning of samples and tips conditioned with 100% acetonitrile; followed by 30% methanol, 1% TFA followed by 90% isopropanol, 1% TFA with each wash spun through the column at $1,000 \times g$ for 3 min. Acidified isopropanol/peptide mixtures were loaded onto the SDB-RPS columns and spun through tips before being washed with 90% isopropanol, 1% TFA followed by 1% TFA in Milli-Q water. Peptide samples were eluted with 80% acetonitrile, 5% ammonium hydroxide, and dried by vacuum centrifugation at room temperature and stored at -20°C .

ZIC-HILIC enrichment of glycopeptides

ZIC-HILIC enrichment was performed as previously described with minor modifications (82). ZIC-HILIC Stage-tips (81) were created by packing 0.5 cm of 10 μm ZIC-HILIC resin (Millipore/Sigma) into 200 μL tips containing a frit of C8 Empore (Sigma) material. Prior to use, the columns were washed with Milli-Q water, followed by 95% acetonitrile, and then equilibrated with 80% acetonitrile, 1% TFA. Digested proteome samples were resuspended in 80% acetonitrile, 1% TFA. Samples were adjusted to a concentration of 5 $\mu\text{g}/\mu\text{L}$ then loaded onto equilibrated ZIC-HILIC columns. ZIC-HILIC columns were washed with 20 bed volumes of 80% acetonitrile, 1% TFA to remove non-glycosylated peptides and bound peptides eluted with 10 bed volumes of Milli-Q water. Eluted peptides were dried by vacuum centrifugation at room temperature and stored at -20°C .

LC-MS analysis of ZIC-HILIC based glycopeptide enrichments

ZIC-HILIC enriched samples were re-suspended in Buffer A* and separated using a two-column chromatography set-up composed of a PepMap100 C18 20 mm \times 75 μm trap and a PepMap C18 500 mm \times 75 μm analytical column (Thermo Fisher Scientific). Samples were concentrated onto the trap column at 5 $\mu\text{L}/\text{min}$ for 5 min with Buffer A (0.1% formic acid, 2% DMSO) and then infused into an Orbitrap Fusion Lumos Tribrid Mass Spectrometer (Thermo Fisher Scientific) at 300 nL/min via the analytical column using a Dionex Ultimate 3000 UPLC (Thermo Fisher Scientific). 185-min analytical runs were undertaken by altering the buffer composition from 2% Buffer B (0.1% formic acid, 77.9% acetonitrile, 2% DMSO) to 28% B over 155 min, then from 22% B to 45% B over 10 min, then from 45% B to 80% B over 2 min. The composition was held at 80% B for 3 min, and then dropped to 2% B over 5 min before being held at 2% B for another 10 min. The Lumos Mass Spectrometer was operated in a data-dependent mode automatically switching between the acquisition of a single Orbitrap MS scan (maximal injection time of 50 ms, an Automated Gain Control (AGC) set to a maximum of 1×10^6 ions and a resolution of 120k) every 3 s and Orbitrap MS/MS HCD scans of precursors (NCE 28% with 5% Stepping, maximal injection time of 60 ms, an AGC set to a maximum of 1×10^5 ions and a resolution of 15k). Each biological replicate was analyzed two times using two different MS1 mass ranges of 350–1,800 m/z and 500–2,000 m/z to improve glycoproteome coverage. MS/MS Scans containing HexNAc-associated oxonium ions (204.0867; 138.0545 or 366.1396 m/z) triggered three additional product-dependent MS/MS scans (83) of potential glycopeptides; an Orbitrap EThcD scan (NCE 15%, maximal injection time of 250 ms, AGC set to a maximum of 2×10^5 ions with a resolution of 30k and using the extended mass range setting to improve the detection of high mass glycopeptide fragment ions (84); an ion trap CID scan (NCE 35%, maximal injection time of 40 ms, an AGC set to a maximum of 5×10^4 ions) and a stepped collision energy HCD scan (using NCE 35% with 8% Stepping, maximal injection time of 150 ms, an AGC set to a maximum of 2×10^5 ions and a resolution of 30k).

LC-MS analysis of FAIMS-based glycopeptide enrichments

SDB-RPS cleaned-up samples were re-suspended in Buffer A* and separated using a two-column chromatography set-up composed of a PepMap100 C18 20 mm \times 75 μm trap and a PepMap C18 500 mm \times 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). 145-min gradients 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 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 CVs -25 , -45 and -65 as previously described (43). For each FAIMS CV a single Orbitrap MS scan (450–2,000 m/z , maximal injection time of 50 ms, an AGC of maximum of 4×10^5 ions and a resolution of 60k)

was acquired every 1.2 s, followed by Orbitrap MS/MS HCD scans of precursors (NCE 30%, maximal injection time of 80 ms, an AGC set to a maximum of 1.25×10^5 ions and a resolution of 30k). HCD scans containing HexNAc-associated oxonium ions (204.0867, 138.0545, and 366.1396 m/z) triggered three additional product-dependent MS/MS scans (83) of potential glycopeptides; an Orbitrap EThcD scan (NCE 15%, maximal injection time of 250 ms, AGC set to a maximum of 2×10^5 ions with a resolution of 30k using the extended mass range setting to improve the detection of high mass glycopeptide fragment ions (84); an ion trap CID scan (NCE 35%, maximal injection time of 40 ms, an AGC set to a maximum of 5×10^4 ions) and a stepped collision energy HCD scan (using NCE 35% with 5% Stepping, maximal injection time of 250 ms, an AGC set to a maximum of 2×10^5 ions and a resolution of 30k).

LC-MS analysis of DIA samples

SDB-RPS cleaned-up samples were resuspended in Buffer A* and separated using a two-column chromatography set-up composed of a PepMap100 C18 20 mm \times 75 μ m trap and a PepMap C18 500 mm \times 75 μ m analytical column (Thermo Fisher Scientific). Samples were concentrated onto the trap column at 5 μ L/min for 5 min with Buffer A and then infused into an Orbitrap Eclipse (Thermo Fisher Scientific) at 300 nL/min via the analytical column using a Dionex Ultimate 3000 UPLC (Thermo Fisher Scientific). 95-min analytical runs were undertaken by altering the buffer composition from 3% Buffer B to 25% B over 82 min, then from 25% B to 40% B over 4 min, and then from 40% B to 80% B over 1 min. The composition was held at 80% B for 3 min, and then dropped to 3% B over 1 min before being held at 3% B for another 4 min. Data were collected in a data-independent manner with a single MS1 event (120k resolution, AGC 1×10^6 , 350–1,400 m/z) and 50 MS2 scans (NCE 30%, 30k resolution, AGC 1×10^6 , 200–2,000 m/z and a maximal injection time of 55 ms) of a width of 13.7 m/z collected over the mass range of 360 to 1,033.5 m/z .

Glycopeptide identifications

ZIC-HILIC and FAIMS enrichments were analyzed using MSfragger-Glyco within MSFragger (versions 17.1) (85–87). Samples were searched with a Tryptic specificity allowing a maximum of two missed cleavage events and Carbamidomethyl set as a fixed modification of Cysteine while oxidation of Methionine allowed as a variable modification. The *Burkholderia* glycans HexHexNAc₂ (elemental composition: C₂₂O₁₅H₃₆N₂, mass: 568.2115 Da) and Suc-HexHexNAc₂ (elemental composition: C₂₆O₁₈H₄₀N₂, mass: 668.2276 Da) were included as variable modifications at Serine in line with the strong preference for PgL glycosylation at Serine residues (38). The glycan fragments of ions were defined as 204.0866, 186.0760, 168.0655, 366.1395, 144.0656, 138.0550, 466.1555, and 407.1594. For HCD searches, y, b, Y, y~/b~ (y and b ion with the addition of HexNAc) were allowed, while for EThcD searches, the ions c, z, y and y~ ions were allowed. A maximum mass precursor tolerance of 20 ppm was allowed at both the MS1 and MS2 levels. Samples were searched against the *B. cenocepacia* reference proteome J2315 (UniProt accession: [UP000001035](https://doi.org/10.1002/UniProt/J2315), 6,993 proteins, downloaded 24 July 2020) (88) to enable the mapping of proteins to consensus gene accessions. HCD and EThcD searches were performed separately on each data set with the resulting “psm.tsv” combined using R with the best identification for each scan event retained and only glycopeptides with an MSfragger Hyperscore >20 used for data visualization undertaken using ggplot2 (89) using R and RStudio. To ensure high data quality, assigned glycopeptides were manually assessed and the HCD and EThcD spectra assigned to each unique glycopeptide annotated with the Interactive Peptide Spectral Annotator (90) (<http://www.interactivepeptidespectralannotator.com/PeptideAnnotator.html>) with annotated glycopeptides provided within Supplementary Data 1 deposited within Figshare (<https://doi.org/10.6084/m9.figshare.25492774.v1>). To assess the uniqueness of identified glycoproteins within this study we curated a list of all previously known *B. cenocepacia* glycoproteins (38, 39, 41, 43). Unique glycopeptides identified using the

glycopeptide identification tool Byonic [Protein Metrics Inc (91)]. yet not previously manually confirmed from studies (39, 43) were annotated to ensure correctness resulting in a total of 129 unique glycoproteins across all studies (38, 39, 41, 43) (Table S2) with unique glycopeptides from these previous Byonic searches provided within Supplementary Data 2 deposited within Figshare (<https://doi.org/10.6084/m9.figshare.25492774.v1>).

Quantitative proteomics analysis using Spectronaut

DIA data sets were searched using DIA-library free analysis within Spectronaut (version 15.5, 17.6, or 18.5). Data files were searched against the *B. cenocepacia* K56-2 proteome (UniProt accession: [UP000011196](#), 7467 proteins, downloaded 13 October 2020) (92) and *B. cenocepacia* strain J2315 (UniProt accession: [UP000001035](#)) (88), effectively merging these proteomes enabling the matching of proteins to both J2315 and K56-2 accessions. Oxidation of Methionine was allowed as a variable modification, Carbamidomethyl set as a fixed modification of Cysteine and the protease specificity set to Trypsin. Protein quantitation was undertaken using MaxLFQ (93) based analysis. The precursor PEP was altered to 0.01 from the default 0.2 to improve quantitative accuracy with all single peptide protein matches excluded. Statistical analysis was undertaken using Perseus (94) with missing values imputed based on the total observed protein intensities with a range of 0.3 σ and a downshift of 2.5 σ . Biological replicates were grouped together and either ANOVA analysis was used to compare changes across all groups or student *t* tests were used to compare individual groups with a minimum fold change of \pm considered for further analysis. Multiple hypothesis correction was undertaken using a permutation-based FDR approach allowing an FDR of 5%. Enrichment analysis using Fisher exact tests was undertaken in Perseus using gene ontology (GO) terms obtained from UniProt (*B. cenocepacia* strain J2315 proteome: [UP000001035](#)). Data visualization was undertaken using ggplot2 (89) within RStudio.

DDA confirmation of independent Δ fliF, Δ motB, and Δ BCAM0505 *B. cenocepacia* mutants

In-StageTip prepared proteome digests as above were re-suspended in Buffer A* and separated using a two-column chromatography setup composed of a PepMap100 C18 20 mm \times 75 μ m trap and a PepMap C18 500 mm \times 75 μ m analytical column (Thermo Fisher Scientific). Samples were concentrated onto the trap column at 5 μ L/min for 5 min with Buffer A and then infused into either an Orbitrap Q-Exactive plus, an Orbitrap Eclipse or Orbitrap Fusion Lumos equipped with a FAIMS Pro interface (Thermo Fisher Scientific) at 300 nL/min via the analytical column using Dionex Ultimate 3000 UPLCs (Thermo Fisher Scientific). 125-min analytical runs were undertaken by altering the buffer composition from 2% Buffer B to 22% B over 95 min, then from 22% B to 40% B over 10 min, and then from 40% B to 80% B over 5 min. The composition was held at 80% B for 5 min, and then dropped to 2% B over 2 min before being held at 2% B for another 8 min. The Mass Spectrometers were operated in a data-dependent mode automatically switching between the acquisition of a single Orbitrap MS scan and up to 15 Orbitrap MS/MS HCD scans of precursors on the Orbitrap Q-Exactive plus (Stepped NCE of 28%, 32%, and 38%, a maximal injection time of 100 ms, an AGC set to a maximum of 2×10^5 ions and a resolution of 35k) or up to 3 s of MS/MS HCD scans of precursors on the Orbitrap Lumos and Eclipse (Stepped NCE of 25%, 30%, and 40%, a maximal injection time of 55 ms, an AGC set to a maximum of 2×10^5 ions and a resolution of 30k).

DDA proteomic data analysis

Identification and LFQ analysis to confirm *B. cenocepacia* mutants was accomplished using MaxQuant (v1.6.17.0) (95) using the *B. cenocepacia* K56-2 proteome (UniProt accession: [UP000011196](#)) (92) and *B. cenocepacia* strain J2315 (UniProt accession: [UP000001035](#)) (88) allowing for oxidation on Methionine, Carbamidomethyl set as a fixed modification on Cysteine and the protease specificity set to Trypsin. Prior to MaxQuant

analysis data sets acquired on the Fusion Lumos were separated into individual FAIMS fractions using the FAIMS MzXML Generator (96). The LFQ and “Match Between Run” options were enabled to allow comparisons between samples. The resulting data files were processed using Perseus (v1.4.0.6) (94) using Student's *t* tests to compare proteomic changes between independent mutants with data visualization undertaken using ggplot2 (89) within RStudio.

Western blotting

Cells grown overnight with or without L-rhamnose (0.1%; stationary phase, OD_{600nm} 1.5) were resuspended in 1× SDS-PAGE sample buffer, boiled for 10 min and then run on a NuPAGE 4–12% Bis-Tris gradient polyacrylamide gel in MES buffer (Thermo Fisher Scientific). Proteins were transferred to a nitrocellulose membrane using the iBlot 2 Western transfer system (Thermo Fisher Scientific) and blocked for 1 h in 5% skim milk in TBS-T (Tris-buffered saline with 0.1% Tween-20). Nitrocellulose membranes were probed with a monoclonal mouse αHistidine tag primary antibody (1:1,000; AD1.1.10, Bio-Rad) overnight at 4°C, washed with TBS-T, then incubated with horseradish peroxidase (HRP) conjugated mouse antibody (1:5,000; NEF82200-1EA, PerkinElmer) for 1 h at room temperature. Proteins were detected using the Clarity Western ECL Substrate (Bio-Rad), and images were obtained using an Amersham imager 600 or 800 (GE Healthcare Life Sciences). Membranes were Stripped with Restore Plus Stripping Buffer for 10 min before being re-blocked in 5% skim milk in TBS-T and probed with α-*E. coli* RNA polymerase primary antibody (1:5,000; 4RA2, Neoclone), followed by HRP conjugated mouse and imaged as above. All antibodies were diluted in TBS-T with 1% bovine serum albumin (BSA; Sigma-Aldrich). Raw uncropped and protein marker-associated Western blotting images are provided in Fig. S24.

Swimming motility assays

Motility assays were performed with 0.3% L-agar as previously described (26) with minor modifications. Briefly, 20% sterile L-rhamnose or an equivalent volume of sterile water was added to molten 0.3% L-agar media for a final concentration of 0.1% rhamnose, 20 mL aliquoted into 9 mm Petri dishes, and plates dried briefly once set. The dried plates were inoculated with 1 μL of stationary phase cultures standardized to an OD_{600nm} of 1.0 (paired inoculation of plates with and without rhamnose) and plates were incubated inside sealed bags at 37°C for 24 h. Following incubation, the diameter of turbidity was measured, and plates were imaged.

RESULTS

Analysis of the *B. cenocepacia* O-glycoproteome across different growth media reveals previously unrecognized glycoproteins

While our previous work using multiple proteases has revealed 98 proteins are glycosylated undergrowth in LB within *B. cenocepacia* K56-2 (38), the true size of the *B. cenocepacia* glycoproteome remains unknown. To improve our understanding of this glycoproteome, we reasoned that growth under varying culturing conditions may enable the identification of novel glycoproteins. To assess this we examined the glycoproteome of *B. cenocepacia* K56-2 (97) under growth in two common culturing media (LB and TSB) and the CF sputum-like medium ASM. As our recent work has shown improvements in the coverage of bacterial glycoproteomes using complementary enrichment approaches (43), we utilized both ZIC-HILIC (50) and FAIMS (43) based glycopeptide enrichments. Glycopeptides were identified against the J2315 proteome with our previous work highlighting the high sequence identity within the surrounding amino acids of glycosylation sites within *B. cenocepacia* (38); however, this approach is insensitive to the detection of glycopeptides unique to K56-2. Analysis of the glycopeptides observed across these growth conditions demonstrated most glycopeptides were shared across conditions yet revealed unique glycopeptides observable in each medium [Fig. 1A; Table

S1; Supplementary Data 1 (<https://doi.org/10.6084/m9.figshare.25492774.v1>]). A total of 74 distinct glycopeptides were identified corresponding to 54 unique glycoproteins of which 35 and 32 glycoproteins were observed in ZIC-HILIC and FAIMS enrichments, respectively (Fig. 1B). Consistent with previous analysis ZIC-HILIC and FAIMS glycopeptide enrichments provide access to complementary glycopeptide subsets (43) (Fig. S2A through D). To assess the uniqueness of the glycoproteins identified across multiple growth conditions, we manually curated a list of all known *B. cenocepacia* glycoproteins (38, 39, 41, 43) [129 glycoproteins in total, Table S2 and Supplementary Data 2 (<https://doi.org/10.6084/m9.figshare.25492774.v1>)] revealing 12 of the 54 glycoproteins identified within this study correspond to previously unrecognized *B. cenocepacia* glycoproteins (Fig. 1C; Fig. S3). These novel glycoproteins, such as the Lipase chaperone (BCAM0950, Fig. 1D), raise the number of known glycoproteins to 141 with 50% of glycoproteins defined as uncharacterised or putative proteins and >40% lacking any GO terms (Fig. S4; Table S2).

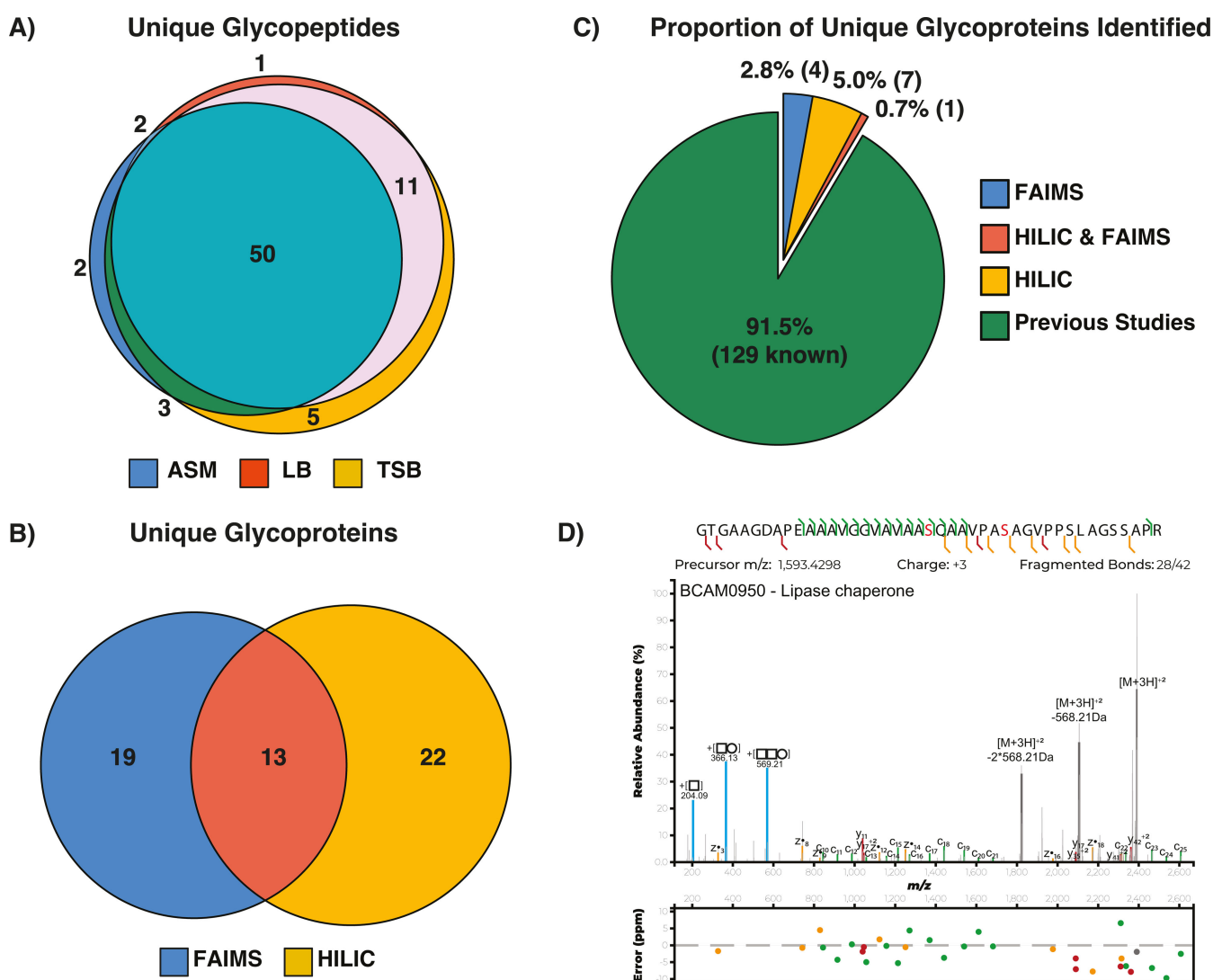


FIG 1 Expanding the glycoproteome of *B. cenocepacia* K56-2. (A) Glycopeptide analysis reveals the majority of *B. cenocepacia* K56-2 glycopeptides appear detectable across multiple culturing conditions. (B) Venn diagram of the overlap within the identified glycoproteins across different growth conditions supports FAIMS and HILIC enrichments and provide access to complementary regions of the glycoproteome. (C) Pie chart of identified glycoproteins reveals a total of 141 glycoproteins have now been identified across *B. cenocepacia* K56-2 and J2315. (D) EthcD MS2 of the novel glycoprotein BCAM0950 reveals two glycosylation sites within the glycopeptide ³⁹GTGAAGDAPEIAAAMGGVAVAASQAAVPASAGVPPSLAGSSAPR⁸¹ with glycan-associated ions (204.09, 366.13, and 569.21 *m/z*, highlighted in blue) denoted according to the Symbol Nomenclature for Glycans.

The glycoproteome is largely stable across growth phase and culturing conditions

While the identification of additional novel glycoproteins expands the known *B. cenocepacia* glycoproteome, the high degree of overlap between the glycoproteome observed under differing media and growth phases supports that a significant proportion of glycoproteins may be expressed under a range of conditions. To explore the stability of glycoprotein abundance in a robust manner, we utilized a DIA-based approach to track the known glycoproteome of *B. cenocepacia* at log and stationary phases across LB, TSB, and ASM culturing media. DIA enabled the identification of 3,923 proteins across all replicates, revealing clear clustering based on growth phases as well as culturing conditions (Fig. S5A and B; Table S3). Within the observable proteome, 106 of the 141 known glycoproteins were identified (Fig. 2A), with >90% of these glycoproteins observed in each replicate (Fig. S6A and B; Table S3). Comparing growth phases between each culturing media reveals less than 30% of the known glycoproteome appears altered in response to bacterial growth phases, while the majority of the glycoproteome appears stable across growth (Fig. 2B through E). These dynamics mirror the total proteome (Fig. S7) and support that most glycoproteins are constitutively expressed being observable across all conditions examined. Consistent with only modest alterations in glycoprotein levels across conditions, examination of the abundance of PgIL (BCAL0960) and the

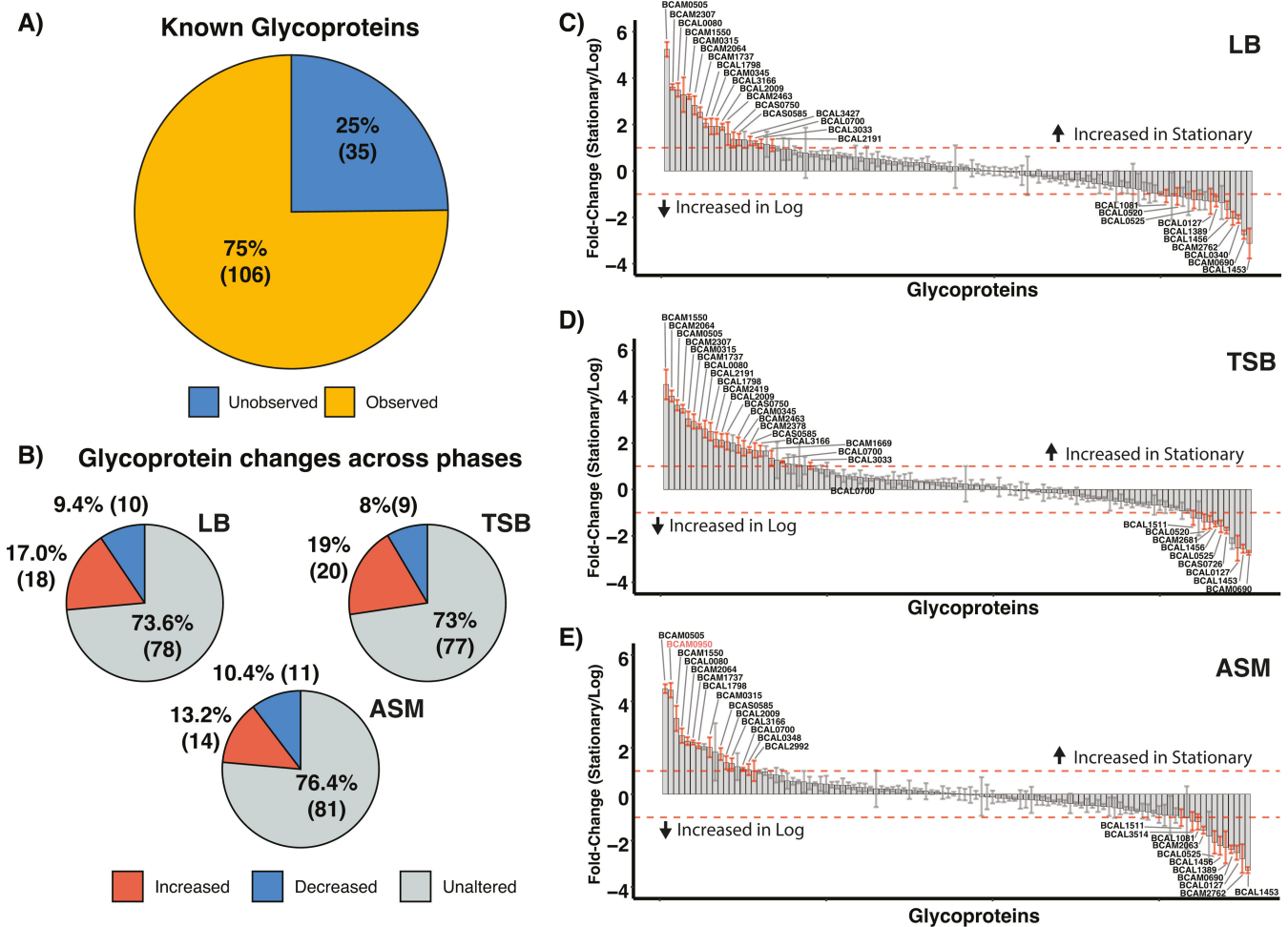


FIG 2 DIA analysis of the *B. cenocepacia* K56-2 glycoproteome. (A) DIA analysis allows the assessment of >100 known glycoproteins across different media and growth phases. (B) Quantitative analysis of identified glycoproteins across culture media reveals most glycoproteins are unaltered across growth phases based on ± 1 fold (\log_2) and a $-\log_{10}(P \text{ value}) > 2$. (C–E) Rank order plots of glycoprotein levels observed across the log and stationary growth phases demonstrate only a subset of glycoproteins, denoted with red error bars, appear to be altered in abundance across growth phases.

observed members of the Ogc (BCAL3115-BCAL3117) (40) reveals constant protein levels between conditions and only modest differences ($<1 \log_2$) in the abundance of these proteins between growth phases (Fig. S8) supporting that glycosylation machinery abundance remains stable across growth phases and culturing conditions. While these observations support that the abundance of glycosylation substrates is mostly constant across growth phases, there were some notable exceptions such as the Lipase chaperone BCAM0950 that appeared elevated across stationary growth in ASM media (highlighted in red, Fig. 2C through E). Thus, these findings support that the majority of the glycoproteome and the associated glycosylation machinery do not appear to be impacted by changes in growth conditions or growth phases.

The loss of glycosylation within *B. cenocepacia* K56-2 impacts only a subset of glycoproteins

Previous work has demonstrated that the loss of glycosylation results in reduced protein levels of two known glycoproteins within *B. cenocepacia* K56-2 (BCAL1086 and BCAL2974) (55). As this initial analysis only considered 21 glycoproteins known at the time, we sought to apply our DIA approach to improve the characterization of changes in the glycoproteome in the absence of glycosylation across the now known 141 glycoproteins. Examining the proteome of *B. cenocepacia* K56-2 wild type (WT), $\Delta pglL$ and complemented $\Delta pglL$ [$\Delta pglL\Delta amrAB::native-promotor-pglL-his$ (55)] across differing culturing media (LB, TSB, and ASM) at stationary phase resulted in the identification of 3,783 proteins including 110 glycoproteins (Table S4; Fig. S9 and 10). Consistent with our previous work (55), we observed dramatic proteome changes between glycosylation null and glycosylation competent strains which are highly correlative across all three media conditions examined (Fig. S11). Analysis of the observable glycoproteome reveals only a subset of glycoproteins are impacted by the loss of glycosylation across all culturing media (Fig. 3A) in addition to the expected loss of AmrAB (BCAL1674-5) within the complement due to the integration of *pglL-his* into the *amrAB* locus (68, 98) (Fig. 3A). Across the glycoproteins impacted by loss of glycosylation multiple proteins appeared to undergo reductions in protein abundance, yet surprisingly two proteins, BCAL1917 and BCAL1093, also appeared to increase in abundance (Fig. 3A). Analysis of the peptides utilized for protein quantitation for BCAL1917 and BCAL1093 reveals these increases were driven by changes in peptide abundances of unmodified forms of known glycopeptide supporting this increase in abundance is an artifact resulting from the observation of peptides within $\Delta pglL$ that are glycosylated within WT and the *pglL*-complement strains (Fig. S12). Focusing on glycoproteins reduced in the absence of glycosylation across multiple growth media revealed the loss of five glycoproteins BCAL1086, BCAL2974, FliF (BCAL0525), BCAM0505, and MotB (BCAL0127) (Fig. 3B). Examination of the peptides quantified for these glycoproteins supports the reduction of these proteins across all growth conditions and the restoration of these proteins upon complementation (Fig. 3C). Thus, these findings support that the abundance of most glycoproteins is unaffected by the loss of glycosylation, with only five glycoproteins undergoing dramatic reductions in abundance in the absence of glycosylation.

Loss of MotB and FliF mirrors the proteomic impacts observed in *B. cenocepacia* K56-2 $\Delta pglL$

As our previous studies have demonstrated that the disruption of *BCAL1086* or *BCAL2974* does not drive the proteome changes observed in $\Delta pglL$ (55), we focused on the glycoproteins MotB, FliF, and BCAM0505. To assess if the loss of *motB*, *fliF*, and BCAM0505 contributes to the proteomic impact seen in the absence of glycosylation, we generated $\Delta motB$, $\Delta fliF$, and $\Delta BCAM0505$ strains, and using DDA proteomic analysis confirmed the loss of these three proteins within mutants (Fig. S13 to S15; Table S5). Additional DIA analysis of $\Delta motB$, $\Delta fliF$, and $\Delta BCAM0505$ compared to the WT and $\Delta pglL$ strains at stationary phase in LB allowed the identification of 3,934 proteins with individual biological groups demonstrating clustering by Pearson correlation and PCA analysis

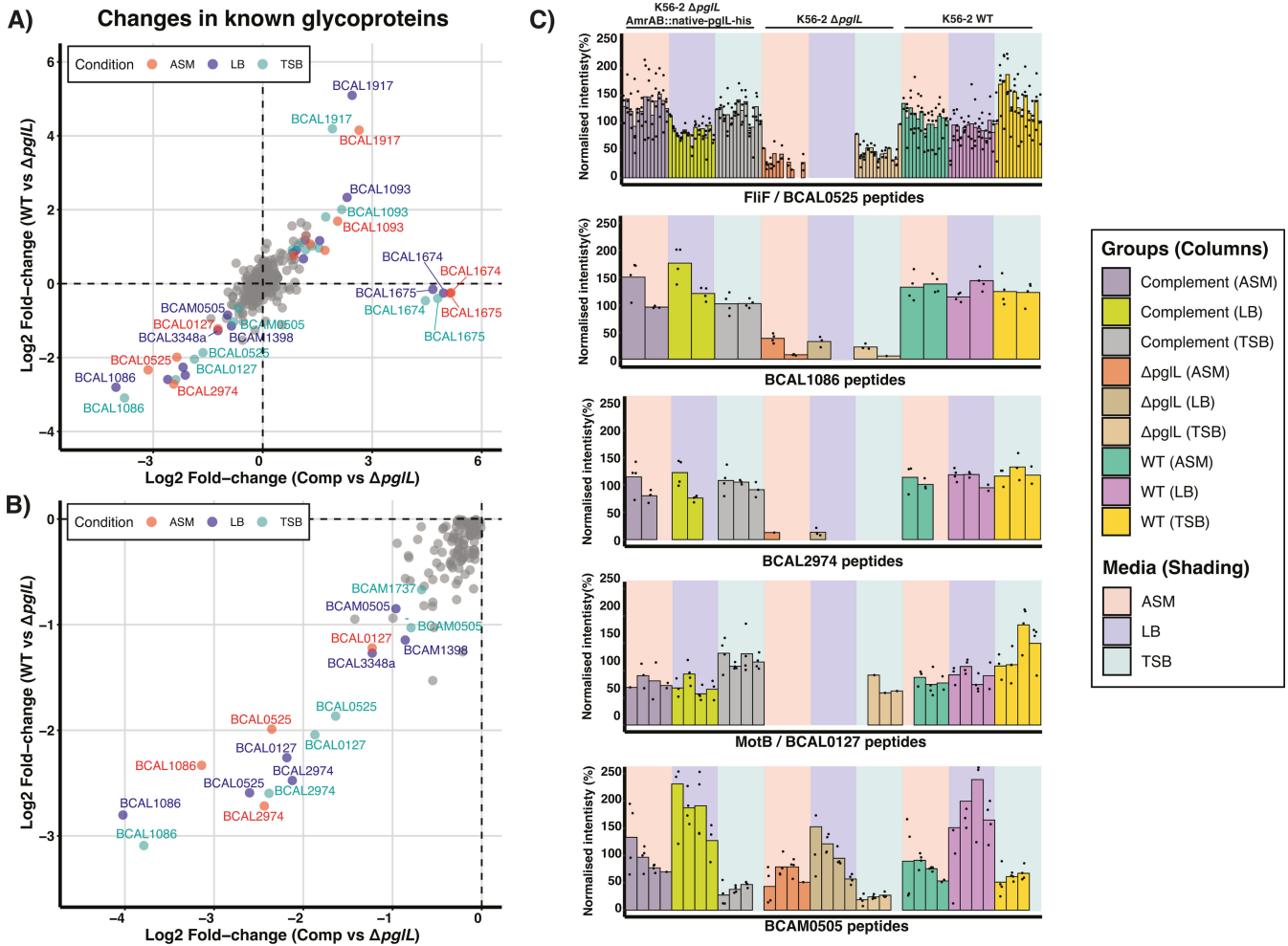


FIG 3 Impact of the loss of glycosylation on the *B. cenocepacia* K56-2 glycoproteome. (A and B) 2D scatter plots comparing glycoprotein levels observed within WT versus Δ pglL and Δ pglL *amrAB::native-pglL-his* versus Δ pglL. Glycoproteins observed to undergo alterations are color-coded according to the growth condition if the observed alteration has a fold change greater than ± 1 fold (\log_2) and a $-\log_{10}(P \text{ value}) > 2$. (C) Peptide-centric analysis of glycoproteins FliF, BCAL1086, BCAL2974, MotB, and BCAM0505 observed to undergo alterations in response to the loss of glycosylation reveals a marked decrease in peptide intensities within these proteins which is restored by the re-introduction of glycosylation within the complement Δ pglL *amrAB::native-pglL-his*.

(Fig. S16 and 17; Table S6). Across strains, we observed that the loss of *motB*, *fliF*, and BCAM0505 leads to observable proteome alterations compared to WT [310, 117, and 46 protein alterations with a fold change of ± 1 fold (\log_2) and a $-\log_{10}(P \text{ value}) > 2.0$ respectively] with many of these alterations observed within proteins also impacted by the loss of *pglL* (Fig. 4A; Fig. S18). Enrichment analysis of the proteins altered within Δ *motB* and Δ *fliF* highlights the high degree of overlap within proteome changes observed within Δ pglL (Fig. 4B and C; Table S7). Additionally, enrichment analysis of GO terms associated with protein alterations within Δ *fliF* supports changes in chemotaxis and flagella-associated motility proteins (Fig. 4D; Table S7) consistent with alterations previously observed within Δ pglL (55). Interestingly, enrichment analysis of alterations observed within Δ *motB* did reveal functional enrichments, yet these were unique to those observed with Δ pglL (Fig. S19A and B; Table S7). In contrast, no enrichment in GO terms or proteins observed altered within Δ pglL were observed with Δ BCAM0505. Examination of the individual proteins associated with the GO terms Chemotaxis [GO:0006935] and Bacterial-type flagellum-dependent cell motility [GO:0071973] supports similar changes between Δ pglL and Δ *fliF* compared to WT yet that the overlap in altered proteins is only partial (Fig. S19C and D). Combined, these

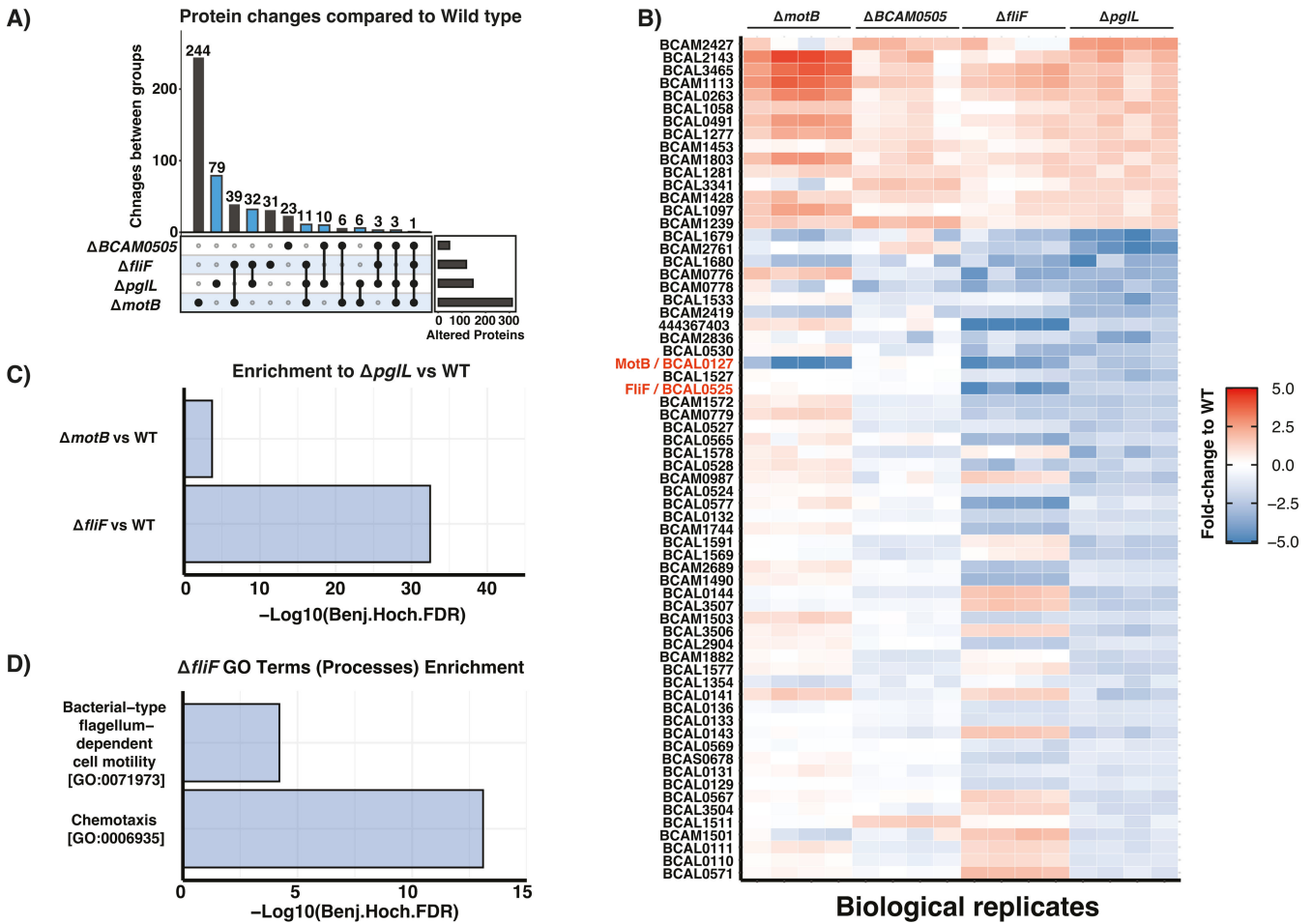


FIG 4 Proteome analysis of $\Delta motB$, $\Delta fliF$, and $\Delta BCAM0505$ compared to WT and $\Delta pglL$. Proteome analysis of the strains lacking the glycoproteins MotB (BCAL0127), FliF (BCAL0525), and BCAM0505 compared to WT and $\Delta pglL$ supports the loss of MotB and/or FliF mirrors proteomic impacts observed within $\Delta pglL$. (A) Upset analysis of the proteins observed altered between strains based on ± 1 fold (\log_2) and a $-\log_{10}(P \text{ value}) > 2.0$ reveals the overlap in protein alterations observed within $\Delta pglL$, denoted in blue, within mutants lacking MotB, FliF, and BCAM0505. (B) Heatmap of fold changes in proteins observed altered within $\Delta pglL$ versus WT reveals similarities in the directionality of protein alterations observed within $\Delta fliF$ and $\Delta motB$ strains. (C) Enrichment analysis demonstrates a high degree of overlap within the proteins altered within $\Delta pglL$ versus WT compared to $\Delta motB$ versus WT and $\Delta fliF$ versus WT. (D) Enrichment analysis of gene ontology (biological process) terms enriched within protein alterations observed within $\Delta fliF$.

findings support that the loss of *fliF* and *motB* results in similar proteomic impacts as those observed within $\Delta pglL$. To understand the importance of glycosylation for the function of *fliF* and *motB* these proteins were selected for further analysis using site-directed mutagenesis.

Glycosylation at S₃₅₈ within FliF and S₃₃₁ within MotB are dispensable for motility

The observed proteomic changes within $\Delta motB$ and $\Delta fliF$, as well as the loss of these proteins in the absence of PglL suggests glycosylation may influence the stability and/or functionality of these proteins. In light of our previous work demonstrating reductions in motility associated with loss of glycosylation (41, 55) as well as FliF/MotB homologs being known to be associated with flagella-based motility (99), we confirmed the requirement of FliF and MotB for motility within *B. cenocepacia* (Fig. 5A and B). Plasmid-based complementation using the inducible vector pScRhaB2 (71) allowed the partial restoration of motility within $\Delta fliF$ pScRhaB2-*fliF* upon induction, while motility was restored within $\Delta motB$ pScRhaB2-*motB* without induction (Fig. 5A and B). Surprisingly,

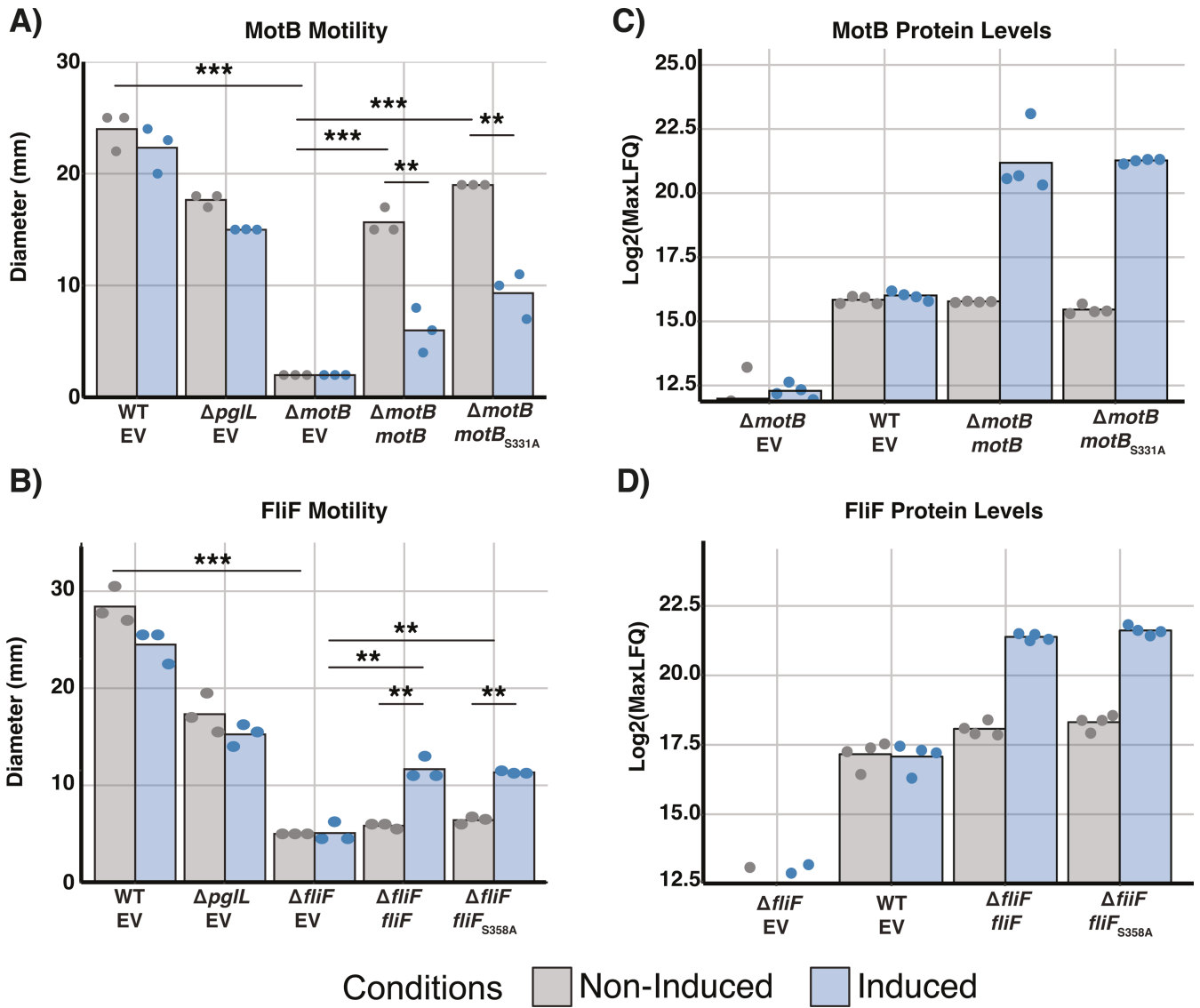


FIG 5 Complementation of Δ fliF and Δ motB partially restores motility and protein levels. (A and B) Motility assays of complemented Δ fliF and Δ motB demonstrate the restoration of motility by the expression of wild-type fliF or motB regardless of the absence or presence of the known glycosylation sites S₃₅₈ and S₃₃₁. (C and D) Proteomic analysis confirms the restoration of MotB and FliF protein levels within complemented strains. Induced samples are denoted by blue boxes, while non-induced samples are denoted by gray boxes. The plasmid contained within each strain is shown below the strain name. EV—empty pSCrhaB2, motB—pSCrhaB2-motB, motB_{S331A}—pSCrhaB2-motB_{S331A}, fliF—pSCrhaB2-fliF, and fliF_{S358A}—pSCrhaB2-fliF_{S358A} with the lower boundaries of the MaxLFQ shown defined by the imputed values observed within Δ fliF and Δ motB EV, respectively.

within Δ motB pSCrhaB2-motB induction reduced motility (Fig. 5B) with reductions in motility also observed within the parental strain *B. cenocepacia* K56-2 in response to MotB induction (Fig. S20), suggesting the overexpression of MotB is detrimental to motility. To assess motB/fliF complementation, proteomic analysis was undertaken (Tables S8 and S9; Fig. S21 and 22) confirming the restoration of MotB with pSCrhaB2-motB in Δ motB and FliF with pSCrhaB2-fliF in Δ fliF (Fig. 5C and D). Interestingly despite the observation of near wild-type levels of FliF even without induction by proteomic analysis, induction was required to partially restore motility within Δ fliF (Fig. 5B and D). The proteomic analysis supports that while FliF is restored without induction additional motility-associated proteins including MotA (BCAL0126) and MotB within Δ fliF were only restored, albeit partially, with induction (Fig. S21). To assess the role of glycosylation sites within FliF and MotB, Alanine substitutions were created based on known glycosylation

sites (FliF_{S358A} and MotB_{S331A}, Table S2), and the impact on motility was examined. As above, complementation of $\Delta motB$ with pSCrhaB2-*motB*_{S331A} restored motility even without induction (Fig. 5A and B), while $\Delta fliF$ with pSCrhaB2-*fliF*_{S358A} only restored motility with induction (Fig. 5C and D). Combined these results support while FliF and MotB are required for motility the known O-linked glycosylation sites of these proteins appear dispensable for motility.

FliF and MotB contribute to motility independent of their glycosylation status

Finally, to confirm the loss of glycosylation within FliF_{S358A} and MotB_{S331A} as well as investigate the requirement of these proteins for motility in the absence of glycosylation, we created $\Delta pgII\Delta fliF$ and $\Delta pgII\Delta motB$. Consistent with the requirement of FliF and MotB for motility within glycosylation null backgrounds we observe marked reductions in motility within both $\Delta pgII\Delta fliF$ and $\Delta pgII\Delta motB$ compared to wild-type and $\Delta pgII$ strains with a partial restoration in motility observed upon complementation with pSCrhaB2-*fliF*/pSCrhaB2-*motB* respectively (Fig. 6A and B). Consistent with complementation Western analysis reveals comparable levels of FliF/MotB upon induction within $\Delta fliF/\Delta motB$ compared to $\Delta pgII\Delta fliF/\Delta pgII\Delta motB$, respectively (Fig. 6C and D). Analysis of FliF_{S358A} and MotB_{S331A} expressed within $\Delta fliF/\Delta motB$ strains demonstrates that FliF_{S358A}/MotB_{S331A} possess comparable sizes to FliF/MotB expressed within

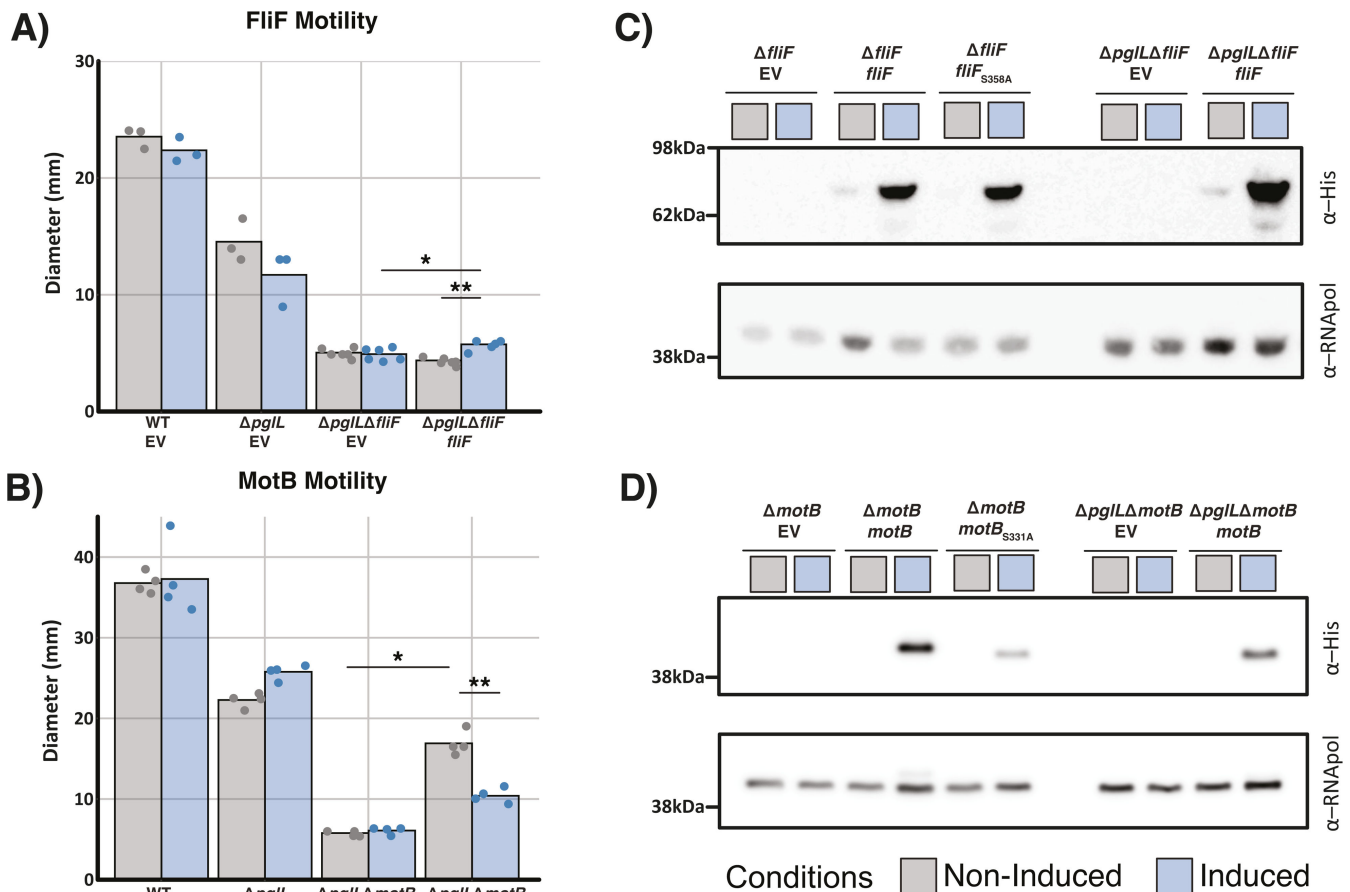


FIG 6 FliF and MotB contribute to motility in the absence of glycosylation. (A and B) Motility assays of complemented $\Delta pgII\Delta fliF$ and $\Delta pgII\Delta motB$ compared to WT and $\Delta pgII$ demonstrate the restoration of motility by the expression of wild-type *fliF* or *motB* compared to EV. (C and D) Western analysis complemented strains demonstrate mobility shifts with FliF or MotB Alanine substitutions as well as FliF or MotB expressed within $\Delta pgII$ backgrounds. Induced samples are denoted by blue boxes, while non-induced samples are denoted by gray boxes. The plasmid contained within each strain is shown below the strain name. EV—empty pSCrhaB2, *motB*—pSCrhaB2-*motB*, *motB*_{S331A}—pSCrhaB2-*motB*_{S331A}, *fliF*—pSCrhaB2-*fliF*, and *fliF*_{S358A}—pSCrhaB2-*fliF*_{S358A}.

glycosylation null backgrounds supporting the abolishment of glycosylation within these point-mutants (Fig. 6C and D). Additionally, we also note that the complementation of ΔpgL with pSCrhaB2-*motB*/pSCrhaB2-*fliF* is insufficient to restore motility defects associated with ΔpgL (Fig. S23). Combined these results support the loss of glycosylation within FliF_{S358A} and MotB_{S331A} and highlight that these proteins contribute to motility independent of their glycosylation status.

DISCUSSION

Within this study, we have expanded and curated the known glycoproteome of *B. cenocepacia* revealing at least 141 proteins or ~2% of the proteome is subjected to glycosylation [Supplementary Data 1 and 2 (<https://doi.org/10.6084/m9.figshare.25492774.v1>)] and, by leveraging this resource assessed the impact of glycosylation on two proteins associated with motility in *B. cenocepacia*, FliF and MotB. While several studies (55, 100, 101) have explored the impact of glycosylation on the broader proteomes, the dynamics of bacterial glycoproteomes, both across growth phases and in response to the loss of glycosylation, have largely been unexplored due to the technical challenges associated with tracking bacterial glycoproteins. The recent refinements to MS approaches such as DIA analysis (57, 58), now enable quantitative assessments of bacterial proteomes to be assessed at scale. While MS approaches to undertake such studies are now widely accessible, to our knowledge, only a single study within *Campylobacter jejuni* has leveraged similar approaches to specifically track glycoprotein occupancy in response to the loss of N-linked glycosylation utilizing these insights to uncover previously unrecognized glycosylation sites (102). The recent identification of hundreds of glycosylation events within several Gram-negative species (102–105) highlights the need for robust tools to track bacterial glycoproteomes with this work highlighting how dynamic information can be used to prioritize the identification of potentially important glycoproteins and sites.

The identification of only a modest increase in the known glycoproteome (<10%, Fig. 1) across differing culturing conditions as well as growth phases coupled with our measurements of >100 glycoproteins across these conditions (Fig. 2) supports the notion that the glycoproteome is a constitutive and invariant feature of the *B. cenocepacia* proteome. Consistent with this, the enzymes associated with O-linked glycosylation are produced at similar protein levels across all growth phases/culturing media examined within this study (Fig. S8). Across the 141 glycoproteins confirmed within *B. cenocepacia*, we found that despite large proteomic changes, the absence of glycosylation only impacts the abundance of five observable glycoproteins (Fig. 3; Fig. S11). While this work recapitulates our previous observations of the loss of BCAL1086 and BCAL2974 (55), it supports the loss of glycoproteins in response to the disruption of glycosylation are the exception, not the rule within the *B. cenocepacia* glycoproteome. Consistent with this concept, recent work on *C. jejuni* N-linked glycosylation has demonstrated that glycosylation events can be required for the stabilization of protein complexes exerting their effects independent of changes in protein levels (106). These observations support the assessments of protein properties invisible to standard proteomic assays, such as exploring changes within the interactome, or protein folding, are likely required to dissect the underlying alterations driving the large proteomic impact observed by the loss of glycosylation.

Within this work, we explored the link between *B. cenocepacia* glycosylation, and the reduced motility observed within ΔpgL (41, 55). Combining our glycoproteome and proteomic data sets, we identified two motility-associated glycoproteins, MotB and FliF, that are dramatically reduced in response to the loss of glycosylation with the disruption of these proteins observed to drive proteomic effects mirroring those seen within ΔpgL (Fig. 3 and 4). However, our analysis of the role of glycosylation within these proteins using site-directed mutagenesis and glycosylation null strains demonstrates that glycosylation does not appear to directly impact the role of these proteins (Fig. 5 and 6). Interestingly, while we did observe complementation within the $\Delta pgL\Delta fliF$

and $\Delta pgII\Delta motB$ backgrounds, the observed levels of motility restoration were reduced compared to complementation within the $\Delta fliF$ and $\Delta motB$ backgrounds (Fig. 5 and 6). We hypothesize that the widespread proteomic impacts observed from the loss of O-linked glycosylation may reduce the amount of functional FliF and MotB despite complementation, leading to similar protein levels between $\Delta pgII\Delta fliF$ and $\Delta fliF$, as well as between $\Delta pgII\Delta motB$ and $\Delta motB$ (Fig. 6). Regardless, our observations suggest that the O-linked glycosylation sites of S358 and S331 within FliF and MotB, respectively, are not integral for the function of these glycoproteins in motility. While this does not exclude the potential importance of these sites for functions beyond motility, it is notable that our proteomic analysis supports no clear proteomic differences from the absence or presence of these glycosylation events (Tables S8 and S9; Fig. S21 and S22). However, it should be noted that the observation of the dispensable nature of these sites may only occur under the ideal conditions examined here which are dramatically different from *B. cenocepacia's* natural reservoir. This said, within *C. jejuni*, it has been noted that multiple N-linked glycoproteins possess glycosylation sites that are dispensable for protein function such as Cj1496c (107) and Cj0143c (108). In *C. jejuni*, site-directed mutagenesis of glycosylation events in Cj1496c and Cj0143c demonstrate glycosylation sites in these proteins do not impact chicken colonization, yet loss of these proteins is detrimental to colonization (107, 108). These findings highlight that the presence of glycosylation in bacterial systems alone is not an indicator of the importance of a given glycosylation event and that the glycosylation events of FliF and MotB appear dispensable within *B. cenocepacia*.

In summary, this work furthers our understanding of the breadth and dynamics of *Burkholderia* glycoproteins revealing the *B. cenocepacia* glycoproteome composition is constitutively expressed and largely invariant within the proteome. Using DIA proteomics, we demonstrate that while the loss of glycosylation drives large proteome alterations in *B. cenocepacia* few glycoproteins display significant alterations in abundance. These observations support the loss of glycosylation likely impacts protein properties invisible to direct proteomic analysis, such as protein interactions. Combined our identification of 141 glycoproteins within *B. cenocepacia* represents the largest known PglL-mediated O-linked glycoproteome identified in Gram-negative species to date providing a roadmap for systematic characterization of *B. cenocepacia* glycoproteins to understand the role of these proteins in physiology and virulence.

ACKNOWLEDGMENTS

The authors thank the Melbourne Mass Spectrometry and Proteomics Facility of The Bio21 Molecular Science and Biotechnology Institute for access to MS instrumentation.

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

AUTHOR AFFILIATION

¹Department of Microbiology and Immunology, University of Melbourne at the Peter Doherty Institute for Infection and Immunity, Melbourne, Australia

PRESENT ADDRESS

Jessica M. Lewis, School of Life Sciences, University of Warwick, Coventry, United Kingdom

Pauline M. L. Coulon, Australian Institute for Microbiology and Immunology, Faculty of Science, University of Technology Sydney, Ultimo, Australia

AUTHOR ORCID^s

Nicholas E. Scott  <http://orcid.org/0000-0003-2556-8316>

FUNDING

Funder	Grant(s)	Author(s)
Australian Research Council	FT200100270, DP210100362	Nichollas E. Scott

DATA AVAILABILITY

Mass spectrometry data (RAW files, FragPipe outputs, Spectronaut outputs and experiment files, MaxQuant search outputs, Rmarkdown scripts, and output tables) have been deposited into the PRIDE ProteomeXchange repository (109, 110). All PRIDE accession numbers, descriptions of the associated experiments and experiment types (DIA or DDA) are provided within Table S13. Supplementary Data set 1 and 2 have been deposited within Figshare (<https://doi.org/10.6084/m9.figshare.25492774.v1>). All pride proteomic datasets are available at accessions [PXD043184](#), [PXD043186](#), [PXD043187](#), [PXD043188](#), [PXD043193](#), [PXD043194](#), [PXD043197](#), [PXD043221](#), [PXD043227](#), [PXD043280](#), and [PXD045246](#).

ADDITIONAL FILES

The following material is available [online](#).

Supplemental Material

Supplemental material (Spectrum00346-24-S0001.pdf). Tables S10 to S13; Fig. S1 to S24.

Table S1 (Spectrum00346-24-S0002.xlsx). Glycopeptide identifications within *B. cenocepacia* K56-2 identified across growth conditions using FAIMS and ZIC-HILIC enrichments.

Table S2 (Spectrum00346-24-S0003.xlsx). Glycoproteins identified across *B. cenocepacia*.

Table S3 (Spectrum00346-24-S0004.xlsx). DIA Protein level analysis of *B. cenocepacia* K56-2 grown in LB, TSB and ASM culturing media (Stationary and Log growth phases).

Table S4 (Spectrum00346-24-S0005.xlsx). DIA Protein level analysis of *B. cenocepacia* K56-2 wildtype, Δ pglL and Δ pglL Δ amrAB::native-promotor-pglL-his grown to stationary phase in LB, TSB and ASM culturing media.

Table S5 (Spectrum00346-24-S0006.xlsx). DDA protein level LFQ analysis of *B. cenocepacia* Δ motB candidates; Δ fliF candidates and Δ BCAM0505 candidates compared to WT at stationary phase in LB.

Table S6 (Spectrum00346-24-S0007.xlsx). DIA Protein level analysis of Δ motB, Δ fliF, Δ BCAM0505 and Δ pglL compared to WT at stationary phase in LB.

Table S7 (Spectrum00346-24-S0008.xlsx). Enrichment analysis of proteome changes observed within Δ motB, Δ fliF and Δ pglL compared to WT at stationary phase in LB.

Table S8 (Spectrum00346-24-S0009.xlsx). DIA Protein level analysis of Δ motB and WT containing pScrhaB2, pScrhaB2-motB or pScrhaB2-motB5331A with / without induction (0.1% rhamnose) at stationary phase in LB.

Table S9 (Spectrum00346-24-S0010.xlsx). DIA Protein level analysis of Δ fliF containing pScrhaB2, pScrhaB2-fliF or pScrhaB2-fliF5358A and WT containing pScrhaB2 with / without induction (0.1% rhamnose) at stationary phase in LB.

Open Peer Review

PEER REVIEW HISTORY (review-history.pdf). An accounting of the reviewer comments and feedback.

REFERENCES

- Depoorter E, Bull MJ, Peeters C, Coenye T, Vandamme P, Mahenthiralingam E. 2016. *Burkholderia*: an update on taxonomy and biotechnological potential as antibiotic producers. *Appl Microbiol Biotechnol* 100:5215–5229. <https://doi.org/10.1007/s00253-016-7520-x>
- Jin Y, Zhou J, Zhou J, Hu M, Zhang Q, Kong N, Ren H, Liang L, Yue J. 2020. Genome-based classification of *Burkholderia cepacia* complex provides new insight into its taxonomic status. *Biol Direct* 15:6. <https://doi.org/10.1186/s13062-020-0258-5>
- Vanlaere E, Baldwin A, Gevers D, Henry D, De Brandt E, LiPuma JJ, Mahenthiralingam E, Speert DP, Dowson C, Vandamme P. 2009. Taxon K, a complex within the *Burkholderia cepacia* complex, comprises at least two novel species, *Burkholderia contaminans* sp. nov. and *Burkholderia lata* sp. nov. *Int J Syst Evol Microbiol* 59:102–111. <https://doi.org/10.1099/ijs.0.001123-0>
- Mahenthiralingam E, Urban TA, Goldberg JB. 2005. The multifarious, multireplicon *Burkholderia cepacia* complex. *Nat Rev Microbiol* 3:144–156. <https://doi.org/10.1038/nrmicro1085>
- Sajjan U, Corey M, Humar A, Tullis E, Cutz E, Ackerley C, Forstner J. 2001. Immunolocalisation of *Burkholderia cepacia* in the lungs of cystic fibrosis patients. *J Med Microbiol* 50:535–546. <https://doi.org/10.1099/0022-1317-50-6-535>
- Courtney JM, Dunbar KEA, McDowell A, Moore JE, Warke TJ, Stevenson M, Elborn JS. 2004. Clinical outcome of *Burkholderia cepacia* complex infection in cystic fibrosis adults. *Journal of Cystic Fibrosis* 3:93–98. <https://doi.org/10.1016/j.jcf.2004.01.005>
- Jones AM, Dodd ME, Govan JRW, Barcus V, Doherty CJ, Morris J, Webb AK. 2004. *Burkholderia cenocepacia* and *Burkholderia multivorans*: influence on survival in cystic fibrosis. *Thorax* 59:948–951. <https://doi.org/10.1136/thx.2003.017210>
- Chaparro C, Maurer J, Gutierrez C, Krajdenc M, Chan C, Winton T, Keshavjee S, Scavuzzo M, Tullis E, Hutcheon M, Kesten S. 2001. Infection with *Burkholderia cepacia* in cystic fibrosis: outcome following lung transplantation. *Am J Respir Crit Care Med* 163:43–48. <https://doi.org/10.1164/ajrccm.163.1.9811076>
- Horsley A, Webb K, Bright-Thomas R, Govan J, Jones A. 2011. Can early *Burkholderia cepacia* complex infection in cystic fibrosis be eradicated with antibiotic therapy? *Front Cell Infect Microbiol* 1:18. <https://doi.org/10.3389/fcimb.2011.00018>
- Marciano BE, Spalding C, Fitzgerald A, Mann D, Brown T, Osgood S, Yockey L, Darnell DN, Barnhart L, Daub J, et al. 2015. Common severe infections in chronic granulomatous disease. *Clin Infect Dis* 60:1176–1183. <https://doi.org/10.1093/cid/ciu1154>
- Zlosnik JE, et al. 2014. *Burkholderia* spp. infections in cystic fibrosis patients in British Columbia, Canada: 30 years' experience. *Ann Am Thorac Soc*. <https://doi.org/10.1513/AnnalsATS.201408-395OC>
- Regan KH, Bhatt J. 2016. Eradication therapy for *Burkholderia cepacia* complex in people with cystic fibrosis. *Cochrane Database Syst Rev* 11:CD009876. <https://doi.org/10.1002/14651858.CD009876.pub3>
- Rhodes KA, Schweizer HP. 2016. Antibiotic resistance in *Burkholderia* species. *Drug Resist Updat* 28:82–90. <https://doi.org/10.1016/j.drug.2016.07.003>
- Isles A, Maclusky I, Corey M, Gold R, Prober C, Fleming P, Levison H. 1984. *Pseudomonas cepacia* infection in cystic fibrosis: an emerging problem. *J Pediatr* 104:206–210. [https://doi.org/10.1016/s0022-3476\(84\)80993-2](https://doi.org/10.1016/s0022-3476(84)80993-2)
- Somayaji R, Yau YCW, Tullis E, LiPuma JJ, Ratjen F, Waters V. 2020. Clinical outcomes associated with *Burkholderia cepacia* complex infection in patients with cystic fibrosis. *Ann Am Thorac Soc* 17:1542–1548. <https://doi.org/10.1513/AnnalsATS.202003-204OC>
- Loutet SA, Valvano MA. 2010. A decade of *Burkholderia cenocepacia* virulence determinant research. *Infect Immun* 78:4088–4100. <https://doi.org/10.1128/IAI.00212-10>
- Valvano MA. 2015. Intracellular survival of *Burkholderia cepacia* complex in phagocytic cells. *Can J Microbiol* 61:607–615. <https://doi.org/10.1139/cjm-2015-0316>
- Schwab U, Abdullah LH, Perlmutter OS, Albert D, Davis CW, Arnold RR, Yankaskas JR, Gilligan P, Neubauer H, Randell SH, Boucher RC. 2014. Localization of *Burkholderia cepacia* complex bacteria in cystic fibrosis lungs and interactions with *Pseudomonas aeruginosa* in hypoxic mucus. *Infect Immun* 82:4729–4745. <https://doi.org/10.1128/IAI.01876-14>
- Hare NJ, Soe CZ, Rose B, Harbour C, Codd R, Manos J, Cordwell SJ. 2012. Proteomics of *Pseudomonas aeruginosa* Australian epidemic strain 1 (AES-1) cultured under conditions mimicking the cystic fibrosis lung reveals increased iron acquisition via the siderophore pyochelin. *J Proteome Res* 11:776–795. <https://doi.org/10.1021/pr200659h>
- Scott NE, Hare NJ, White MY, Manos J, Cordwell SJ. 2013. Secretome of transmissible *Pseudomonas aeruginosa* AES-1R grown in a cystic fibrosis lung-like environment. *J Proteome Res* 12:5357–5369. <https://doi.org/10.1021/pr4007365>
- Palmer KL, Aye LM, Whiteley M. 2007. Nutritional cues control *Pseudomonas aeruginosa* multicellular behavior in cystic fibrosis sputum. *J Bacteriol* 189:8079–8087. <https://doi.org/10.1128/JB.01138-07>
- Jaiyesimi OA, McAvoy AC, Fogg DN, Garg N. 2021. Metabolomic profiling of *Burkholderia cenocepacia* in synthetic cystic fibrosis sputum medium reveals nutrient environment-specific production of virulence factors. *Sci Rep* 11:21419. <https://doi.org/10.1038/s41598-021-00421-4>
- Yoder-Himes DR, Chain PSG, Zhu Y, Wurtzel O, Rubin EM, Tiedje JM, Sorek R. 2009. Mapping the *Burkholderia cenocepacia* niche response via high-throughput sequencing. *Proc Natl Acad Sci U S A* 106:3976–3981. <https://doi.org/10.1073/pnas.0813403106>
- Drevinek P, Holden MTG, Ge Z, Jones AM, Ketchell I, Gill RT, Mahenthiralingam E. 2008. Gene expression changes linked to antimicrobial resistance, oxidative stress, iron depletion and retained motility are observed when *Burkholderia cenocepacia* grows in cystic fibrosis sputum. *BMC Infect Dis* 8:121. <https://doi.org/10.1186/1471-2334-8-121>
- Yoder-Himes DR, Constantinidis KT, Tiedje JM. 2010. Identification of potential therapeutic targets for *Burkholderia cenocepacia* by comparative transcriptomics. *PLoS One* 5:e8724. <https://doi.org/10.1371/journal.pone.0008724>
- Kumar B, Cardona ST. 2016. Synthetic cystic fibrosis sputum medium regulates flagellar biosynthesis through the flhF gene in *Burkholderia cenocepacia*. *Front Cell Infect Microbiol* 6:65. <https://doi.org/10.3389/fcimb.2016.00065>
- Hamlin JNR, Bloodworth RAM, Cardona ST. 2009. Regulation of phenylacetic acid degradation genes of *Burkholderia cenocepacia* K56-2. *BMC Microbiol* 9:222. <https://doi.org/10.1186/1471-2180-9-222>
- Liu H, Ibrahim M, Qiu H, Kausar S, Ilyas M, Cui Z, Hussain A, Li B, Waheed A, Zhu B, Xie G. 2015. Protein profiling analyses of the outer membrane of *Burkholderia cenocepacia* reveal a niche-specific proteome. *Microb Ecol* 69:75–83. <https://doi.org/10.1007/s00248-014-0460-z>
- Koomey M. 2019. O-linked protein glycosylation in bacteria: snapshots and current perspectives. *Curr Opin Struct Biol* 56:198–203. <https://doi.org/10.1016/j.sbi.2019.03.020>
- Nothhaft H, Szymanski CM. 2010. Protein glycosylation in bacteria: sweeter than ever. *Nat Rev Microbiol* 8:765–778. <https://doi.org/10.1038/nrmicro2383>
- Nothhaft 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 Proteom* 11:1203–1219. <https://doi.org/10.1074/mcp.M112.021519>
- Jervis AJ, Butler JA, Lawson AJ, Langdon R, Wren BW, Linton D. 2012. Characterization of the structurally diverse N-linked glycans of *Campylobacter* species. *J Bacteriol* 194:2355–2362. <https://doi.org/10.1128/JB.00042-12>
- Hadjineophytou C, Anonsen JH, Wang N, Ma KC, Viburiene R, Vik Å, 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>

34. Anonsen JH, Vik Å, Børud B, Viburieni 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>
35. Vik A, Aas FE, Anonsen JH, Billsborough 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>
36. Børud B, Aas FE, Vik A, Winther-Larsen HC, Egge-Jacobsen W, Koomey M. 2010. Genetic, structural, and antigenic analyses of glycan diversity in the O-linked protein glycosylation systems of human *Neisseria* species. *J Bacteriol* 192:2816–2829. <https://doi.org/10.1128/JB.00101-10>
37. Hadjineophytou C, Anonsen JH, Svingerud T, Mortimer TD, Grad YH, Scott NE, Koomey M. 2022. Sculpting the bacterial O-glycoproteome: functional analyses of orthologous oligosaccharyltransferases with diverse targeting specificities. *mBio* 13:e0379721. <https://doi.org/10.1128/mbio.03797-21>
38. 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>
39. Ahmad Izaham AR, Scott NE. 2020. Open database searching enables the identification and comparison of bacterial glycoproteomes without defining glycan compositions prior to searching. *Mol Cell Proteomics* 19:1561–1574. <https://doi.org/10.1074/mcp.TIR120.002100>
40. Fathy Mohamed Y, Scott NE, Molinaro A, Creuzenet C, Ortega X, Lertmemonkolchai G, Tunney MM, Green H, Jones AM, DeShazer D, Currie BJ, Foster LJ, Ingram R, De Castro C, Valvano MA. 2019. A general protein O-glycosylation machinery conserved in *Burkholderia* species improves bacterial fitness and elicits glycan immunogenicity in humans. *J Biol Chem* 294:13248–13268. <https://doi.org/10.1074/jbc.RA119.009671>
41. Lithgow KV, Scott NE, Iwashkiw JA, Thomson ELS, Foster LJ, Feldman MF, Dennis JJ. 2014. A general protein O-glycosylation system within the *Burkholderia cenocepacia* complex is involved in motility and virulence. *Mol Microbiol* 92:116–137. <https://doi.org/10.1111/mmi.12540>
42. Willcocks SJ, Denman C, Cia F, McCarthy E, Cuccui J, Wren BW. 2020. Virulence of the emerging pathogen, *Burkholderia pseudomallei*, depends upon the O-linked oligosaccharyltransferase, PglL. *Future Microbiol* 15:241–257. <https://doi.org/10.2217/fmb-2019-0165>
43. Ahmad Izaham AR, Ang C-S, 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>
44. Saleh S, Staes A, Deborggraeve S, Gevaert K. 2019. Targeted proteomics for studying pathogenic bacteria. *Proteomics* 19:e1800435. <https://doi.org/10.1002/pmic.201800435>
45. Semanjski M, Macek B. 2016. Shotgun proteomics of bacterial pathogens: advances, challenges and clinical implications. *Expert Rev Proteomics* 13:139–156. <https://doi.org/10.1586/14789450.2016.1132168>
46. Aebersold R, Mann M. 2016. Mass-spectrometric exploration of proteome structure and function. *Nature* 537:347–355. <https://doi.org/10.1038/nature19949>
47. Bagdonaite I. 2022. Glycoproteomics. *Nat Rev Methods Primers* 2. <https://doi.org/10.1038/s43586-022-00140-8>
48. Scott NE, Kinsella RL, Edwards AVG, 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>
49. Iwashkiw JA, Seper A, Weber BS, Scott NE, Vinogradov E, Stratilo C, Reiz B, Cordwell SJ, Whittall 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>
50. Scott NE, Parker BL, Connolly AM, Paulech J, Edwards AVG, Crosssett B, Falconer L, Kolarich D, Djordjevic SP, Højrup P, Packer NH, Larsen MR, Cordwell SJ. 2011. Simultaneous glycan-peptide characterization using hydrophilic interaction chromatography and parallel fragmentation by CID, higher energy collisional dissociation, and electron transfer dissociation MS applied to the N-linked glycoproteome of *Campylobacter jejuni*. *Mol Cell Proteomics* 10:M000031–MCP201. <https://doi.org/10.1074/mcp.M000031-MCP201>
51. Hadjineophytou C, Loh E, Koomey M, Scott NE. 2024. Combining FAIMS based glycoproteomics and DIA proteomics reveals widespread proteome alterations in response to glycosylation occupancy changes in *Neisseria gonorrhoeae*. *Proteomics*:e2300496. <https://doi.org/10.1002/pmic.202300496>
52. Knoot CJ, Wantuch PL, Robinson LS, Rosen DA, Scott NE, Harding CM. 2023. Discovery and characterization of a new class of O-linking oligosaccharyltransferases from the *Moraxellaceae* family. *Glycobiology* 33:57–74. <https://doi.org/10.1093/glycob/cwac070>
53. Krug K, Carpy A, Behrends G, Matic K, Soares NC, Macek B. 2013. Deep coverage of the *Escherichia coli* proteome enables the assessment of false discovery rates in simple proteogenomic experiments. *Mol Cell Proteomics* 12:3420–3430. <https://doi.org/10.1074/mcp.M113.029165>
54. Cain JA, Dale AL, Sumer-Bayraktar Z, Solis N, Cordwell SJ. 2020. Identifying the targets and functions of N-linked protein glycosylation in *Campylobacter jejuni*. *Mol Omics* 16:287–304. <https://doi.org/10.1039/d0mo00032a>
55. Oppy CC, Jebeli L, Kuba M, Oates CV, Strugnell R, Edgington-Mitchell LE, Valvano MA, Hartland EL, Newton HJ, Scott NE. 2019. Loss of O-linked protein glycosylation in *Burkholderia cenocepacia* impairs biofilm formation, siderophore activity and alters transcriptional regulators. *mSphere* 4:e00660–00619. <https://doi.org/10.1128/mSphere.00660-19>
56. Lazar C, Gatto L, Ferro M, Bruley C, Burger T. 2016. Accounting for the multiple natures of missing values in label-free quantitative proteomics data sets to compare imputation strategies. *J Proteome Res* 15:1116–1125. <https://doi.org/10.1021/acs.jproteome.5b00981>
57. Pino LK, Just SC, MacCoss MJ, Searle BC. 2020. Acquiring and analyzing data independent acquisition proteomics experiments without spectrum libraries. *Mol Cell Proteomics* 19:1088–1103. <https://doi.org/10.1074/mcp.P119.001913>
58. Ludwig C, Gillet L, Rosenberger G, Amon S, Collins BC, Aebersold R. 2018. Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial. *Mol Syst Biol* 14:e8126. <https://doi.org/10.15252/msb.20178126>
59. Ting YS, Egertson JD, Bollinger JG, Searle BC, Payne SH, Noble WS, MacCoss MJ. 2017. PECAN: library-free peptide detection for data-independent acquisition tandem mass spectrometry data. *Nat Methods* 14:903–908. <https://doi.org/10.1038/nmeth.4390>
60. Röst HL, Rosenberger G, Navarro P, Gillet L, Miladinović SM, Schubert OT, Wolski W, Collins BC, Malmström J, Malmström L, Aebersold R. 2014. OpenSWATH enables automated, targeted analysis of data-independent acquisition MS data. *Nat Biotechnol* 32:219–223. <https://doi.org/10.1038/nbt.2841>
61. Demichev V, Messner CB, Vernardis SI, Lilley KS, Ralser M. 2020. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat Methods* 17:41–44. <https://doi.org/10.1038/s41592-019-0638-x>
62. Bekker-Jensen DB, Bernhardt OM, Hogrebe A, Martinez-Val A, Verbeke L, Gandhi T, Kelstrup CD, Reiter L, Olsen JV. 2020. Rapid and site-specific deep phosphoproteome profiling by data-independent acquisition without the need for spectral libraries. *Nat Commun* 11:787. <https://doi.org/10.1038/s41467-020-14609-1>
63. Huang T, Bruderer R, Muntel J, Xuan Y, Vitek O, Reiter L. 2020. Combining precursor and fragment information for improved detection of differential abundance in data independent acquisition. *Mol Cell Proteomics* 19:421–430. <https://doi.org/10.1074/mcp.RA119.001705>
64. Fabre B, Combier JP, Plaza S. 2021. Recent advances in mass spectrometry-based peptidomics workflows to identify short-open-reading-frame-encoded peptides and explore their functions. *Curr Opin Chem Biol* 60:122–130. <https://doi.org/10.1016/j.cbpa.2020.12.002>
65. Price EP, Viberg LT, Kidd TJ, Bell SC, Currie BJ, Sarovich DS. 2018. Transcriptomic analysis of longitudinal *Burkholderia pseudomallei* infecting the cystic fibrosis lung. *Microb Genom* 4:e000194. <https://doi.org/10.1099/mgen.0.000194>
66. Sriramulu DD, Lünsdorf H, Lam JS, Römmling U. 2005. Microcolony formation: a novel biofilm model of *Pseudomonas aeruginosa* for the

- cystic fibrosis lung. *J Med Microbiol* 54:667–676. <https://doi.org/10.1099/jmm.0.45969-0>
67. Fung C, Naughton S, Turnbull L, Tingpej P, Rose B, Arthur J, Hu H, Harmer C, Harbour C, Hassett DJ, Whitchurch CB, Manos J. 2010. Gene expression of *Pseudomonas aeruginosa* in a mucin-containing synthetic growth medium mimicking cystic fibrosis lung sputum. *J Med Microbiol* 59:1089–1100. <https://doi.org/10.1099/jmm.0.019984-0>
 68. Aubert DF, Hamad MA, Valvano MA. 2014. A markerless deletion method for genetic manipulation of *Burkholderia cenocepacia* and other multidrug-resistant gram-negative bacteria. *Methods Mol Biol* 1197:311–327. https://doi.org/10.1007/978-1-4939-1261-2_18
 69. Flannagan RS, Linn T, Valvano MA. 2008. A system for the construction of targeted unmarked gene deletions in the genus *Burkholderia*. *Environ Microbiol* 10:1652–1660. <https://doi.org/10.1111/j.1462-2920.2008.01576.x>
 70. Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA 3rd, Smith HO. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* 6:343–345. <https://doi.org/10.1038/nmeth.1318>
 71. Cardona ST, Valvano MA. 2005. An expression vector containing a rhamnose-inducible promoter provides tightly regulated gene expression in *Burkholderia cenocepacia*. *Plasmid* 54:219–228. <https://doi.org/10.1016/j.plasmid.2005.03.004>
 72. Sambrook J, Fritsch EF, Maniatis T. 1990. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor.
 73. Darling P, Chan M, Cox AD, Sokol PA. 1998. Siderophore production by cystic fibrosis isolates of *Burkholderia cepacia*. *Infect Immun* 66:874–877. <https://doi.org/10.1128/IAI.66.2.874-877.1998>
 74. Figurski DH, Helinski DR. 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. *Proc Natl Acad Sci U S A* 76:1648–1652. <https://doi.org/10.1073/pnas.76.4.1648>
 75. Hamad MA, Di Lorenzo F, Molinaro A, Valvano MA. 2012. Aminoarabinose is essential for lipopolysaccharide export and intrinsic antimicrobial peptide resistance in *Burkholderia cenocepacia* (dagger). *Mol Microbiol* 85:962–974. <https://doi.org/10.1111/j.1365-2958.2012.08154.x>
 76. Kulak NA, Pichler G, Paron I, Nagaraj N, Mann M. 2014. Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat Methods* 11:319–324. <https://doi.org/10.1038/nmeth.2834>
 77. Masuda T, Tomita M, Ishihama Y. 2008. Phase transfer surfactant-aided trypsin digestion for membrane proteome analysis. *J Proteome Res* 7:731–740. <https://doi.org/10.1021/pr700658q>
 78. Rogers LD, Fang Y, Foster LJ. 2010. An integrated global strategy for cell lysis, fractionation, enrichment and mass spectrometric analysis of phosphorylated peptides. *Mol Biosyst* 6:822–829. <https://doi.org/10.1039/b915986j>
 79. Humphrey SJ, Karayel O, James DE, Mann M. 2018. High-throughput and high-sensitivity phosphoproteomics with the EasyPhos platform. *Nat Protoc* 13:1897–1916. <https://doi.org/10.1038/s41596-018-0014-9>
 80. Harney DJ, Hutchison AT, Hatchwell L, Humphrey SJ, James DE, Hocking S, Heilbronn LK, Larance M. 2019. Proteomic analysis of human plasma during intermittent fasting. *J Proteome Res* 18:2228–2240. <https://doi.org/10.1021/acs.jproteome.9b00090>
 81. 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>
 82. Mysling S, Palmisano G, Højrup P, Thaysen-Andersen M. 2010. Utilizing ion-pairing hydrophilic interaction chromatography solid phase extraction for efficient glycopeptide enrichment in glycoproteomics. *Anal Chem* 82:5598–5609. <https://doi.org/10.1021/ac100530w>
 83. 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>
 84. Čaval T, Zhu J, Heck AJR. 2019. Simply extending the mass range in electron transfer higher energy collisional dissociation increases confidence in N-glycopeptide identification. *Anal Chem* 91:10401–10406. <https://doi.org/10.1021/acs.analchem.9b02125>
 85. Polasky DA, Geiszler DJ, Yu F, Nesvizhskii AI. 2022. Multi-attribute glycan identification and FDR control for glycoproteomics. *Mol Cell Proteomics* 21:100205. <https://doi.org/10.1016/j.mcpro.2022.100205>
 86. Kong AT, Leprevost FV, Avtonomov DM, Mellacheruvu D, Nesvizhskii AI. 2017. MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. *Nat Methods* 14:513–520. <https://doi.org/10.1038/nmeth.4256>
 87. Polasky DA, Yu F, Teo GC, Nesvizhskii AI. 2020. Fast and comprehensive N- and O-glycoproteomics analysis with MSFragger-Glyco. *Nat Methods* 17:1125–1132. <https://doi.org/10.1038/s41592-020-0967-9>
 88. Holden MTG, Seth-Smith HMB, Crossman LC, Sebahia M, Bentley SD, Cerdeño-Tárraga AM, Thomson NR, Bason N, Quail MA, Sharp S, et al. 2009. The genome of *Burkholderia cenocepacia* J2315, an epidemic pathogen of cystic fibrosis patients. *J Bacteriol* 191:261–277. <https://doi.org/10.1128/JB.01230-08>
 89. Wickham H. 2016. ggplot2: elegant graphics for data analysis. Springer-Verlag New York, Cham.
 90. Brademan DR, Riley NM, Kwiecien NW, Coon JJ. 2019. Interactive peptide spectral annotator: a versatile web-based tool for proteomic applications. *Mol Cell Proteomics* 18:S193–S201. <https://doi.org/10.1074/mcp.TIR118.001209>
 91. Bern M, Kil YJ, Becker C. 2012. Byonic: advanced peptide and protein identification software. *Curr Protoc Bioinformatics* 40:13. <https://doi.org/10.1002/0471250953.bi1320s40>
 92. Varga JJ, Losada L, Zelazny AM, Kim M, McCorrison J, Brinkac L, Sampaio EP, Greenberg DE, Singh I, Heiner C, Ashby M, Nierman WC, Holland SM, Goldberg JB. 2013. Draft genome sequences of *Burkholderia cenocepacia* Et12 lineage strains K56-2 and BC7. *Genome Announc* 1:e00841-13. <https://doi.org/10.1128/genomeA.00841-13>
 93. 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>
 94. 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>
 95. 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>
 96. Hebert AS, Prasad S, Belford MW, Bailey DJ, McAlister GC, Abbatiello SE, Huguet R, Wouters ER, Dunyach J-J, 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>
 97. García-Romero I, Valvano MA. 2020. Complete genome sequence of *Burkholderia cenocepacia* K56-2, an opportunistic pathogen. *Microbiol Resour Announc* 9:e01015-20. <https://doi.org/10.1128/MRA.01015-20>
 98. Hamad MA, Skeldon AM, Valvano MA. 2010. Construction of aminoglycoside-sensitive *Burkholderia cenocepacia* strains for use in studies of intracellular bacteria with the gentamicin protection assay. *Appl Environ Microbiol* 76:3170–3176. <https://doi.org/10.1128/AEM.03024-09>
 99. Armitage JP, Berry RM. 2020. Assembly and dynamics of the bacterial flagellum. *Annu Rev Microbiol* 74:181–200. <https://doi.org/10.1146/annurev-micro-090816-093411>
 100. Cain JA, Dale AL, Niewold P, Klare WP, Man L, White MY, Scott NE, Cordwell SJ. 2019. Proteomics reveals multiple phenotypes associated with N-linked glycosylation in *Campylobacter jejuni*. *Mol Cell Proteom* 18:715–734. <https://doi.org/10.1074/mcp.RA118.001199>
 101. Abouelhadi S, North SJ, Hitchen P, Vohra P, Chintoan-Uta C, Stevens M, Dell A, Cuccui J, Wren BW. 2019. Quantitative analyses reveal novel roles for N-glycosylation in a major enteric bacterial pathogen. *mBio* 10:e00297-19. <https://doi.org/10.1128/mBio.00297-19>
 102. Cain JA, Dale AL, Cordwell SJ. 2021. Exploiting pglB oligosaccharyltransferase-positive and -negative *Campylobacter jejuni* and a multiprotease

- digestion strategy to identify novel sites modified by N-linked protein glycosylation. *J Proteome Res* 20:4995–5009. <https://doi.org/10.1021/acs.jproteome.1c00482>
103. Veith PD, Gorasia DG, Reynolds EC. 2023. Characterization of the O-glycoproteome of *Flavobacterium johnsoniae*. *J Bacteriol* 205:e0009323. <https://doi.org/10.1128/jb.00093-23>
104. Veith PD, Scott NE, Reynolds EC. 2021. Characterization of the O-glycoproteome of *Tannerella forsythia*. *mSphere* 6:e0064921. <https://doi.org/10.1128/mSphere.00649-21>
105. Veith PD, Shoji M, Scott NE, Reynolds EC. 2022. Characterization of the O-glycoproteome of *Porphyromonas gingivalis*. *Microbiol Spectr* 10:e0150221. <https://doi.org/10.1128/spectrum.01502-21>
106. Abouelhadid S, Raynes J, Bui T, Cucui J, Wren BW. 2020. Characterization of posttranslationally modified multidrug efflux pumps reveals an unexpected link between glycosylation and antimicrobial resistance. *mBio* 11:e02604-20. <https://doi.org/10.1128/mBio.02604-20>
107. Kakuda T, DiRita VJ. 2006. Cj1496c encodes a *Campylobacter jejuni* glycoprotein that influences invasion of human epithelial cells and colonization of the chick gastrointestinal tract. *Infect Immun* 74:4715–4723. <https://doi.org/10.1128/IAI.00033-06>
108. Davis LM, Kakuda T, DiRita VJ. 2009. A *Campylobacter jejuni* znuA orthologue is essential for growth in low-zinc environments and chick colonization. *J Bacteriol* 191:1631–1640. <https://doi.org/10.1128/JB.01394-08>
109. Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, Inuganti A, Griss J, Mayer G, Eisenacher M, Pérez 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>
110. Vizcaino JA, Csordas A, del-Toro N, Dianas JA, Griss J, Lavidas I, Mayer G, Perez-Riverol Y, Reisinger F, Ternent T, Xu Q-W, Wang R, Hermjakob H. 2016. Update of the PRIDE database and its related tools. *Nucleic Acids Res* 44:D447–D456. <https://doi.org/10.1093/nar/gkv1145>