What are we missing by using hydrophilic enrichment? Improving bacterial glycoproteome coverage using total proteome and FAIMS analysis.

Ameera Raudah Ahmad Izaham¹ Ching-Seng Ang², Shuai Nie², Lauren E. Bird¹, Nicholas A. Williamson² and Nichollas E. Scott^{1#}

¹Department of Microbiology and Immunology, University of Melbourne at the Peter Doherty Institute for Infection and Immunity, Melbourne 3000, Australia

²Melbourne Mass Spectrometry and Proteomics Facility, The Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, Parkville, VIC, 3010, Australia.

Running title: identifying HILIC inaccessible glycopeptides within bacterial glycoproteomes

Keywords: Glycoproteomics, FAIMS, HILIC, Glycopeptide enrichment, Bacterial glycosylation, *Burkholderia*

ABSTRACT

Hydrophilic Interaction Liquid Chromatography (HILIC) glycopeptide enrichment is an indispensable tool for the high-throughput characterisation of glycoproteomes. Despite its utility, HILIC enrichment is associated with a number of shortcomings including requiring large amounts of starting material, potentially introducing chemical artefacts such as formylation, and biasing/under-sampling specific classes of glycopeptides. Here we investigate HILIC enrichment independent approaches for the study of bacterial glycoproteomes. Using three *Burkholderia* species (*B. cenocepacia*, *B. dolosa* and *B. ubonensis*) we demonstrate that short aliphatic *O*-linked glycopeptides are typically absent from HILIC enrichments yet are readily identified in whole proteome samples. Using Field Asymmetric Waveform IMS (FAIMS) fractionation we show that at low compensation voltages (CVs) short aliphatic glycopeptides can be enriched from complex samples providing an alternative means to identify glycopeptides recalcitrant to hydrophilic based enrichment. Combining whole proteome and FAIMS analysis we show that the observable glycoproteome of these *Burkholderia* species is at least 30% larger than initially thought. Excitingly, the ability to enrich glycopeptides using FAIMS appears generally applicable, with the *N*-linked glycopeptides of *Campylobacter fetus subsp. fetus* also enrichable at low FAIMS CVs. Taken together, these results demonstrate that FAIMS provides an alternative means to access glycopeptides and is a valuable tool for glycoproteomic analysis.

INTRODUCTION

Glycosylation is a ubiquitous protein modification found in all domains of life [1]. Over the last two decades significant progress has been made in our understanding of glycosylation targets leading to the establishment of the field of glycoproteomics [2, 3]. The growth of glycoproteomics has been driven by the increasing sensitivity and speed of mass spectrometry (MS) instrumentation as well as dramatic improvements in glycopeptide informatics [2-4]. These advancements now enable intact glycopeptide analysis – the study of glycopeptides decorated with their native [4], native-like [5, 6] or truncated [7, 8] glycoforms – and enables the qualitative and quantitative assessment of glycosylation at a site-specific resolution. Intact glycopeptide analysis has been critical for the study of glycosylation where the sites of modification are not predictable based on amino acid sequence such as mucin O-linked glycosylation [7, 8], O-mannosylation [9, 10] and O-glcNAcylation [11]. Although powerful, a key requirement for intact glycopeptide analysis is the ability to enrich/isolate glycopeptides for downstream analysis.

To date, MS based characterisation of intact glycopeptides has been heavily dependent on glycan centric glycopeptide enrichment methods [2-4, 12]. Enrichment has been proposed to be critical for concentrating glycopeptides in even simple protein digests [13] to overcome the suppression of glycopeptide ions by co-eluting non-glycosylated peptides [12]. A range of methods have been developed to enrich glycopeptides from mixtures [14] including lectin-affinity based approaches [15-17], reversible chemical coupling resins such as boronic acid based derivatives [18] and irreversible chemical strategies such as hydrazine based chemistry followed by carbohydrate selective hydrolysis [19]. One of the most widely used glycopeptide enrichment approaches is

hydrophilic interaction liquid chromatography (HILIC) enrichment using zwitterionic hydrophilic interaction liquid chromatography (ZIC-HILIC) resins [20]. Enrichment using ZIC-HILIC is the result of hydrophilic partitioning of glycopeptides to an adsorbed “water layer” on the surface of the ZIC-HILIC resin [21, 22]. For glycopeptides, hydrophilic partitioning is thought to be driven largely by glycans making ZIC-HILIC enrichment compatible with a range of glycopeptides provided their glycans are sufficiently hydrophilic [23]. This enrichment approach has been used for nearly two decades [20, 24] for both eukaryotic [25-28] and prokaryotic [29-32] glycoproteomic studies. Although ZIC-HILIC enrichment is considered by some to be unbiased for eukaryotic *N*-linked glycoproteomics [12, 33], it has been previously shown to strongly favour large glycans [23] and show poor retention of *O*-linked glycopeptides [34]. This bias has important ramifications especially for the study of glycosylated substrates decorated with short glycans such as the glycoproteomes of *Burkholderia* species [31, 35].

Protein glycosylation is increasingly recognised as a highly conserved feature of bacterial physiology [36-39]. To date, studies of bacterial glycoproteomes have extensively used ZIC-HILIC enrichment [29-32, 40-43] to characterise glycosylation substrates and glycan diversity. Using ZIC-HILIC enrichment, we have previously shown over 100 proteins are glycosylated within *Burkholderia* species [44]. Despite this insight, the complete repertoire of most *Burkholderia* glycoproteomes are still unknown, hampering our understanding of the impacts observed when glycosylation is disrupted [31, 45, 46]. Within members of the *Burkholderia* genus, two linear trisaccharides composed of β -Gal-(1,3)- α -GalNAc-(1,3)- β -GalNAc and Suc- β -Gal-(1,3)- α -GalNAc-(1,3)- β -GalNAc (where Suc is Succinyl) are utilized for protein glycosylation [31, 35]. The short

length of these two glycans raises notable concerns about potential bias in the observable glycoproteome within ZIC-HILIC enrichments [23, 34]. This concern, coupled with our recent identification of extensive formylation artefacts as a result of ZIC-HILIC enrichment [44], suggests alternative approaches – such as orthogonal fractionation – are required for the analysis of bacterial glycopeptides.

Ion Mobility Spectrometry (IMS) is an increasingly utilised gas-phase fractionation approach which enables the separation of analytes based on a combination of charge and collisional cross-section [47]. Within proteomics, Field Asymmetric Waveform IMS (FAIMS) is one of the most widely used IMS approaches [48] and has been used to increase proteome coverage [49-51], reduce ratio compression within isobaric labelling experiments [52], improve the localisation of PTMs [53, 54] as well as enrich modified peptides such as sumoylated [55] and cross-linked [56] peptides. Recently a new FAIMS interface was introduced for Orbitrap instruments: the FAIMS Pro interface [49-51]. This new implementation of FAIMS improves sensitivity and enables rapid compensation voltage (CV) switching, making it ideal for proteomics studies and gas-phase fractionation [49, 50]. By applying multiple CVs in a single experiment, referred to from here in as stepped FAIMS, this has been shown to increase protein identifications by up to 55% vs without FAIMS implementation [49-51]. Previous iterations of FAIMS have been shown to enable the separation of glycopeptide isomers [57, 58] as well as improve the mapping of bacterial glycosylation sites within single proteins [59]. Due to the large collisional cross-section of glycopeptides compared to unmodified peptides [60, 61], IMS has been suggested to provide a

unique avenue for the sorting/enrichment of glycopeptides [60, 62]. However, to our knowledge
FAIMS has not been previously used for glycoproteomics studies.

In this work we explore ZIC-HILIC independent approaches for the characterisation of bacterial
glycoproteomes. By comparing the glycopeptides identified within three Burkholderia species (*B.*
cenocepacia, *B. dolosa* and *B. ubonensis*) with and without ZIC-HILIC enrichment, we find that
short aliphatic glycopeptides are under-represented within ZIC-HILIC enrichments. To improve
the identification of these glycopeptides we explore the use of FAIMS and demonstrate that at
low FAIMS CVs short/aliphatic glycopeptides can be readily enriched from total proteome
samples. By combining multiple FAIMS CVs within a single analytical method and the analysis of
total proteome samples, we demonstrate that the glycoproteome of these three Burkholderia
species are at least 30% larger than previously thought. Similarly, glycopeptides decorated with
large glycans such as the *N*-linked hexasaccharides of *Campylobacter fetus subsp. fetus* are
equally enriched at low FAIMS CVs with the use of FAIMS providing comparable coverage of the
glycoproteome using only a fraction of the starting material. The ability to access complementary
glycoproteomes as well as enrich glycopeptides from low input amounts makes FAIMS a powerful
tool for glycoproteomic analysis.

EXPERIMENTAL PROCEDURES

Bacterial strains and growth conditions: [63]. *B. cenocepacia* J2315; *B. dolosa* AU0158 and *B. ubonensis* MSMB22 strains were grown overnight on LB agar at 37°C as previously described [31]. *C. fetus subsp. fetus* NCTC 10842 was grown on Brain-Heart Infusion medium (Hardy Diagnostics) with 5% defibrinated horse blood (Hemostat, Dixon, CA) under microaerobic conditions (10% CO₂, 5% O₂, 85% N₂) at 37°C as previously reported [64]. Details on the strains, their origins, references and proteome databases are provided within Table 1.

Generation of bacterial lysates for glycoproteome analysis: Bacterial strains were grown to confluency before flooding plates with 5 mL of pre-chilled sterile phosphate-buffered saline (PBS) and bacterial cells collected by scraping. Cells were washed 3 times in PBS to remove media contaminants, then collected by centrifugation at 10,000 x *g* at 4°C for 10 min, and then snap frozen. Snap frozen cells were resuspended in 4% SDS, 100mM Tris pH 8.0, 20mM Dithiothreitol (DTT) and boiled at 95°C with shaking at 2000rpm for 10 min. Samples were clarified by centrifugation at 17,000 x *g* for 10 min, the supernatants were then collected, and protein concentrations determined by a bicinchoninic acid assay (Thermo Fisher Scientific, Waltham, MA, USA). 1mg of protein from each sample was acetone precipitated by mixing one volume of sample with 4 volumes of ice-cold acetone. Samples were precipitated overnight at -20°C and then spun down at 16,000 x *g* for 10 min at 0°C. The precipitated protein pellets were resuspended in 80% ice-cold acetone and precipitated for an additional 4 hours at -20°C. Samples were centrifuged at 17,000 x *g* for 10 min at 0°C, the supernatant discarded, and excess acetone driven off at 65°C for 5 min.

154

155 **Digestion of protein samples:** Protein digestion was undertaken as previously described with
 156 minor alterations [29]. Briefly, dried protein pellets were resuspended in 6 M urea, 2 M thiourea
 157 in 40 mM NH_4HCO_3 then reduced for 1 hour with 20mM DTT, followed by alkylation with 40mM
 158 chloroacetamide for 1 hour in the dark. Samples were then digested with Lys-C (Wako Chemicals,
 159 Japan; 1/200 w/w) for 3 hours before being diluted with 5 volumes of 40 mM NH_4HCO_3 and
 160 digested overnight with sequencing grade modified trypsin (Promega, Madison, WI, USA; 1/50
 161 w/w). Digested samples were acidified to a final concentration of 0.5% formic acid and desalted
 162 with 50 mg tC18 Sep-Pak columns (Waters corporation, Milford, MA, USA) according to the
 163 manufacturer's instructions. tC18 Sep-Pak columns were conditioned with 10 bed volumes of
 164 Buffer B (0.1% formic acid, 80% acetonitrile), then equilibrated with 10 bed volumes of Buffer A*
 165 (0.1% TFA, 2% acetonitrile) before use. Samples were loaded on to equilibrated columns and then
 166 columns washed with at least 10 bed volumes of Buffer A* before bound peptides were eluted
 167 with Buffer B. Eluted peptides were aliquoted into samples for ZIC-HILIC enrichment or total
 168 proteome analysis, then dried by vacuum centrifugation and stored at -20°C .

169

170 **Reverse phase LC-MS/MS:** Proteome samples were re-suspended in Buffer A* and separated
 171 using a two-column chromatography set up composed of a PepMap100 C18 20 mm x 75 μm trap
 172 and a PepMap C18 500 mm x 75 μm analytical column (Thermo Fisher Scientific). Samples were
 173 concentrated onto the trap column at 5 $\mu\text{L}/\text{min}$ for 5 minutes with Buffer A (0.1% formic acid, 2%
 174 DMSO) and then infused into an Orbitrap Fusion™ Lumos™ Tribrid™ Mass Spectrometer (Thermo
 175 Fisher Scientific) at 300 nL/minute via the analytical column using a Dionex Ultimate 3000 UPLC

(Thermo Fisher Scientific). For stepped CV and whole proteome experiments, 185-minute 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 150 minutes, then from 28% B to 40% B over 10 minutes, then from 40% B to 100% B over 2 minutes. The composition was held at 100% B for 3 minutes, and then dropped to 2% B over 5 minutes before being held at 2% B for another 15 minutes. These conditions are identical to the published ZIC-HILIC glycopeptide datasets [44] to enable the direct comparison of glycopeptide coverage. For static FAIMS CVs and associated no FAIMS control experiments, 125-minute 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 90 minutes, then from 28% B to 40% B over 10 minutes, then from 40% B to 100% B over 2 minutes. The composition was held at 100% B for 3 minutes, and then dropped to 2% B over 5 minutes before being held at 2% B for another 15 minutes. For all experiments, the Lumos™ Mass Spectrometer was operated in a data-dependent mode, automatically switching between the acquisition of a Orbitrap MS scan (120,000 resolution) every 3 seconds and Orbitrap MS/MS HCD scans of precursors (NCE 30%, maximal injection time of 60 ms, with an AGC of 200% and a resolution of 15,000). Glycan fragment ions (204.087; 366.1396 and 138.0545 m/z) product-dependent MS/MS analysis [65] was used to trigger three additional scans of potential glycopeptides; a Orbitrap EThcD scan (NCE 15%, maximal injection time of 250 ms with an AGC of 500% and a resolution of 30,000 using the extended mass range setting to improve the detection of high mass glycopeptide fragment ions [61]); an ion trap CID scan (NCE 35%, maximal injection time of 40 ms with an AGC of 200%) and a stepped collision energy HCD scan (using NCE 30%, 35%, 45% with a maximal injection time of 250 ms with an AGC of 500% and a resolution of

30,000). For FAIMS experiments, static FAIMS analysis was undertaken using CVs of -20; -30; -40; -50; -60; -70; -80 and -90 while stepped FAIMS experiments were undertaken switching between the CVs -25; -35 and -45. All experiments were collected in triplicate.

Data Analysis - Glycopeptide identification: Raw data files were batch processed using Byonic v3.5.3 (Protein Metrics Inc. [66]) with the proteome databases denoted within Table 1. Data was searched on a desktop with two 3.00GHz Intel Xeon Gold 6148 processors, a 2TB SSD and 128 GB of RAM using a maximum of 16 cores for a given search. For all searches, a semi-tryptic N-ragged specificity was set, and a maximum of two missed cleavage events allowed. Carbamidomethyl was set as a fixed modification of cystine while oxidation of methionine was included as a variable modification. A maximum mass precursor tolerance of 5 ppm was allowed while a mass tolerance of up to 10 ppm was set for HCD fragments and 20 ppm for EThcD fragments. *Burkholderia* species searches were conducted allowing two O-linked glycans: Hex-HexNAc-HexNAc (elemental composition: C₂₂O₁₅H₃₆N₂, mass: 568.2115) and Suc-Hex-HexNAc-HexNAc (elemental composition: C₂₆O₁₈H₄₀N₂, mass: 668.2276), while *C. fetus fetus* searches were conducted allowing two N-linked glycans: Hex-HexNAc₄-diNAcBac (elemental composition: C₄₈O₂₉H₇₈N₆, mass: 1202.4813) and HexNAc₅-diNAcBac (elemental composition: C₅₀O₂₉H₈₁N₇, mass: 1243.5078) where diNAcBac is the bacterial specific sugar 2,4-diacetamido-2,4,6-trideoxyglucopyranose [64]. To ensure high data quality, technical replicates were combined using R (<https://www.r-project.org/>) and only glycopeptides with a Byonic score >300 were used for further analysis. This cut-off is in line with previous reports highlighting that score thresholds greater than at least 150 are required for robust glycopeptide assignments with Byonic [28, 67].

It should be noted that a score threshold of above 300 resulted in false discovery rates of less than 1% for all combined datasets.

Data Analysis - iBAQ analysis: To assess the relative abundance of proteins, iBAQ analysis was undertaken. Unenriched proteome datasets were processed using MaxQuant v1.6.3.4 [68] searching against the reference *B. cenocepacia* strain J2315. Carbamidomethylation of cysteine was set as a fixed modification while oxidation of methionine and acetylation of protein N-termini were allowed as variable modifications. An enzyme specificity of “trypsin/P” was set with a maximum of 2 missed cleavage events. The resulting outputs were processed in the Perseus v1.5.0.9 [69] analysis environment to remove reverse matches and common protein contaminants prior to further analysis.

Peptide analysis, data visualization and data sharing: The previously published ZIC-HILIC glycopeptide datasets [44] were used to compare the differences in glycopeptides observed with and without ZIC-HILIC enrichment. It should be noted that the proteome digests within this work are the same digests used for ZIC-HILIC enrichments in [44]. Analysis and visualization was undertaken using R (<https://www.r-project.org/>) with the glycopeptide physicochemical properties determined using the Peptides package [70] and data visualized using ggplot2 [71] and ggpubr packages. The aliphatic index within the Peptides package is calculated by assessing the relative volume occupied by aliphatic side chains within a peptide [72]. All mass spectrometry proteomics data (Raw data files, Byonic search outputs and R Scripts) have been deposited into the PRIDE ProteomeXchange Consortium repository [73, 74] with the dataset identifier:

242 PXD020442 and are accessible with the reviewer account details: Username:
243 reviewer56784@ebi.ac.uk Password: aoATFQMQ.

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

RESULTS

Discrete glycopeptide populations are observed within whole cell digests compared to ZIC-HILIC enriched samples

Recently we noted glycopeptides could be readily identified without enrichment across a range of bacterial proteomes [23, 78]. Given this, we sought to compare the glycopeptides observed with and without enrichment for bacterial species focusing on members of the Burkholderia genus such as *B. cenocepacia* J2315. Within whole cell digests of *B. cenocepacia* J2315 we find that only ~2% of all high-confidence PSMs (Byonic™ scores >300) correspond to glycopeptides (Supplementary Figure 1A, Supplementary Table 1). Although a low proportion of all PSMs this still corresponds to hundreds of individual glycopeptide PSMs identified within analytical runs (Supplementary Figure 1B, Supplementary Table 1). Intuitively, we assumed that within these whole cell digests of *B. cenocepacia* J2315, the observed glycopeptides would correspond to previously identified glycopeptides within ZIC-HILIC enrichments, yet this was not the case. Surprisingly, the majority of glycopeptides identified within whole cell digests corresponded to novel glycoproteins/glycopeptides (Figure 1A and B), with manual inspection of previously unreported glycopeptides supporting the correctness of these assignments (Supplementary Figure 2A-H). These results demonstrate that different subsets of the *B. cenocepacia* J2315 glycoproteome are observed with and without ZIC-HILIC enrichment.

To understand the differences in the glycopeptides identified with and without ZIC-HILIC enrichment we first compared the relative abundances of glycoproteins using iBAQ values derived from whole cell digests (Figure 1C, Supplementary Table 2). We reasoned that the

dramatic difference in the glycopeptide profiles may be linked to glycoprotein abundances with ZIC-HILIC enrichment potentially accessing glycopeptides from lower abundance glycoproteins. Although the iBAQ values of glycoproteins observed within enriched samples are slightly lower ($P = 0.0036$; mean iBAQ value of glycoproteins identified in enriched samples = 28.045 compared to the mean iBAQ value of glycoproteins identified within unenriched samples = 28.918, Supplementary Table 3), the difference in the abundance is modest (Figure 1C), suggesting that the difference in the *B. cenocepacia* glycopeptides profiles observed between samples is independent of protein abundance. As previous reports have suggested that peptide sequences can influence ZIC-HILIC glycopeptide enrichment [23, 75] we examined the physiochemical properties of the peptide sequences revealing marked differences in the glycopeptides observed between samples (Figure 1D). We found that glycopeptides observed within whole cell digests are both more aliphatic and typically shorter in length, while the majority of ZIC-HILIC enriched glycopeptides are both longer and less aliphatic ($P = 5.10 \times 10^{-15}$ and 6.68×10^{-15} aliphaticity and length respectively, Supplementary Table 4). Combined these results support that differences in the physiochemical properties of glycopeptides influences their ability to be enriched by ZIC-HILIC enrichment.

ZIC-HILIC incompatible glycopeptides are readily accessible using FAIMS fractionation

The findings that ZIC-HILIC enrichment provides access to only a subset of Burkholderia glycopeptides suggests that alternative approaches are required for comprehensive analysis of Burkholderia glycoproteomes. As glycopeptides were readily identified within whole cell digests of *B. cenocepacia*, we reasoned that an orthogonal fractionation approach may identify

previously overlooked glycopeptides. Since FAIMS enables gas-phase fractionation based on charge/collision cross section [49, 50], we examined the charge state distributions of glycopeptides identified from whole cell digests compared to ZIC-HILIC enrichments (Figure 2A and B). We noted that glycopeptides observed within the whole cell digest were predominantly of low charge states ($<+3$). As previous work [50] has shown that +3 and +2 peptides are identified across CVs of -20 to -120 with apexes at -80 and -45 respectively this suggested low FAIMS CVs would enable enrichment of potential low charge state glycopeptides. To assess this, we assayed the ability of *B. cenocepacia* glycopeptides to be identified under differing FAIMS CVs from -20 to -90 (Figure 2C, Supplementary Table 5 and 6). Consistent with our hypothesis we found that FAIMS CVs of less than -50 readily enabled the isolation of glycopeptides (Figure 2C, Supplementary Figure 3A). Interestingly, we also found that within individual analytical runs (denoted with dots in Figure 2C) with FAIMS CVs of -20, -30 and -40 more unique glycopeptides were identified per run than analysis without FAIMS. However, when the results of triplicate runs were combined (denoted by the bars in Figure 2C), only the FAIMS CV -30 outperformed analysis without FAIMS (Figure 2C). Similar to whole proteome digests, FAIMS-fractionated glycopeptides typically corresponded to novel glycoproteins/glycopeptides that were not captured by ZIC-HILIC enrichments (Figure 2D and E) and were again shorter and more aliphatic than ZIC-HILIC enriched glycopeptides (Supplementary Figure 3B). Within FAIMS CVs we noted that, consistent with previous reports [50], the overlap between 10 CV fractions was modest with a large proportion of glycopeptides unique to a single CV (Figure 3A and B). Taken together these results suggest that low FAIMS CVs provide access to ZIC-HILIC inaccessible *B. cenocepacia* glycopeptides yet the coverage of the glycoproteome within a single FAIMS CV is limited.

Using stepped FAIMS improves the glycoproteome coverage of Burkholderia species

Since individual FAIMS CVs provided access to only a subset of *B. cenocepacia* glycopeptides, we reasoned that using stepped FAIMS may improve glycopeptide identifications. Comparing the performance of stepped FAIMS, unenriched and ZIC-HILIC enriched digests we find that stepped FAIMS leads to a >30% increase in unique glycopeptides compared to unenriched samples (Figure 4A, Supplementary Table 7). As with static FAIMS CVs this improvement was predominantly within glycopeptides not previously observed within ZIC-HILIC enrichments with >60% of glycopeptides unique to stepped FAIMS (Figure 4B). Interestingly, we find that the glycopeptides identified in both stepped FAIMS and unenriched samples provide access to unique glycopeptide subsets compared to ZIC-HILIC enrichments (Figure 4C). Analysis of the physiochemical properties of glycopeptides observed using each approach highlights that the glycopeptides observed within *B. cenocepacia* with stepped FAIMS and unenriched samples are largely similar, favouring shorter and more aliphatic glycopeptides compared to the larger and less aliphatic glycopeptides observed with ZIC-HILIC enrichment (Figure 4D, Supplementary Table 8). The modest overlap of unique glycopeptides observed with stepped FAIMS and analysis without FAIMS fractionation illustrates that the glycoproteome of *B. cenocepacia* is larger than initially thought [31, 44], and that ZIC-HILIC enrichment fails to capture a significant proportion of the Burkholderia glycoproteome.

To further confirm this bias within ZIC-HILIC enrichments we assessed two additional Burkholderia species recently assessed using ZIC-HILIC based enrichment: *B. dolosa* and *B. ubonensis* [44]. Consistent with our observations within *B. cenocepacia* both stepped FAIMS and

unenriched samples revealed multiple unique glycopeptides/glycoproteins absent within ZIC-HILIC enrichments (Figure 5A to B; Supplementary Tables 9-13). As with *B. cenocepacia*, these newly observed glycopeptides in each Burkholderia species are typically shorter and/or more aliphatic than glycopeptides observed within ZIC-HILIC enrichments (Supplementary Figure 4). These results demonstrate that at least 30% of the glycoproteome of Burkholderia species are not observed within ZIC-HILIC enrichments, making orthogonal approaches such as stepped FAIMS valuable tools to access uncharacterised regions of Burkholderia glycoproteomes.

FAIMS fractionation improves the identification of bacterial N-linked glycoproteomes

Finally, as FAIMS improved the identification of glycopeptides within Burkholderia glycoproteomes, we wondered if other bacterial glycoproteomes would also benefit from FAIMS fractionation. To assess this we used *C. fetus subsp. fetus* NCTC 10842 which is known to N-glycosylate proteins with two glycans composed of β -GlcNAc- α 1,3-[GlcNAc1,6-]GlcNAc- α 1,4-GalNAc- α 1,4-GalNAc- α 1,3-diNAcBac and β -GlcNAc- α 1,3-[Glc1,6-]GlcNAc- α 1,4-GalNAc- α 1,4-GalNAc- α 1,4-diNAcBac [64]. As with the trisaccharide decorated glycopeptides of Burkholderia, we noted a marked enrichment of *C. fetus subsp. fetus* glycopeptides within low CVs FAIMS fractions compared to matching analytical runs without FAIMS (Figure 6A, Supplementary Table 14 and 15). As observed within Burkholderia glycoproteomes, the use of stepped FAIMS improved the detection of unique glycopeptides compared to unenriched samples with matching analytical gradients (Figure 6B, Supplementary Table 16). Excitingly, we find that stepped FAIMS enabled the identification of a large proportion of previously identified glycoproteins and glycopeptides from ZIC-HILIC enrichments (Figure 6C, Supplementary Table 17) while requiring

significantly less material (2µg per stepped FAIMS run compared to 300µg per ZIC-HILIC enrichment). These results demonstrate that large glycan-containing glycopeptides are enriched at low FAIMS CVs in a similar manner to glycopeptides decorated with shorter glycans, and this supports the general utilisation of FAIMS fractionation for isolating bacterial glycopeptides.

Discussion

Over the last decade, glycopeptide-based enrichment approaches have increasingly been used to track and characterise microbial glycosylation events [42, 44, 64, 76]. Although enrichment strategies enable high-throughput analysis of glycosylation, we demonstrate here that hydrophilic enrichment provides only partial access to some microbial glycoproteomes. In light of these findings we evaluated the application of FAIMS fractionation using the FAIMS Pro interface [49-51] as an orthogonal approach for probing bacterial glycosylation. We find that for bacterial *N*- and *O*-linked glycans, low FAIMS CVs enables the enrichment of glycopeptides from whole proteome samples. By utilising a stepped FAIMS approach in concert with direct analysis of whole cell digests we show that the glycoproteomes of three Burkholderia species are >30% larger than initially thought and that the majority of the known glycoproteome of *C. fetus subsp. fetus* can be directly monitored from whole proteome digests.

To date, multiple studies have shown that glycosylation events can be readily detected within proteome samples using variable modification searches [77, 78] or open searching approaches [79-81]. These previous findings and our work here highlight that a wealth of glycoproteome information can be gained from unenriched samples, despite the known issue of glycopeptide

ion suppression [12]. Experimentally, it has been shown that glycopeptide signals can be suppressed up to ten-fold compared to unmodified forms of the identical peptide [82] yet this issue is addressable. As current state of the art MS instruments enable the identification of potential glycopeptides by the presence of known sugar oxonium ions [65] this allows glycopeptide spectra to be re-collected with altered settings generating high-quality MS/MS data for even low abundance signals. Since FAIMS fractionation occurs post ionisation, it should be noted that ion suppression is not prevented but is likely being mitigated by the improvement in the signal to noise of analytes. As the amount of ion suppression is impacted by the complexity of the attached glycans of glycopeptides [82], it is reasonable to assume that some glycoproteomes will be more amenable to FAIMS analysis than others. Our work here demonstrates that glycopeptides of Burkholderia species and *C. fetus subsp. fetus* appear readily amenable to FAIMS enrichment, although further work will be required to establish general trends for specific classes of glycosylation.

The ability to enrich glycopeptides independent of off-line glycan-centric enrichment strategies opens up exciting possibilities for the study of previously recalcitrant glycoproteomes. Previous comparative genomic [83] as well as *ex vivo* studies [84] have identified a number of potential bacterial glycosylation systems for which the native substrates and glycans have yet to be defined. For example, in *Vibrio cholerae* a functional *O*-linked glycosylation system has been identified [84] and shown to be required for cholera toxin secretion [85], yet the glycoproteins of this system are still unknown. Due to the diversity of glycans used for bacterial protein glycosylation [29, 30, 35, 41, 42, 59, 64, 76, 86, 87] it is not surprising that a single approach such

as ZIC-HILIC enrichment is inapplicable to all glycosylation systems or substrates within a single glycoproteome. Indeed, within this study we note that the bias within Burkholderia ZIC-HILIC enrichments does not appear to occur within ZIC-HILIC enrichments of *C. fetus subsp. fetus* (Supplementary Figure 5, Supplementary Table 17). This difference is likely due to the larger more hydrophilic glycans of *C. fetus subsp. fetus* which are retained more effectively during ZIC-HILIC enrichment. Thus, the diversity within microbial glycosylation systems means that multiple approaches are likely needed, with FAIMS providing an orthogonal approach for undertaking high throughput microbial glycoproteomic analysis.

The observation that FAIMS enables the enrichment of glycopeptides at a proteome scale has not explicitly been shown before although multiple IMS studies have highlighted this potential. In 2013 the Hill group demonstrated the ability of IMS to separate glycosylated and non-glycosylated components based on differences in collisional cross section [62], noting that this could provide a novel strategy for high throughput glycopeptide identification. This concept has been echoed in more recent studies which have shown that the ion mobility of low charge state glycopeptides can be dramatically impacted by the presence of glycans [60]. Our finding that FAIMS enables access to novel glycopeptides is consistent with previous studies of single glycoproteins [59] and we demonstrate that this can be extended to the analysis of glycoproteomes. A clear benefit of the recent FAIMS Pro interface is the rapid CV switching enabling multiple CVs to be accessible within a single analytical framework. We find stepped FAIMS approaches improves the detection of unique glycopeptides (Figure 4A and 6A) yet it should be noted that the optimal stepping regimes for glycopeptide identification is still unclear.

It is plausible that further improvements in glycoproteome coverage may be possible by optimising the CV switching to ensure minimal overlap between steps, while still focusing on the low CV region where glycopeptides are enriched. Our results provide support for the notion that FAIMS can improve our ability to study previously overlooked regions of glycoproteomes, warranting further investigation of these methods for the field of glycoproteomics.

Figures

Figure 1: Discrete glycopeptides are observed within unenriched and ZIC-HILIC enriched bacterial samples. A and B) The overlap of glycoproteins (A) and glycopeptides (B) identified without and with ZIC-HILIC enrichment demonstrates the majority of glycopeptides/glycoproteins observed within unenriched samples are unique to those observed within ZIC-HILIC enrichments. **C)** iBAQ measurements demonstrate only modest differences in the relative abundance of glycoproteins observed with and without ZIC-HILIC enrichment. The mean iBAQ values for each group are denoted by the dotted lines. **D)** The physiochemical properties of glycopeptides observed with and without ZIC-HILIC enrichment demonstrates glycopeptides observed without enrichment are shorter and more aliphatic than glycopeptides observed with ZIC-HILIC enrichment.

Figure 2. FAIMS fractionation enables the enrichment of glycopeptides within *B. cenocepacia* J2315. A and B) Charge state distributions of glycopeptides observed without ZIC-HILIC enrichment (A) and glycopeptides observed with ZIC-HILIC enrichment (B). **C)** 2 hr analytical gradients of FAIMS fractionation at different CVs demonstrate glycopeptides are enriched at low FAIMS CVs compared to 2 hr analytical gradients of analysis without FAIMS fractionation. Red dots denote the numbers of unique glycopeptides in individual replicates at each FAIMS CV. Blue dots denote the numbers of unique glycopeptides in individual replicates from analysis without FAIMS. Bars represent the total number of unique glycopeptides identified under each condition. **D and E)** FAIMS fractionated samples enable the identification of glycoproteins (D) and glycopeptides (E) not observed within ZIC-HILIC enrichments.

Figure 3. FAIMS CV fractions provide access to unique glycopeptide pools. A) Analysis of static CVs observed to contain glycopeptides demonstrates that a large proportion of glycopeptides are unique to a single CV fraction. **B)** Pie chart of unique glycopeptides observed within static FAIMS CV fractions. The majority of all glycopeptides are observed in CVs below -50.

Figure 4. Stepped FAIMS Improves the identification of glycopeptides within *B. cenocepacia*. A) Comparison of 3 hr analytical gradients of ZIC-HILIC enrichments, stepped FAIMS and unenriched digests. Stepped FAIMS improves the identification of unique glycopeptides compared to unenriched samples. Individual replicates are denoted by dots and the combined number of unique glycopeptides within an approach are denoted by bars **B)** Comparison of the overlap in glycoproteins and glycopeptides identified with stepped FAIMS compared to ZIC-HILIC enrichments reveals stepped FAIMS enables access to glycopeptides/glycoproteins not observed within ZIC-HILIC enrichments. **C)** Overlap of glycopeptides observed within ZIC-HILIC enrichments, stepped FAIMS and analysis without FAIMS fractionation identifies that unique pools of glycopeptides are identified within each approach. **D)** The physiochemical properties of glycopeptides demonstrate that stepped FAIMS and analysis of unenriched digests allows the identification of typically more aliphatic and shorter glycopeptides.

Figure 5. Comparison of glycoproteome coverage of Burkholderia species with ZIC-HILIC enrichment, Stepped FAIMS and analysis of unenriched samples. Analysis of *B. dolosa* and *B. ubonensis* demonstrate that multiple novel glycopeptides/glycoproteins are readily identified using stepped FAIMS and unenriched whole proteome samples.

Figure 6. FAIMS fractionation of *C. fetus subsp. fetus* N-linked glycopeptides. A) Static FAIMS CVs demonstrate glycopeptides are enriched at low FAIMS CVs compared to analysis without FAIMS fractionation within 2hr analytical runs. Red dots denote the numbers of unique glycopeptides in individual replicates in each FAIMS CV. Blue dots denote the numbers of unique glycopeptides in individual replicates from analysis without FAIMS. Bars represent the total number of unique glycopeptides identified under each condition. **B)** Comparison of 3hr analytical runs of ZIC-HILIC enrichment, stepped FAIMS and analysis of unenriched whole proteome samples demonstrates an improvement in unique glycopeptides identified with stepped FAIMS compared to analysis without FAIMS fractionation. **C)** Comparison of the overlap in glycoproteins and glycopeptides identified with stepped FAIMS compared to ZIC-HILIC enrichments. At both the glycopeptide and glycoprotein level, stepped FAIMS enables identification of >60% of the known glycoproteome.

Table 1. Strain list

Strains	Source	Reference	Proteome database
<i>B. cenocepacia</i> LMG 16656 / J2315	Human clinical specimen, United Kingdom, 1989	[88]	Uniprot database: UP000001035
<i>B. dolosa</i> AU0158	Human clinical specimen, USA unknown	[89]	Uniprot database: UP000032886
<i>B. ubonensis</i> MSMB22	Soil isolate, Australia, 2001	[89]	Burkholderia Genome Database [90], Strain number: 3404
<i>C. fetus subsp. fetus</i> NCTC 10842	Brain of sheep fetus, France, 1952	[91]	Uniprot database: UP000001035

ASSOCIATED CONTENT

Supporting Information

Supplementary document: Containing Supplementary figures 1 to 5

Supplementary Figure 1: Identification of glycopeptides within *B. cenocepacia* J2315 whole proteome digests. A) Pie charts showing all unique PSMs identified with scores above 300 B) Number of glycopeptides PSMs (Byonic Score >300) identified within each unenriched sample.

Supplementary Figure 2: Manual assessment of glycopeptides identified with *B. cenocepacia* J2315 whole proteome digests. HCD and ETD fragmentation of novel glycoproteins A-B) B4E7L5; C-D) B4EBZ3; E-F) B4E732 and G-H) B4EB92.

Supplementary Figure 3: Analysis of *B. cenocepacia* glycopeptide identified within static FAIMS fractions. A) Scatter plot of CV vs m/z of glycopeptide identified. B) Peptide properties of glycopeptide identified across static FAIMS CVs vs HILIC enriched glycopeptides

Supplementary Figure 4: Comparisons of glycopeptide length vs aliphaticity across the Burkholderia species *B. dolosa* and *B. ubonensis*. Within strains examined difference in aliphaticity properties or length are observed from glycopeptides identified within stepped FAIMS and unenriched samples compared to ZIC-HILIC enrichments.

Supplementary Figure 5: Comparisons of glycopeptide length vs aliphaticity within *C. fetus* subsp. *fetus* glycopeptide. No differences in the aliphatic properties or length are observed from glycopeptides identified within stepped FAIMS and unenriched samples compared to ZIC-HILIC enrichments.

Supplementary Tables: All Supplementary tables are provided within a single zip file

Supplementary table 1: Combined Byonic searches of *B. cenocepacia* J2315 whole proteome digests. The combined Byonic results of three technical replicates of *B. cenocepacia* J2315 showing all glycopeptides identified with a Byonic score over 300.

Supplementary table 2: Maxquant analysis of *B. cenocepacia* J2315. Maxquant analysis of three technical replicates of *B. cenocepacia* J2315 whole proteome digests.

Supplementary table 3: iBAQ measurements of *B. cenocepacia* J2315. iBAQ values for all glycoproteins identified within enriched and unenriched samples of *B. cenocepacia* J2315 whole proteome digests.

Supplementary table 4: Peptide properties of *B. cenocepacia* J2315 glycopeptides identified within and without ZIC-HILIC enrichments. For each glycopeptide identified within unenriched and ZIC-HILIC enriched samples the length and aliphatic index is provided.

Supplementary table 5: Combined Byonic searches of *B. cenocepacia* J2315 without FAIMS. The combined Byonic results of three technical replicates of *B. cenocepacia* J2315 showing all glycopeptides identified with a Byonic score over 300.

Supplementary table 6: Combined Byonic searches of *B. cenocepacia* J2315 with static FAIMS CVs. The combined Byonic results of each CV from -20 to -90 (three technical replicates of each) for *B. cenocepacia* J2315 showing all glycopeptides identified with a Byonic score over 300.

Supplementary table 7: Combined Byonic searches of *B. cenocepacia* J2315 with stepped FAIMS. The combined Byonic results for three technical replicates of *B. cenocepacia* J2315 with stepped FAIMS showing all glycopeptides identified with a Byonic score over 300.

Supplementary table 8: Peptide properties of glycopeptides identified within *B. cenocepacia* J2315 across analysis approaches. For each glycopeptide identified within stepped FAIMS, unenriched and ZIC-HILIC enriched samples the length and aliphatic index are provided.

Supplementary table 9: Combined Byonic searches of *B. ubonensis* MSMB22 with stepped FAIMS. The combined Byonic results for three technical replicates of *B. ubonensis* MSMB22 with stepped FAIMS showing all glycopeptides identified with a Byonic score over 300.

Supplementary table 10: Combined Byonic searches of *B. ubonensis* MSMB22 without FAIMS. The combined Byonic results of three technical replicates of *B. ubonensis* MSMB22 showing all glycopeptides identified with a Byonic score over 300.

Supplementary table 11: Combined Byonic searches of *B. dolosa* AU0158 with stepped FAIMS. The combined Byonic results for three technical replicates of *B. dolosa* AU0158 with stepped FAIMS showing all glycopeptides identified with a Byonic score over 300.

Supplementary table 12: Combined Byonic searches of *B. dolosa* AU0158 without FAIMS. The combined Byonic results of three technical replicates of *B. dolosa* AU0158 showing all glycopeptides identified with a Byonic score over 300.

Supplementary table 13: Peptide properties of glycopeptides identified within *B. ubonensis* MSMB22 and *B. dolosa* AU0158 across analysis approaches. For each glycopeptide identified within stepped FAIMS, unenriched and ZIC-HILIC enriched samples the length and Aliphatic index is provided.

Supplementary table 14: Combined Byonic searches of *C. fetus subsp. fetus* NCTC 10842 with static FAIMS CVs. The combined Byonic results of each CV from -20 to -90 (three technical replicates) for *C. fetus subsp. fetus* NCTC 10842 showing all glycopeptides identified with a Byonic score over 300.

Supplementary table 15: Combined Byonic searches of *C. fetus subsp. fetus* NCTC 10842 without FAIMS. The combined Byonic results of three technical replicates of *C. fetus subsp. fetus* NCTC 10842 showing all glycopeptides identified with a Byonic score over 300.

Supplementary table 16: Combined Byonic searches of *C. fetus subsp. fetus* NCTC 10842 with stepped FAIMS. The combined Byonic results for three technical replicates of *C. fetus subsp. fetus* NCTC 10842 with stepped FAIMS showing all glycopeptides identified with a Byonic score over 300.

Supplementary table 17: Peptide properties of glycopeptides identified within *C. fetus subsp. fetus* NCTC 10842 across analysis approaches. For each glycopeptide identified within stepped FAIMS, unenriched and ZIC-HILIC enriched samples the length and Aliphatic index is provided.

AUTHOR INFORMATION

Corresponding Author

Dr Nichollas E. Scott

Email: Nichollas.scott@unimelb.edu.au

Address: Department of Microbiology and Immunology, University of Melbourne at the Peter Doherty Institute for Infection and Immunity, Melbourne 3000, Australia

Author contributions

A.R.A.I prepared all sample for MS analysis. C.S.A, S.N and N.A.W maintained MS instrumentation and provided access to critical resources. L.E.B assisted in manuscript preparation. N.E.S conceived, designed all experiments, performed analysis, and undertook the preparation of the manuscript.

Funding Sources

This work was supported by a National Health and Medical Research Council of Australia (NHMRC) project grant awarded to NES (APP1100164).

ACKNOWLEDGEMENTS

We thank the Melbourne Mass Spectrometry and Proteomics Facility of The Bio21 Molecular Science and Biotechnology Institute for access to MS instrumentation and Byonic. We would like to thank Christine Szymanski and Justin Duma for the kind gift of *Campylobacter fetus fetus* NCTC

10842 lysates. We thank Ben Parker for feedback on the manuscript. We would also like to thank Morten Thaysen-Andersen and Daniel Kolarich for their fruitful discussions and highlighting that ZIC-HILIC enrichment was likely missing numerous Burkholderia glycopeptides. These conversations were instrumental in leading to this work.

ABBREVIATIONS

Hydrophilic Interaction Liquid Chromatography (HILIC); Mass spectrometry (MS); Compensation voltage (CV); Zwitterionic hydrophilic interaction liquid chromatography (ZIC-HILIC); galactose (Gal); Glucose (Glc); N-Acetylgalactosamine (GalNAc); N-Acetylglucosamine (GlcNAc); Succinyl (Suc); 2,4-diacetamido-2,4,6 trideoxyglucopyranose (diNAcBac); N-acetylhexoseamine (HexNAc); Hexose (Hex); Ion Mobility Spectrometry (IMS); Field Asymmetric Waveform IMS (FAIMS); Luria Bertani (LB); Phosphate-buffered saline (PBS); Sodium dodecyl sulfate (SDS); Normalized collisional energy (NCE); Tris(hydroxymethyl)aminomethane (Tris); Trifluoroacetic acid (TFA); Dithiothreitol (DTT); Higher-energy collision dissociation (HCD); Collision-induced dissociation (CID); Electron-transfer/higher-energy collision dissociation (ET_hCD); Automatic Gain Control (AGC); Dimethyl sulfoxide (DMSO); intensity-based absolute quantification (iBAQ); Peptide spectrum matches (PSMs)

References

1. Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, et al., editors. Essentials of Glycobiology. 3rd ed: Cold Spring Harbor; 2015.
2. Yu A, Zhao J, Peng W, Banazadeh A, Williamson SD, Goli M, et al. Advances in mass spectrometry-based glycoproteomics. Electrophoresis. 2018;39(24):3104-22. Epub 2018/09/12. doi: 10.1002/elps.201800272. PubMed PMID: 30203847; PubMed Central PMCID: PMCPMC6375712.
3. Xiao H, Sun F, Suttapitugsakul S, Wu R. Global and site-specific analysis of protein glycosylation in complex biological systems with Mass Spectrometry. Mass Spectrom Rev. 2019;38(4-5):356-79. Epub 2019/01/04. doi: 10.1002/mas.21586. PubMed PMID: 30605224; PubMed Central PMCID: PMCPMC6610820.
4. Thaysen-Andersen M, Packer NH, Schulz BL. Maturing Glycoproteomics Technologies Provide Unique Structural Insights into the N-glycoproteome and Its Regulation in Health and Disease. Molecular & cellular proteomics : MCP. 2016;15(6):1773-90. doi: 10.1074/mcp.O115.057638. PubMed PMID: 26929216; PubMed Central PMCID: PMCPMC5083109.
5. Choi J, Wagner LJS, Timmermans S, Malaker SA, Schumann B, Gray MA, et al. Engineering Orthogonal Polypeptide GalNAc-Transferase and UDP-Sugar Pairs. J Am Chem Soc. 2019;141(34):13442-53. Epub 2019/08/03. doi: 10.1021/jacs.9b04695. PubMed PMID: 31373799; PubMed Central PMCID: PMCPMC6813768.
6. Schumann B, Malaker SA, Wisnovsky SP, Debets MF, Agbay AJ, Fernandez D, et al. Bump-and-Hole Engineering Identifies Specific Substrates of Glycosyltransferases in Living Cells. Mol Cell. 2020. Epub 2020/04/24. doi: 10.1016/j.molcel.2020.03.030. PubMed PMID: 32325029.
7. Steentoft C, Vakhrushev SY, Vester-Christensen MB, Schjoldager KT, Kong Y, Bennett EP, et al. Mining the O-glycoproteome using zinc-finger nuclease-glycoengineered SimpleCell lines. Nature methods. 2011;8(11):977-82. Epub 2011/10/11. doi: 10.1038/nmeth.1731. PubMed PMID: 21983924.
8. Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, Schjoldager KT, et al. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. EMBO J. 2013;32(10):1478-88. Epub 2013/04/16. doi: 10.1038/emboj.2013.79. PubMed PMID: 23584533; PubMed Central PMCID: PMCPMC3655468.
9. Vester-Christensen MB, Halim A, Joshi HJ, Steentoft C, Bennett EP, Levery SB, et al. Mining the O-mannose glycoproteome reveals cadherins as major O-mannosylated glycoproteins. Proceedings of the National Academy of Sciences of the United States of America. 2013;110(52):21018-23. Epub 2013/10/09. doi: 10.1073/pnas.1313446110. PubMed PMID: 24101494; PubMed Central PMCID: PMCPMC3876253.
10. Halim A, Larsen IS, Neubert P, Joshi HJ, Petersen BL, Vakhrushev SY, et al. Discovery of a nucleocytoplasmic O-mannose glycoproteome in yeast. Proceedings of the National Academy of Sciences of the United States of America. 2015;112(51):15648-53. doi: 10.1073/pnas.1511743112. PubMed PMID: 26644575; PubMed Central PMCID: PMCPMC4697373.
11. Woo CM, Lund PJ, Huang AC, Davis MM, Bertozzi CR, Pitteri SJ. Mapping and Quantification of Over 2000 O-linked Glycopeptides in Activated Human T Cells with Isotope-Targeted Glycoproteomics (Isotag). Molecular & cellular proteomics : MCP. 2018;17(4):764-75.

- Epub 2018/01/21. doi: 10.1074/mcp.RA117.000261. PubMed PMID: 29351928; PubMed Central PMCID: PMC5880114.
12. Wohlgemuth J, Karas M, Eichhorn T, Hendriks R, Andrecht S. Quantitative site-specific analysis of protein glycosylation by LC-MS using different glycopeptide-enrichment strategies. *Anal Biochem.* 2009;395(2):178-88. Epub 2009/08/25. doi: 10.1016/j.ab.2009.08.023. PubMed PMID: 19699707.
13. Hart-Smith G, Raftery MJ. Detection and characterization of low abundance glycopeptides via higher-energy C-trap dissociation and orbitrap mass analysis. *J Am Soc Mass Spectrom.* 2012;23(1):124-40. Epub 2011/11/16. doi: 10.1007/s13361-011-0273-y. PubMed PMID: 22083589.
14. Chen CC, Su WC, Huang BY, Chen YJ, Tai HC, Obena RP. Interaction modes and approaches to glycopeptide and glycoprotein enrichment. *Analyst.* 2014;139(4):688-704. Epub 2013/12/18. doi: 10.1039/c3an01813j. PubMed PMID: 24336240.
15. Zielinska DF, Gnad F, Schropp K, Wisniewski JR, Mann M. Mapping N-glycosylation sites across seven evolutionarily distant species reveals a divergent substrate proteome despite a common core machinery. *Mol Cell.* 2012;46(4):542-8. Epub 2012/05/29. doi: 10.1016/j.molcel.2012.04.031. PubMed PMID: 22633491.
16. Zielinska DF, Gnad F, Wisniewski JR, Mann M. Precision mapping of an in vivo N-glycoproteome reveals rigid topological and sequence constraints. *Cell.* 2010;141(5):897-907. Epub 2010/06/01. doi: 10.1016/j.cell.2010.04.012. PubMed PMID: 20510933.
17. Trinidad JC, Schoepfer R, Burlingame AL, Medzihradszky KF. N- and O-glycosylation in the murine synaptosome. *Molecular & cellular proteomics : MCP.* 2013;12(12):3474-88. Epub 2013/07/03. doi: 10.1074/mcp.M113.030007. PubMed PMID: 23816992; PubMed Central PMCID: PMC3861701.
18. Xiao H, Chen W, Smeekens JM, Wu R. An enrichment method based on synergistic and reversible covalent interactions for large-scale analysis of glycoproteins. *Nat Commun.* 2018;9(1):1692. Epub 2018/04/29. doi: 10.1038/s41467-018-04081-3. PubMed PMID: 29703890; PubMed Central PMCID: PMC5923262.
19. Nilsson J, Ruetschi U, Halim A, Hesse C, Carlsohn E, Brinkmalm G, et al. Enrichment of glycopeptides for glycan structure and attachment site identification. *Nature methods.* 2009;6(11):809-11. Epub 2009/10/20. doi: 10.1038/nmeth.1392. PubMed PMID: 19838169.
20. Hagglund P, Bunkenborg J, Elortza F, Jensen ON, Roepstorff P. A new strategy for identification of N-glycosylated proteins and unambiguous assignment of their glycosylation sites using HILIC enrichment and partial deglycosylation. *Journal of proteome research.* 2004;3(3):556-66. Epub 2004/07/16. PubMed PMID: 15253437.
21. Jandera P. Stationary and mobile phases in hydrophilic interaction chromatography: a review. *Anal Chim Acta.* 2011;692(1-2):1-25. Epub 2011/04/20. doi: 10.1016/j.aca.2011.02.047. PubMed PMID: 21501708.
22. Jandera P, Janas P. Recent advances in stationary phases and understanding of retention in hydrophilic interaction chromatography. A review. *Anal Chim Acta.* 2017;967:12-32. Epub 2017/04/10. doi: 10.1016/j.aca.2017.01.060. PubMed PMID: 28390482.
23. Mysling S, Palmisano G, Hojrup P, Thaysen-Andersen M. Utilizing ion-pairing hydrophilic interaction chromatography solid phase extraction for efficient glycopeptide enrichment in glycoproteomics. *Anal Chem.* 2010;82(13):5598-609. Epub 2010/06/12. doi: 10.1021/ac100530w. PubMed PMID: 20536156.

24. Hagglund P, Matthiesen R, Elortza F, Hojrup P, Roepstorff P, Jensen ON, et al. An enzymatic deglycosylation scheme enabling identification of core fucosylated N-glycans and O-glycosylation site mapping of human plasma proteins. *Journal of proteome research*. 2007;6(8):3021-31. Epub 2007/07/20. doi: 10.1021/pr0700605. PubMed PMID: 17636988.
25. Parker BL, Thaysen-Andersen M, Solis N, Scott NE, Larsen MR, Graham ME, et al. Site-specific glycan-peptide analysis for determination of N-glycoproteome heterogeneity. *Journal of proteome research*. 2013;12(12):5791-800. Epub 2013/10/05. doi: 10.1021/pr400783j. PubMed PMID: 24090084.
26. Parker BL, Thaysen-Andersen M, Fazakerley DJ, Holliday M, Packer NH, James DE. Terminal Galactosylation and Sialylation Switching on Membrane Glycoproteins upon TNF-Alpha-Induced Insulin Resistance in Adipocytes. *Molecular & cellular proteomics : MCP*. 2016;15(1):141-53. Epub 2015/11/06. doi: 10.1074/mcp.M115.054221. PubMed PMID: 26537798; PubMed Central PMCID: PMCPMC4762517.
27. Loke I, Ostergaard O, Heegaard NHH, Packer NH, Thaysen-Andersen M. Paucimannose-Rich N-glycosylation of Spatiotemporally Regulated Human Neutrophil Elastase Modulates Its Immune Functions. *Molecular & cellular proteomics : MCP*. 2017;16(8):1507-27. Epub 2017/06/21. doi: 10.1074/mcp.M116.066746. PubMed PMID: 28630087; PubMed Central PMCID: PMCPMC5546201.
28. Lee LY, Moh ES, Parker BL, Bern M, Packer NH, Thaysen-Andersen M. Toward Automated N-Glycopeptide Identification in Glycoproteomics. *Journal of proteome research*. 2016;15(10):3904-15. Epub 2016/08/16. doi: 10.1021/acs.jproteome.6b00438. PubMed PMID: 27519006.
29. Scott NE, Parker BL, Connolly AM, Paulech J, Edwards AV, Crossett B, et al. 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*. *Molecular & cellular proteomics : MCP*. 2011;10(2):M000031-MCP201. doi: 10.1074/mcp.M000031-MCP201. PubMed PMID: 20360033; PubMed Central PMCID: PMCPMC3033663.
30. Scott NE, Nothaft H, Edwards AV, Labbate M, Djordjevic SP, Larsen MR, et al. Modification of the *Campylobacter jejuni* N-linked glycan by EptC protein-mediated addition of phosphoethanolamine. *The Journal of biological chemistry*. 2012;287(35):29384-96. Epub 2012/07/05. doi: 10.1074/jbc.M112.380212. PubMed PMID: 22761430; PubMed Central PMCID: PMCPMC3436159.
31. Lithgow KV, Scott NE, Iwashkiw JA, Thomson EL, Foster LJ, Feldman MF, et al. A general protein O-glycosylation system within the *Burkholderia cepacia* complex is involved in motility and virulence. *Molecular microbiology*. 2014;92(1):116-37. doi: 10.1111/mmi.12540. PubMed PMID: 24673753.
32. Iwashkiw JA, Seper A, Weber BS, Scott NE, Vinogradov E, Stratilo C, et al. Identification of a general O-linked protein glycosylation system in *Acinetobacter baumannii* and its role in virulence and biofilm formation. *PLoS pathogens*. 2012;8(6):e1002758. doi: 10.1371/journal.ppat.1002758. PubMed PMID: 22685409; PubMed Central PMCID: PMCPMC3369928.
33. Xue Y, Xie J, Fang P, Yao J, Yan G, Shen H, et al. Study on behaviors and performances of universal N-glycopeptide enrichment methods. *Analyst*. 2018;143(8):1870-80. Epub 2018/03/21. doi: 10.1039/C7AN02062G. PubMed PMID: 29557479.

34. Neue K, Mormann M, Peter-Katalinic J, Pohlentz G. Elucidation of glycoprotein structures by unspecific proteolysis and direct nanoESI mass spectrometric analysis of ZIC-HILIC-enriched glycopeptides. *Journal of proteome research*. 2011;10(5):2248-60. Epub 2011/03/30. doi: 10.1021/pr101082c. PubMed PMID: 21443200.
35. Fathy Mohamed Y, Scott NE, Molinaro A, Creuzenet C, Ortega X, Lertmemongkolchai G, et al. A general protein O-glycosylation machinery conserved in Burkholderia species improves bacterial fitness and elicits glycan immunogenicity in humans. *The Journal of biological chemistry*. 2019. Epub 2019/07/28. doi: 10.1074/jbc.RA119.009671. PubMed PMID: 31350337.
36. Nothaft H, Szymanski CM. New discoveries in bacterial N-glycosylation to expand the synthetic biology toolbox. *Curr Opin Chem Biol*. 2019;53:16-24. Epub 2019/07/10. doi: 10.1016/j.cbpa.2019.05.032. PubMed PMID: 31288139.
37. Szymanski CM, Wren BW. Protein glycosylation in bacterial mucosal pathogens. *Nature reviews Microbiology*. 2005;3(3):225-37. doi: 10.1038/nrmicro1100. PubMed PMID: 15738950.
38. Koomey M. O-linked protein glycosylation in bacteria: snapshots and current perspectives. *Curr Opin Struct Biol*. 2019;56:198-203. Epub 2019/05/13. doi: 10.1016/j.sbi.2019.03.020. PubMed PMID: 31078896.
39. Joshi HJ, Narimatsu Y, Schjoldager KT, Tytgat HLP, Aebi M, Clausen H, et al. SnapShot: O-Glycosylation Pathways across Kingdoms. *Cell*. 2018;172(3):632- e2. Epub 2018/01/27. doi: 10.1016/j.cell.2018.01.016. PubMed PMID: 29373833.
40. Elhenawy W, Scott NE, Tondo ML, Orellano EG, Foster LJ, Feldman MF. Protein O-linked glycosylation in the plant pathogen *Ralstonia solanacearum*. *Glycobiology*. 2016;26(3):301-11. doi: 10.1093/glycob/cwv098. PubMed PMID: 26531228; PubMed Central PMCID: PMC4736539.
41. Harding CM, Nasr MA, Kinsella RL, Scott NE, Foster LJ, Weber BS, et al. *Acinetobacter* strains carry two functional oligosaccharyltransferases, one devoted exclusively to type IV pilin, and the other one dedicated to O-glycosylation of multiple proteins. *Molecular microbiology*. 2015;96(5):1023-41. Epub 2015/03/03. doi: 10.1111/mmi.12986. PubMed PMID: 25727908.
42. Scott NE, Kinsella RL, Edwards AV, Larsen MR, Dutta S, Saba J, et al. Diversity within the O-linked protein glycosylation systems of *acinetobacter* species. *Molecular & cellular proteomics : MCP*. 2014;13(9):2354-70. doi: 10.1074/mcp.M114.038315. PubMed PMID: 24917611; PubMed Central PMCID: PMC4159654.
43. Scott NE, Marzook NB, Cain JA, Solis N, Thaysen-Andersen M, Djordjevic SP, et al. Comparative proteomics and glycoproteomics reveal increased N-linked glycosylation and relaxed sequon specificity in *Campylobacter jejuni* NCTC11168 O. *Journal of proteome research*. 2014;13(11):5136-50. Epub 2014/08/06. doi: 10.1021/pr5005554. PubMed PMID: 25093254.
44. Ahmad Izaham AR, Scott NE. Open database searching enables the identification and comparison of bacterial glycoproteomes without defining glycan compositions prior to searching. *Molecular & cellular proteomics : MCP*. 2020. Epub 2020/06/25. doi: 10.1074/mcp.TIR120.002100. PubMed PMID: 32576591.
45. Oppy CC, Jebeli L, Kuba M, Oates CV, Strugnelli R, Edgington-Mitchell LE, et al. Loss of O-linked protein glycosylation in *Burkholderia cenocepacia* impairs biofilm formation, siderophore activity and alters transcriptional regulators mSphere. 2019;(Accepted).
46. Willcocks SJ, Denman C, Cia F, McCarthy E, Cuccui J, Wren BW. Virulence of the emerging pathogen, *Burkholderia pseudomallei*, depends upon the O-linked oligosaccharyltransferase, PglL. *Future Microbiol*. 2020;15:241-57. Epub 2020/04/10. doi: 10.2217/fmb-2019-0165. PubMed PMID: 32271107.

47. Cumeras R, Figueras E, Davis CE, Baumbach JJ, Gracia I. Review on ion mobility spectrometry. Part 1: current instrumentation. *Analyst*. 2015;140(5):1376-90. Epub 2014/12/04. doi: 10.1039/c4an01100g. PubMed PMID: 25465076; PubMed Central PMCID: PMC4331213.
48. Winter DL, Wilkins MR, Donald WA. Differential Ion Mobility-Mass Spectrometry for Detailed Analysis of the Proteome. *Trends Biotechnol*. 2019;37(2):198-213. Epub 2018/09/09. doi: 10.1016/j.tibtech.2018.07.018. PubMed PMID: 30193737.
49. Pfammatter S, Bonneil E, McManus FP, Prasad S, Bailey DJ, Belford M, et al. A Novel Differential Ion Mobility Device Expands the Depth of Proteome Coverage and the Sensitivity of Multiplex Proteomic Measurements. *Molecular & cellular proteomics : MCP*. 2018;17(10):2051-67. Epub 2018/07/17. doi: 10.1074/mcp.TIR118.000862. PubMed PMID: 30007914; PubMed Central PMCID: PMC6166672.
50. Hebert AS, Prasad S, Belford MW, Bailey DJ, McAlister GC, Abbatiello SE, et al. Comprehensive Single-Shot Proteomics with FAIMS on a Hybrid Orbitrap Mass Spectrometer. *Anal Chem*. 2018;90(15):9529-37. Epub 2018/07/04. doi: 10.1021/acs.analchem.8b02233. PubMed PMID: 29969236; PubMed Central PMCID: PMC6145172.
51. Bekker-Jensen DB, Martinez-Val A, Steigerwald S, Ruther P, Fort KL, Arrey TN, et al. A Compact Quadrupole-Orbitrap Mass Spectrometer with FAIMS Interface Improves Proteome Coverage in Short LC Gradients. *Molecular & cellular proteomics : MCP*. 2020;19(4):716-29. Epub 2020/02/14. doi: 10.1074/mcp.TIR119.001906. PubMed PMID: 32051234; PubMed Central PMCID: PMC67124470.
52. Schweppe DK, Prasad S, Belford MW, Navarrete-Perea J, Bailey DJ, Huguet R, et al. Characterization and Optimization of Multiplexed Quantitative Analyses Using High-Field Asymmetric-Waveform Ion Mobility Mass Spectrometry. *Anal Chem*. 2019;91(6):4010-6. Epub 2019/01/24. doi: 10.1021/acs.analchem.8b05399. PubMed PMID: 30672687; PubMed Central PMCID: PMC6993951.
53. Bridon G, Bonneil E, Muratore-Schroeder T, Caron-Lizotte O, Thibault P. Improvement of phosphoproteome analyses using FAIMS and decision tree fragmentation. application to the insulin signaling pathway in *Drosophila melanogaster* S2 cells. *Journal of proteome research*. 2012;11(2):927-40. Epub 2011/11/09. doi: 10.1021/pr200722s. PubMed PMID: 22059388.
54. Shvartsburg AA, Zheng Y, Smith RD, Kelleher NL. Separation of variant methylated histone tails by differential ion mobility. *Anal Chem*. 2012;84(15):6317-20. Epub 2012/07/21. doi: 10.1021/ac301541r. PubMed PMID: 22812477; PubMed Central PMCID: PMC3418378.
55. Pfammatter S, Bonneil E, McManus FP, Thibault P. Gas-Phase Enrichment of Multiply Charged Peptide Ions by Differential Ion Mobility Extend the Comprehensiveness of SUMO Proteome Analyses. *J Am Soc Mass Spectrom*. 2018;29(6):1111-24. Epub 2018/04/07. doi: 10.1007/s13361-018-1917-y. PubMed PMID: 29623662.
56. Schnirch L, Nadler-Holly M, Siao SW, Frese CK, Viner R, Liu F. Expanding the depth and sensitivity of cross-link identification by differential ion mobility using FAIMS. *Anal Chem*. 2020. Epub 2020/07/10. doi: 10.1021/acs.analchem.0c01273. PubMed PMID: 32643919.
57. Creese AJ, Cooper HJ. Separation and identification of isomeric glycopeptides by high field asymmetric waveform ion mobility spectrometry. *Anal Chem*. 2012;84(5):2597-601. Epub 2012/01/28. doi: 10.1021/ac203321y. PubMed PMID: 22280549; PubMed Central PMCID: PMC3295202.

58. Pathak P, Baird MA, Shvartsburg AA. High-Resolution Ion Mobility Separations of Isomeric Glycoforms with Variations on the Peptide and Glycan Levels. *J Am Soc Mass Spectrom*. 2020. Epub 2020/06/06. doi: 10.1021/jasms.0c00183. PubMed PMID: 32501708.
59. Ulasi GN, Creese AJ, Hui SX, Penn CW, Cooper HJ. Comprehensive mapping of O-glycosylation in flagellin from *Campylobacter jejuni* 11168: A multienzyme differential ion mobility mass spectrometry approach. *Proteomics*. 2015;15(16):2733-45. Epub 2015/04/18. doi: 10.1002/pmic.201400533. PubMed PMID: 25884275; PubMed Central PMCID: PMC4975691.
60. Gelb AS, Lai R, Li H, Dodds ED. Composition and charge state influence on the ion-neutral collision cross sections of protonated N-linked glycopeptides: an experimental and theoretical deconstruction of coulombic repulsion vs. charge solvation effects. *Analyst*. 2019;144(19):5738-47. Epub 2019/08/28. doi: 10.1039/c9an00875f. PubMed PMID: 31453603; PubMed Central PMCID: PMC6756964.
61. Glaskin RS, Khatri K, Wang Q, Zaia J, Costello CE. Construction of a Database of Collision Cross Section Values for Glycopeptides, Glycans, and Peptides Determined by IM-MS. *Anal Chem*. 2017;89(8):4452-60. Epub 2017/03/23. doi: 10.1021/acs.analchem.6b04146. PubMed PMID: 28323417; PubMed Central PMCID: PMC5766263.
62. Li H, Bendiak B, Siems WF, Gang DR, Hill HH, Jr. Ion Mobility-Mass Correlation Trend Line Separation of Glycoprotein Digests without Deglycosylation. *Int J Ion Mobil Spectrom*. 2013;16(2):105-15. Epub 2013/08/06. doi: 10.1007/s12127-013-0127-3. PubMed PMID: 23914139; PubMed Central PMCID: PMC3728182.
63. Bzdyl NM, Scott NE, Norville IH, Scott AE, Atkins T, Pang S, et al. Peptidyl-Prolyl Isomerase ppiB Is Essential for Proteome Homeostasis and Virulence in *Burkholderia pseudomallei*. *Infection and immunity*. 2019;87(10). Epub 2019/07/25. doi: 10.1128/IAI.00528-19. PubMed PMID: 31331957; PubMed Central PMCID: PMC6759293.
64. Nothaft H, Scott NE, Vinogradov E, Liu X, Hu R, Beadle B, et al. Diversity in the protein N-glycosylation pathways within the *Campylobacter* genus. *Molecular & cellular proteomics : MCP*. 2012;11(11):1203-19. doi: 10.1074/mcp.M112.021519. PubMed PMID: 22859570; PubMed Central PMCID: PMC3494190.
65. Saba J, Dutta S, Hemenway E, Viner R. Increasing the productivity of glycopeptides analysis by using higher-energy collision dissociation-accurate mass-product-dependent electron transfer dissociation. *Int J Proteomics*. 2012;2012:560391. doi: 10.1155/2012/560391. PubMed PMID: 22701174; PubMed Central PMCID: PMC3369405.
66. Bern M, Kil YJ, Becker C. Byonic: advanced peptide and protein identification software. *Curr Protoc Bioinformatics*. 2012;Chapter 13:Unit13 20. Epub 2012/12/21. doi: 10.1002/0471250953.bi1320s40. PubMed PMID: 23255153; PubMed Central PMCID: PMC3545648.
67. Lenco J, Khalikova MA, Svec F. Dissolving Peptides in 0.1% Formic Acid Brings Risk of Artificial Formylation. *Journal of proteome research*. 2020;19(3):993-9. Epub 2020/02/19. doi: 10.1021/acs.jproteome.9b00823. PubMed PMID: 32068400.
68. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol*. 2008;26(12):1367-72. doi: 10.1038/nbt.1511. PubMed PMID: 19029910.
69. Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature methods*. 2016;13(9):731-40. Epub 2016/06/28. doi: 10.1038/nmeth.3901. PubMed PMID: 27348712.

70. Osorio D, Rondón-Villarreal P, Torres R. Peptides: A Package for Data Mining of Antimicrobial Peptides. *The R Journal*. 2015;7(1):10.
71. Wickham H. *ggplot2: Elegant Graphics for Data Analysis*: Springer-Verlag New York; 2016.
72. Ikai A. Thermostability and aliphatic index of globular proteins. *J Biochem*. 1980;88(6):1895-8. Epub 1980/12/01. PubMed PMID: 7462208.
73. Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res*. 2019;47(D1):D442-D50. Epub 2018/11/06. doi: 10.1093/nar/gky1106. PubMed PMID: 30395289; PubMed Central PMCID: PMC6323896.
74. Vizcaino JA, Csordas A, del-Toro N, Dianes JA, Griss J, Lavidas I, et al. 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res*. 2016;44(D1):D447-56. doi: 10.1093/nar/gkv1145. PubMed PMID: 26527722; PubMed Central PMCID: PMC4702828.
75. Alagesan K, Khilji SK, Kolarich D. It is all about the solvent: on the importance of the mobile phase for ZIC-HILIC glycopeptide enrichment. *Anal Bioanal Chem*. 2017;409(2):529-38. Epub 2016/12/03. doi: 10.1007/s00216-016-0051-6. PubMed PMID: 27909778; PubMed Central PMCID: PMC5203826.
76. Hadjineophytou C, Anonsen JH, Wang N, Ma KC, Viburiene R, Vik A, et al. Genetic determinants of genus-Level glycan diversity in a bacterial protein glycosylation system. *PLoS Genet*. 2019;15(12):e1008532. Epub 2019/12/24. doi: 10.1371/journal.pgen.1008532. PubMed PMID: 31869330.
77. Hahne H, Gholami AM, Kuster B. Discovery of O-GlcNAc-modified proteins in published large-scale proteome data. *Molecular & cellular proteomics : MCP*. 2012;11(10):843-50. Epub 2012/06/05. doi: 10.1074/mcp.M112.019463. PubMed PMID: 22661428; PubMed Central PMCID: PMC3494142.
78. Hahne H, Kuster B. Discovery of O-GlcNAc-6-phosphate modified proteins in large-scale phosphoproteomics data. *Molecular & cellular proteomics : MCP*. 2012;11(10):1063-9. Epub 2012/07/25. doi: 10.1074/mcp.M112.019760. PubMed PMID: 22826440; PubMed Central PMCID: PMC3494138.
79. Chick JM, Kolippakkam D, Nusinow DP, Zhai B, Rad R, Huttlin EL, et al. A mass-tolerant database search identifies a large proportion of unassigned spectra in shotgun proteomics as modified peptides. *Nat Biotechnol*. 2015;33(7):743-9. doi: 10.1038/nbt.3267. PubMed PMID: 26076430; PubMed Central PMCID: PMC4515955.
80. Li Q, Shortreed MR, Wenger CD, Frey BL, Schaffer LV, Scalf M, et al. Global Post-Translational Modification Discovery. *Journal of proteome research*. 2017;16(4):1383-90. Epub 2017/03/02. doi: 10.1021/acs.jproteome.6b00034. PubMed PMID: 28248113; PubMed Central PMCID: PMC5387672.
81. Solntsev SK, Shortreed MR, Frey BL, Smith LM. Enhanced Global Post-translational Modification Discovery with MetaMorpheus. *Journal of proteome research*. 2018;17(5):1844-51. Epub 2018/03/27. doi: 10.1021/acs.jproteome.7b00873. PubMed PMID: 29578715.
82. Stavenhagen K, Hinneburg H, Thaysen-Andersen M, Hartmann L, Varon Silva D, Fuchser J, et al. Quantitative mapping of glycoprotein micro-heterogeneity and macro-heterogeneity: an evaluation of mass spectrometry signal strengths using synthetic peptides and glycopeptides. *J Mass Spectrom*. 2013;48(6):627-39. Epub 2013/06/01. doi: 10.1002/jms.3210. PubMed PMID: 23722953.

83. Schulz BL, Jen FE, Power PM, Jones CE, Fox KL, Ku SC, et al. Identification of bacterial protein O-oligosaccharyltransferases and their glycoprotein substrates. *PloS one*. 2013;8(5):e62768. Epub 2013/05/10. doi: 10.1371/journal.pone.0062768. PubMed PMID: 23658772; PubMed Central PMCID: PMC3643930.
84. Gebhart C, Ielmini MV, Reiz B, Price NL, Aas FE, Koomey M, et al. Characterization of exogenous bacterial oligosaccharyltransferases in *Escherichia coli* reveals the potential for O-linked protein glycosylation in *Vibrio cholerae* and *Burkholderia thailandensis*. *Glycobiology*. 2012;22(7):962-74. doi: 10.1093/glycob/cws059. PubMed PMID: 22391990.
85. Vorkapic D, Mitterer F, Pressler K, Leitner DR, Anonsen JH, Liesinger L, et al. A Broad Spectrum Protein Glycosylation System Influences Type II Protein Secretion and Associated Phenotypes in *Vibrio cholerae*. *Front Microbiol*. 2019;10:2780. Epub 2019/12/19. doi: 10.3389/fmicb.2019.02780. PubMed PMID: 31849912; PubMed Central PMCID: PMC6901666.
86. Jervis AJ, Wood AG, Cain JA, Butler JA, Frost H, Lord E, et al. Functional analysis of the *Helicobacter pullorum* N-linked protein glycosylation system. *Glycobiology*. 2018;28(4):233-44. Epub 2018/01/18. doi: 10.1093/glycob/cwx110. PubMed PMID: 29340583; PubMed Central PMCID: PMC6025236.
87. Jervis AJ, Butler JA, Lawson AJ, Langdon R, Wren BW, Linton D. Characterization of the structurally diverse N-linked glycans of *Campylobacter* species. *Journal of bacteriology*. 2012;194(9):2355-62. Epub 2012/03/06. doi: 10.1128/JB.00042-12. PubMed PMID: 22389484; PubMed Central PMCID: PMC3347071.
88. Vandamme P, Holmes B, Vancanneyt M, Coenye T, Hoste B, Coopman R, et al. Occurrence of multiple genomovars of *Burkholderia cepacia* in cystic fibrosis patients and proposal of *Burkholderia multivorans* sp. nov. *Int J Syst Bacteriol*. 1997;47(4):1188-200. Epub 1997/10/23. doi: 10.1099/00207713-47-4-1188. PubMed PMID: 9336927.
89. Johnson SL, Bishop-Lilly KA, Ladner JT, Daligault HE, Davenport KW, Jaissle J, et al. Complete genome sequences for 59 *Burkholderia* isolates, both pathogenic and near neighbor. *Genome Announc*. 2015;3(2). Epub 2015/05/02. doi: 10.1128/genomeA.00159-15. PubMed PMID: 25931592; PubMed Central PMCID: PMC4417688.
90. Winsor GL, Khaira B, Van Rossum T, Lo R, Whiteside MD, Brinkman FS. The *Burkholderia* Genome Database: facilitating flexible queries and comparative analyses. *Bioinformatics*. 2008;24(23):2803-4. Epub 2008/10/10. doi: 10.1093/bioinformatics/btn524. PubMed PMID: 18842600; PubMed Central PMCID: PMC2639269.
91. VÉRON MC, R. Taxonomic Study of the Genus *Campylobacter* Sebald and Véron and Designation of the Neotype Strain for the Type Species, *Campylobacter fetus* (Smith and Taylor) Sebald and Véro. *INTERNATIONAL JOURNAL OF SYSTEMATIC BACTERIOLOGY*. 1973;23(3):122-34.

A) Glycoproteins **B) Glycopeptides** **C) iBAQ analysis** **D) Peptide Properties**

A) Charge states of unenriched glycopeptides

B) Charge states of enriched glycopeptides

Charge States

C) Unique Glycopeptides vs CV

D) Glycoproteins

E) Glycopeptides

FAIMS only
Identified in both
Enriched only

A)**Unique to CVs** **B)****Unique glycopeptides observed across all CVs**

A) Unique Glycopeptides

B) Stepped FAIMS vs ZIC-HILIC

C) Glycopeptides Overlap

D) Peptide Properties

A) *B. dolosa* AU0158
Glycoproteome

Glycopeptides

B) *B. ubonensis* MSMB22
Glycoproteome

Glycopeptides