Journal of proteome research

Article

pubs.acs.org/jpr

Confident Assignment of Site-Specific Glycosylation in Complex Glycoproteins in a Single Step

Kshitij Khatri,[†] Gregory O. Staples,[§] Nancy Leymarie,[†] Deborah R. Leon,[†] Lilla Turiák,[†] Yu Huang,[†] Shun Yip,[‡] Han Hu,[†,‡] Christian F. Heckendorf,[†] and Joseph Zaia*[,†]

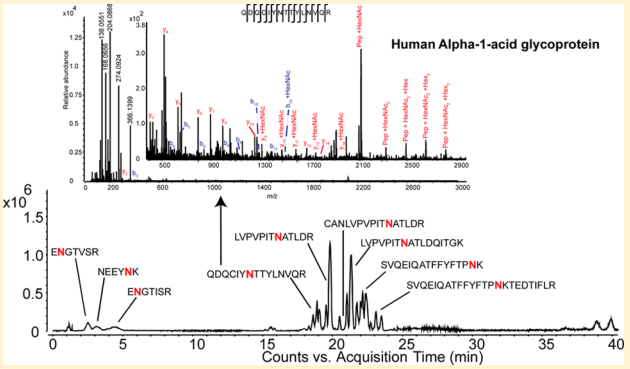
[†]Center for Biomedical Mass Spectrometry, Department of Biochemistry, Boston University School of Medicine, Boston, Massachusetts 02118, United States

[‡]Bioinformatics Program, Boston University, Boston, Massachusetts 02215, United States

[§]Agilent Technologies, Santa Clara, California 95051, United States

Supporting Information

ABSTRACT: A glycoprotein may contain several sites of glycosylation, each of which is heterogeneous. As a consequence of glycoform diversity and signal suppression from non-glycosylated peptides that ionize more efficiently, typical reversed-phase LC−MS and bottom−up proteomics database searching workflows do not perform well for identification of site-specific glycosylation for complex glycoproteins. We present an LC−MS system for enrichment, separation, and analysis of glycopeptides from complex glycoproteins (>4 *N*-glycosylation sequons) in a single step. This system uses an online HILIC enrichment trap prior to reversed-phase C18-MS analysis. We demonstrated the effectiveness of the system using a set of glycoproteins including human transferrin (2 sequons), human alpha-1-acid glycoprotein (5 sequons), and influenza A virus hemagglutinin (9 sequons). The online enrichment renders glycopeptides the most abundant ions detected, thereby facilitating the generation of high-quality data-dependent tandem mass spectra. The tandem mass spectra exhibited product ions from both glycan and peptide backbone dissociation for a majority of the glycopeptides tested using collisionally activated dissociation that served to confidently assign site-specific glycosylation. We demonstrated the value of our system to define site-specific glycosylation using a hemagglutinin containing 9 *N*-glycosylation sequons from a single HILIC-C18-MS acquisition.



KEYWORDS: *Glycopeptides, tandem MS, CAD, HILIC*

INTRODUCTION

Occurring through a series of regulated biosynthetic events in the endoplasmic reticulum and Golgi apparatus, the extent of glycosylation of mature proteins depends on numerous factors. Thus, the glycan structures assembled onto a given protein core are regulated according to cell type, developmental stage, and spatial considerations. It is therefore not surprising that distinct glycoprotein glycosylation patterns characterize normal versus disease states.[1−3] As has been described in a recent report by the National Research Council,[4] methods for the study of glycoprotein structures are essential to progress in biomedicine. Many glycoproteins have several sites of glycosylation, each with a range of glycoforms. Glycoproteins pose analytical challenges based both on their chemical properties and the nature of the information produced. Thus, analytical methods used for proteomics and glycomics do not suffice for analysis of glycoproteins.[5]

Mass spectrometry workflows for discovery and targeted proteomics are applicable to post-translational modifications (PTM) with defined molecular weight values, including phosphorylation, acetylation, methylation, ubiquitination, and *β*-O-GlcNAcylation.[6,7] Such workflows have been applied to analysis of protein *N*-glycosylation site occupancy only following enzymatic release of glycans.[8] Thus, because complex glycosylation is beyond the scope of traditional proteomics workflows, there is a paucity of information regarding the structures of *N*-glycans that occupy each protein site.[9] In order to address this problem, algorithms that allow users to define an unlimited number of peptide modifications have been developed.[10]

In order to determine the site-specific glycosylation structure of mature glycoproteins, it is necessary to acquire accurate proteolytic peptide maps that include the range of glycan structures present at each glycosylated amino acid. In practice, accurate determination of the site-specific glycan structures present on complex glycoproteins (those containing more than one glycosylation site) requires purified starting material. This

Received: May 25, 2014
Published: August 25, 2014

is demonstrated by a recent interlaboratory study[11] in which prostate specific antigen, a glycoprotein with a single site of *N*-glycosylation, was analyzed. Although there was a consensus in the interlaboratory results regarding the distribution of the most abundant *N*-glycan compositions present, it was clear that a more complex glycoprotein would pose serious challenges to investigators using analytical methods in current practice. The reason for this is that glycosylation multiplies the number of molecular forms of each glycosylated peptide by more than 10-fold. In addition, the most often used reversed-phase chromatography methods in proteomics workflows lack an enrichment step and therefore do not perform well for mapping glycopeptides from glycoproteins that contain several glycosylation sites.

Glycosylated proteins and peptides ionize poorly relative to nonglycosylated molecules. In order to produce confident identification of peptide backbone and glycan composition, it is desirable to analyze intact glycopeptides. Intact analysis eliminates the possibility of false positive identifications that result from detection of deglycosylated peptides as surrogates for glycopeptides. Specifically, peptides that become deamidated, either through chemical or enzymatic processes, may be confused with formerly glycosylated peptides.[12] On the other hand, direct MS analysis of glycopeptides is made challenging by their aforementioned heterogeneity and poor ionization properties. Fortunately, hydrophilic interaction liquid chromatography (HILIC) has proven to be effective for enrichment of glycopeptides based on the hydrophilic glycan moiety.[13−16] However, these enrichment strategies, often in the form of solid-phase extraction cartridges, add a number of steps to the sample preparation workflow. We conceived that it would be possible to integrate HILIC enrichment of glycopeptides directly to the reversed-phase separation. HILIC has been directly interfaced to revered-phase separation in previous reports for 2D separations.[17,18] In these cases, significant dilution of the fractions is necessary in order to ensure proper sample binding on the second dimension phase. This complication arises from the fact that HILIC separations typically begin in high-percent organic mobile phase, and reversed-phased separations, in high-percent aqueous mobile phase. The design of a platform to automate HILIC enrichment followed by reversed-phase separation is considerably simpler in architecture than a 2D HILIC-reversed-phase separation.

We found that HILIC enrichment could be performed online with reversed-phase separation as long as the enrichment column volume was kept sufficiently small compared to the separation column volume. Here, we present the use of online HILIC enrichment prior to C18 analytical separation with MS detection (HILIC-C18-MS). While we used a microfluidic chromatography chip to perform the experiments, such online enrichment and separation arrangement can be achieved using any HPLC system with appropriate valves for flow control and plumbing for low dead volume, see, for example, Supporting Information Figure S-4. Using this system, the most abundant ions detected corresponded to glycopeptides.

As has been pointed out by Desaire and Hua,[19] it is difficult to assign glycopeptides based solely on MS data, even with the use of high-resolution Fourier transform-MS instruments, with high mass accuracy. As the number of glycosylation sites, variable modifications, and peptide miscleavages increases, the usefulness of high-resolution MS data becomes limited in analyzing complex glycoproteins. A tandem mass spectrometry step therefore becomes necessary in order to map glycopeptides with confidence. A recent review of software tools for glycopeptide analysis also highlights the importance of tandem fragmentation data for high-confidence glycopeptide assignments.[20] Thus, the need to improve quality of glycopeptide tandem mass spectra acquired during LC−MS (liquid chromatography−mass spectrometry) experiments also drove the present work. Beam-type collision-activated dissociation (CAD) of glycopeptides produces abundant oxonium ions, the m/z values of which serve as useful indicators of glycosylation.[21] Under typical CAD conditions, ions produced from dissociation of the peptide backbone are not detected.[22] As a result, separate electron activated dissociation (ExD) methods are often used to produce abundant peptide backbone dissociation while leaving the glycan intact.[23,24] The online HILIC-C18-MS system we describe increases the abundances of glycopeptide precursor ions and is, in principle, applicable to any tandem MS dissociation method.

Using elevated collision energies, we observed peptide backbone fragments from glycopeptides, in addition to oxonium ions and intact peptide ions with varying numbers of saccharide units attached (stub glycopeptides). An and colleagues,[25] who used elevated collision energies to fragment all ions entering the mass spectrometer without ion isolation/selection (known as MS^E), have also reported glycopeptide backbone fragmentation, and others have reported similar glycopeptide backbone fragmentation using higher energy collisional (C-trap) dissociation (HCD).[26,27] We demonstrate that HILIC-C18-MS enriches proteolytic glycopeptides from complex glycoproteins efficiently and thereby improves the ability to acquire data-dependent MS/MS on these molecules. As a result, the use of HILIC-C18-MS/MS generated significantly greater informational value than did C18-MS/MS. We demonstrated this method by analyzing glycoproteins including human transferrin and human alpha-1 acid glycoprotein (AGP) and influenza A virus (IAV) hemagglutinin (HA). HA is a viral coat protein that is heavily glycosylated with the number of glycosylation sites varying between 5 and 12, depending on the strain. It binds sialylated glycoproteins present in host airway epithelia to facilitate infection and presents as a target for both innate and adaptive host response. HA glycosylation, therefore, represents an attractive target for researchers studying the pathogenicity of IAV.[28−31]

EXPERIMENTAL SECTION

Samples and Reagents

Human transferrin and human alpha-1-acid glycoprotein were purchased from Sigma-Aldrich (St. Louis, MO). Recombinant HA1 from *A/USSR/90/1977 (H1N1)* (amino acids 18−345) (Genebank no. ABD60933) was purchased from Immune Technology Corp. (New York, NY). Sequencing grade trypsin was purchased from Promega Corp. (Madison, WI). LC−MS grade solvents were purchased from Fisher Scientific (Pittsburgh, PA). All other chemicals and reagents were purchased from Sigma-Aldrich, unless otherwise stated.

Methods

Glycoproteins were reduced, alkylated, and digested using trypsin, as described in Supporting Information Section S-1. The glycoprotein digests were subjected to HILIC-C18 and C18 HPLC-chip LC−MS/MS using an Agilent 1200 series nanoflow HPLC coupled to an Agilent 6550 Q-TOF system with a chip-cube nano-ESI source (Agilent Technologies, Santa Clara, CA). The loading/trapping conditions for HILIC-C18

Table 1. Total Number of Identified Peptides and Glycopeptides from LC−MS Data Sets Using GlycReSoft (MS Only)[a]

glycoprotein	no. of theoretical peptides + glycopeptides	total no. of nonglycosylated peptides identified (MS1 only)		no. of theoretical glycopeptides (search space)	no. of ambiguous matches (more than one match)		total no. of glycopeptides identified (MS1 only)		total % relative abundance of glycopeptides	
		HILIC-C18	C18		HILIC-C18	C18	HILIC-C18	C18	HILIC-C18	C18
transferrin	1576	2	77	1230	6	5	54	38	97.67%	0.86%
AGP	4009	2	19	3936	16	23	155	141	99.93%	3.40%
hemagglutinin	57 358	4	39	56 886	725	278	1220	530	99.97%	2.52%

[a]Only the glycopeptides/peptides found in all three replicate data were included. Relative abundances were calculated as a percentage of total abundances of all identified compounds. Note that deamidation was considered as a variable modification only in the case of HA.

were 80% acetonitrile modified with 0.1% TFA, which allowed binding of hydrophilic glycan chains of the glycopeptides while the nonglycosylated peptides were not retained.[15] Trapped glycopeptides were eluted from the HILIC trapping column to the C18 column, where they were resolved using a reversed-phase gradient. Transfer of the glycopeptides from the HILIC column to the C18 column was automatically controlled by switching the chip-cube rotor from enrichment to analysis mode. The high aqueous initial conditions of the C18 separation were sufficient to elute the glycopeptides from the HILIC trapping column and to focus them onto the head of the C18 separation column. Results were compared to those acquired using a reversed-phase HPLC-chip that had C18 packing material in both enrichment and analytical columns. The loading and trapping steps were performed under high aqueous mobile phase conditions on the C18 HPLC-chip, followed by the same reversed-phase analytical gradient that was used on the HILIC-C18 HPLC-chip. Tandem MS was performed on the top 5 most abundant precursor ions in each MS scan. The mass spectrometer was programmed to calculate collision energies using a linear equation, based on the m/z of selected precursor ions. For additional details of sample preparation and LC−MS methods, please refer to Supporting Information Section S-1.

LC−MS Data Analysis

For automated data analysis, entire LC−MS data files were deconvoluted using DeconTools.[32] Deconvoluted data were matched and scored against theoretical peptide and glycopeptide neutral mass lists, using GlycReSoft,[33] with a 10 ppm mass tolerance. Lists of glycopeptides and glycoforms used for searching have been provided in Supporting Information Section S-3 and Tables S-1 and S-2.

For quantitative evaluation of the ability to perform tandem MS on glycopeptides, LC−MS data files were also searched for precursor ions, which produced oxonium ions. The search was performed using software written in-house that generates a list with the total number of precursor ions in a LC−MS/MS data file and precursors that produce oxonium ions. Data files were searched for the following oxonium ion masses[34] with a 0.01 unit mass tolerance: 204.08635, 186.07539, 168.065, 138.05423, 366.13942, 274.0920, and 292.1026. Only the precursor ions that produced 3 or more of the listed oxonium ions were identified as glycopeptide precursor ions. The number of unique glycopeptide precursors selected for data-dependent tandem MS were determined in this way and compared among HILIC-C18 and C18 data.

Tandem mass spectra were analyzed both manually and using a prototype software tool, written in-house, to confirm peptide backbone fragment ions and glycan compositions for manually assigned glycopeptides. Theoretical peptide fragment ion masses were calculated and matched, using a 20 ppm mass error tolerance. For manual matching of tandem MS peaks, a minimum signal threshold of 5 counts was used.

RESULTS AND DISCUSSION

Glycopeptide Identification Based on Accurate Intact Mass

We compared the usefulness of LC−MS data acquired using HILIC-C18-MS versus C18-MS for profiling glycoprotein proteolytic peptides. For this purpose, the neutral mass values extracted from the data sets were searched against a set of theoretical glycopeptide masses, with a 10 ppm mass-error tolerance. The theoretical masses consisted of the protein proteolytic peptides with up to two missed cleavage sites. Masses for those peptides containing an *N*-glycosylation sequon were calculated as a set of glycosylation variants using *N*-glycosylation compositions ranging from core *N*-glycan structures to penta-antennary complex-type *N*-glycans containing *N*-acetylneuraminic acid and high-mannose *N*-glycans (shown in Supporting Information Section S-3). The GlycReSoft program[33] was used to match and score deconvoluted masses from an LC−MS experiment against this list of theoretical compound accurate masses. Because the complexity in glycopeptide data increases with the number of glycosylation sites, intact mass assignments suffer from ambiguous matches and false positives that cannot be verified in the absence of tandem MS.

We acquired LC−MS profiling data for three glycoproteins with a range of complexities. Transferrin is known to have two *N*-linked glycosylation sites with complex *N*-glycans, and the glycan heterogeneity for this protein is fairly limited with [5,4,0,2,0] [Hex, HexNAc, dHex, NeuAc, NeuGc] contributing to over 90% of the glycan distribution.[35] Thus, this glycoprotein presents low complexity, which makes the search space small and minimizes the chances for false positives and ambiguous assignments. AGP, by contrast, has 5 known *N*-linked glycosylation sites[36] with multiple genetic variants and a more diverse set of complex glycans that make this a relatively more complex glycoprotein. Recombinant hemagglutinin from *A/USSR/90/1977 (H1N1)*, referred to herein as HA, was the most complex of the three glycoproteins analyzed, with 9 consensus glycosylation sites. In addition, theoretical tryptic cleavage of HA produced 2 peptides that presented more than 1 putative glycosylation site on a single peptide. Thus, expected ambiguity and false positives scaled with increasing number of glycosylation sites in transferrin, AGP, and HA. The number of theoretical glycopeptide compositions used for each glycoprotein is given in Table 1.

dx.doi.org/10.1021/pr500506z | J. Proteome Res. 2014, 13, 4347−4355

We observed a significantly higher number of glycopeptide matches using HILIC-C18 LC−MS data over C18 LC−MS data, for all glycoproteins tested. Table 1 shows the number of glycopeptides and peptides identified based on intact glycopeptide masses. Even with the limited possibility of false matches when searching solely based on theoretical masses, it was evident that HILIC-C18 enriched glycopeptides efficiently. Data were acquired as analytical triplicates, and peptide/glycopeptide compositions found in all three replicates were considered matches.

Oxonium Ion Distributions

In order to confirm the ion compositions from the MS data, we used data-dependent LC−tandem mass spectrometry. Collisional dissociation of glycopeptides leads to formation of oxonium ions that are useful as features for identifying glycopeptides in LC−MS/MS chromatograms.[21] We therefore used the absence of oxonium ions to disqualify a precursor ion as glycopeptide. On the basis of this rationale, we compared the abundances of oxonium ions in extracted ion chromatograms (EICs) of glycopeptides observed using C18-MS/MS versus HILIC-C18-MS/MS data. Figure 1 compares the HexNAc oxonium ion (m/z 204) EIC with the BPC (base peak chromatogram) for hemagglutinin. Oxonium ion profiles for the rest of the proteins analyzed are presented in Supporting Information Figure S-1. When using HILIC-C18, we observed that the abundances of the HexNAc (m/z 204) oxonium ion profile were similar to those of the BPC, consistent with the fact that most of the detected ions corresponded to glycopeptides. For the C18 LC−MS experiments, the abundances of HexNAc oxonium ion were 10-fold lower than the BPC, indicating that most of the detected ions corresponded to unglycosylated peptides.

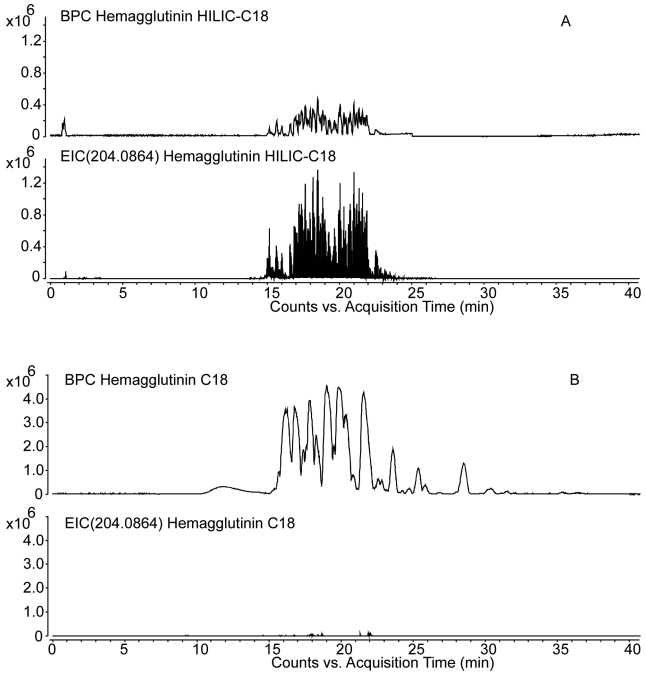


Figure 1. Comparison of tryptic glycopeptide abundances using (A) HILIC-C18-MS and (B) C18-MS. Each panel shows the base peak chromatogram (BPC) and the extracted ion chromatogram (EIC) for the HexNAc oxonium ion (m/z 204.08). BPC and EIC are shown using the same y scale.

Using C18-MS, unmodified peptides were the most abundant ions selected by the acquisition software for tandem MS based on abundance. By contrast, for HILIC-C18 data, glycopeptides were the most abundant ions present and were selected automatically for tandem MS. The oxonium ion distributions were consistent with the results obtained from the LC−MS data, shown in Table 1, indicating the contribution of glycopeptides to total ion abundances in C18 and HILIC-C18 runs.

Oxonium ions detected early in the chromatogram using HILIC-C18-MS resulted from a small fraction of the enriched glycopeptides from the HILIC trapping column getting washed away with the dead volume containing high organic mobile phase. This was a fixed and unbiased loss that accounted for less that 2% of the total sample abundance and did not affect the relative abundances of the compounds being retained on the C18 analytical columns.

Figure 2 shows the percent of precursor ions identified as glycopeptides based on formation of oxonium ions in data-dependent LC−MS/MS runs using HILIC-C18-MS versus C18-MS. We concluded that use of HILIC-C18-MS resulted in significant increase in the ability to analyze glycopeptides using data-dependent LC−MS/MS. These data demonstrate that the HILIC-C18-MS increased abundances of glycopeptides relative to unglycosylated peptides and thus the quality of data-dependent LC−tandem MS data. The improved data quality increased confidence in assignments and decreased false identifications.

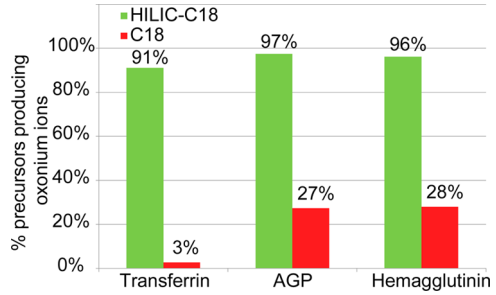


Figure 2. Comparison of percent of precursor ions that generate diagnostic oxonium ions between HILIC-C18-MS (green) and C18-MS (red). Results are shown for the glycoproteins labeled on the x axis.

Two AGP glycopeptides (ENGTISR/ENGTVSR and NEEYNK) were detected only using HILIC-C18-MS. This was due to the fact that tryptic cleavage of AGP produces glycopeptides that are only 5−7 amino acids long and are not retained during the trapping step when using C18-MS. These glycopeptides were trapped on the HILIC enrichment column when using HILIC-C18-MS and eluted soon after the solvent front on the analytical chromatogram, as seen in Supporting Information Figure S-1(b), which allowed them to undergo MS and MS/MS analysis. The C18 trapping step can be eliminated to retain any shorter glycopeptides in a sample. However, in our experience, an online trapping step is a useful means to eliminate salts and other contaminants while minimizing manual manipulation.

Tandem MS of Glycopeptides

While the presence of oxonium ions from dissociation of glycosidic bonds is a useful feature, the formation of peptide backbone product ions is necessary for unambiguous assign-

dx.doi.org/10.1021/pr500506z | J. Proteome Res. 2014, 13, 4347−4355

ment of the peptide sequence and glycan composition. We therefore investigated the extent to which the increased glycopeptide precursor ion abundances obtained using HILIC-C18-MS improved the quality of the resulting glycopeptide tandem mass spectra. Collisional dissociation using typical conditions for peptide tandem MS results in formation of abundant ions from dissociation of the glycan with those from peptide backbone dissociation often not detected. In the interest of maximizing the information produced from a CAD LC−MS/MS experiment, we investigated use of higher collision energies for glycopeptides. Figures 3 and S-2 show examples of glycopeptide tandem mass spectra from each of the three glycoproteins studied; collision energies were calculated as per the equations described in Supporting Information Section S-1. As expected, glycopeptide dissociation produced abundant oxonium ions in the low m/z range that confirmed the presence of monosaccharides in the precursor ion composition. In addition, peptide backbone product ions were detected that enabled direct identification of the peptide. This information increased confidence in true matches and decreased those in incorrect matches. This was important in cases where more than one theoretical glycopeptide matched an observed mass value. Complete peptide backbone or parts of the peptide backbone could be sequenced for glycopeptide precursors, as shown in the annotated spectra. Intact peptide and glycopeptides + saccharide (referred to as stub glycopeptide) ions were also detected in the higher m/z range of the spectra that matched the masses of peptide with N-glycan core structures. The presence of intact peptide or stub glycopeptide ions significantly increased confidence of assignments over those made from MS-only data. In the majority of LC−tandem mass spectra, ions corresponding to peptide backbone product ions plus a HexNAc residue were detected, which confirmed the site of glycosylation, as shown in Supporting Information Figure S-2 and Tables 2 and S-2.

Although the relative abundance of a glycopeptide precursor ion had an effect on the quality of tandem MS by affecting the ability to select the ion in a data-dependent LC−MS/MS experiment, the absolute abundance did not appear to play a role in generation of useful fragment ions for confident

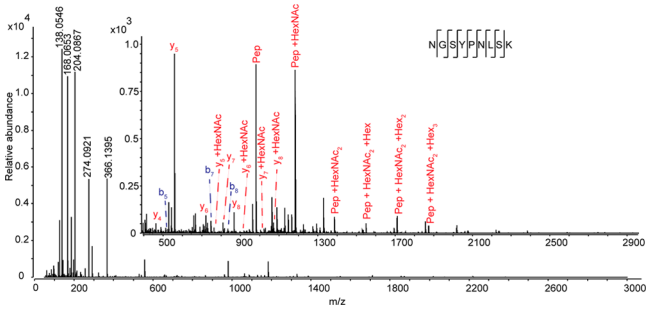


Figure 3. Tandem mass spectrum of hemagglutinin glycopeptide precursor ion 1013.7474 $[M + 3H]^{3+}$, identified as NGSYPNLSK-[5,4,1,1,0]. Glycopeptide identifier is listed as peptide-[a,b,c,d,e], where a = number of hexoses; b = number of *N*-acetylhexosamines; c = number of deoxyhexoses; d = number of *N*-acetylneuraminic acids; e = number of *N*-glycolylneuraminic acids.

Table 2. Summary of Glycopeptide Tandem Mass Spectra for Intact-Mass Matched Compounds[a]

Glycoprotein	Glycresoft assignments based on intact neutral mass				Precursor ion m/z and charge	Manual interpretation of tandem MS															
						HILIC-C18									C18						
	HILIC-C18 Compositions	Avg. % relative volume (out of matched compounds)	C18 Compositions	Avg. % relative volume (out of matched compounds)		Total precursor ion abundance (area)	Selected for Tandem MS	Oxonium ions	Intact peptide	Stub glycopeptides	Peptide sequence	% y-ion coverage	% b-ion coverage	Glycosylation site (b/y-ion with HexNAc)	Total precursor ion abundance (area)	Selected for Tandem MS	Oxonium ions	Intact peptide	Stub glycopeptides	Peptide sequence	% y-ion coverage	% b-ion coverage	Glycosylation site (b/y-ion with HexNAc)
Transferrin	QQQHLFGSNVTDCSGNFCLFR-[5,4,0,2,0]	30.04%	QQQHLFGSNVTDCSGNFCLFR-[5,4,0,2,0]	0.27%	1180.7328 (4+)	1.02E+07	√	√	√	√	√	95%	70%	√	7.77E+06	√	√	√	√	√	85%	70%	√
	CGLVPVLAENYNK-[5,4,0,2,0]	18.57%	CGLVPVLAENYNK-[5,4,0,2,0]	0.25%	921.1389 (4+)	5.76E+06	√	√	√	√	√	67%	67%	√	6.66E+06	√	√	√	√	√	67%	67%	√
	QQQHLFGSNVTDCSGNFCLFR-[5,4,1,2,0]	0.50%	QQQHLFGSNVTDCSGNFCLFR-[5,4,1,2,0]	0.02%	1217.2468 (4+)	1.10E+06	√	√	√	√	√	85%	55%	√	8.87E+05	x	x	x	x	x	0%	0%	x
	QQQHLFGSNVTDCSGNFCLFR-[6,5,1,3,0]	0.91%	QQQHLFGSNVTDCSGNFCLFR-[6,5,1,3,0]	0.00%	1381.3025 (4+)	3.21E+05	√	√	√	√	√	80%	50%	√	1.11E+05	x	x	x	x	x	0%	0%	x
AGP	ITGKWFYIASAFRNEEYNK-[4,6,1,1,0]	10.32%	ITGKWFYIASAFRNEEYNK-[4,6,1,1,0]	0.76%	1161.0091 (4+)	1.56E+07	√	√	x	x	x	0%	0%	x	1.58E+07	√	√	x	x	x	0%	0%	x
	SVQEIQATFFYFTPNK-[7,6,0,4,0] or QNQCFYNSSYLNVQRENGTVSR-[6,6,0,2,0]	10.80%	SVQEIQATFFYFTPNK-[7,6,0,4,0] or QNQCFYNSSYLNVQRENGTVSR-[6,6,0,2,0]	0.02%	1088.2441 (5+)	1.49E+07	√	√	√	√	√	93%	80%	√	0.00E+00	x	x	x	x	x	0%	0%	x
	QNQCFYNSSYLNVQRENGTVSR-[4,4,0,1,0]	4.46%	QNQCFYNSSYLNVQRENGTVSR-[4,4,0,1,0]	0.27%	1104.7291 (4+)	1.22E+07	√	√	x	x	x	0%	0%	x	5.88E+06	√	√	x	x	x	0%	0%	x
	QDQCIYNTTYLNVQR-[6,5,0,3,0]	1.94%	QDQCIYNTTYLNVQR-[6,5,0,3,0]	0.01%	1194.9868 (4+)	6.04E+06	√	√	√	√	√	100%	57%	√	1.73E+06	√	√	√	√	√	43%	14%	√
	NEEYNK-[5,4,0,2,0]	2.77%	Not found	0.00%	751.0409 (4+)	4.74E+06	√	√	√	√	√	100%	80%	√	0.00E+00	x	x	x	x	x	0%	0%	x
	ENGTISR-[7,6,0,4,0]	1.18%	ENGTISR-[7,6,0,4,0]	0.00%	1074.1627 (4+)	3.97E+06	√	√	√	√	√	100%	83%	√	0.00E+00	x	x	x	x	x	0%	0%	x
	SVQEIQATFFYFTPNKTEDTIFLR-[7,6,0,4,0]	1.91%	Not found	0.00%	1283.3416 (5+)	3.80E+06	√	√	√	√	√	74%	70%	√	0.00E+00	x	x	x	x	x	0%	0%	x
	ENGTVSR-[7,6,0,4,0]	0.24%	Not found	0.00%	1070.6586 (4+)	6.95E+05	√	√	√	√	√	100%	33%	√	0.00E+00	x	x	x	x	x	0%	0%	x
	ENGTISR-[6,5,1,3,0]	0.19%	Not found	0.00%	1261.8232 (3+)	2.90E+05	√	√	√	√	x	17%	0%	x	0.00E+00	x	x	x	x	x	0%	0%	x
Hemagglutinin	NGSYPNLSK-[5,4,1,1,0]	0.02%	NGSYPNLSK-[5,4,1,1,0]	0.01%	1013.7474 (3+)	1.04E+06	√	√	√	√	√	88%	63%	√	4.80E+05	√	√	√	√	√	100%	63%	x
	NVTVTHSVNLLEDSHNGK(1-Deamidated)-[7,6,1,3,0] NGSYPNLSKSYVNNK(2Deamidated)-[8,8,3,1,0]	0.02%	NVTVTHSVNLLEDSHNGK(1-Deamidated)-[7,6,1,3,0] NGSYPNLSKSYVNNK(2Deamidated)-[8,8,3,1,0]	0.03%	1068.2338 (5+)	6.38E+05	√	√	√	√	√	88%	88%	√	8.60E+04	x	x	x	x	x	0%	0%	x
	GFGSGIITSNASMDECDTK-[7,6,1,3,0]	0.03%	GFGSGIITSNASMDECDTK-[7,6,1,3,0]	0.01%	1341.2684 (4+)	6.35E+05	√	√	√	√	√	83%	61%	√	8.00E+04	x	x	x	x	x	0%	0%	x
	SWSYIAETPNSENGTCYPGYFADYEELR-[7,6,1,3,0] or NGSYPNLSKSYVNNKEK1Deamidated-14,8,0,3,0] or NGSYPNLSKSYVNNK2Deamidated-[12,13,3,0,0]	0.22%	SWSYIAETPNSENGTCYPGYFADYEELR-[7,6,1,3,0] or NGSYPNLSKSYVNNKEK1Deamidated-14,8,0,3,0] or NGSYPNLSKSYVNNK2Deamidated-[12,13,3,0,0]	0.01%	1339.1270 (5+)	3.09E+05	√	√	√	√	√	81%	59%	√	3.23E+04	x	x	x	x	x	0%	0%	x
	GFGSGIITSNASMDECDTK-[7,6,1,4,0]	0.06%	GFGSGIITSNASMDECDTK-[7,6,1,4,0]	0.01%	1131.4390 (5+)	1.68E+05	√	√	√	√	√	78%	67%	√	2.27E+04	x	x	x	x	x	0%	0%	x

[a]The features used in identification are listed in the table as yes/present (check mark) or no/absent (×). Abundances refer to the composite values of all charge states identified and matched for a given glycopeptide, presented as a percent of the total volume of all matched compounds. The absolute precursor ion abundances were determined from extracted ion chromatograms. Peptide backbone coverage was reported as percent product ions detected of total possible peptide product ions. Glycopeptide compositions in red indicate false/incorrect match. Glycopeptide nomenclature has been described in the text, and the peptide sequences are listed in Supporting Information Table S-1. Detailed mass tables confirming glycopeptide assignments have been presented in Supporting Information Section S-4. Glycopeptide identifiers are listed as peptide-[a,b,c,d,e], where a = number of hexoses; b = number of *N*-acetylhexosamines; c = number of deoxyhexoses; d = number of *N*-acetylneuraminic acids; e = number of *N*-glycolylneuraminic acids.

dx.doi.org/10.1021/pr500506z | *J. Proteome Res.* 2014, 13, 4347−4355

assignment of the glycopeptides. In order to demonstrate that the detection of peptide backbone product ions was a universal phenomenon, we analyzed glycopeptides with different peptide backbones and varying absolute abundances from each of the three glycoproteins studied. Table 2 summarizes the types of product ions observed using HILC-C18 vs C18 LC−tandem MS. Mass tables with observed and calculated fragment ion masses are shown in Supporting Information Section S-4.

For all glycopeptide assignments that were confirmed by tandem MS, the following criteria were used: (1) intact glycopeptide mass match, (2) presence of oxonium ions, (3) presence of either an abundant protonated peptide ion and/or a peptide + HexNAc ion, and (4) presence of significant peptide backbone product ion coverage. It is significant that some of the glycopeptide compositions that were assigned to high-abundance precursor ions based on intact mass that also produced oxonium ions were rejected (shown in red font) because the tandem MS stub glycopeptide or peptide backbone ions were not consistent with the assignment. This emphasizes the need for tandem MS on glycopeptides for confident assignments. Although the ions shown in Table 2 were quite abundant in the mixtures analyzed, many were not selected for data-dependent tandem MS when using C18-MS due to the presence of more abundant nonglycosylated peptides. In addition, GlycReSoft failed to find and match some of these compositions in C18 data, which was probably due to very low abundances or presence of overlapping isotopic peaks. HILIC-C18-MS improved the abundances of glycopeptides relative to nonglycosylated peptides, which allowed these ions to be matched by the GlycReSoft software. Data-dependent acquisition of tandem mass spectra led to systematic identification of these glycopeptides and allowed confident assignments due to the presence of peptide backbone product ions. The abundant product ions containing stub glycopeptides and oxonium ions were useful for confirming peptide identities.

Results for Transferrin. Transferrin presents two glycosylation sites, with the most abundant *N*-linked glycan of composition [5,4,0,2,0] at each of those two sequons. In agreement with this, glycopeptides QQQHLFGSNVTDCSGNFCLFR-[5,4,0,2,0] and CGLVPVLAENYNK-[5,4,0,2,0] could be identified using tandem MS in both HILIC-C18 and C18 analyses. Other transferrin glycopeptides, although identified by GlycReSoft in the C18 LC−MS data, did not get selected for data-dependent tandem MS due to presence of abundant nonglycosylated peptides. However, these glycopeptides underwent tandem MS and could be confidently assigned in case of HILIC-C18.

Results for Alpha-1-acid Glycoprotein. Compared to transferrin, which has only two glycosylation sites and limited glycan heterogeneity, AGP is more complex with five different glycosylation sites and differences in peptide backbone arising from genetic variants. The HILIC-C18 method was significantly more effective at acquiring data-dependent tandem mass spectra of sufficient quality to enable identification of the peptide backbone. We observed formation of useful peptide backbone ions, even for precursor ions with moderate abundances (~1−2 × 10^5 counts). Tandem mass spectra were useful in discriminating single amino acid genetic variants; for example, in Table 2, two AGP glycopeptides have similar peptide sequences, ENGTVSR and ENGTISR, for which we observed complete peptide sequences in the glycopeptide tandem mass spectra of precursors 1070.6586 (4+) and 1074.1627 (4+). These peptides were not retained using the C18-MS system.

Results for Influenza A Virus Hemagglutinin. HA presents a considerably greater analytical challenge than AGP. Due to the presence of 9 putative glycosylation sites with a wide distribution of possible glycan compositions at each site, the number of theoretical glycopeptides (the search space) was more than 10-fold greater than that of AGP, as shown in Table 1. In addition, the number of glycopeptide glycoforms detected was greater, and as a result of higher glycoform heterogeneity, the corresponding relative abundances of each glycopeptide precursor ion were generally lower than those observed for transferrin and AGP. In Table 2, we compared 5 different glycopeptide compositions that were identified by GlycReSoft in both HILIC-C18 and C18 data from HA analyses. Out of the 5 glycopeptides assigned by intact mass, only 1, NGSYPNLSK-[5,4,1,1,0], underwent tandem MS with C18 analysis. However, in the case of HILIC-C18, useful tandem MS data were acquired on all 5 glycopeptides, leading to formation of not only peptide backbone ions but also backbone ions with an attached HexNAc, facilitating identification of the glycosylation site. This was important in case of NGSYPNLSK-[5,4,1,1,0], where two glycosylation sequons, NGS and NLS, were present on the same glycopeptide; HILIC-C18 tandem MS data allowed us to identify NLS as the occupied sequon, as shown in Supporting Information Table S-3(c) and Figure 3. Tandem MS was also useful in identifying other modifications. For example, the exact site of deamidation could be identified in case of HA glycopeptide NVTVTHSVNLLEDSH**N**GK(1-Deamidation)-[7,6,1,3,0].

As shown in Table 1, measurement of glycopeptide mass did not suffice to identify the peptide and glycan composition because of the size of the search space. Thus, observed masses consistent with more than one glycopeptide were not uncommon. We showed that tandem MS could be used to resolve ambiguities. As shown in Table 2, HA precursor ion *m*/*z* 1339.1270 (5+) matched three different compositions, SWSYIAETPNSENGTCYPGYFADYEELR-[7,6,1,3,0], NGSYPNLSKSYVNNKEK (1 deamidation)-[14,8,0,3,0], and NGSYPNLSKSYVNNK (2 deamidation)-[12,13,3,0,0], within a 10 ppm mass error tolerance. Peptide backbone and stub glycopeptide ions proved useful in assigning SWSYIAETPNSENGTCYPGYFADYEELR-[7,6,1,3,0] as the correct composition. In case of AGP, SVQEIQATFFYFTPNK-[7,6,0,4,0] or QNQCFYNSSYLNVQRENGTVSR-[6,6,0,2,0] also corresponded to the same precursor mass 1088.2441 (5+), and SVQEIQATFFYFTPNK-[7,6,0,4,0] was assigned as the correct composition based on tandem MS data. The fact that HILIC-C18-MS resulted in relative precursor ion abundances sufficient to allow selection for tandem MS and consequently make these assignments demonstrates the value of this approach.

Confident Site-Specific Glycan Profiling

We used the glycopeptides assignments to construct a site-specific glycosylation map for HA and AGP (Figures 4 and S-3, respectively). Glycan compositions were classified, based on the number of HexNAc units, as high-mannose (2 HexNAc), biantennary (3 or 4 HexNAc), triantennary (5 HexNAc), tetra-antennary (6 HexNAc), and penta-antennary (7 HexNAc). Only glycopeptides identified with at least 40% peptide backbone coverage from tandem MS data were included in the glycan profile maps. In some cases, particularly for HA, two putative glycosylation sites appeared on a single tryptic peptide.

dx.doi.org/10.1021/pr500506z | *J. Proteome Res.* 2014, 13, 4347−4355

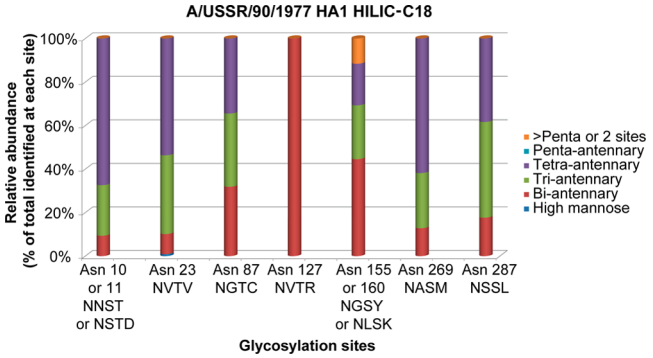


Figure 4. Site-specific glycan profile of hemagglutinin from influenza A virus.

The peptide backbone fragment ions with an attached HexNAc helped resolve the ambiguity in such cases by helping to identify the occupied site. Such site-specific glycan profile assignments were possible only when using HILIC-C18-MS/MS, whereby at least 7 times more glycopeptides underwent tandem-MS, from evaluation of precursors generating oxonium ions. It is evident from the data that the usefulness of HILIC-C18-MS scales with the number of glycosylation sites and glycopeptides in the theoretical search space. For transferrin, where the glycopeptide and glycoform diversity is very limited, the C18 and HILIC-C18 platforms yield similar results. However, with AGP and hemagglutinin, there is an obvious improvement in data quality with the use of HILIC-C18 over C18. We made these assignments using assumptions regarding the protein sequences, glycosylation sites, glycoform distributions, and post-translational modifications. It is, therefore, possible that modified peptides not included in our assumptions exist. Since the goal of this study was to compare performances of the HILIC-C18 and C18 systems, our glycopeptide analyses sufficed to reach clear conclusions despite the absence of exhaustive discovery-mode proteomics data interpretation.

CONCLUSIONS

We assert that even for purified glycoproteins, a tandem MS step is necessary in order to confidently assign glycopeptide compositions with respect to peptide and glycan. Because the traditional data-dependent mass spectrometry algorithms used in proteomics select the most abundant precursor ions, most glycopeptides are not selected for tandem MS when using C18 chromatography. We therefore developed a method in which glycopeptides are enriched using a HILIC trapping column online prior to a C18 analytical column. Using this method, we observed that glycopeptides were the most abundant ions in the data sets and were therefore preferentially selected for data-dependent LC−MS/MS. The HILIC trapping column provides efficient enrichment of glycopeptides. The C18 analytical column provides high chromatographic resolution. Thus, we demonstrated that online enrichment and separation can be conveniently performed without manual sample manipulation. This combination enables facile glycopeptide enrichment, separation, and analysis with online LC−MS/MS using a commercially available instrument system.

CAD-based tandem MS of glycopeptides is useful for determining peptide sequence and glycan composition; it does not determine glycan structure in detail. We found that by programming the MS system to acquire CAD tandem MS using elevated collision energy, we observed peptide backbone product ions for the majority of the glycopeptides selected for tandem MS. Even for the lowest abundance glycopeptide precursor ions so selected, we observed glycan-modified peptide stubs that were useful for confirming the mass of the peptide moiety. These tandem mass spectra improve the informational value of an LC−MS/MS experiment, from the point-of-view of glycopeptide analysis. In principle, however, use of online enrichment of glycopeptides with HILIC-C18-MS will improve tandem MS results for any dissociation method.

ASSOCIATED CONTENT

Supporting Information

Section S-1: Details of sample preparation and LC−MS/MS methods. Figure S-1: Oxonium ion distributions overlaid with base peak chromatograms, of glycoprotein digests, analyzed by C18 and HILIC-C18 LC−MS/MS. Figure S-2: Representative tandem mass spectra of glycopeptides. Figure S-3: Site-specific glycan profile of AGP. Figure S-4: HILIC-C18-MS of transferrin digest on 2.1 mm scale column setup. Tables S-1 and S-2: Theoretical lists of peptides and glycopeptides used for searching LC−MS data (downloadable spreadsheets). Table S-3: Mass-lists showing theoretical and observed product ion masses from manual glycopeptide product ion matching. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Phone: 617-638-6762. Fax: 617-638-6761. E-mail: jzaia@bu.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Funding for this work was provided by NIH grant P41GM104603.

REFERENCES

(1) Boersema, P. J.; Geiger, T.; Wisniewski, J. R.; Mann, M. Quantification of the *N*-glycosylated secretome by super-SILAC during breast cancer progression and in human blood samples. *Mol. Cell. Proteomics* **2013**, *12*, 158−171.

(2) Rudd, P. M.; Endo, T.; Colominas, C.; Groth, D.; Wheeler, S. F.; Harvey, D. J.; Wormald, M. R.; Serban, H.; Prusiner, S. B.; Kobata, A.; Dwek, R. A. Glycosylation differences between the normal and pathogenic prion protein isoforms. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 13044−13049.

(3) Turner, G. A. *N*-Glycosylation of serum-proteins in disease and its investigation using lectins. *Clin. Chim. Acta* **1992**, *208*, 149−171.

(4) National Research Council (U.S.). Committee on Assessing the Importance and Impact of Glycomics and Glycosciences; *Transforming Glycoscience: A Roadmap for the Future*; National Academies Press: Washington, DC, 2012.

(5) Leymarie, N.; Zaia, J. Effective use of mass spectrometry for glycan and glycopeptide structural analysis. *Anal. Chem.* **2012**, *84*, 3040−3048.

(6) Ivens, A. C.; Peacock, C. S.; Worthey, E. A.; Murphy, L.; Aggarwal, G.; Berriman, M.; Sisk, E.; Rajandream, M. A.; Adlem, E.; Aert, R.; Anupama, A.; Apostolou, Z.; Attipoe, P.; Bason, N.; Bauser, C.; Beck, A.; Beverley, S. M.; Bianchettin, G.; Borzym, K.; Bothe, G.; Bruschi, C. V.; Collins, M.; Cadag, E.; Ciarloni, L.; Clayton, C.; Coulson, R. M.; Cronin, A.; Cruz, A. K.; Davies, R. M.; De Gaudenzi, J.; Dobson, D. E.; Duesterhoeft, A.; Fazelina, G.; Fosker, N.; Frasch, A.

C.; Fraser, A.; Fuchs, M.; Gabel, C.; Goble, A.; Goffeau, A.; Harris, D.; Hertz-Fowler, C.; Hilbert, H.; Horn, D.; Huang, Y.; Klages, S.; Knights, A.; Kube, M.; Larke, N.; Litvin, L.; Lord, A.; Louie, T.; Marra, M.; Masuy, D.; Matthews, K.; Michaeli, S.; Mottram, J. C.; Muller-Auer, S.; Munden, H.; Nelson, S.; Norbertczak, H.; Oliver, K.; O'Neil, S.; Pentony, M.; Pohl, T. M.; Price, C.; Purnelle, B.; Quail, M. A.; Rabbinowitsch, E.; Reinhardt, R.; Rieger, M.; Rinta, J.; Robben, J.; Robertson, L.; Ruiz, J. C.; Rutter, S.; Saunders, D.; Schafer, M.; Schein, J.; Schwartz, D. C.; Seeger, K.; Seyler, A.; Sharp, S.; Shin, H.; Sivam, D.; Squares, R.; Squares, S.; Tosato, V.; Vogt, C.; Volckaert, G.; Wambutt, R.; Warren, T.; Wedler, H.; Woodward, J.; Zhou, S.; Zimmermann, W.; Smith, D. F.; Blackwell, J. M.; Stuart, K. D.; Barrell, B.; Myler, P. J. The genome of the kinetoplastid parasite, *Leishmania major*. *Science* **2005**, *309*, 436−42.

(7) Choudhary, C.; Mann, M. Decoding signalling networks by mass spectrometry-based proteomics. *Nat. Rev. Mol. Cell Biol.* **2010**, *11*, 427−439.

(8) Kirsch, S.; Zarei, M.; Cindric, M.; Muthing, J.; Bindila, L.; Peter-Katalinic, J. On-line nano-HPLC/ESI QTOF MS and tandem MS for separation, detection, and structural elucidation of human erythrocytes neutral glycosphingolipid mixture. *Anal. Chem.* **2008**, *80*, 4711−22.

(9) Thaysen-Andersen, M.; Packer, N. H. Site-specific glycoproteomics confirms that protein structure dictates formation of *N*-glycan type, core fucosylation and branching. *Glycobiology* **2012**, *22*, 1440−1452.

(10) Bern, M.; Kil, Y. J.; Becker, C. Byonic: advanced peptide and protein identification software. *Curr. Protoc. Bioinf.* **2012**, *40*, 13.20.1−13.20.14.

(11) Leymarie, N.; Griffin, P. J.; Jonscher, K.; Kolarich, D.; Orlando, R.; McComb, M.; Zaia, J.; Aguilan, J.; Alley, W. R.; Altmann, F.; Ball, L. E.; Basumallick, L.; Bazemore-Walker, C. R.; Behnken, H.; Blank, M. A.; Brown, K. J.; Bunz, S.-C.; Cairo, C. W.; Cipollo, J. F.; Daneshfar, R.; Desaire, H.; Drake, R. R.; Go, E. P.; Goldman, R.; Gruber, C.; Halim, A.; Hathout, Y.; Hensbergen, P. J.; Horn, D. M.; Hurum, D.; Jabs, W.; Larson, G.; Ly, M.; Mann, B. F.; Marx, K.; Mechref, Y.; Meyer, B.; Möginger, U.; Neusüss, C.; Nilsson, J.; Novotny, M. V.; Nyalwidhe, J. O.; Packer, N. H.; Pompach, P.; Reiz, B.; Resemann, A.; Rohrer, J. S.; Ruthenbeck, A.; Sanda, M.; Schulz, J. M.; Schweiger-Hufnagel, U.; Sihlbom, C.; Song, E.; Staples, G. O.; Suckau, D.; Tang, H.; Thaysen-Andersen, M.; Viner, R. I.; An, Y.; Valmu, L.; Wada, Y.; Watson, M.; Windwarder, M.; Whittal, R.; Wuhrer, M.; Zhu, Y.; Zou, C. Interlaboratory study on differential analysis of protein glycosylation by mass spectrometry: the ABRF glycoprotein research multi-institutional study 2012. *Mol. Cell. Proteomics* **2013**, *12*, 2935−2951.

(12) Palmisano, G.; Parker, B. L.; Engholm-Keller, K.; Lendal, S. E.; Kulej, K.; Schulz, M.; Schwämmle, V.; Graham, M. E.; Saxtorph, H.; Cordwell, S. J.; Larsen, M. R. A novel method for the simultaneous enrichment, identification, and quantification of phosphopeptides and sialylated glycopeptides applied to a temporal profile of mouse brain development. *Mol. Cell. Proteomics* **2012**, *11*, 1191−1202.

(13) Wuhrer, M.; Catalina, M. I.; Deelder, A. M.; Hokke, C. H. Glycoproteomics based on tandem mass spectrometry of glycopeptides. *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.* **2007**, *849*, 115−28.

(14) Ongay, S.; Boichenko, A.; Govorukhina, N.; Bischoff, R. Glycopeptide enrichment and separation for protein glycosylation analysis. *J. Sep. Sci.* **2012**, *35*, 2341−2372.

(15) 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*, 5598−5609.

(16) Wuhrer, M.; de Boer, A. R.; Deelder, A. M. Structural glycomics using hydrophilic interaction chromatography (HILIC) with mass spectrometry. *Mass Spectrom. Rev.* **2009**, *28*, 192−206.

(17) Wilson, S. R.; Jankowski, M.; Pepaj, M.; Mihailova, A.; Boix, F.; Truyols, G. V.; Lundanes, E.; Greibrokk, T. 2D LC separation and determination of bradykinin in rat muscle tissue dialysate with on-line SPE-HILIC-SPE-RP-MS. *Chromatographia* **2007**, *66*, 469−474.

(18) Boersema, P. J.; Divecha, N.; Heck, A. J. R.; Mohammed, S. Evaluation and optimization of ZIC-HILIC-RP as an alternative MudPIT strategy. *J. Proteome Res.* **2007**, *6*, 937−946.

(19) Desaire, H.; Hua, D. When can glycopeptides be assigned based solely on high-resolution mass spectrometry data? *Int. J. Mass Spectrom.* **2009**, *287*, 21−26.

(20) Dallas, D. C.; Martin, W. F.; Hua, S.; German, J. B. Automated glycopeptide analysis—review of current state and future directions. *Briefings Bioinf.* **2013**, *14*, 361−374.

(21) Carr, S. A.; Huddleston, M. J.; Bean, M. F. Selective identification and differentiation of N- and O-linked oligosaccharides in glycoproteins by liquid chromatography-mass spectrometry. *Protein Sci.* **1993**, *2*, 183−196.

(22) Mayampurath, A.; Yu, C.-Y.; Song, E.; Balan, J.; Mechref, Y.; Tong, H. Computational framework for identification of intact glycopeptides in complex samples. *Anal. Chem.* **2014**, *86*, 453−463.

(23) Catalina, M. I.; Koeleman, C. A.; Deelder, A. M.; Wuhrer, M. Electron transfer dissociation of *N*-glycopeptides: loss of the entire *N*-glycosylated asparagine side chain. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 1053−61.

(24) Alley, W. R.; Mechref, Y.; Novotny, M. V. Characterization of glycopeptides by combining collision-induced dissociation and electron-transfer dissociation mass spectrometry data. *Rapid Commun. Mass Spectrom.* **2009**, *23*, 161−170.

(25) An, Y.; Rininger, J. A.; Jarvis, D. L.; Jing, X.; Ye, Z.; Aumiller, J. J.; Eichelberger, M.; Cipollo, J. F. Comparative glycomics analysis of influenza hemagglutinin (H5N1) produced in vaccine relevant cell platforms. *J. Proteome Res.* **2013**, *12*, 3707−3720.

(26) Scott, N. E.; Parker, B. L.; Connolly, A. M.; Paulech, J.; Edwards, A. V. G.; Crossett, B.; Falconer, L.; Kolarich, D.; Djordjevic, S. P.; Hojrup, P.; Packer, N. H.; Larsen, M. R.; Cordwell, S. J. 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* **2011**, *10*, M000031-MCP201.

(27) Segu, Z. M.; Mechref, Y. Characterizing protein glycosylation sites through higher-energy C-trap dissociation. *Rapid Commun. Mass Spectrom.* **2010**, *24*, 1217−1225.

(28) Nelson, M. I.; Holmes, E. C. The evolution of epidemic influenza. *Nat. Rev. Genet.* **2007**, *8*, 196−205.

(29) Klenk, H. D.; Wagner, R.; Heuer, D.; Wolff, T. Importance of hemagglutinin glycosylation for the biological functions of influenza virus. *Virus Res.* **2002**, *82*, 73−75.

(30) Abe, Y.; Takashita, E.; Sugawara, K.; Matsuzaki, Y.; Muraki, Y.; Hongo, S. Effect of the addition of oligosaccharides on the biological activities and antigenicity of influenza A/H3N2 virus hemagglutinin. *J. Virol.* **2004**, *78*, 9605−9611.

(31) An, Y.; Rininger, J. A.; Jarvis, D. L.; Jing, X.; Ye, Z.; Aumiller, J. J.; Eichelberger, M.; Cipollo, J. F. Comparative glycomics analysis of influenza Hemagglutinin (H5N1) produced in vaccine relevant cell platforms. *J. Proteome Res.* **2013**, *12*, 3707−20.

(32) Jaitly, N.; Mayampurath, A.; Littlefield, K.; Adkins, J. N.; Anderson, G. A.; Smith, R. D. Decon2LS: An open-source software package for automated processing and visualization of high resolution mass spectrometry data. *BMC Bioinf.* **2009**, *10*, 87.

(33) Maxwell, E.; Tan, Y.; Tan, Y.; Hu, H.; Benson, G.; Aizikov, K.; Conley, S.; Staples, G. O.; Slysz, G. W.; Smith, R. D.; Zaia, J. GlycReSoft: a software package for automated recognition of glycans from LC/MS data. *PLoS One* **2012**, *7*, e45474.

(34) Zhao, P.; Viner, R.; Teo, C. F.; Boons, G.-J.; Horn, D.; Wells, L. Combining high-energy C-trap dissociation and electron transfer dissociation for protein O-GlcNAc modification site assignment. *J. Proteome Res.* **2011**, *10*, 4088−4104.

(35) Yamashita, K.; Ideo, H.; Ohkura, T.; Fukushima, K.; Yuasa, I.; Ohno, K.; Takeshita, K. Sugar chains of serum transferrin from patients with carbohydrate-deficient glycoprotein syndrome—evidence of asparagine-*N*-linked oligosaccharide transfer deficiency. *J. Biol. Chem.* **1993**, *268*, 5783−5789.

(36) Treuheit, M. J.; Costello, C. E.; Halsall, H. B. Analysis of the five glycosylation sites of human alpha 1-acid glycoprotein. *Biochem. J.* **1992**, *283*, 105−112.