



Published in final edited form as:

*J Am Soc Mass Spectrom.* 2017 September ; 28(9): 1751–1764. doi:10.1007/s13361-017-1701-4.

## Electron-Transfer/Higher-Energy Collision Dissociation (EThcD)-enabled Intact Glycopeptide/Glycoproteome Characterization

Qing Yu<sup>1</sup>, Bowen Wang<sup>2</sup>, Zhengwei Chen<sup>3</sup>, Go Urabe<sup>2</sup>, Matthew S. Glover<sup>1,4</sup>, Xudong Shi<sup>2</sup>, Lian-Wang Guo<sup>2</sup>, K. Craig Kent<sup>5</sup>, and Lingjun Li<sup>1,3,4,\*</sup>

<sup>1</sup>School of Pharmacy, University of Wisconsin, Madison, WI 53705, USA

<sup>2</sup>Department of Surgery, Wisconsin Institute for Medical Research, Madison, WI 53705, USA

<sup>3</sup>Department of Chemistry, University of Wisconsin, Madison, WI 53706, USA

<sup>4</sup>Cardiovascular Research Center Training Program in Translational Cardiovascular Science, University of Wisconsin-Madison, Madison, WI 53705, USA

<sup>5</sup>The Ohio State University Wexner Medical Center, Columbus, OH 43210, USA

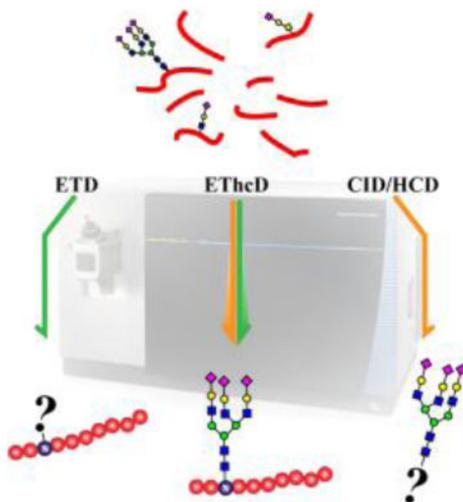
### Abstract

Protein glycosylation, one of the most heterogeneous post-translational modifications, can play a major role in cellular signal transduction and disease progression. Traditional mass spectrometry (MS)-based large-scale glycoprotein sequencing studies heavily rely on identifying enzymatically released glycans and their original peptide backbone separately, as there is no efficient fragmentation method to produce unbiased glycan and peptide product ions simultaneously in a single spectrum and can be conveniently applied to high throughput glycoproteome characterization, especially for *N*-glycopeptides which can have much more branched glycan side chains than relatively less complex *O*-linked glycans. In this study a re-defined electron-transfer/higher-energy collision dissociation (EThcD) fragmentation scheme is applied to incorporate both glycan and peptide fragments in one single spectrum, enabling complete information to be gathered and great microheterogeneity details to be revealed. Fetuin was first utilized to prove the applicability with 19 glycopeptides and corresponding 5 glycosylation sites identified. Subsequent experiments tested its utility for human plasma *N*-glycoproteins. Large-scale studies explored *N*-glycoproteomics in rat carotid over the course of restenosis progression to investigate potential role of glycosylation. The integrated fragmentation scheme provides a powerful tool for the analysis of intact *N*-glycopeptides and *N*-glycoproteomics. We also anticipate this approach can be readily applied to large-scale *O*-glycoproteome characterization.

### Graphical Abstract

\*Address reprint requests to: Dr. Lingjun Li, School of Pharmacy & Department of Chemistry, University of Wisconsin, 777 Highland Ave, Madison, WI 53705. lingjun.li@wisc.edu. Phone: (608)265-8491, Fax: (608)262-5345.

Supplemental Information: Additional information available as noted in the text.



## Introduction

Protein glycosylation is the covalent attachment of complex carbohydrates or oligosaccharides, collectively called glycans, to specific amino acid residues of the polypeptide backbone of proteins. Being one of the most common post-translational modifications (PTMs), glycosylation is estimated to occur in more than half of all eukaryotic proteins [1] with two major types being *N*-glycosylation and *O*-glycosylation [2]. It is involved in biological processes such as cell adhesion, signaling, inflammatory response as well as a wide variety of pathological states [2–4].

The ability to accurately characterize the structure of glycoproteins is of great importance to unravel the diverse functions of glycosylation, especially for *N*-glycosylation that can be highly heterogeneous. However, *N*-glycosylated proteins are most often highly complex by presenting a multitude of diverse sugar–amino acid combinations and characterization with MS remains a great analytical challenge due to this heterogeneity [4, 5]. Structural characterization encompasses *N*-glycoprotein/glycopeptide identification, locations of attachment sites, and evaluation of glycosylation site micro-heterogeneity. Most of the previous *N*-glycoproteomic studies involved enzymatic removal of glycans, and separate identifications of glycans and amino acid sequence [6–10]. Although there are multiple well-established protocols, detaching glycans from their peptide backbones accompanies loss of multiple layers of information (i.e., amino acid site-specific glycoform and micro-heterogeneity) [11, 12].

The ideal way to comprehensively characterize glycopeptides, especially in a large-scale glycoproteomics context, is to preserve the glycan side chain on the peptide and collect glycan and peptide fragmentation information simultaneously [13, 14]. However, due to its heterogeneous nature, fragmenting an intact glycopeptide has always been difficult and yet no single fragmentation technique is able to generate a complete picture in a single MS/MS spectrum [6, 7, 14, 15]. Peaks resulting from glycosidic bond cleavages dominate spectra generated by collision-induced dissociation (CID) with little knowledge of glycosylation

sites and amino acid sequences, whereas c/z-ion series in ETD type of experiments yield the glycosylation site and peptide identity with little information on glycan side chain composition [16, 17]. Higher-energy collision dissociation produces abundant diagnostic oxonium ions and partial glycopeptide information as the collision energy is more evenly distributed, enabling intact glycopeptide identification in some cases, most of which are studies on purified target proteins [16–18]. Although there have been multiple reports showing the capabilities of using CID or ETD alone to sequence sugar-modified peptides, most often the data were collected only with a few protein standards that were modified by less complex glycation or *O*-glycosylation. It is not applicable to large-scale characterization of much more branched and complex glycosylated peptides, such as *N*-glycopeptides where glycan fragments dominate MS/MS spectra, and it is especially challenging to work with complex proteome samples [19–24]. An alternative approach for a more detailed characterization of glycopeptides combines MS/MS and MS<sup>3</sup> experiments with CID or HCD. In this approach, the glycopeptide ion is selected and fragmented resulting in a variety of fragment ions predominantly due to the cleavage of glycosidic linkages. The peptide ion carrying a single HexNAc, is subjected to a second ion isolation/fragmentation cycle, resulting in fragmentation of the peptide moiety [25–28]. However, this also requires prior knowledge of the targeted peptides to select for MS<sup>3</sup>, thus limiting its throughput [25–28]. Several other reports alternate collision energies and collect sequential MS/MS on the same precursor, based on the observation that lower energy preferentially fragment glycans and higher energy can fragment peptide backbones [29, 30]. However, acquiring or averaging sequential scans significantly slows down instruments and special custom-made bioinformatics tools are required to process such data. As a consequence, the authors have only demonstrated the utility with a few protein standards, but have not yet been able to apply to real complex samples [29, 30]. Accurate mass matching has also been employed in the studies involving glycoprotein standards; however, its applicability to complex mixtures is questionable due to the existence of thousands of possibilities for one *m/z* value [31, 32]. Additional studies identify *N*-glycopeptides with HCD using the spectra acquired for intact glycopeptides as well as de-glycopeptides to explore the heterogeneity assisted by a custom-made software platform [33, 34]. Yet its successful application still relies on extra enzymatic PNGase-F treatment and requires dedicated data processing tools.

To overcome these limitations, researchers have built a multifaceted method, running sequential HCD (or CID) and ETD in hope of assembling pieces of information from both fragmentation techniques [35–37]. Later on, a more targeted approach, the product-triggered method took advantage of the abundant oxonium ion upon HCD by triggering subsequent ETD events only if certain oxonium ions were recognized in the prior HCD MS/MS [38, 39]. This method allowed more instrument time to be spent on glycopeptides by avoiding non-glycosylated peptides to be analyzed by MS/MS. Despite providing complementary data, sequential HCD and ETD acquisitions consume more duty cycle and require complex post-acquisition data processing since two types of information need to be considered respectively. Frese et al. combined ETD with HCD and developed a hybrid dissociation method, termed EThcD [40]. A supplemental energy is applied to all ions formed by ETD to generate more informative spectra [41, 42]. It also showed great promise in glycopeptide studies [43, 44].

In this study, we established a redefined EThcD approach for improved analysis of glycopeptides, mainly *N*-glycopeptides. Instead of using calibrated ETD reaction times and using HCD as a supplemental energy to improve ETD spectra populated with mostly *c/z* ions as Parker et al presented [43], our approach integrated HCD as the second dimension of fragmentation. It enabled both HCD and ETD type of fragments to be collected in a single spectrum, enabling improved instrumental speed by greatly shortening ETD reaction times for each charge state [43], complete product ion generation, consisting of glycan fragments, *c/z* and *b/y* ions from peptide backbones and therefore intact glycopeptide characterization. Following the method validation with fetuin, a slightly more complex human serum sample was studied where 331 unique *N*-glycopeptides were sequenced and 66 glycosylation sites mapped. Large scale experiments explored *N*-glycoproteomics from rat carotids collected over the course of restenosis progression with over 2000 glycopeptides identified. Detailed information on microheterogeneity was gained to guide further biological studies.

## Experimental Section

### Materials

Fetuin from fetal bovine serum, concanavalin A (ConA), wheat germ agglutinin(WGA), *Ricinuscommunis* agglutinin (RCA<sub>120</sub>), iodoacetamide (IAA), *N*-acetyl-D-glucosamine, D-lactose, methyl  $\alpha$ -D-mannopyranoside and manganese dichloride were obtained from Sigma-Aldrich (St. Louis, MO). Tris base, urea, sodium chloride, human serum and calcium chloride were obtained from Fisher Scientific (Pittsburgh, PA). C18 OMIX tips were obtained from Agilent (Santa Clara, CA). Dithiothreitol (DTT) and sequencing grade trypsin were supplied from Promega (Madison, WI). Hydrophilic interaction chromatography material was obtained from PolyLC (Columbia, MD).

### Rat Carotid Artery Sample Collection

To induce restenosis in experimental model, carotid artery balloon angioplasty was performed in male Sprague–Dawley rats (Charles River; 350–400 g) as previously described [45]. Briefly, rats were anesthetized with isoflurane (5% for inducing and 2.5% for maintaining anesthesia). A longitudinal incision was made in the neck and carotid arteries were exposed. A 2-F balloon catheter (Edwards Lifesciences, Irvine, CA) was inserted through an arteriotomy on the left external carotid artery. To produce arterial injury, the balloon was inflated at a pressure of 2 atm and withdrawn to the carotid bifurcation and this action was repeated three times. The external carotid artery was then permanently ligated, and blood flow was resumed. At day 3, day 7, and day 14 post surgery, injured carotid arteries as well as uninjured ones of ~1.5cm in length were collected following systemic perfusion of the animals with cold PBS solution. Peri-adventitial tissues were carefully dissected and excised in cold PBS, and remaining blood in the lumen was cleaned. The arteries were then immediately snap-frozen in liquid nitrogen and stored in cryogenic vials until ready for downstream processing. For protein extraction, frozen arterial segments were pulverized with micro-pestle over liquid nitrogen bath.

## Trypsin Digestion

Fetuin and human serum proteins were dissolved in 8 M urea, reduced (5 mM DTT, 1 h at room temperature) and alkylated (15 mM IAA, 30 min at room temperature in the dark). Sample was diluted with 50 mM Tris-HCl (pH=8) to lower urea to 0.9 M, trypsin was added in a 1:50 (w/w) ratio and incubated for 18 h at 37°C. Digestion was quenched with 10% TFA to a final concentration of 0.3%. Rat carotids were processed with filter-aided sample preparation (FASP) technique [46]. Details can be found in Supplemental Information.

## Lectin Enrichment

Human plasma glycopeptide enrichment was performed with a modified filter-aided protocol [47–49]. Tryptic plasma peptides were loaded onto Microcon filters YM-30 with 80  $\mu$ L binding buffer (1 mM  $\text{CaCl}_2$ , 1 mM  $\text{MnCl}_2$ , 0.5 M NaCl in 20 mM TrisHCl, pH 7.3). 90 mg ConA, 90 mg WGA, and 71.5 mg RCA<sub>120</sub> were mixed in 36  $\mu$ L 2  $\times$  binding buffer and loaded onto the filter unit. After incubation for 1 h, the unbound peptides were eluted by centrifugation at 14,000 xg at 18°C for 10 min. The captured peptides were washed four times with 200  $\mu$ L binding buffer and then eluted two times with elution solution (300 mM *N*-acetyl-D-glucosamine, D-lactose, methyl  $\alpha$ -D-mannopyranoside in 200  $\mu$ L binding buffer). Final solution was acidified with 10% TFA and desalted using C18 OMIX tips. Tips were first wetted with ACN and equilibrated with water containing 0.1% TFA. Samples were applied onto tip, washed with water containing 0.1% TFA and then eluted with 50% and 75% aqueous ACN containing 0.1% TFA. Samples were dried down and loaded on LC-MS.

## HILIC Enrichment

Rat carotid glycopeptides were enriched using HILIC beads (PolyHYDROXYETHYL A, PolyLC Inc., Columbia, MD) following a previously reported procedure with minor modification [50]. HILIC beads were first activated with 200  $\mu$ L of elution buffer (0.1% TFA, 99.9% H<sub>2</sub>O) for 30 min and then washed with binding buffer twice. About 140  $\mu$ g tryptic peptides mixture was dissolved in 300  $\mu$ L of binding buffer (0.1% TFA, 19.9% H<sub>2</sub>O, 80% ACN) and mixed with 7 mg activated ZIC-HILIC resin at a 1:50 peptide-to-material mass ratio in a microcentrifuge tube. The tube was shaken over a vortex mixer for 1 h and supernatant was removed by centrifugation. The beads were washed with 70  $\mu$ L binding buffer (6X) and glycopeptides were eluted with 70  $\mu$ L elution buffer (5X).

## NanoLC MS/MS Analysis

The desalted peptide mixtures were analyzed on the Orbitrap Fusion<sup>TM</sup> Lumos<sup>TM</sup> Tribrid<sup>TM</sup> Mass Spectrometer (Thermo Fisher Scientific, San Jose, CA) coupled to a Dionex UPLC system. A binary solvent system composed of H<sub>2</sub>O containing 0.1% formic acid (A) and MeCN containing 0.1% formic acid (B) was used for all analyses. Peptides were loaded and separated on a 75  $\mu$ m  $\times$  15 cm homemade column packed with 1.7  $\mu$ m, 150 Å, BEH C18 material obtained from a Waters UPLC column (part no. 186004661). A gradient ramping 3% to 30% solvent B in 30 min was used at a flow rate of 300 nL/min to separate bovine fetuin digest. A same solvent composition with an extended 120 min gradient was used to separate enriched human serum glycopeptides.

The mass spectrometer was operated in data dependent mode to automatically switch between MS and MS/MS acquisition. Survey full scan MS spectra (from  $m/z$  300 to 1800) were acquired in the Orbitrap with resolution of 120,000 at  $m/z$  200. The ten precursors with the highest charge states were selected for MS/MS in an order of intensity.

Three consecutive scans with CID, ETD and EThcD were acquired on the same precursor respectively for bovine fetuin peptides. CID was carried out at collision energy at 35% in the ion trap. ETD was performed with calibrated charge dependent reaction time [51](Table S1) supplemented by 15% HCD activation and ion species was analyzed in the Orbitrap at resolution of 60,000 FWHM (at  $m/z$  200). EThcD was performed with user defined charge dependent reaction time (Table S1) supplemented by 33% HCD activation and ion species was analyzed in the Orbitrap at resolution of 60,000 FWHM (at  $m/z$  200).  $1.0e^4$  and  $3.0e^5$  were set as MS/MS AGC target for ion trap and Orbitrap scans respectively. HCD alone experiments with normal human serum was performed with a 33% normalized collision energy while the other parameters remained changed as used in EThcD experiments.

Lectin-enriched human serum glycopeptides and HILIC-enriched rat carotid glycopeptides were only analyzed in the EThcD mode with user defined charge dependent reaction times (Table S1) supplemented by 33% HCD activation.

## Data Analysis

All spectra were analyzed with Byonic (Protein Metrics, San Carlos, CA)[52] with corresponding protein databases (i.e., bovine fetuin or UniProt *Homo sapiens* proteome (April 12, 2016; 16764 entries) or *Rattus norvegicus* proteome (August 27, 2016; 36991 entries)). Trypsin was selected with maximum two missed cleavages allowed. Static modifications consisted of carbamidomethylation of cysteine residues (+57.0215 Da) and dynamic oxidation of methionine (+15.9949 Da).

Glycan searches were conducted on fetuin, all with default mammalian *N*- and *O*-glycan databases. Enriched serum glycopeptide spectra were searched against human serum glycan database or human glycan database respectively. Glycopeptides from carotid arteries were searched with mammalian *N*-glycan database. Precursor ion tolerance of 10 ppm and a product ion mass tolerance of 0.01 Da were allowed. Results were filtered at 1% FDR. Fetuin data was filtered with Byonic score>300 while plasma and rat carotid artery data were filtered with the normalized  $|\text{Log Prob}|>2$  and further validation was performed manually.

## Results and Discussion

Traditional glycoproteomic studies usually opt to enzymatically release glycans from their attached peptides, due to their heterogeneous nature and thereby not amenable to a single fragmentation method [6, 14, 16]. However, certain information is lost during such a process and creates a gap between glycomics, proteomics and glycoproteomics. A universal MS method is in great demand and has driven multiple studies to characterize intact glycopeptides rather than breaking them apart. In recent years, sequential MS/MS acquisition on the same precursor with multiple fragmentation methods (i.e., CID/HCD/

ETD) has become a popular approach as the consensus exists that CID/HCD primarily produces glycan products whereas ETD cleaves peptide backbones [31, 35, 43, 53]. However, running sequential scans undoubtedly lowers instrument duty cycle and therefore sensitivity and sequence coverage, especially with ETD scans consuming milliseconds [51, 54]. In addition, it often requires tedious sample preparation and complex data processing since a complete picture is divided and distributed among several spectra [31, 55]. To address these existing limitations, this study aims to employ a hybrid EThcD approach to characterize intact glycopeptides in a one-spectrum fashion with much improved instrument duty cycle.

### Identification of Glycopeptides from Trypsin Digest of Bovine Fetuin

Analysis of a rather simple mixture of modified and native peptides was performed using an enzymatic digest of fetuin, which is highly glycosylated [56, 57]. To illustrate the difference between CID, ETD with supplemental HCD activation and our EThcD approach, three sequential MS/MS scans were recorded on the same precursor with each of the three activation types (Figure 1). As annotated, most of the dominant peaks formed during CID fragmentation of glycopeptide KLCPCPLLLAPLNDNR resulted from the cleavage of glycosidic bonds without breaking amide bonds. Additionally, diagnostic oxonium ions at the lower mass region were lost during fragmentation, which arose from decreased stability of fragment ions with  $m/z$  values less than 30% of the  $m/z$  for the precursor peptide selected for fragmentation in a linear ion trap device, commonly known as the “one-third rule”. Therefore, as stated above, for glycopeptides CID should preferentially be used for glycan sequencing but not for glycosylation site determination and amino acid sequencing. ETD spectrum was also recorded (Figure 1a, inset), using calibrated ETD reaction time plus 15% HCD supplemental activation. However, barely any sequence information was provided, likely due to inefficiency of ETD on doubly or triply charged ions. In sharp contrast to CID and ETD, our EThcD scheme generated a variety of product ions, originating from cleaving both amide bonds and glycosidic bonds (Figure 1b). The spectrum was manually deconvoluted with Xtract (Thermo Fisher Scientific, San Jose, CA) to reveal more fragments. As manually annotated, CID and ETD provided 22 pieces of fragments altogether with neither of them informative enough to complete the sequencing, whereas EThcD alone revealed over 30 unique fragments with much more balanced information for sequencing the glycopeptide. The sequence was also verified by Byonic search.

Using this strategy, various glycoforms of *N*- and *O*-glycopeptides were identified (Table 1) and assigned with their MS/MS. All glycopeptide spectrum matches were filtered to ensure mass accuracy better than 10 ppm and scores above 300 [52]. The observed glycoforms were in good agreement with previous reports [31, 57–61]. Even though some of these prior studies reported a few additional glycoforms, it is worth noting that these studies required extra experiments, extensive fractionation, extended acquisition time, accurate mass matching or a lot more materials to work with, whereas our current method has greatly simplified workflow only requiring trypsin digestion. Compared to CID and ETD, our approach provided improved accuracy and sensitivity towards intact glycopeptide sequencing and site localization.

## Identification of *N*-Glycopeptides Enriched from Trypsin Digestion of Normal Human Serum

We also applied our method to a large-scale *N*-glycopeptide identification with lectin enriched human serum, in order to further investigate its utility. Regular HCD and our EThcD were first evaluated in terms of their effectiveness in identifying glycopeptides. Data was searched against human plasma glycan database (57 glycan entities). Results from a duplicate comparison clearly suggested that EThcD outperformed HCD, by enabling more unique glycoform identifications as well as glycosylation sites (Figure 2a, b). Even some of their identifications overlapped, EThcD provided better spectral quality which is more informative and therefore allowing more confident sequence determination (Figure 2c, Figure 3).

Presented in Figure 3 are spectra of the same glycopeptide with different charge states and fragmentation methods. Although HCD permits successful sequencing (Figure 3a), it was quite obvious that EThcD outperforms HCD with detection of more complete ion series (Figure 3b) for 3+ charged precursors. Apart from several oxonium ions ( $m/z$  163.06, 204.08 and 366.14) and signature HexNAc fragments ( $m/z$  138.05, 168.06 and 186.07), there were only two ions with one or two HexNAc still attached in HCD, whereas EThcD spectrum recorded cleavage at every single glycosidic bond and entire glycan sidechain products. This was even more obvious with EThcD on 4+ charged precursor (Figure 3c). In general, EThcD peptide spectral matches (PSMs) were higher scored than HCD (Figure 2c). Overall, after filtering with several criteria (1% FDR, Log Prob>3 and mass error <10 ppm), EThcD enabled successful identifications of 184 unique glycoforms, 49 glycosylation sites from 29 proteins and HCD identified 106 unique glycoforms, 34 glycosylation sites from 23 proteins (Table S2).

As there are several glycan databases available in Byonic, we reasoned that it is worth comparing a couple of them in terms of any influence that different databases may have on identifications [52]. With identical searching parameters, our data was searched against the human plasma *N*-glycan database (HPD) which included 57 glycan entities and human *N*-glycan database (HD) which included 182 glycan entities respectively. HPD resulted in 223 unique glycoforms that met our filter criteria, whereas that number with HD was 288. Analysis of the results revealed 170 glycoforms identified by both databases at first, 53 being unique with HPD and 118 with HD. However, further investigation revealed that ten unique glycoforms from each side were ambiguously assigned different glycan sequences, though the peptide backbone stayed the same. As shown in Figure 4c, d, HexNAc(5)Hex(6)NeuAc(2) was assigned by HPD and HexNAc(5)Hex(5)Fuc(1)NeuAc(2) was assigned by HD, mass difference (~16 Da) was attributed to methionine oxidation. A closer look at the spectrum suggested that database searching against HD not only enabled assignment of more product ions, but also, and most importantly, annotated  $y_{11}$  and  $z_{11}$  ions that contained oxidized methionine, increasing the confidence of assignment. Furthermore, Jia et al reported this particular site to be core fucosylated, which was in good agreement with HD assignment [62]. This ambiguity happened because certain glycan entities were not included in HPD. Comparing Byonic score ratios indicated all these ten ambiguous assignments tended to have better annotations with HD (Figure 4b) and therefore enriching

glycan database would certainly increase search confidence in general. After manually removing ambiguity, 180 glycoforms were shared between HPD and HD. The rest of the 43 unique ones with HPD were primarily (24 in 43) due to sodium adducts not included in HD. On the other hand, 100 out of 108 unique ones were solely sequenced with HD since those glycan entities were not included in HPD.

All *N*-glycopeptides from human plasma had their glycans on 66 different asparagine residues from 41 proteins (Table 2). Most of these glycosylation sites (~91%) were very well documented by UniProt and previous publications [62–68]. It also resulted in six novel *N*-glycosylation sites that required further validation. Moreover, one of the major advantages of sequencing intact glycopeptides rather than sequencing glycans and peptides separately is the capability of comprehensive characterization of glycoforms. Instead of only knowing what glycans are in the sample and what potential peptides they are attached to, we are able to unequivocally list multiple glycoforms with individual glycosite mapped.

### Identification of *N*-Glycopeptides Enriched from Trypsin Digestion of Rat Carotid Arteries

Restenosis is the re-narrowing of a blood vessel following surgical interventions that aim to treat stenosis. While the underlying mechanisms have been studied from multiple perspectives [69, 70], the *N*-glycoproteomics has not been very well characterized, despite that *N*-glycosylation is critical for correct folding, stability and mediating cell attachment [71]. Extracellular matrix (ECM) proteins has long been implicated in the pathogenesis of restenosis, and the fact that more than 90% of ECM proteins are glycosylated makes it more relevant to study carotid artery glycoproteomics [71, 72]. We harvested rat carotid arteries along the restenosis progression at 0, 3, 7, 14 day time point and enriched *N*-glycopeptides with a HILIC method. Four biological replicates were analyzed and glycopeptides were only counted if they were identified in at least two replicates with  $|\text{Log Prob}| > 2$ .

Combining results from all four time points resulted in the identification of total 2092 glycopeptides, 387 glycosylation sites and 226 glycoproteins. By analyzing data from each time point, we managed to observe temporal dynamics of glycosylation events along with the progression of the restenotic lesion, with a gradual increase in glycosylation in the early phase of the vascular responses post angioplasty (i.e., from day 0 to day 7), followed by a decrease in the late phase of restenosis progression (i.e., from day 7 to day 14) (Figure 5a, b). Indeed, it has been widely acknowledged that the acute and subacute window following angioplasty features the most active cellular events in the vessel wall, such as cell proliferation, migration, and apoptosis, and day 7 post procedure is frequently used for marking the peak of vascular smooth muscle cells' proliferative state [73]. Subsequently, vascular cells become relatively quiescent, and restenosis progression reaches a plateau during the late phase of vascular repair post injury, which would explain the observed reduction of general glycosylations during the transition from day 7 to day 14. As intact glycopeptides were analyzed, details about microheterogeneity were gained with multiple glycoforms mapped on the same glycosylation site. While most peptides (80.0%) identified had less than 5 glycan compositions at a given site, there were a few (8.6%) that came with more than 10 glycan compositions (Figure 5c). For example, 83 unique glycan masses were identified at lumican Asn 127 (Table S3) from four time points, suggesting a highly diverse

microheterogeneity. Specifically, 76 glycoforms were observed in healthy rats and after a sharp drop to 53 at day 3, the number gradually went up to 59 at day 14, suggesting varied extent of glycosylation and possible functional modulation. Lumican is an extracellular matrix protein extensively modified by keratan sulfate in cornea [74] but occurs predominantly as a glycoprotein with little or no sulfate in arteries [75–77]. Therefore, the large number of glycoforms could be derived from various lengths of lactosaminoglycan with or without fucose and NeuAc. Some of these glycopeptides co-eluting with identical retention time might also be derived from in-source fragmentation. The accumulation of lumican deposition has long been established as an indicator for cardiovascular diseases such as atherosclerosis, restenosis, and aneurysm [78, 79]. Detailed information on its primary structure with glycosylation would surely help elucidate associated functions. Another interesting protein was LRP1, the homologues to human prolow-density lipoprotein receptor-related protein 1. Twenty-two glycosites were identified and they were either previously documented sites or identical sites as in human LRP1 [80, 81]. LRP1 is a key suppressor in preventing atherosclerosis and restenosis [82, 83] and *N*-linked glycosylation is known to influence its proper folding, and more importantly, its resistance to  $\gamma$ -secretase [84, 85]. LRP1's atheroprotective function heavily relies on its cleavage by  $\gamma$ -secretase, as its anti-inflammatory signaling is entirely mediated by its intracellular domain cleaved and released from the membrane bound form of intact LRP1. Previous studies have suggested that hyperglycosylation, particularly *N*-linked glycosylation, would render LRP1 resistant to  $\gamma$ -secretase [86]. We observed a hyperglycosylation trend with the largest number of glycoforms being identified at day 7, indicating that LRP1 protein folding pattern as well as resistance to secretase could be altered as restenosis progressed. Further biological study is underway to determine how these differentially glycosylated proteins could affect diseased phenotypes of vascular cells, and ultimately evaluate their potential as intervention targets for restenosis treatment.

## Conclusions

Enabled by the recently introduced EThcD fragmentation technique and bioinformatics tool, we manage to move one step forward towards large-scale characterizing intact glycopeptides. While ETD, CID/HCD have their own merits, we develop a re-defined EThcD approach, combining their unique features, in pursuit of intact glycopeptide sequencing. This hybrid approach features much more informative spectra as both amide bond and glycosidic bond are cleaved, yielding dual information about peptide sequence and residue-specific glycosite attachment simultaneously. Duty cycle is improved as EThcD approach takes less than half of the reaction time that traditional ETD uses and only one spectrum is required, instead of acquiring several consecutive CID/HCD and ETD scans in order to collect sufficient information. This one-spectrum approach also accompanies simplified data processing by avoiding combining information from multiple different scans. Its feasibility is demonstrated with simple fetuin standard, complex human plasma and rat carotid artery samples. Identification of previously documented glycosylation proves its reliability and identification of over a thousand *N*-glycopeptides in rat carotid artery during restenosis progression demonstrates its utility in large-scale glycoproteomics studies and provides novel targets for future investigation. Effects of using databases with different *N*-

glycan compositions/inclusions are explored. As expected, more comprehensive glycan database will facilitate more accurate and reliable glycopeptide discovery. We believe that such a workflow can be coupled with the HCD-product dependent-ETHcD function to further improve sensitivity and specificity by performing ETHcD only if diagnostic oxonium ions are detected in the prior HCD scan. The integration of these advanced workflows will accelerate the pace of in-depth glycopeptide mapping.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

This research was supported in part by the National Institutes of Health grants R21AG055377, R01 DK071801 (to LL), R01 HL133665 (to LWG), and NIH R01 HL068673 (to KCK). The Orbitrap instruments were purchased through the support of an NIH shared instrument grant (NIH-NCRR S10RR029531) and Office of the Vice Chancellor for Research and Graduate Education at the University of Wisconsin-Madison. LL acknowledges a Vilas Distinguished Achievement Professorship with funding provided by the Wisconsin Alumni Research Foundation and University of Wisconsin-Madison School of Pharmacy. MSG acknowledges a postdoctoral fellowship supported by the National Institutes of Health, under Ruth L. Kirschstein National Research Service Award T32 HL 007936 from the National Heart Lung and Blood Institute to the University of Wisconsin-Madison Cardiovascular Research Center. We would like to acknowledge Dr. Andrew Alpert from PolyLC Inc. for generous gift of the PolyHYDROXYETHYL A material.

## References

1. Apweiler R, Hermjakob H, Sharon N. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochimica et biophysica acta*. 1999; 1473:4–8. [PubMed: 10580125]
2. Spiro RG. Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology*. 2002; 12:43R–56R.
3. Haltiwanger RS, Lowe JB. Role of glycosylation in development. *Annu Rev Biochem*. 2004; 73:491–537. [PubMed: 15189151]
4. Moremen KW, Tiemeyer M, Nairn AV. Vertebrate protein glycosylation: diversity, synthesis and function. *Nature reviews Molecular cell biology*. 2012; 13:448–462. [PubMed: 22722607]
5. Varki, A. 2. Cold Spring Harbor Laboratory Press; Cold Spring Harbor, N.Y.: 2009.
6. Thaysen-Andersen M, Packer NH. Advances in LC-MS/MS-based glycoproteomics: getting closer to system-wide site-specific mapping of the N- and O-glycoproteome. *Biochimica et biophysica acta*. 2014; 1844:1437–1452. [PubMed: 24830338]
7. Lazar IM, Deng J, Ikenishi F, Lazar AC. Exploring the glycoproteomics landscape with advanced MS technologies. *Electrophoresis*. 2015; 36:225–237. [PubMed: 25311661]
8. Chen W, Smeekens JM, Wu R. A universal chemical enrichment method for mapping the yeast N-glycoproteome by mass spectrometry (MS). *Mol Cell Proteomics*. 2014; 13:1563–1572. [PubMed: 24692641]
9. Tian Y, Zhou Y, Elliott S, Aebersold R, Zhang H. Solid-phase extraction of N-linked glycopeptides. *Nat Protoc*. 2007; 2:334–339. [PubMed: 17406594]
10. Sun S, Shah P, Eshghi ST, Yang W, Trikannad N, Yang S, Chen L, Aiyetan P, Hoti N, Zhang Z, Chan DW, Zhang H. Comprehensive analysis of protein glycosylation by solid-phase extraction of N-linked glycans and glycosite-containing peptides. *Nat Biotechnol*. 2016; 34:84–88. [PubMed: 26571101]
11. Kolarich D, Jensen PH, Altmann F, Packer NH. Determination of site-specific glycan heterogeneity on glycoproteins. *Nat Protoc*. 2012; 7:1285–1298. [PubMed: 22678432]

12. Kaji H, Yamauchi Y, Takahashi N, Isobe T. Mass spectrometric identification of N-linked glycopeptides using lectin-mediated affinity capture and glycosylation site-specific stable isotope tagging. *Nat Protocols*. 2007; 1:3019–3027.
13. Zhu Z, Desaire H. Carbohydrates on Proteins: Site-Specific Glycosylation Analysis by Mass Spectrometry. *Annu Rev Anal Chem (Palo Alto Calif)*. 2015; 8:463–483. [PubMed: 26070719]
14. 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. *Mol Cell Proteomics*. 2016; 15:1773–1790. [PubMed: 26929216]
15. Alley WR Jr, Mann BF, Novotny MV. High-sensitivity analytical approaches for the structural characterization of glycoproteins. *Chem Rev*. 2013; 113:2668–2732. [PubMed: 23531120]
16. Dalpathado DS, Desaire H. Glycopeptide analysis by mass spectrometry. *Analyst*. 2008; 133:731–738. [PubMed: 18493671]
17. Leymarie N, Zaia J. Effective use of mass spectrometry for glycan and glycopeptide structural analysis. *Anal Chem*. 2012; 84:3040–3048. [PubMed: 22360375]
18. 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; 23:124–140. [PubMed: 22083589]
19. Jahouh F, Hou SJ, Kovac P, Banoub JH. Determination of the glycation sites of Bacillus anthracis neoglycoconjugate vaccine by MALDI-TOF/TOF-CID-MS/MS and LC-ESI-QqTOF-tandem mass spectrometry. *J Mass Spectrom*. 2011; 46:993–1003. [PubMed: 22012665]
20. Demian WL, Kottari N, Shiao TC, Randell E, Roy R, Banoub JH. Direct targeted glycation of the free sulfhydryl group of cysteine residue (Cys-34) of BSA. Mapping of the glycation sites of the anti-tumor Thomsen-Friedenreich neoglycoconjugate vaccine prepared by Michael addition reaction. *J Mass Spectrom*. 2014; 49:1223–1233. [PubMed: 25476939]
21. Zhu Z, Su X, Clark DF, Go EP, Desaire H. Characterizing O-linked glycopeptides by electron transfer dissociation: fragmentation rules and applications in data analysis. *Anal Chem*. 2013; 85:8403–8411. [PubMed: 23909558]
22. Chandler KB, Pompach P, Goldman R, Edwards N. Exploring Site-Specific N-Glycosylation Microheterogeneity of Haptoglobin Using Glycopeptide CID Tandem Mass Spectra and Glycan Database Search. *J Proteome Res*. 2013; 12:3652–3666. [PubMed: 23829323]
23. Bourgoin-Voillard S, Leymarie N, Costello CE. Top-down tandem mass spectrometry on RNase A and B using a Qh/FT-ICR hybrid mass spectrometer. *Proteomics*. 2014; 14:1174–1184. [PubMed: 24687996]
24. Nicolardi S, van der Burgt YE, Dragan I, Hensbergen PJ, Deelder AM. Identification of new apolipoprotein-CIII glycoforms with ultrahigh resolution MALDI-FTICR mass spectrometry of human sera. *J Proteome Res*. 2013; 12:2260–2268. [PubMed: 23527852]
25. Giangrande C, Auberger N, Rentier C, Papini AM, Mallet JM, Lavielle S, Vinh J. Multi-Stage Mass Spectrometry Analysis of Sugar-Conjugated beta-Turn Structures to be Used as Probes in Autoimmune Diseases. *J Am Soc Mass Spectrom*. 2016; 27:735–747. [PubMed: 26729456]
26. Hsiao HH, Urlaub H. Pseudo-neutral-loss scan for selective detection of phosphopeptides and N-glycopeptides using liquid chromatography coupled with a hybrid linear ion-trap/orbitrap mass spectrometer. *Proteomics*. 2010; 10:3916–3921. [PubMed: 20925059]
27. Segu ZM, Mechref Y. Characterizing protein glycosylation sites through higher-energy C-trap dissociation. *Rapid Commun Mass Sp*. 2010; 24:1217–1225.
28. Medzihradzky KF, Gillece-Castro BL, Settineri CA, Townsend RR, Masiarz FR, Burlingame AL. Structure determination of O-linked glycopeptides by tandem mass spectrometry. *Biomedical & environmental mass spectrometry*. 1990; 19:777–781. [PubMed: 1708302]
29. Hinneburg H, Stavenhagen K, Schweiger-Hufnagel U, Pengelley S, Jabs W, Seeberger PH, Silva DV, Wuhler M, Kolarich D. The Art of Destruction: Optimizing Collision Energies in Quadrupole-Time of Flight (Q-TOF) Instruments for Glycopeptide-Based Glycoproteomics. *J Am Soc Mass Spectrom*. 2016; 27:507–519. [PubMed: 26729457]
30. Cao L, Tolic N, Qu Y, Meng D, Zhao R, Zhang QB, Moore RJ, Zink EM, Lipton MS, Paga-Tolic L, Wu S. Characterization of intact N- and O-linked glycopeptides using higher energy collisional dissociation. *Anal Biochem*. 2014; 452:96–102. [PubMed: 24440233]

31. Ye HP, Boyne MT, Buhse LF, Hill J. Direct Approach for Qualitative and Quantitative Characterization of Glycoproteins Using Tandem Mass Tags and an LTQ Orbitrap XL Electron Transfer Dissociation Hybrid Mass Spectrometer. *Anal Chem.* 2013; 85:1531–1539. [PubMed: 23249142]
32. Zhou W, Yao N, Yao G, Deng C, Zhang X, Yang P. Facile synthesis of aminophenylboronic acid-functionalized magnetic nanoparticles for selective separation of glycopeptides and glycoproteins. *Chemical communications.* 2008:5577–5579. [PubMed: 18997957]
33. Yang W, Shah P, Toghi Eshghi S, Yang S, Sun S, Ao M, Rubin A, Jackson JB, Zhang H. Glycoform analysis of recombinant and human immunodeficiency virus envelope protein gp120 via higher energy collisional dissociation and spectral-aligning strategy. *Anal Chem.* 2014; 86:6959–6967. [PubMed: 24941220]
34. Cheng K, Chen R, Seebun D, Ye ML, Figeys D, Zou HF. Large-scale characterization of intact N-glycopeptides using an automated glycoproteomic method. *Journal of proteomics.* 2014; 110:145–154. [PubMed: 25182382]
35. Zhao P, Viner R, Teo CF, Boons GJ, 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. [PubMed: 21740066]
36. Darula Z, Sherman J, Medzihradzky KF. How to dig deeper? Improved enrichment methods for mucin core-1 type glycopeptides. *Mol Cell Proteomics.* 2012; 11:O111.016774.
37. Hogan JM, Pitteri SJ, Chrisman PA, McLuckey SA. Complementary structural information from a tryptic N-linked glycopeptide via electron transfer ion/ion reactions and collision-induced dissociation. *J Proteome Res.* 2005; 4:628–632. [PubMed: 15822944]
38. Singh C, Zampronio CG, Creese AJ, Cooper HJ. Higher energy collision dissociation (HCD) product ion-triggered electron transfer dissociation (ETD) mass spectrometry for the analysis of N-linked glycoproteins. *J Proteome Res.* 2012; 11:4517–4525. [PubMed: 22800195]
39. Saba J, Dutta S, EH. Increasing the Productivity of Glycopeptides Analysis by Using Higher-Energy Collision Dissociation-Accurate Mass-Product-Dependent Electron Transfer Dissociation. *International Journal of Proteomics.* 2012; 2012:7.
40. Frese CK, Altelaar AF, van den Toorn H, Nolting D, Griep-Raming J, Heck AJ, Mohammed S. Toward full peptide sequence coverage by dual fragmentation combining electron-transfer and higher-energy collision dissociation tandem mass spectrometry. *Anal Chem.* 2012; 84:9668–9673. [PubMed: 23106539]
41. Frese CK, Zhou HJ, Taus T, Altelaar AFM, Mechter K, Heck AJR, Mohammed S. Unambiguous Phosphosite Localization using Electron-Transfer/Higher-Energy Collision Dissociation (EThcD). *J Proteome Res.* 2013; 12:1520–1525. [PubMed: 23347405]
42. Mommen GP, Frese CK, Meiring HD, van Gaans-van den Brink J, de Jong AP, van Els CA, Heck AJ. Expanding the detectable HLA peptide repertoire using electron-transfer/higher-energy collision dissociation (EThcD). *Proc Natl Acad Sci U S A.* 2014; 111:4507–4512. [PubMed: 24616531]
43. 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. *Mol Cell Proteomics.* 2016; 15:141–153. [PubMed: 26537798]
44. Marino F, Bern M, Mommen GP, Leney AC, van Gaans-van den Brink JA, Bonvin AM, Becker C, van Els CA, Heck AJ. Extended O-GlcNAc on HLA Class-I-Bound Peptides. *J Am Chem Soc.* 2015; 137:10922–10925. [PubMed: 26280087]
45. Guo LW, Wang B, Goel SA, Little C, Takayama T, Shi XD, Roenneburg D, DiRenzo D, Kent KC. Halofuginone stimulates adaptive remodeling and preserves re-endothelialization in balloon-injured rat carotid arteries. *Circulation Cardiovascular interventions.* 2014; 7:594–601. [PubMed: 25074254]
46. Wisniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. *Nat Methods.* 2009; 6:359–362. [PubMed: 19377485]
47. Zielinska DF, Gnadt F, Wisniewski JR, Mann M. Precision mapping of an in vivo N-glycoproteome reveals rigid topological and sequence constraints. *Cell.* 2010; 141:897–907. [PubMed: 20510933]

48. 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. *Molecular cell*. 2012; 46:542–548. [PubMed: 22633491]
49. Deeb SJ, Cox J, Schmidt-Supprian M, Mann M. N-linked glycosylation enrichment for in-depth cell surface proteomics of diffuse large B-cell lymphoma subtypes. *Mol Cell Proteomics*. 2014; 13:240–251. [PubMed: 24190977]
50. Zhang C, Ye Z, Xue P, Shu Q, Zhou Y, Ji Y, Fu Y, Wang J, Yang F. Evaluation of Different N-Glycopeptide Enrichment Methods for N-Glycosylation Sites Mapping in Mouse Brain. *J Proteome Res*. 2016; 15:2960–2968. [PubMed: 27480293]
51. Rose CM, Rush MJ, Riley NM, Merrill AE, Kwiecien NW, Holden DD, Mullen C, Westphall MS, Coon JJ. A calibration routine for efficient ETD in large-scale proteomics. *J Am Soc Mass Spectrom*. 2015; 26:1848–1857. [PubMed: 26111518]
52. Bern M, Kil YJ, Becker C. Byonic: advanced peptide and protein identification software. *Current protocols in bioinformatics / editorial board, Andreas D. Baxevanis ... [et al.]*. 2012; Chapter 13(Unit13 20)
53. Medzihradsky KF, Kaasik K, Chalkley RJ. Tissue-Specific Glycosylation at the Glycopeptide Level. *Mol Cell Proteomics*. 2015; 14:2103–2110. [PubMed: 25995273]
54. Wu SW, Pu TH, Viner R, Khoo KH. Novel LC-MS(2) product dependent parallel data acquisition function and data analysis workflow for sequencing and identification of intact glycopeptides. *Anal Chem*. 2014; 86:5478–5486. [PubMed: 24796651]
55. Hoffmann M, Marx K, Reichl U, Wührer M, Rapp E. Site-specific O-Glycosylation Analysis of Human Blood Plasma Proteins. *Mol Cell Proteomics*. 2016; 15:624–641. [PubMed: 26598643]
56. 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. [PubMed: 20536156]
57. Peterman SM, Mulholland JJ. A novel approach for identification and characterization of glycoproteins using a hybrid linear ion trap/FT-ICR mass spectrometer. *J Am Soc Mass Spectr*. 2006; 17:168–179.
58. 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. *J Proteome Res*. 2004; 3:556–566. [PubMed: 15253437]
59. Ritchie MA, Gill AC, Deery MJ, Lilley K. Precursor ion scanning for detection and structural characterization of heterogeneous glycopeptide mixtures. *J Am Soc Mass Spectrom*. 2002; 13:1065–1077. [PubMed: 12322954]
60. Windwarder M, Altmann F. Site-specific analysis of the O-glycosylation of bovine fetuin by electron-transfer dissociation mass spectrometry. *Journal of proteomics*. 2014; 108:258–268. [PubMed: 24907489]
61. Huang LJ, Lin JH, Tsai JH, Chu YY, Chen YW, Chen SL, Chen SH. Identification of protein O-glycosylation site and corresponding glycans using liquid chromatography-tandem mass spectrometry via mapping accurate mass and retention time shift. *J Chromatogr A*. 2014; 1371:136–145. [PubMed: 25456591]
62. Jia W, Lu Z, Fu Y, Wang HP, Wang LH, Chi H, Yuan ZF, Zheng ZB, Song LN, Han HH, Liang YM, Wang JL, Cai Y, Zhang YK, Deng YL, Ying WT, He SM, Qian XH. A strategy for precise and large scale identification of core fucosylated glycoproteins. *Mol Cell Proteomics*. 2009; 8:913–923. [PubMed: 19139490]
63. Hagglund P, Matthiesen R, Elortza F, Hojrup P, Roepstorff P, Jensen ON, Bunkenborg J. An enzymatic deglycosylation scheme enabling identification of core fucosylated N-glycans and O-glycosylation site mapping of human plasma proteins. *J Proteome Res*. 2007; 6:3021–3031. [PubMed: 17636988]
64. Liu T, Qian WJ, Gritsenko MA, Camp DG 2nd, Monroe ME, Moore RJ, Smith RD. Human plasma N-glycoproteome analysis by immunoaffinity subtraction, hydrazide chemistry, and mass spectrometry. *J Proteome Res*. 2005; 4:2070–2080. [PubMed: 16335952]

65. Wang J, Zhou C, Zhang W, Yao J, Lu H, Dong Q, Zhou H, Qin L. An integrative strategy for quantitative analysis of the N-glycoproteome in complex biological samples. *Proteome science*. 2014; 12:4. [PubMed: 24428921]
66. Zhao J, Qiu WL, Simeone DM, Lubman DM. N-linked glycosylation profiling of pancreatic cancer serum using capillary liquid phase separation coupled with mass spectrometric analysis. *J Proteome Res*. 2007; 6:1126–1138. [PubMed: 17249709]
67. Nilsson J, Ruetschi U, Halim A, Hesse C, Carlsohn E, Brinkmalm G, Larson G. Enrichment of glycopeptides for glycan structure and attachment site identification. *Nat Methods*. 2009; 6:809–811. [PubMed: 19838169]
68. Clerc F, Reiding KR, Jansen BC, Kammeijer GS, Bondt A, Wuhrer M. Human plasma protein N-glycosylation. *Glycoconjugate journal*. 2015
69. Goel SA, Guo LW, Liu B, Kent KC. Mechanisms of post-intervention arterial remodelling. *Cardiovasc Res*. 2012; 96:363–371. [PubMed: 22918976]
70. Forte A, Rinaldi B, Berrino L, Rossi F, Galderisi U, Cipollaro M. Novel potential targets for prevention of arterial restenosis: insights from the pre-clinical research. *Clinical science*. 2014; 127:615–634. [PubMed: 25072327]
71. Lynch M, Barallobre-Barreiro J, Jahangiri M, Mayr M. Vascular proteomics in metabolic and cardiovascular diseases. *Journal of internal medicine*. 2016; 280:325–338. [PubMed: 26940365]
72. Yu Q, Shi X, Greer T, Lietz CB, Kent KC, Li L. Evaluation and Application of Dimethylated Amino Acids as Isobaric Tags for Quantitative Proteomics of the TGF-beta/Smad3 Signaling Pathway. *J Proteome Res*. 2016; 15:3420–3431. [PubMed: 27457343]
73. Wang B, Zhang M, Takayama T, Shi X, Roenneburg DA, Kent KC, Guo LW. BET Bromodomain Blockade Mitigates Intimal Hyperplasia in Rat Carotid Arteries. *EBioMedicine*. 2015; 2:1650–1661.
74. Chen S, Birk DE. The regulatory roles of small leucine-rich proteoglycans in extracellular matrix assembly. *The FEBS journal*. 2013; 280:2120–2137. [PubMed: 23331954]
75. Chen R, Wang F, Tan Y, Sun Z, Song C, Ye M, Wang H, Zou H. Development of a combined chemical and enzymatic approach for the mass spectrometric identification and quantification of aberrant N-glycosylation. *Journal of proteomics*. 2012; 75:1666–1674. [PubMed: 22202184]
76. Funderburgh JL, Funderburgh ML, Mann MM, Conrad GW. Arterial lumican. Properties of a corneal-type keratan sulfate proteoglycan from bovine aorta. *J Biol Chem*. 1991; 266:24773–24777. [PubMed: 1761572]
77. Hultgardh-Nilsson A, Boren J, Chakravarti S. The small leucine-rich repeat proteoglycans in tissue repair and atherosclerosis. *Journal of internal medicine*. 2015; 278:447–461. [PubMed: 26477596]
78. Onda M, Ishiwata T, Kawahara K, Wang R, Naito Z, Sugisaki Y. Expression of lumican in thickened intima and smooth muscle cells in human coronary atherosclerosis. *Experimental and molecular pathology*. 2002; 72:142–149. [PubMed: 11890723]
79. Gu GR, Wan F, Xue Y, Cheng WZ, Zheng HY, Zhao Y, Fan F, Han Y, Tong CY, Yao CL. Lumican as a novel potential clinical indicator for acute aortic dissection: A comparative study, based on multi-slice computed tomography angiography. *Experimental and therapeutic medicine*. 2016; 11:923–928. [PubMed: 26998013]
80. Li X, Jiang J, Zhao X, Wang J, Han H, Zhao Y, Peng B, Zhong R, Ying W, Qian X. N-glycoproteome analysis of the secretome of human metastatic hepatocellular carcinoma cell lines combining hydrazide chemistry, HILIC enrichment and mass spectrometry. *Plos One*. 2013; 8:e81921. [PubMed: 24324730]
81. Apweiler R, Bairoch A, Wu CH, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M, Martin MJ, Natale DA, O'Donovan C, Redaschi N, Yeh LS. UniProt: the Universal Protein knowledgebase. *Nucleic acids research*. 2004; 32:D115–119. [PubMed: 14681372]
82. Muratoglu SC, Belgrave S, Lillis AP, Migliorini M, Robinson S, Smith E, Zhang L, Strickland DK. Macrophage LRP1 suppresses neo-intima formation during vascular remodeling by modulating the TGF-beta signaling pathway. *Plos One*. 2011; 6:e28846. [PubMed: 22174911]
83. Boucher P, Gotthardt M, Li WP, Anderson RG, Herz J. LRP: role in vascular wall integrity and protection from atherosclerosis. *Science*. 2003; 300:329–332. [PubMed: 12690199]

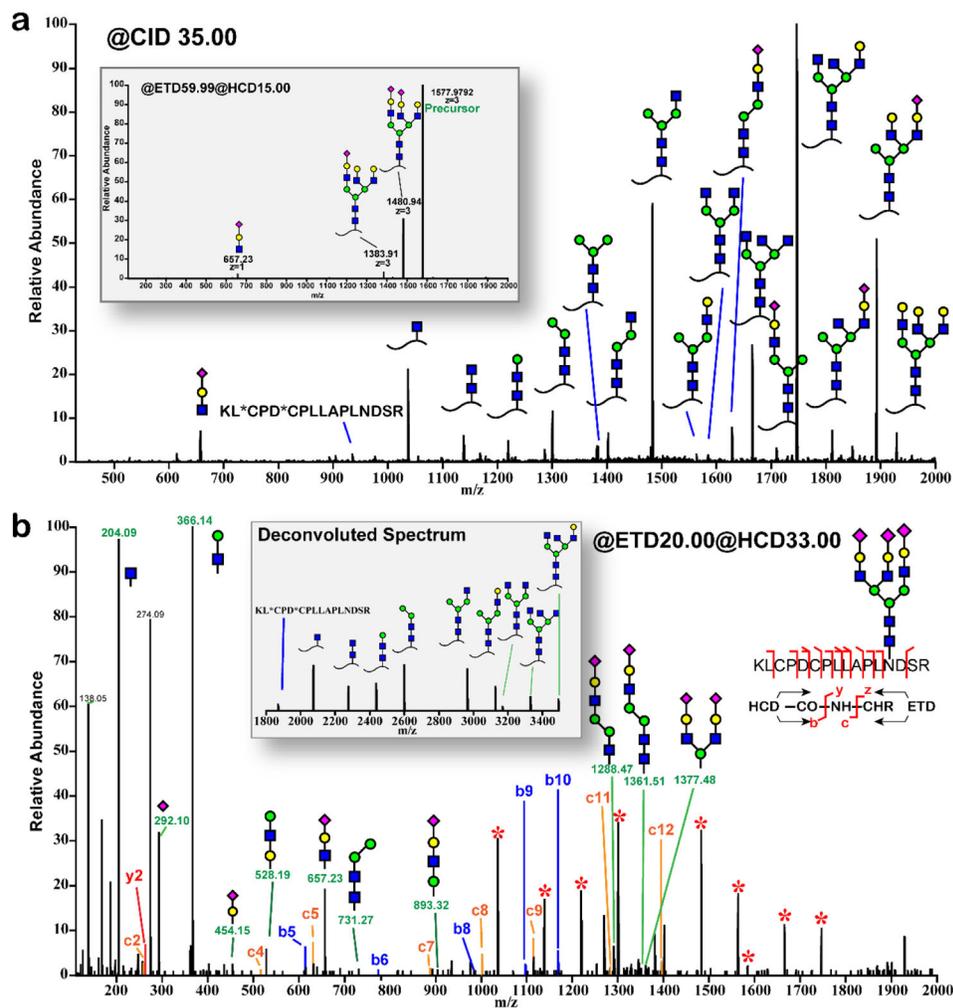
84. McCormick LM, Urade R, Arakaki Y, Schwartz AL, Bu G. Independent and cooperative roles of N-glycans and molecular chaperones in the folding and disulfide bond formation of the low-density lipoprotein (LDL) receptor-related protein. *Biochemistry*. 2005; 44:5794–5803. [PubMed: 15823038]
85. May P, Bock HH, Nimpf J, Herz J. Differential glycosylation regulates processing of lipoprotein receptors by gamma-secretase. *J Biol Chem*. 2003; 278:37386–37392. [PubMed: 12871934]
86. Zurhove K, Nakajima C, Herz J, Bock HH, May P. gamma-Secretase Limits the Inflammatory Response Through the Processing of LRP1. *Science signaling*. 2008; 1

Author Manuscript

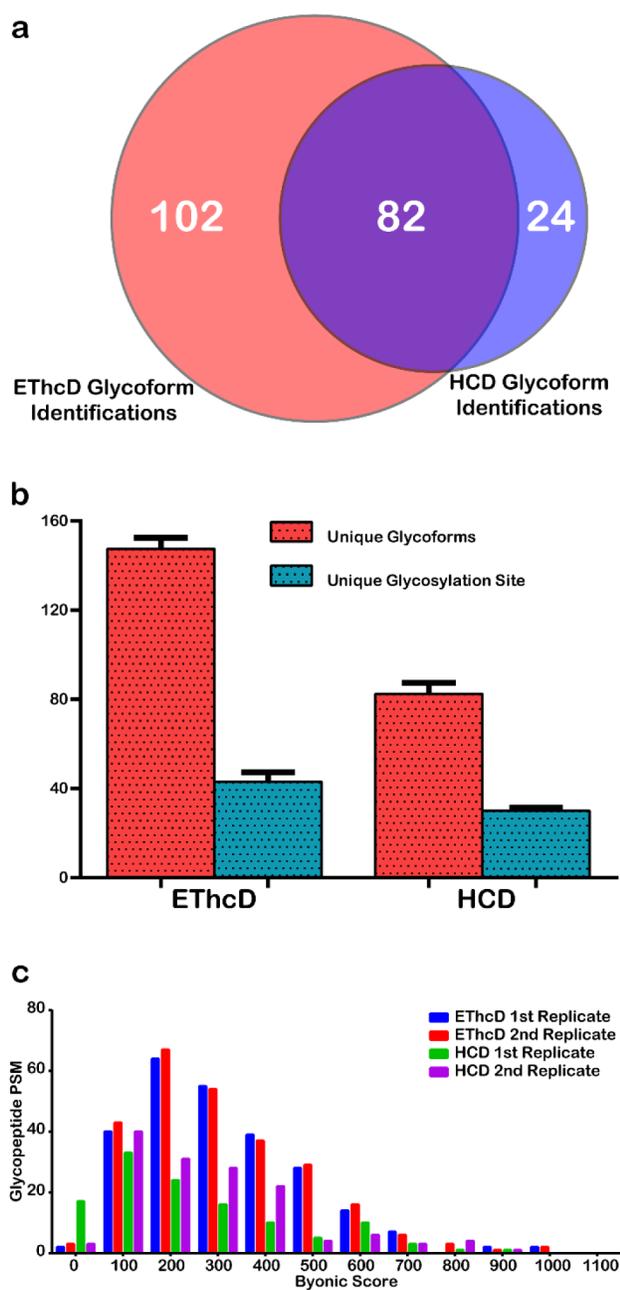
Author Manuscript

Author Manuscript

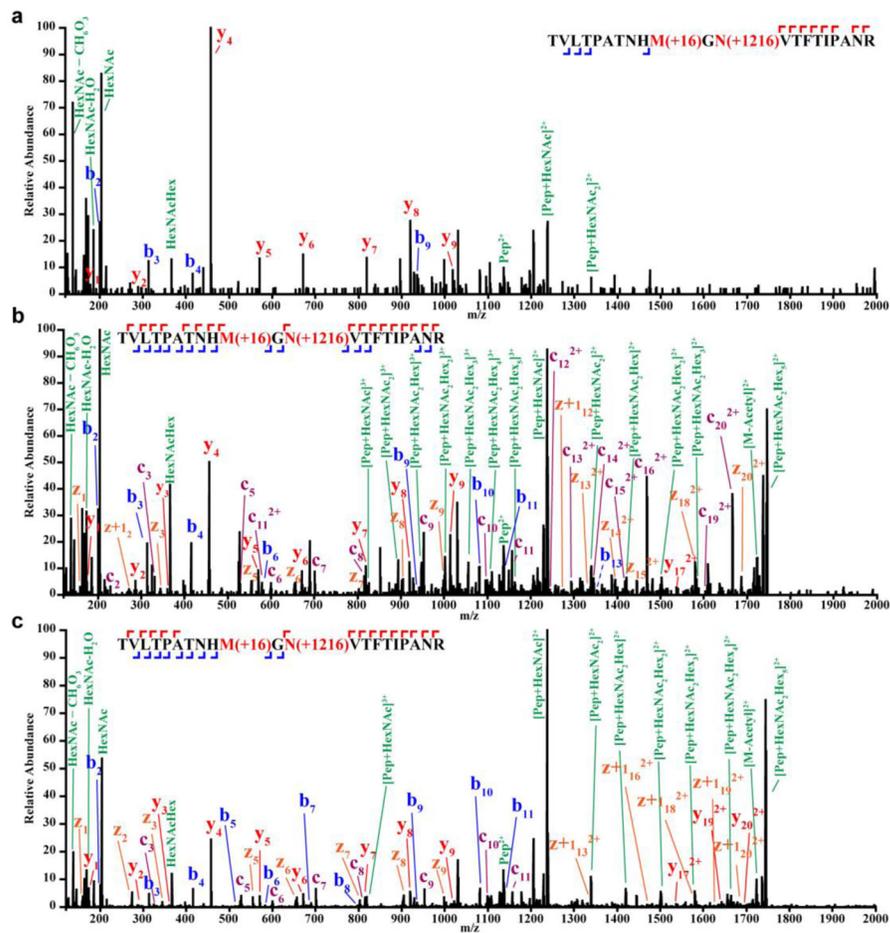
Author Manuscript



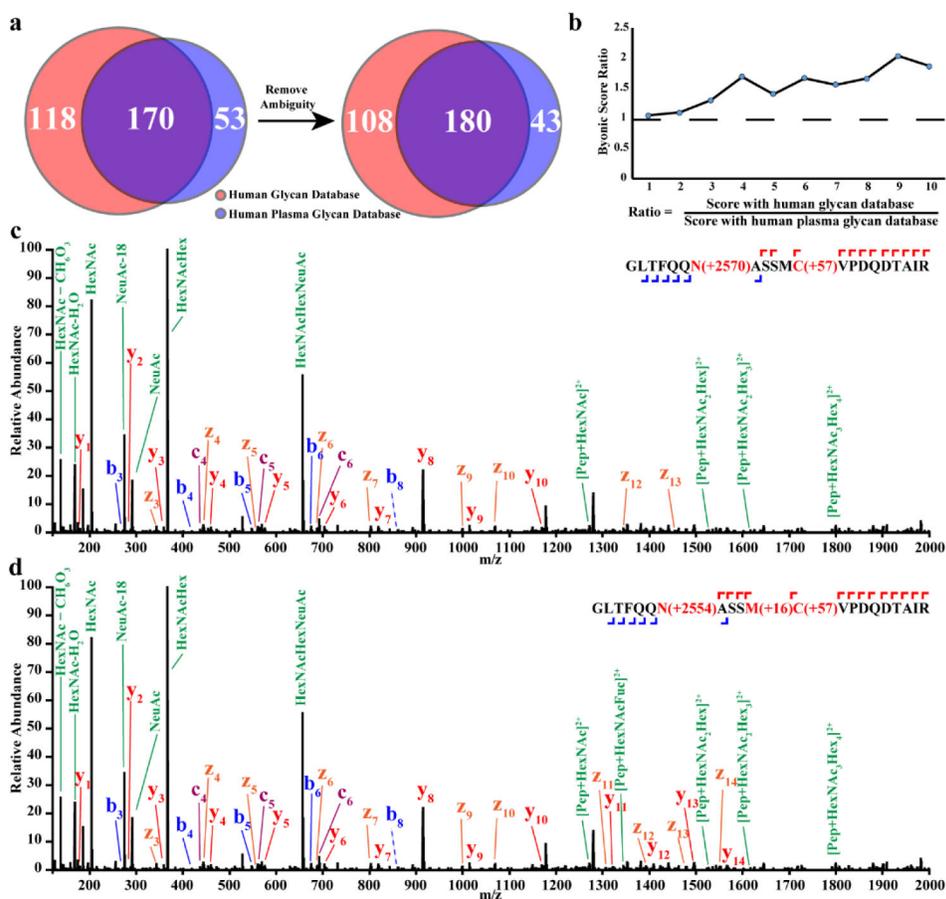
**Figure 1.** MS/MS of 3+ charge state precursor ion at  $m/z$  1577.9 of bovine fetuin triantennary *N*-glycopeptide KLCPCDLLAPLNSDR (AA 126–141). Alternating between CID/ETD/ETDC resulted in different sets of ions. (a) CID and ETD spectra (inset). \* in the peptide sequence indicates carbamidomethylation. (b) EThcD spectrum. Starred peaks (\*) in the spectra were deconvoluted and annotated in the inset.



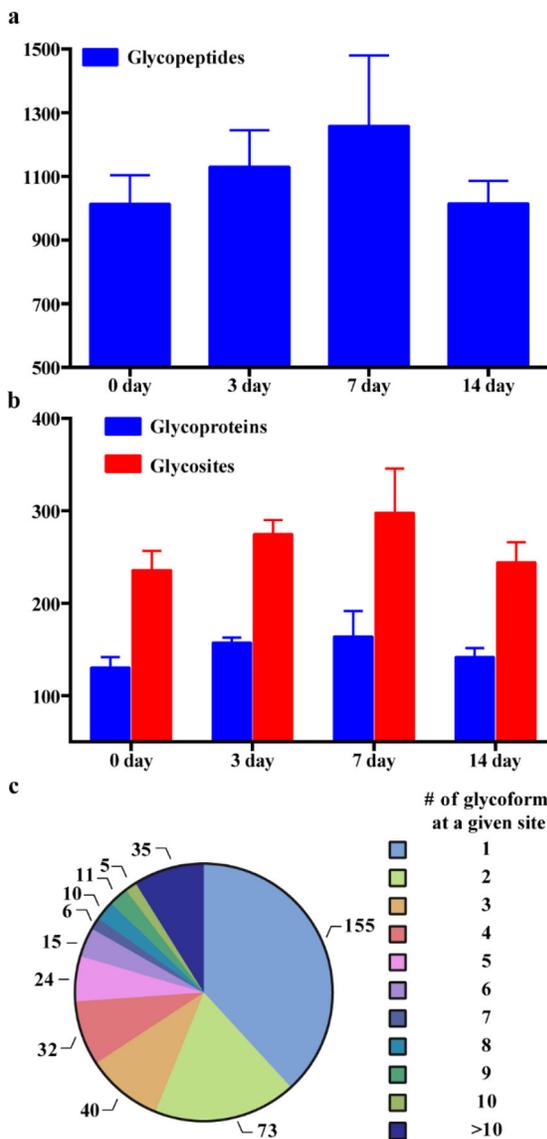
**Figure 2.** Comparison between HCD and EThcD. (a, b) Number of identified glycopeptide and glycosylation site comparison. (c) distribution of the Byonic scores of identified glycopeptide peptide spectral matches.



**Figure 3.** MS/MS of *N*-glycopeptide TVLTPATNHMGNVTFITIPANR (AA 74–94) from human complement C3. (a) HCD and (b) EThcD spectra of 3+ charge state precursor at *m/z* 1163.20. (c) EThcD spectrum of 4+ charge state precursor at *m/z* 872.65.



**Figure 4.** Evaluation of human glycan database and human plasma glycan database. (a) Unique glycopeptide identification between search results with two databases, before and after manual validation of ambiguous spectra. (b) Byonic ratio of ambiguous peptide spectral matches. The ratio is calculated as the score with human glycan database divided by plasma glycan database. (c) Assigned MS/MS of a 4+ charged precursor at  $m/z$  1228.00 with human plasma glycan database and (d) human glycan database.



**Figure 5.** Glycosylation identifications with four biological replicates from uninjured and injured carotid arteries at 3, 7, and 14 days post injury. (a) Unique glycopeptides and (b) unique glycosites and glycoproteins. Each column represents mean + S.D. (c) Pie chart of number of glycoforms mapped to each glycosite. Each segment represents how many glycosites have a particular number of glycoforms.

Table 1

Glycopeptides and corresponding glycoforms identified by EThcD for the *N*- and *O*-linked oligosaccharides from the enzymatic digestion of bovine fetuin.

Glycopeptide	Sequence with glycan composition	Position	Experimental mass [M+H]	Theoretical mass [M+H]	Mass error (ppm)	Log Prob	Score
<i>N</i> -glycans							
<b>LC</b> <sup>3</sup> <b>PDC</b> <sup>3</sup> <b>PLLAPLN</b> <sup>4</sup> <b>DSR</b>	HexNAc(4)Hex(5)NeuAc(1)	127–141	3654.5264	3654.5177	2.37	7.53	519.5
	HexNAc(4)Hex(5)NeuAc(2)		3945.6106	3945.6132	-0.64	9.95	657.5
	HexNAc(5)Hex(6)NeuAc(2)		4310.7576	4310.7453	2.85	6.86	368.6
	HexNAc(5)Hex(6)NeuAc(3)		4601.8436	4601.8408	0.61	7.22	523.2
	HexNAc(5)Hex(6)NeuAc(4)		4892.9505	4892.9362	2.93	6.72	536.7
	HexNAc(4)Hex(6)Fuc(1)NeuAc(1)		3962.6387	3962.6285	2.58	5.74	334.9
<b>KLC</b> <sup>3</sup> <b>PDC</b> <sup>3</sup> <b>PLLAPLN</b> <sup>4</sup> <b>DSR</b>		126–141					
	HexNAc(4)Hex(5)NeuAc(2)		4073.7105	4073.7081	0.58	8.11	472.0
	HexNAc(5)Hex(6)NeuAc(3)		4729.9359	4729.9357	2.20	5.12	193.7
<b>VVHAVEVALAFNAESN</b> <sup>4</sup> <b>GSYLQLVEISR</b>		142–169					
	HexNAc(5)Hex(6)NeuAc(3)		5877.5920	5877.5738	3.65	12.53	861.5
	HexNAc(5)Hex(6)NeuAc(4)		6167.6774	6168.6692	1.87	4.69	447.6
<b>RPTGEVYDIEDLETTC</b> <sup>4</sup> <b>HVLDPPLAN</b> <sup>4</sup> <b>C</b> <sup>4</sup> <b>SVR</b>		54–58					
	HexNAc(4)Hex(5)NeuAc(2)		5876.5524	5876.5404	2.05	15.26	905.9
	HexNAc(5)Hex(6)NeuAc(2)		6242.7043	6241.6726	4.55	7.24	646.8
	HexNAc(5)Hex(6)NeuAc(3)		6532.7837	6532.7680	3.03	13.29	824.6
	HexNAc(5)Hex(6)NeuAc(4)		6823.8892	6823.8634	3.78	4.87	530.2
	HexNAc(6)Hex(6)Fuc(1)NeuAc(2)		6589.7734	6590.8098	-5.02	2.67	550.3
	HexNAc(4)Hex(5)NeuAc(1)NeuAc(1)		5893.5612	5892.5353	3.84	7.24	613.5
<i>O</i> -glycans							
<b>TPVGGQPS</b> <sup>4</sup> <b>IPGGCPVIR</b>		316–323					
	HexNAc(1)Hex(1)NeuAc(1)		2131.0665	2131.0652	0.59	11.51	723.0
	HexNAc(1)Hex(1)		1839.9708	1839.9698	0.52	10.38	611.9

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Glycopeptide Sequence with glycan composition	Position	Experimental mass [M+H]	Theoretical mass [M+H]	Mass error (ppm)	Log Prob	Score
HexNAc(1)		1677.9223	1677.9170	3.13	12.38	711.7

\* Cysteine carbamidomethylation.

*a,b,c* Glycosylation site.

**Table 2**

Glycosylation site identifications using human plasma and human glycan databases.

UniProt Accession	Protein Name	Sites	
		Plasma Glycan Database	Human Glycan Database
O75882	Attractin	<i>U</i> <sub>N300</sub> ; N416; N731; N1054; N1073	N416; N731; N1054; N1073
P00450	Ceruloplasmin	N138; N358; N397; N762	N138; N358; N397; N762
P00734	Prothrombin	N121; N143	N121; N143
P00738	Haptoglobin	N184; N207; N211; N241	N184; N207; N211; N241
P01008	Antithrombin-III	N128; N224	N128; N224
P01009	Alpha-1-antitrypsin	N271	N271
P01019	Angiotensinogen		<i>U</i> <sub>N47</sub>
P01024	Complement C3	N85	N85
P01042	Kininogen-1	M169; N205; N294	N169; N205; N294
P01591	Immunoglobulin J chain	N71	N71
P01857	Ig gamma-1 chain C region	N180	N180
P01859	Ig gamma-2 chain C region	N176	N176
P01861	Ig gamma-4 chain C region	N177	N177
P01871	Ig mu chain C region	N46; N209; N439	N46; N209; N439
P01876	Ig alpha-1 chain C region	N340	N340
P01877	Ig alpha-2 chain C region	N205	N205
P02749	Beta-2-glycoprotein	N162; N253	N162; N253
P02763	Alpha-1-acid glycoprotein 1	N56; N93	N56; N93
P02765	Alpha-2-HS-glycoprotein	N156; N176	N156; N176
P02787	Serotransferrin	N432; N630	N432; N630
P02790	Hemopexin	N187; N453	N187; N453
P04220	Ig mu heavy chain disease protein	N147	N147
P05156	Complement factor I	N103	N103
P05543	Thyroxine-binding globulin	N36	N36
P06681	Complement C2	N621; N651	N621; N651
P08185	Corticosteroid-binding globulin	N96	N96
P19652	Alpha-1-acid glycoprotein 2		<i>U</i> <sub>N93</sub>
P19823	Inter-alpha-trypsin inhibitor heavy chain H2	N118	N118
P22792	Carboxypeptidase N subunit 2	N74	N74
P25311	Zinc-alpha-2-glycoprotein	N112; N128	N112; N128
P43652	Afamin	N33; N109; N402	N33; N109; N402
Q06033	Inter-alpha-trypsin inhibitor heavy chain H3	N580	N580
Q14624	Inter-alpha-trypsin inhibitor heavy chain H4	N517	N517
Q5SVZ6	Zinc finger MYM-type protein 1		<i>U</i> <sup>*</sup> <sub>N1006</sub>
Q86XJ1	GAS2-like protein 3		<i>U</i> <sup>*</sup> <sub>N401</sub>
Q3SY69	Mitochondrial 10-formyltetrahydrofolate dehydrogenase	<i>U</i> <sup>*</sup> <sub>N142</sub>	

UniProt Accession	Protein Name	Sites	
		Plasma Glycan Database	Human Glycan Database
Q8WYG9	G-protein coupled receptor 98	N3794	N3794
Q92833	Protein Jumonji		<i>U</i> *N1201
Q96N64	PWWP domain-containing protein 2A		<i>U</i> *N563
Q99459	Cell division cycle 5-like protein		<i>U</i> *N436
Q9NZP8	Complement C1r subcomponent-like protein	N242	N242

*U* Unique identification with one database.

\* Novel glycosylation site.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript