

The Fine Art of Destruction: A Guide to In-Depth Glycoproteomic Analyses – Exploiting the Diagnostic Potential of Fragment Ions

Marcus Hoffmann¹, Markus Pioch¹, Alexander Pralow¹, René Hennig^{1,2}, Robert Kottler^{1,2}, Udo Reichl^{1,3}, Erdmann Rapp^{1,2}

¹ Max Planck Institute for Dynamics of Complex Technical Systems, Bioprocess Engineering, 39106 Magdeburg, Germany.

² glyXera GmbH, 39120 Magdeburg, Germany.

³ Otto von Guericke University Magdeburg, Chair of Bioprocess Engineering, 39106 Magdeburg, Germany.

Corresponding author: Dr. Erdmann Rapp (rapp@mpi-magdeburg.mpg.de)

Keywords: *Glycoproteomics / N-Glycosylation / O-Glycosylation / Mass spectrometry / IgG / Fibrinogen / Lactotransferrin / Ribonuclease B / HCD / Proteinase K / HILIC / Oxonium Ions*

ABSTRACT: The unambiguous mass spectrometric identification and characterization of glycopeptides is crucial to elucidate the micro- and macroheterogeneity of glycoproteins. Here, we propose combining lower and stepped collisional energy fragmentation for the in-depth and site-specific analysis of *N*- and *O*-glycopeptides. Using a set of four representative and biopharmaceutically-relevant glycoproteins (IgG, fibrinogen, lactotransferrin, and ribonuclease B), we highlight the benefits and limitations of the developed workflow and provide a state-of-the-art blueprint for conducting high-quality in-depth *N*- and *O*-glycoproteomic analyses. Further, we describe a modified and improved version of cotton HILIC-based solid phase extraction for glycopeptide enrichment. For the unambiguous identification of *N*-glycopeptides we propose the use of a conserved fragmentation signature $[M_{\text{peptide}} + \text{H} + {}^{0,2}\text{X GlcNAc}]^+$, that has rarely been employed in glycoproteomic analyses up to now. We show for the first time that this fragmentation signature can consistently be found across all *N*-glycopeptides, but not on *O*-glycopeptides. Moreover, we have systematically and comprehensively evaluated the use of the relative abundance of oxonium ions to retrieve glycan structure information, e.g. differentiation of hybrid- and high-mannose-type *N*-glycans or differentiation between antenna GlcNAc and bisecting GlcNAc. Our findings may increase confidence and comprehensiveness in manual and software-assisted glycoproteomics.

STATEMENT OF SIGNIFICANCE: With the outcome of our study, we provide a universal blueprint for performing high-quality and in-depth *N*- and *O*-glycoproteomics. By combining stepped with lower-energy HCD fragmentation the presented workflow enables the unambiguous and site-specific identification and characterization of *N*- and *O*-glycopeptides generated by specific as well as non-specific digestion. An improved version of cotton HILIC-SPE glycopeptide enrichment is presented. Our version is easier to use, allows for much higher loading and throughput and is more cost-efficient. We comprehensively investigated the conserved glycopeptide fragmentation signature and show for the first time that this signature occurs for all *N*-glycopeptides, but not for *O*-glycopeptides. Moreover, we describe how this signature enables to unambiguously identify the peptide moiety of *N*-glycopeptides, and thus significantly increases identification reliability. We also report on unique oxonium ion patterns enabling discrimination of hybrid- and high-mannose-type *N*-glycans, of antenna and bisecting GlcNAc, of antenna and core fucosylation, or differentiation between *N*- and *O*-glycopeptides. The manuscript also features new glycosylation sites and regions for human fibrinogen and lactotransferrin. The acquired new insights into the HCD fragmentation behaviour of *N*- and *O*-glycopeptides will significantly improve reliability and depth of manual as well as software-assisted glycoproteomics.

INTRODUCTION: The reliable and unambiguous mass spectrometric identification and characterization of *N*- and *O*-glycopeptides is indispensable for in-depth glycoproteomic studies. Fields of application comprise structure-function studies, pharmacodynamics, clinical studies, and production of biopharmaceuticals. Such in-depth analyses require a detailed description of all potential glycosylation sites (site occupancy, i.e. macroheterogeneity) with their respective glycoforms (glycan composition, i.e. microheterogeneity). Usually the method of choice for such in-depth glycoproteomic analyses involves the proteolytic digest of glycoproteins and the subsequent analysis of the resulting intact glycopeptides by reversed-phase liquid chromatography coupled to electrospray tandem mass spectrometry (RP-LC-ESI-MS/MS). Owing to their inherent complexity, lower abundance and lower ionization efficiency, the analysis of glycopeptides poses significant challenges towards sample preparation, measurement, and software-assisted data analysis compared to non-glycosylated peptides. In recent years, several key advances were made rendering *N*- and *O*-glycoproteomic analyses more reliable, more comprehensive, and less time-consuming^[1-3]. One such improvement is the application of proteolytic enzymes with broad cleavage specificity, like proteinase K^[4-8] or pronase^[7, 9], in addition to the

conventionally used trypsin^[10]. At times, the peptide moiety of tryptic glycopeptides can be too long and can harbor more than one potential glycosylation site, rendering the identification and characterization of those glycopeptides difficult, if not impossible. Proteases such as pronase and proteinase K, can generate short and distinct glycopeptides – even in densely glycosylated regions as found in mucin-type *O*-glycosylated glycoproteins^[11]. The use of these enzymes can thus enable the identification and characterization of glycosylation sites and respective glycoforms that otherwise would have evaded the analysis. Further improvements were also made in the selective enrichment of glycopeptides – an essential step in many workflows to overcome sensitivity issues caused by the lower abundance and lower ionization efficiency of glycopeptides. Here different liquid chromatography (LC) and solid phase extraction (SPE) based methods can be employed, such as hydrophilic interaction liquid chromatography (HILIC)^[12-15]. Arguably the biggest leap in the field of glycoproteomics can be attributed to advances in mass spectrometry instrumentation^[16] and software development^[17-19]. The advent of high-resolution mass spectrometers, including orbitrap instruments and next-generation Q-TOF instruments, along with new and complementary fragmentation techniques, namely higher energy collision dissociation (HCD) and electron transfer dissociation (ETD) have tremendously increased the throughput, depth and confidence of glycoproteomic analyses. Using HCD fragmentation with varied collisional energies, both the glycan composition and the peptide sequence can be reliably determined from the same fragment ion spectrum^[9, 20, 21] – a unique and powerful capability that conventional collision-induced dissociation (CID) fragmentation methods normally not provide. Additional pinpointing of occupied glycosylation sites is particularly crucial for the analysis of mucin-type *O*-glycopeptides, since *O*-glycans – as opposed to *N*-glycans – are not attached according to a conserved consensus motif ($\underline{N}|X|S/T/(C/V); X\neq P$), but instead can be attached to virtually every serine or threonine residue of a peptide. Along with these technical advances, more and more appropriate software tools and databases are now available to assist researchers during analysis and interpretation of *N*- and *O*-glycopeptide fragment spectra – a process that is usually very time-consuming and challenging. With the ability to discriminate glycopeptide spectra from non-glycopeptide spectra by means of characteristic fragment ions (glycan-derived B-ions, i.e. oxonium ions, and Y-ions) along with powerful search engines for the identification of the glycan and the peptide moiety, programs like Proteinscape and Byonic now render the analysis and interpretation of *N*- and *O*-glycopeptide more reliable, more comparable, and less time-consuming^[19]. To further increase the confidence in glycopeptide identification, oxonium ions got more and more into focus lately^[22-25]. This is largely due to their potential to predict the type of glycosylation (different forms of *N*-glycosylation as well as *O*-glycosylation) and to confirm certain structural glycan features. In this report, we present a universal glycoproteomic workflow developed for the analysis of *N*- and *O*-glycosylated proteins that enables a detailed characterization of the glycan moiety and its peptide backbone. The workflow is based on the analysis of tryptic and non-tryptic glycopeptides that were enriched by spin-cotton-HILIC SPE and measured by LC-ESI-MS/MS using HCD fragmentation with fixed (HCD.low) as well as stepped collisional energy (HCD.step) on an Orbitrap mass spectrometer. The workflow is demonstrated using a representative set of three human and one bovine glycoprotein (human: immunoglobulin gamma, fibrinogen, lactotransferrin; bovine: ribonuclease B), covering a broad range of protein and glycan features (*N*- and *O*-glycosylation, micro- and macroheterogeneity, different glycoforms). The resulting HCD.low and HCD.step fragment ion spectra were analyzed using the dedicated in-house software glyXtool^{MS} (Pioch et al. 2018, submitted, 2018), which enabled the reliable and in-depth analysis of *N*- and *O*-glycopeptides alike. For the representative glycoproteins we were able to discover and characterize several new glycosylation sites and regions with their respective glycoforms – in addition to confirming already known glycosylation sites and glycoforms. An interesting finding is the presence of a conserved fragmentation signature (I-IV) exclusively detected with *N*-glycosylated peptides: (I) peptide -amine, (II) peptide, (III) peptide +^{0,2}X GlcNAc, and (IV) $[M_{\text{peptide}} + H + \text{GlcNAc}]^+$. Up to now, the occurrence of this pattern using Orbitrap mass spectrometers has not been investigated intensively, and the pattern has only rarely been used for glycoproteomic analyses. In this study for the first time we show that this pattern only occurs for *N*-glycopeptides but not for *O*-glycopeptides and that it can significantly improve the unambiguous determination of the putative peptide mass of *N*-glycopeptides. The latter is particularly important for glycopeptides generated by a non-specific digest. Furthermore, we have systematically and comprehensively evaluated the use of the relative abundance of oxonium ions to obtain a more detailed characterization of the glycan moiety. We have identified unique oxonium ion patterns that provide insights on certain structural glycan features, allowing for instance differentiation between *N*-glycan antenna GlcNAc and bisecting GlcNAc. Overall, the presented analytical workflow along with the analyzed representative glycoproteins and the given insights into diagnostic glycopeptide fragment ions serves as a guide to in-depth glycoproteomic analysis for a broad range of glycoproteins. It will thus facilitate basic glycoproteomic research as well as support biopharmaceutical R&D and QA/QC.

EXPERIMENTAL PROCEDURES: Experimental procedures are schematically shown in figure 1 and described in detail in the supplement.

RESULTS AND DISCUSSION: The developed glycoproteomic workflow for the analysis of *N*- and *O*-glycosylated proteins allows for a detailed characterization of the glycan moiety and the peptide backbone alike. The workflow is focused on the analysis of tryptic and non-tryptic glycopeptides that were enriched by spin-cotton-HILIC-SPE and measured by nano-RP-LC-ESI-OT-OT-MS/MS using HCD fragmentation (Figure 1).

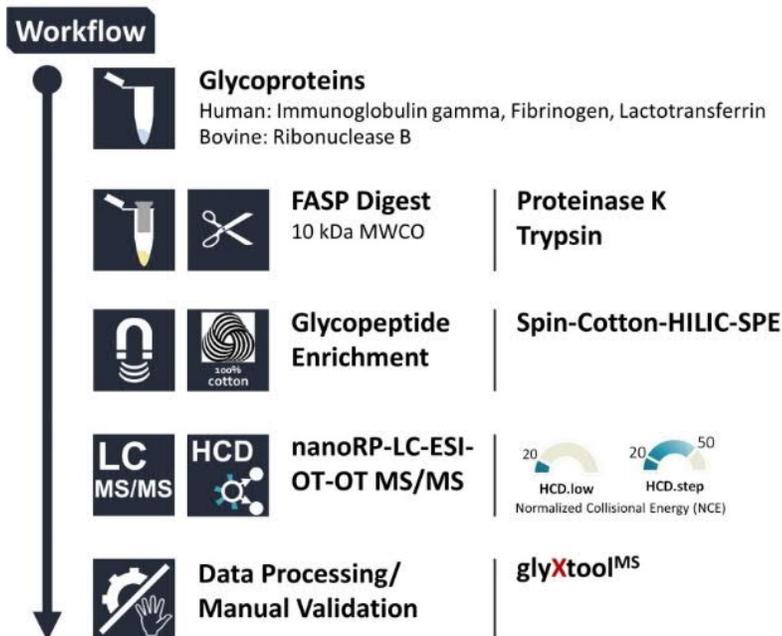


FIGURE 1: Glycoproteomic workflow developed for the in-depth analysis of *N*- and *O*-glycosylated proteins. A set of representative glycoproteins was proteolytically digested either with trypsin or proteinase K using a FASP approach. Resulting peptides were subjected to spin-cotton-HILIC-SPE to enrich *N*- and *O*-glycopeptides. HCD fragment ion spectra using fixed and stepped normalized collisional energy (*HCD.low*, *HCD.step*) were acquired from glycopeptide-enriched HILIC fractions. For data processing and annotations, glyXtool^{MS} was used; validation was done manually. FASP = Filter-aided sample preparation; HCD = Higher-energy collision dissociation; OT = Orbitrap mass analyzer; RP = reversed-phase.

We demonstrate the workflow's efficiency, flexibility, merits and limitations with the help of four representative glycoproteins, comprising IgG, LTF, Fib, and RNase B. This set of glycoproteins covers a broad range of glycan and protein features relevant for many glycoproteomic studies, including different types of glycosylation (*N*- and *O*-glycosylation), different glycoforms (complex-, hybrid-, high-mannose-type *N*-glycans, mucin-type *O*-glycans; microheterogeneity), specific glycan features (core- vs antenna-fucosylation, bisecting *N*-acetylglucosamine, degree of sialylation) as well as multiple glycosylation sites (macroheterogeneity). Moreover, we highlight unique oxonium ion patterns that can be used to characterize the glycan moiety more precisely, i.e. to obtain additional structural information.

Keystone of our analysis workflow is the implementation of two Orbitrap Elite MS HCD fragmentation regimes (MS^2), namely *HCD.low* and *HCD.step*, for the analysis of *N*- and *O*-glycopeptides (Figure 2). The applied normalized collisional energies during this process can be varied individually, such that for instance low- and high-energy-fragmentation can be carried out sequentially on the very same precursor ion. The readout of the resulting fragment ions is done in the Orbitrap mass analyzer – where fragment ion spectra can be acquired with high resolution and without compromising information by a low-mass cut-off. The latter is inherent to (linear quadrupole) ion trap mass analyzers usually used for CID fragmentation and leads to the absence of important fragment ions in the low-molecular range region (fragment ions with m/z values <30% that of the precursor ion are not stable in the ion trap, and therefore not present in the fragment ion spectrum). The low m/z range region of *N*- and *O*-glycopeptide fragment ion spectra is of particular interest for glycopeptide analyses, as it features diagnostic fragment ions derived from the glycan moiety of the glycopeptide, so-called oxonium ions. These singly charged low-molecular weight fragment ions (B-ions, according to Domon and Costello^[26]) can not only be used to discriminate a glycopeptide spectrum from a non-glycopeptide spectrum, they can also provide crucial information on the glycan moiety that is linked to the peptide backbone.

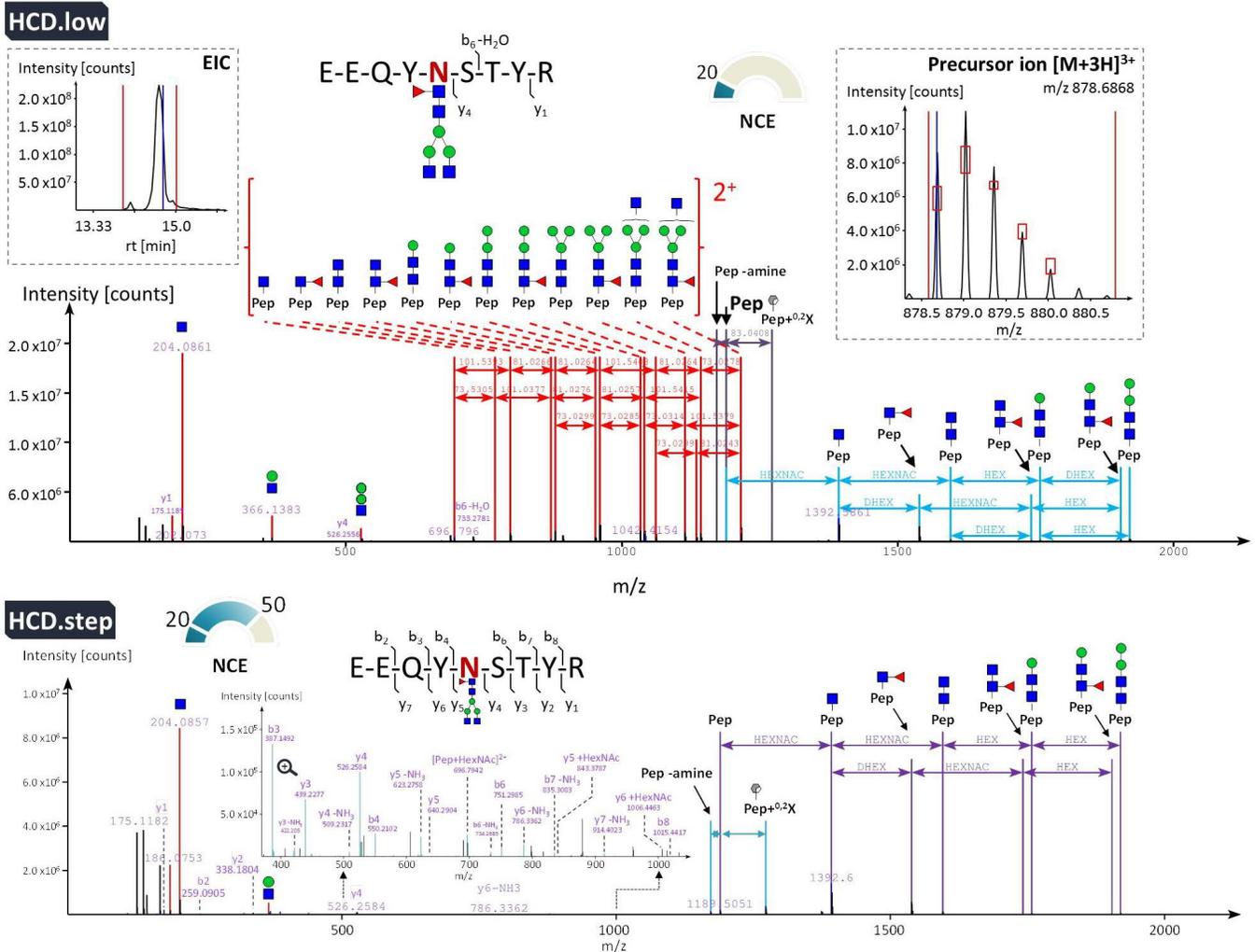


FIGURE 2: HCD fragment ion spectra (MS^2) of tryptic IgG 1 Fc N-glycopeptide ${}_{176}\text{EEQYNSTYR}_{184} + \text{HexNAc}_4\text{Hex}_3\text{Fuc}_1$ (m/z 878.687 $[\text{M}+3\text{H}]^{3+}$) acquired by nano-RP-LC-ESI-OT-OT-MS/MS using varying normalized collisional energies (NCE; positive ion mode, LTQ Orbitrap Elite hybrid mass spectrometer). *Top: HCD.low* – fragment ion spectrum acquired at low NCE (20): enables detailed analysis of the glycan moiety. Putative peptide mass can be deduced via a conserved fragmentation pattern: (I) $[\text{M}_{\text{peptide}} + \text{H} - \text{NH}_3]^+$, (II) $[\text{M}_{\text{peptide}} + \text{H}]^+$, (III) $[\text{M}_{\text{peptide}} + \text{H} + {}^{0,2}\text{X GlcNAc}]^+$, and (IV) $[\text{M}_{\text{peptide}} + \text{H} + \text{GlcNAc}]^+$. *Top (inset, left):* extracted ion chromatogram (EIC). *Top (inset, right):* isotopic pattern of the precursor ion. *Bottom: HCD.step* – fragment ion spectrum acquired using stepped NCE (20, 50): enables unambiguous identification peptide backbone. Again, the peptide mass can be deduced via a conserved fragmentation pattern. For the annotation of the glycan moiety the following abbreviations/symbols were used: N-acetylglucosamine (HexNAc, “blue square”), mannose (Hex, “green circle”), fucose (DHEx, “red triangle”), peptide (Pep).

During *HCD.low* fragmentation a constant NCE of 20 was used. At this low collisional energy, the resulting N- and O-glycopeptide fragment ion spectra resemble those acquired by CID fragmentation performed in an ion trap. Similar to CID, N- and O-glycopeptide fragment ion spectra acquired by *HCD.low* are also mainly populated by fragment ions derived from the glycan moiety (B- and Y-ions), and only rarely feature fragment ions derived from the more stable peptide backbone (b- and y-ions) (Figure 2, top). The most prominent glycan-derived signals in *HCD.low* glycopeptide spectra correspond to the successive neutral loss of monosaccharides from the non-reducing end towards the reducing end of the attached glycan (Y-ions, mostly singly but also doubly charged). These fragment ions along with diagnostic oxonium ions are the basis for the determination of the glycan composition.

For N-glycopeptides the successive fragmentation of the glycan moiety usually leads to its complete loss – leaving the peptide still intact. Thereby, the signal corresponding to the peptide mass is often rather low, or even absent. Nevertheless, the putative peptide mass can still be determined with high reliability in most cases by means of another fragment ion: a main event during *HCD.low* and CID fragmentation is the cleavage of the glycosidic bond in-between the chitobiose core, which results in an intense signal corresponding to the peptide with an additional N-acetylglucosamine (GlcNAc); this signal in turn allows to infer the putative peptide mass (Y₁ fragment ion) and thus to identify the peptide. Still, an unambiguous identification of the peptide is only possible with an appropriate number of peptide fragment ions.

For *O*-glycopeptides, particularly for the most commonly found mucin-type core-1 and core-2 *O*-glycopeptides, the glycan moiety comprises usually only a few monosaccharides (most commonly found glycoform: mucin-type core-1 GalNAc₁Gal₁NeuAc_{1/2})^[27]. This normally results in less complex *HCD.low* fragment ion spectra. As for *N*-glycopeptides, the successive fragmentation of the glycan moiety of *O*-glycopeptides up to the de-glycosylated peptide mass can be observed. In contrast to *N*-glycopeptides, the signal corresponding to the peptide mass is usually rather intense. Additionally, *HCD.low* fragmentation of mucin-type *O*-glycopeptides is more likely to exhibit enough peptide-derived b- and y-ions to allow for an unambiguous identification. The latter can be linked to the smaller glycan moiety typically found for mucin-type *O*-glycopeptides.

Apart from *HCD.low* we have added a second fragmentation regime to our workflow: *HCD.step* was developed to complement *HCD.low* fragmentation, such that for all selected and fragmented *N*- and *O*-glycopeptides always enough peptide fragment ions are generated to allow for an unambiguous peptide identification, without sacrificing too much information of the glycan moiety. *HCD.step* features a two-step fragmentation duty cycle, where every selected precursor ion is fragmented at both, low- and high-collisional energy. As such, the resulting fragment ion spectra are hybrid spectra, combining fragment ions arising from the application of two different collisional energies. For the analysis of *N*- and *O*-glycopeptides this approach proved to be beneficial as it generates glycan-specific fragment ions and peptide backbone fragment ions alike. Similar to *HCD.low*, *HCD.step* also generates intense glycan-derived signals corresponding to oxonium ions and neutral loss fragment ions. The latter, though, are mostly present as singly charged species. Hence, at the given acquisition window of *m/z* 100-2000, the number of glycan-derived fragment ions that can be used to deduce the glycan composition is lower, compared to *HCD.low*. Still, *HCD.step* fragmentation generates enough Y-ions to reliably infer the putative peptide mass. Thereby the signal corresponding to the peptide mass is always present with decent intensity. Likewise, a high number of b- and y-ions are reliably generated – for both, *N*- and *O*-glycopeptides. The detection of these fragment ions along with the putative peptide mass are the basis for a reliable and unambiguous identification of the peptide moiety (Figure 2). Further confidence in the correct peptide identification is provided by the presence of amino acid immonium ions in the low *m/z* range (immonium ions > 100 Da detected by *HCD.step*). These marker ions arise from internal fragmentation (a- and y-ions) of the peptide backbone, and can not only support peptide identification but can also be crucial for *de novo* sequencing^[28, 29]. Another striking feature of *HCD.step* fragmentation is that it generates a conserved fragmentation pattern (I-IV) around the peptide mass of *N*-glycosylated peptides: (I) $[M_{\text{peptide}} + \text{H} - \text{NH}_3]^+$, (II) $[M_{\text{peptide}} + \text{H}]^+$, (III) $[M_{\text{peptide}} + \text{H} + {}^{0.2}\text{X GlcNAc}]^+$ (cross-ring fragment of the innermost GlcNAc), and (IV) $[M_{\text{peptide}} + \text{H} + \text{GlcNAc}]^+$ (Figure 3).

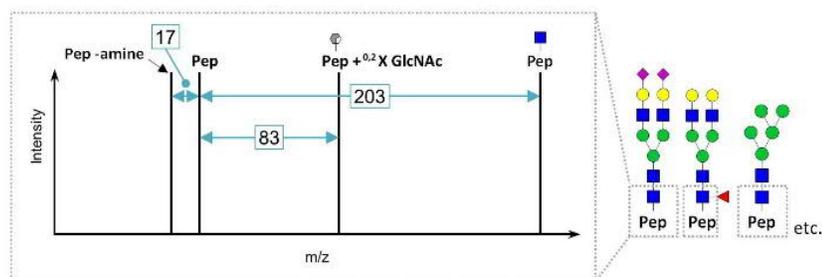


FIGURE 3: Conserved *N*-glycopeptide fragmentation pattern consistently obtained when using stepped normalized collisional energy (*HCD.step*): (I) $[M_{\text{peptide}} + \text{H} - \text{NH}_3]^+$, (II) $[M_{\text{peptide}} + \text{H}]^+$, (III) $[M_{\text{peptide}} + \text{H} + {}^{0.2}\text{X GlcNAc}]^+$, and (IV) $[M_{\text{peptide}} + \text{H} + \text{GlcNAc}]^+$.

This *N*-glycopeptide fragmentation pattern has already been reported for MALDI-TOF/TOF analyses by Wührer *et al.*^[30]. Recently this pattern was also identified for Orbitrap HCD fragmentation by Dong *et al.*^[31] and Stadlmann *et al.*^[32]. However, the detection of this fragment pattern among different *N*-glycosylation types was not systematically evaluated up to now. In our study, this pattern was consistently found in every *N*-glycopeptide fragment ion spectrum acquired by *HCD.step* – for complex-, high-mannose-, and hybrid-type *N*-glycopeptides alike. For *HCD.low* fragmentation the pattern was also found – not consistently and only at low intensity, though. Interestingly, we could not find this pattern with mucin-type *O*-glycopeptides. This fragmentation pattern thus only occurs for peptide+GlcNAc(+Fuc), but not for peptide+GalNAc. Using this conserved fragmentation pattern not only provides additional confidence in the annotation of the peptide mass, and with this, also more confidence regarding sequencing of the peptide backbone and the glycan moiety; it also allows for a more reliable differentiation between *N*-glycopeptides and mucin-type *O*-glycopeptides. Further details on differences and characteristics of *HCD.low*, *HCD.step*, and *HCD.high* (using solely high-energy collisional fragmentation) can be found in supplement 1.1.

Spin-Cotton-HILIC-SPE

As a result of the glycan microheterogeneity, glycopeptides are usually of lower abundance compared to their non-glycosylated counterparts. This, along with their inherent lower ionization efficiency, necessitates the enrichment of glycopeptides in many glycoproteomic studies. In our workflow we have implemented an improved version of the cotton-HILIC-SPE for the selective enrichment of *N*- and *O*-glycopeptides, published by Selman *et al.*^[33]. The modified spin-cotton-HILIC-SPE comprises the following two differences to the original protocol: (I) the amount of cotton filled in the pipette tips was increased by a factor of 12 (before: 0.5 mg; new: 6 mg cotton wool), and (II) centrifugation, instead of pipetting up and down, was employed throughout the entire procedure (equilibration, loading, washing, and elution). While the former (I) allows now for a higher loading capacity – approximately 10 μg (glyco)peptide digest instead of 3.3 μg - the latter (II) renders the workflow less cumbersome and labor-intensive, and at the same time allows for a higher throughput (several samples can be prepared in parallel, depending on the centrifuge employed). Reproducibility and glycopeptide enrichment efficiency of the new spin-cotton-HILIC-SPE was assessed using tryptic peptides derived from IgG and LTF and appeared to be very good (Supplemental Figure 3 and Supplement 1.2).

In the following, the results of the glycoproteomic analyses of the four representative proteins, IgG, LTF, Fib, and RNase B, are shown in detail. Using the FASP approach, tryptic as well as proteinase K generated (glyco)peptides of these proteins were treated by spin-cotton-HILIC-SPE, and the resulting HILIC fractions were subjected to nano-RP-LC-ESI-OT-OT-MS/MS analysis. Thereby the FASP approach combines the benefits of an in-gel digest – purification of the target protein(s) by removal of the background matrix – with those of an in-solution digest – high amount of protein sample and ease of use. Each sample was measured using both fragmentation regimes: *HCD.low* and *HCD.step*. The acquired fragment ion spectra were then analyzed using glyXtool^{MS}. The presence of abundant oxonium ion signals (Supplemental Figure 4, top) together with glycan-derived neutral loss fragment ions allowed for a reliable discrimination of glycosylated and non-glycosylated fragment spectra by glyXtool^{MS}. For each glycopeptide fragment ion spectrum, annotations and identifications suggested by the software were validated manually. In addition, the extracted ion chromatograms (EIC) as well as the isotopic patterns of the precursor ion were checked for each glycopeptide. For accepting a glycopeptide identification a set of twelve criteria was applied (for details please refer to Supplement 1.3).

Human Immunoglobulin Gamma N-Glycoproteomics (Fc-glycosylation)

Human IgG comprises four subclasses (IgG 1, 2, 3, and 4). Our analysis allowed to identify 87 *N*-glycopeptides derived from the constant CH2 region of the heavy chain of IgG, the majority being derived from IgG 1 and 2. Profiling of these *N*-glycopeptides, revealed predominantly core-fucosylated diantennary complex-type *N*-glycans with different degrees of galactosylation, with and without a terminal NeuAc and/or a bisecting-GlcNAc (Supplemental Figure 6). The three main glycoforms, for both IgG 1 and IgG 2, were HexNAc₄Hex₃Fuc₁, HexNAc₄Hex₄Fuc₁, and HexNAc₄Hex₃Fuc₁NeuAc₁ – also known as G0F, G1F, G2FS. This is in very good agreement with previously reported data^[34]. In total, glycan compositions corresponding to 24 different *N*-glycoforms were registered for IgG 1, 2, and 4 (Supplemental Table 2). For details please refer to supplement 1.4.

Human Fibrinogen N- and O-Glycoproteomics

Human fibrinogen features an α -, β -, and γ -chain and is known to be *N*- and *O*-glycosylated^[4, 5, 9, 35]. The α -chain is reported to solely carry mono- and disialylated mucin-type-1 *O*-glycans (T-antigen), while the γ -chain is reported to carry solely diantennary non-, mono- and disialylated complex-type *N*-glycans^[4]. The β -chain, in contrast, is known to carry both types of glycosylation^[4]. In the present study, both the tryptic and the proteinase K digest allowed to reliably identify and characterize *N*-glycopeptides covering already known *N*-glycosylation sites present on the β -, and γ -chain (β -chain: N₃₉₄, γ -chain: N₇₈) plus some lower-abundant *N*-glycopeptides derived from the α -chain (N₆₈₆). The latter has been reported to be occupied before^[36], but, to the best of our knowledge, the attached *N*-glycans have not been described so far (Supplemental Figure 2, pp. 27-31; Supplemental Figure 7; Supplemental Table 3). For the α - and γ -chain exclusively diantennary mono- and disialylated complex-type *N*-glycans without a core-fucose (HexNAc₄Hex₃Fuc₁NeuAc_{1/2} (G2FS1/2)) were detected. The β -chain, in contrast, also features non-sialylated, core-fucosylated, and glycoforms with a bisecting GlcNAc – though at much lower relative abundance. Also a number of *O*-glycopeptides were detected – covering already known but also some new glycosylation sites and regions on the α - and β -chain. The registered *O*-glycopeptides featured non-, mono- and disialylated mucin-type core-1 *O*-glycans (T-antigen). On the α -chain six different *O*-glycosylation regions including one newly discovered region (S₆₀₉, S₆₁₆, S₆₁₈, S₆₁₉) were found when using trypsin. With proteinase K eight *O*-glycosylation sites/regions were detected, of which three regions (I: S₃₂₅, S₃₂₆; II: S₃₅₆, S₃₅₇, S₃₅₉; III: T₄₉₉, T₅₀₁) and three sites (I: T₄₉₉, II: T₅₂₂; III: S₅₃₄) have not been reported so far. For the β -chain, both trypsin and proteinase K generated *O*-glycopeptides covered the already known *O*-glycosylation region (S₅₈, S₆₇). In agreement with previous reports, no *O*-glycopeptides were detected on the fibrinogen γ -chain. For details please refer to supplement 1.5.

Human Lactotransferrin N-Glycoproteomics

Human lactotransferrin is a globular glycoprotein that is reported to be solely *N*-glycosylated. The protein harbors three potential *N*-glycosylation sites (N₁₅₆, N₄₉₇, N₆₄₂), of which only the first and the second are reported to be glycosylated^[37, 38]. In this study we were able to detect *N*-glycans on all three *N*-glycosylation sites (Supplemental Figure 2, pp. 61-90; Supplemental Table 4). The first *N*-glycosylation site, N₁₅₆, features mono- and disialylated core-fucosylated complex- and hybrid-type *N*-glycans that can also have additional antenna fucoses and/or LacNAc extensions (LacNAc = GlcNAc₁Gal₁). The two major glycoforms are HexNAc₄Hex₃Fuc₁NeuAc_{1/2} (G2FS1/2). The second *N*-glycosylation site, N₄₉₇, features non-, mono-, and disialylated complex-type *N*-glycans, that are mainly core-fucosylated, but also non-fucosylated. The two major glycoforms are HexNAc₄Hex₃Fuc₁NeuAc_{1/2} (G2FS1/2). Also here *N*-glycoforms with additional antenna fucoses were also detected. Interestingly, though, no LacNAc extended glycoforms were registered. *N*-glycosylation site N₆₄₂ features mono- and disialylated core-fucosylated complex-type *N*-glycans, that can also have additional antenna fucoses and/or LacNAc extensions. All identified glycoforms exhibited nearly similar abundance; hence, there is no single predominant glycoform on this *N*-glycosylation site. For LTF, no *O*-glycopeptides were detected, which is in agreement with common knowledge. For details please refer to supplement 1.6.

Bovine Ribonuclease B N-Glycoproteomics

RNase B is a globular glycoprotein that is solely *N*-glycosylated. It harbors only one *N*-glycosylation site, which is occupied by high-mannose-type *N*-glycans^[39, 40]. In this study, a total of six tryptic *N*-glycopeptides of RNase B was identified (Supplemental Figure 2, pp. 91-97; Supplemental Table 5). Three different high-mannose-type glycoforms were allocated to these *N*-glycopeptides (HexNAc₂Hex_{5,6,8} – also known as Man5, Man6, Man8). Detection of these glycoforms is in good agreement with literature reports, as these are the major ones present one RNase B^[39, 40] (Supplemental Figure 8). The two lower-abundant glycoforms, Man7 and Man9, were not detected in the present work.

Quantification of Oxonium Ions and Its Potential for N- and O-Glycoproteomics

For the analysis of *N*- and *O*-glycopeptides, oxonium ions represent important diagnostic markers. Derived from the glycan moiety, these singly charged low-molecular weight fragment ions (B-ions) can not only be used to discriminate a glycopeptide spectrum from a non-glycopeptide spectrum, they can also provide crucial information on the glycan moiety that is linked to the peptide backbone (Supplemental Figure 4, top). The presence, absence, and relative abundance of oxonium ions can serve as a fingerprint characteristic for a certain *N*- or *O*-glycoform present on any peptide. To systematically evaluate the diagnostic characteristics of oxonium ions, glycopeptide fragment ion spectra of a set of *N*- and *O*-glycosylated proteins comprising IgG, LTF, Fib, and RNase B, were analyzed. For these four glycoproteins the relative abundance of the detected oxonium ions were investigated with respect to the type of glycosylation, the glycoform and certain glycan features, the peptide moiety, the precursor ion charge state, and the impact of the collisional energy that was applied for the fragmentation of the glycopeptides.

Complex-type N-glycopeptides: General Observations

In the following, characteristics of oxonium ions produced by fragmentation of complex-type *N*-glycopeptides via *HCD.low* and *HCD.step* will be described. The first noticeable characteristic is that relative oxonium ion abundance do not differ between different glycoproteins or *N*-glycosylation sites when considering the same *N*-glycoform and the same collisional energy, as evidenced by comparison of IgG 1 and 2 as well as LTF and IgG 2 (Supplemental Figure 9: Parts A and B). The produced oxonium ion patterns thus seem to be conserved for specific complex-type *N*-glycoforms independent of the peptide backbone. Also the charge state of the *N*-glycopeptide precursor ion seems to have only a minor influence on the produced oxonium ions. In general, relative abundance of oxonium ions produced by *HCD.low* and *HCD.step* differ significantly from each other when comparing the very same complex-type *N*-glycopeptides (Supplemental Figure 9: Part D). With increasing collisional energy, the relative abundance of di- and trisaccharide oxonium ions, such as HexNAc₁Hex₁, HexNAc₁Hex₂ or HexNAc₁Hex₁NeuAc₁ decreases – some di- and trisaccharide oxonium ions even disappear due to decomposition into mono- and disaccharide oxonium ions. Another striking characteristic that differentiates *HCD.step* from *HCD.low* is that *HCD.step* fragmentation of complex-type *N*-glycopeptides always results in the HexNAc oxonium ion peak being the dominant peak among the oxonium ions – independent of the present complex-type *N*-glycoform (relative abundance always >38%, Supplemental Figure 9: (Parts A (VI-XIII), D, E (IV-VI), J (II)). In addition, at higher collisional energy, the relative abundance of corresponding water-loss species is also increased compared to lower energy fragmentation (e.g. average relative abundance for HexNAc-H₂O: *HCD.low* 6%, *HCD.step* 19%) (Supplemental Figure 9: Part J (I-II)). For more information please refer to supplement 1.7.

Complex-type N-glycopeptides: antenna GlcNAc vs bisecting GlcNAc

Comparing the relative oxonium ion abundance of individual fragment ion scans acquired for a particular *N*-glycopeptide with each other, for both *HCD.low* and *HCD.step*, only marginal differences can be detected in most cases, as indicated by the standard deviation (<0.5; Supplemental Figure 9: Parts A, D). Supplemental Figure 10 (Part A) exemplarily shows the individual *HCD.low* fragment ion scans acquired for the IgG 1 *N*-glycopeptide HexNAc₅Hex₃Fuc₁ (G0FN) over time. This *N*-glycopeptide features a non-galactosylated *N*-glycan with a bisecting GlcNAc. The presence of a bisecting GlcNAc gives rise to two characteristic fragment ions when using low-energy fragmentation: (I) peptide+HexNAc₃Hex₁ (Y-ion, [M+H]⁺, [M+2H]²⁺) and (II) HexNAc₂ (oxonium ion, B-ion, [M+H]⁺). These fragment ions were consistently found in all *HCD.low* scans of the IgG 1 *N*-glycopeptide HexNAc₅Hex₃Fuc₁ (Supplemental Figure 10: Part A). Interestingly, *HCD.low* fragmentation of the IgG 1 *N*-glycopeptide HexNAc₃Hex₃Fuc₁ (Supplemental Figure 10: Part D) showed an inconsistent oxonium ion pattern across the acquired scans (which also explains the high standard deviation in Supplemental Figure 10: Part D(I)). In the first four scans the HexNAc₁Hex₁ oxonium ion dominates, which along with the absence of a peptide+HexNAc₃Hex₁ fragment ion suggests a non-galactosylated *N*-glycan with an antenna GlcNAc attached to the peptide (Supplemental Figure 10: Part D, scan #1-4). In scans #5 and #6, however, the HexNAc₁ oxonium ion dominates, which along with the presence of the peptide+HexNAc₃Hex₁ and the HexNAc₂ fragment ion in scan #6, suggests a non-galactosylated *N*-glycan with a bisecting GlcNAc attached (Figure 4 and Supplemental Figure 10: Part D, scan #5-6; another example can be found in Supplemental Figure 10: Part F). This finding enables discrimination between isobaric *N*-glycopeptides featuring either an antenna GlcNAc or a bisecting GlcNAc, based on differences in the retention time, characteristic *HCD.low* oxonium ion pattern, and diagnostic fragment ions. For details on this please refer to supplemental 1.8.

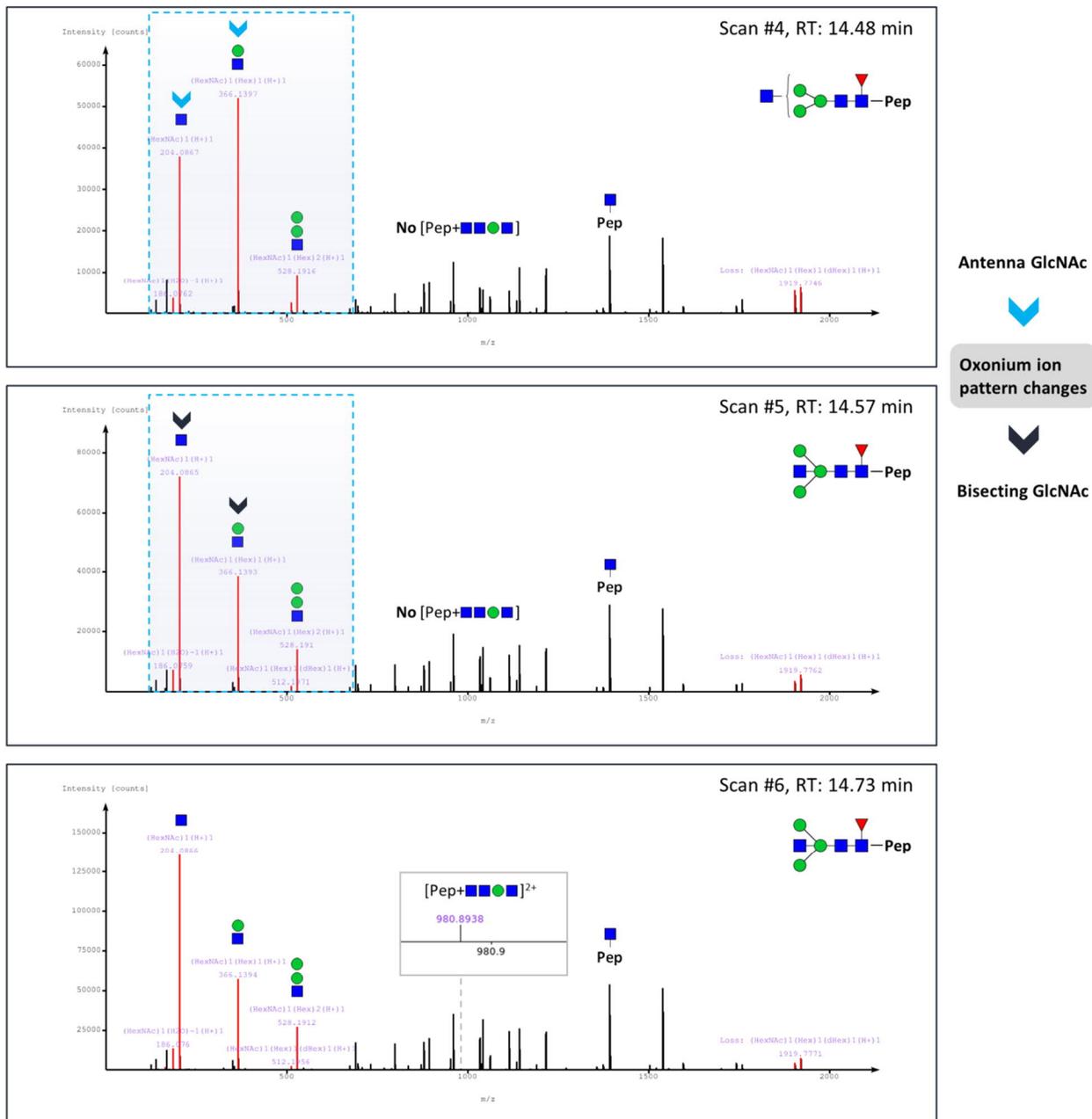


FIGURE 4: Change of the *N*-glycopeptide oxonium ion pattern between associated fragment ion scans allows discrimination of antenna GlcNAc and bisecting GlcNAc *N*-glycoforms. The example shows consecutive *HCD.low* fragment ion scans derived from the IgG1 *N*-glycopeptide ${}_{76}EEQYNSTYR_{184} + HexNAc_3Hex_3Fuc_1$.

Complex-type *N*-glycopeptides: core- vs antenna-fucosylation

The reliable detection and discrimination of core- and antenna-fucosylation is another important objective in many *N*-glycoproteomic studies. In this work, the presence of a core-fucose did not produce any unique and characteristic oxonium ion, neither for *HCD.low* nor for *HCD.step* (Supplemental Figure 9: Part J (I-II)). However, in many cases a $HexNAc_1Hex_1Fuc_1$ oxonium ion was detected for both core- and antenna-fucosylated *N*-glycopeptides. For core-fucosylated *N*-glycopeptides this oxonium ion is present only at very low relative abundance (<1%) – independent of the applied collisional energy – and seems to be an artifact generated by fucose rearrangement^[41]. The presence of an antenna-fucose, instead, produced a much higher signal of the $HexNAc_1Hex_1Fuc_1$ oxonium ion – but only at lower collisional energy (*HCD.low*: 10%; *HCD.step*: 2%) (Supplemental Figure 9: Part F (I-II)). Thus, at lower collisional energy discrimination between core- and antenna-fucosylation based on the relative abundance of the $HexNAc_1Hex_1Fuc_1$ oxonium ion seems to be possible. Interestingly, for *N*-glycopeptides presumed to carry an

antenna-fucose, we also detected a HexNAc₁Hex₁Fuc₁NeuAc₁ oxonium ion when using *HCD.low*, thus giving further evidence for this annotation.

Complex-type N-glycopeptides: type of sialic acid, degree of sialylation

N-acetylneuraminic acid (NeuAc, NANA) and *N*-glycolylneuraminic (NeuGc, NGNA) are the two most commonly found types of sialic acids in vertebrates^[42]. High- as well as low-energy fragmentation of NeuAc or NeuGc containing *N*-glycopeptides produces intense and distinct oxonium ions allowing a clear distinction of these two types of sialic acid: NeuAc oxonium ions (NeuAc; NeuAc -H₂O; HexNAc₁Hex₁NeuAc₁), NeuGc oxonium ions (NeuGc; NeuGc -H₂O; HexNAc₁Hex₁NeuGc₁) (NeuGc not shown in this study). To assess the degree of NeuAc sialylation the relative abundance of oxonium ions derived from mono- and disialylated glycoforms of LTF glycopeptides were compared (Supplemental Figure 9: Part G). With *HCD.low* relative abundance of the NeuAc and NeuAc -H₂O oxonium ions did not differ significantly between mono- and disialylated LTF *N*-glycopeptides. The relative abundance of the HexNAc₁Hex₁NeuAc₁ oxonium ion, though, was found to be higher with disialylated LTF *N*-glycopeptides in most, yet not all, cases (relative abundance HexNAc₁Hex₁NeuAc₁: monosialylated, between 10-18%; disialylated, between 22-28%). With *HCD.step* only slight differences between NeuAc-related oxonium ion abundance were detected. Overall, predicting the degree of NeuAc sialylation based on the relative abundance of NeuAc, NeuAc -H₂O, and HexNAc₁Hex₁NeuAc₁ oxonium ions seems promising, but needs to be further investigated, as for instance the influence of different sialic acid linkages needs to be evaluated, and might have caused the observed inconsistencies. Further information can be found in supplemental 1.9.

High-mannose-type N-glycopeptides

The relative oxonium ion abundance detected for high-mannose-type *N*-glycopeptides are remarkably different from those previously described for complex-type *N*-glycopeptides. *HCD.low* and *HCD.step* generated fragment ion spectra of complex-type *N*-glycopeptide are usually dominated by HexNAc and/or HexNAc₁Hex₁ oxonium ions as well as HexNAc₁Hex₁NeuAc₁ oxonium ions in some cases. In contrast, high-mannose-type *N*-glycopeptide fragment ion spectra, by nature feature only Hex or HexNAc related oxonium ions (HexNAc, HexNAc -H₂O, Hex, Hex -H₂O, HexNAc₁Hex₁, HexNAc₁Hex₂). The most striking feature of high-mannose-type *N*-glycopeptides, however, is the high relative abundance of Hex and Hex -H₂O oxonium ions, which directly reflects the high content of mannoses in the *N*-glycan structures. While for complex-type *N*-glycopeptides, if present at all, the relative abundance of Hex and Hex -H₂O oxonium ions did not exceed 0.5%, when low collisional energy was applied (Supplemental Figure 9: Part J (I-II)) for high-mannose-type *N*-glycopeptides, the relative abundance of the Hex oxonium ion, for instance, ranged from 25-80% when using *HCD.low* fragmentation, as exemplified for RNase B Man-5/6/8 *N*-glycopeptides (Supplemental Figure 9: Part H (I)). A similar effect was also observed for *HCD.step*: here the relative abundance of the Hex oxonium ion ranged from 20-42% for the RNase B Man-5/6/8 *N*-glycopeptides (for complex-type *N*-glycopeptides the relative abundance of Hex and Hex -H₂O oxonium ions was normally below 1%). Compared to *HCD.low*, the increased collisional energy applied with *HCD.step* resulted in an increase of the Hex -H₂O oxonium ion abundance (*HCD.low*: 2-10%; *HCD.step*: 12-20%. Supplemental Figure 9: Part H (I-II)). Overall, these findings suggest that both fragmentation regimes enable the discrimination of high-mannose-type *N*-glycopeptides from complex-type *N*-glycopeptides based on the relative abundance of Hex and Hex -H₂O oxonium ions. Moreover, *HCD.low* results indicate a direct correlation between the degree of mannosylation and the relative abundance of the Hex oxonium ion.

Hybrid-type N-glycopeptides

Hybrid-type *N*-glycopeptides combine features of complex- and high-mannose-type *N*-glycopeptides alike. Comparing the relative oxonium ion abundance detected for an *HCD.low* measurement of an IgG 2 complex-type *N*-glycopeptide (HexNAc₄Hex₅Fuc₁NeuAc₁) with those detected for an IgG 2 hybrid-type *N*-glycopeptide (HexNAc₃Hex₆Fuc₁NeuAc₁), major differences relate to the relative abundance of the HexNAc₁Hex₁ oxonium ion, which reflects the two fully galactosylated antenna of the complex-type *N*-glycan (Supplemental Figure 9: Part I). A minor, though significant, difference relates to the relative abundance of the Hex oxonium ion: for complex-type *N*-glycopeptides this oxonium is either not present or only present at a relative abundance of less than 0.5% at *HCD.low*; for hybrid-type *N*-glycopeptides the relative abundance of this oxonium ion was consistently found to be higher than 1.1%. This finding is in agreement with the observation previously made for high-mannose-type *N*-glycopeptides; also there an increase of the relative abundance of the Hex₁ oxonium ion was observed, correlating with the degree of mannosylation (Supplemental Figure 9: Part I). Overall, the relative abundance of the Hex oxonium ion and the presence of non-high-mannose type oxonium ions, such as HexNAc₁Hex₁NeuAc₁, may allow discriminating hybrid-type *N*-glycopeptides from complex-type or high-mannose-type *N*-glycopeptides. Another interesting finding in this context is the presence of a Hex₂ oxonium ion exclusively in *N*-glycopeptides presumed to be of the hybrid-type, high-mannose-type, or complex-type with not more than one terminal galactose (Hex₂ oxonium ion exclusively found in IgG and RNase B *N*-glycopeptide fragment ion spectra; not present in Fib and LTF) (Supplemental Figure 10: Part H). Produced by *HCD.low* fragmentation, this oxonium ion can give additional confidence in the *N*-glycopeptide annotation.

Mucin-type O-glycopeptides

Analysis of mucin-type *O*-glycopeptides revealed that differentiating mono- and disialylated *O*-glycoforms based on their relative oxonium ion abundance seems to be not possible (Supplemental Figure 9: Part M (I, II)). However, a general comparison of the relative oxonium ion abundance acquired for core-1 mucin-type *O*-glycopeptides with those acquired for *N*-glycopeptides, revealed significant differences that enable differentiation of these two forms of protein glycosylation. The most striking difference is the lack of the HexNAc₁Hex₂ oxonium ion for core-1 mucin-type *O*-glycopeptides. This oxonium ion was found consistently across all analyzed *N*-glycopeptides, independent of the applied collision energy (Supplemental Figure 9: Parts A-J). It appears to represent a

characteristic oxonium ion that occurs upon fragmentation of (I) the chitobiose core – a main fragmentation event during *N*-glycopeptide fragmentation – and (II), optionally, further antenna-directed fragmentation steps, to ultimately generate a fragment ion corresponding to 2nd chitobiose GlcNAc with two attached mannoses (HexNAc₁Hex₂). Additionally, also fragment ions derived from *N*-glycan antennae LacNAc (+mannose) residues can give rise to a HexNAc₁Hex₂ oxonium ion. Another important observation relates to the HexNAc and HexNAc -H₂O oxonium ion ratio. In agreement with Halim *et al.*, using HCD.step this ratio allows to discriminate *N*-glycopeptides from *O*-glycopeptides (HexNAc -H₂O/ HexNAc ratio: for *N*-glycopeptides 0.1-0.3; for *O*-glycopeptides 0.85-1.2). Details can be found in supplemental 1.10.

A summary of all findings related to the relative oxonium ion quantitation is given in Table 1.

Table 1: Diagnostic potential of oxonium ions - Overview

		Supplemental Figure
Complex-type <i>N</i>-glycopeptides		
General Observations	<ul style="list-style-type: none"> ▪ Oxonium ion abundance are characteristic for specific <i>N</i>-glycan features ▪ Oxonium ion abundance of a specific <i>N</i>-glycoform are independent of the peptide backbone ▪ Oxonium ion abundance of a specific <i>N</i>-glycoform are mostly independent of the precursor ion charge state ▪ Oxonium ion abundance depend on applied collisional energy (<i>HCD.low</i>, <i>HCD.step</i>) ▪ <i>HCD.step</i>: decrease/absence of di- and trisaccharide oxonium ions compared to <i>HCD.low</i> ▪ Differentiation of complex-, hybrid- and high-mannose-type <i>N</i>-glycopeptides based on oxonium ion abundance possible (<i>HCD.low</i>) <ul style="list-style-type: none"> ▪ Diagnostic oxonium ion (<i>HCD.low</i>): Hex Relative abundance: high-mannose-type >> hybrid-type > complex-type 	S-F 9: A and B S-F 9: A and B S-F 9: C S-F 9: D S-F 9: E (II, V), J (I-II) S-F 9: I
Antenna GlcNAc vs Bisecting GlcNAc	<ul style="list-style-type: none"> ▪ Differentiation of antenna GlcNAc and bisecting GlcNAc based on oxonium ion abundance and retention time possible (<i>HCD.low</i>) <ul style="list-style-type: none"> ▪ Diagnostic fragment ions for bisecting GlcNAc (<i>HCD.low</i>): <ul style="list-style-type: none"> ▪ Peptide+HexNAc₃Hex₁ (Y-ion, [M+H]⁺, [M+2H]²⁺) ▪ HexNAc₂ (oxonium ion, B-ion, [M+H]⁺) ▪ Retention time: bisecting GlcNAc > antenna GlcNAc 	S-F 10: A-D, and F
Core- vs Antenna-Fucosylation	<ul style="list-style-type: none"> ▪ Differentiation of core- and antenna-fucosylation based on oxonium ion abundance possible (<i>HCD.low</i>) <ul style="list-style-type: none"> ▪ Diagnostic oxonium ion (<i>HCD.low</i>): HexNAc₁Hex₁Fuc₁ Relative abundance: antenna fucose > core fucose 	S-F 9: F (I-II)
Type of Sialic Acid, Degree of Sialylation	<ul style="list-style-type: none"> ▪ Differentiation of the type of sialic acid (NANA, NGNA) based on diagnostic oxonium ions possible (<i>HCD.low</i>, <i>HCD.step</i>) ▪ Predicting the degree of <i>N</i>-acetylneuraminic acid sialylation based on diagnostic oxonium ion abundance seems promising, though further analyses are necessary (<i>HCD.low</i>, <i>HCD.step</i>) 	S-F 2: pp 98-100 S-F 9: G
High-Mannose-Type <i>N</i>-Glycopeptides	<ul style="list-style-type: none"> ▪ Direct correlation between the degree of mannosylation and the relative abundance of the diagnostic Hex oxonium ion (<i>HCD.low</i>) Relative abundance of Hex oxonium ion: e.g. Man₁<Man₅<Man₉ 	S-F 9: H (I-II)
Hybrid-Type <i>N</i>-Glycopeptides	<ul style="list-style-type: none"> ▪ Combination of complex- and high-mannose-type related diagnostic oxonium ions 	S-F 9: I
Mucin-Type <i>O</i>-Glycopeptides	<ul style="list-style-type: none"> ▪ Differentiation of core-1 mucin-type <i>O</i>-glycopeptides and <i>N</i>-glycopeptides based on diagnostic oxonium ion abundance possible (<i>HCD.low</i>, <i>HCD.step</i>) <ul style="list-style-type: none"> ▪ Diagnostic oxonium ion (<i>HCD.low</i>, <i>HCD.step</i>): HexNAc₁Hex₂ HexNAc₁Hex₂ absent in <i>O</i>-glycopeptide fragment ion spectra ▪ HexNAc -H₂O/ HexNAc ratio (<i>HCD.step</i>): for <i>N</i>-glycopeptides: 0.1-0.3; for <i>O</i>-glycopeptides: 0.85-1.2 ▪ Differentiation of mono- and disialylated <i>O</i>-glycopeptides (NANA) based on diagnostic oxonium ion abundance not possible (<i>HCD.low</i>, <i>HCD.step</i>) 	S-F 9: K (I-II) S-F 9: A-J S-F 9: L and M

SUMMARY AND OUTLOOK: Over the last years site-specific mass spectrometry-based *N*- and *O*-glycoproteomic analyses have become more and more advanced, more reliable, and more popular. By introducing novel techniques and software solutions, the detailed characterization of single glycoproteins or mixtures of glycoproteins with different types of glycosylation and multiple glycosylation sites, has become feasible. However, the analysis of intact glycosylated peptides remains challenging and still has potential for improvements, for instance with respect to quantitation and structural analysis of the glycan moiety. Facing multiple layers of complexity by unambiguously identifying two structurally diverse entities – the glycan moiety and the peptide backbone – dedicated analysis workflows along with a basic understanding of the glycopeptide fragmentation process and *a priori* knowledge about the sample to be analyzed are crucial. Here, we present a universal glycoproteomic workflow that enables the in-depth analysis of *N*- and *O*-glycopeptides derived from single proteins but also from complex protein samples (not shown in this study).

To this end, glycopeptides generated by tryptic or proteinase K digest, were enriched by spin-cotton-HILIC-SPE and subjected to reversed-phase liquid chromatography coupled to tandem mass spectrometry analysis. The use of two HCD fragmentation regimes, *HCD.low* and *HCD.step*, enabled the unambiguous identification of the peptide backbone and allowed the detailed analysis of the glycan moiety. Using a set of four representative *N*- and *O*-glycosylated proteins (IgG, Fib, LTF, and RNase B) the versatility, robustness and comprehensiveness of the developed workflow was demonstrated. In addition, diagnostic oxonium ion patterns that are characteristic for certain glycoforms and glycan features are discussed in detail, and we report on a conserved *N*-glycopeptide fragmentation pattern, which enables the reliable determination of the peptide mass. Our findings are applicable for a broad range of glycoproteins, can significantly increase the confidence in manual and software-assisted analysis of *N*- and *O*-glycopeptides and can thus facilitate MS-based glycoanalytics.

ACKNOWLEDGEMENT: The authors gratefully acknowledge the excellent technical assistance of Silvana Fischer. Moreover, the authors like to thank Samanta Čajić and Terry Nguyen-Khuong for support and discussions. MH and ER acknowledge support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under the project “The concert of dolichol-based glycosylation: from molecules to disease models” (grant identifier FOR2509). ER acknowledges support by the European Union (EC) under the project “IMforFuture” (grant no. 721815).

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS: M.H. conceived and designed the method, designed and performed experiments, analyzed data and wrote the manuscript. M.P. developed glyXtool^{MS}, analyzed data, and reviewed and edited the manuscript. A.P. designed the method, and reviewed and edited the manuscript. R.H., R.K. and UR reviewed and edited the manuscript. E.R. conceived and designed the method, and wrote the manuscript.

References:

- [1] M. Thaysen-Andersen, N. H. Packer, B. L. Schulz, *Mol. Cell. Proteomics* 2016, 15, 1773.
- [2] M. Thaysen-Andersen, N. H. Packer, *Biochim. Biophys. Acta* 2014, 1844, 1437.
- [3] S. B. Lavery, C. Steentoft, A. Halim, Y. Narimatsu, H. Clausen, S. Y. Vakhrushev, *Biochim. Biophys. Acta* 2014, 1850, 33.
- [4] G. Zauner, M. Hoffmann, E. Rapp, C. A. Koeleman, I. Dragan, A. M. Deelder, M. Wuhler, P. J. Hensbergen, *J Proteome Res* 2012, 11, 5804.
- [5] M. Hoffmann, K. Marx, U. Reichl, M. Wuhler, E. Rapp, *Mol. Cell. Proteomics* 2016, 15, 624.
- [6] R. Plomp, G. Dekkers, Y. Rombouts, R. Visser, C. A. M. Koeleman, G. S. M. Kammeijer, B. C. Jansen, T. Rispens, P. J. Hensbergen, G. Vidarsson, M. Wuhler, *Mol. Cell. Proteomics* 2015, 14, 1373.
- [7] S. Hua, C. Y. Hu, B. J. Kim, S. M. Totten, M. J. Oh, N. Yun, C. C. Nwosu, J. S. Yoo, C. B. Lebrilla, H. J. An, *J Proteome Res* 2013, 12, 4414.
- [8] G. Zauner, C. A. Koeleman, A. M. Deelder, M. Wuhler, *J Sep Sci* 2010, 33, 903.
- [9] K. Stavenhagen, R. Plomp, M. Wuhler, *Anal. Chem.* 2015, 87, 11691.
- [10] D. Kolarich, P. H. Jensen, F. Altmann, N. H. Packer, *Nat. Protoc.* 2012, 7, 1285.
- [11] N. Jentoft, *Trends Biochem. Sci.* 1990, 15, 291.
- [12] S. Ongay, A. Boichenko, N. Govorukhina, R. Bischoff, *J. Sep. Sci.* 2012, 35, 2341.
- [13] Z. Darula, J. Sherman, K. F. Medzihradzsky, *Mol. Cell. Proteomics* 2012, 11, O111.016774.
- [14] K. Khatri, G. O. Staples, N. Leymarie, D. R. Leon, L. Turiák, Y. Huang, S. Yip, H. Hu, C. F. Heckendorf, J. Zaia, *J Proteome Res* 2014, 13, 4347.
- [15] J. Nilsson, U. Ruetschi, A. Halim, C. Hesse, E. Carlsohn, G. Brinkmalm, G. Larson, *Nat Meth* 2009, 6, 809.
- [16] I. M. Lazar, J. Deng, F. Ikenishi, A. C. Lazar, *Electrophoresis* 2015, 36, 225.
- [17] I. Walsh, S. Zhao, M. Campbell, C. H. Taron, P. M. Rudd, *Curr. Opin. Struct. Biol.* 2016, 40, 70.
- [18] H. Hu, K. Khatri, J. Klein, N. Leymarie, J. Zaia, *Glycoconj. J.* 2016, 33, 285.
- [19] L. Y. Lee, E. S. X. Moh, B. L. Parker, M. Bern, N. H. Packer, M. Thaysen-Andersen, *J. Proteome Res.* 2016, 15, 3904.
- [20] H. Hinneburg, K. Stavenhagen, U. Schweiger-Hufnagel, S. Pengelley, W. Jabs, P. H. Seeberger, D. V. Silva, M. Wuhler, D. Kolarich, *J. Am. Soc. Mass Spectrom.* 2016, 27, 507.
- [21] Q. Cao, X. Zhao, Q. Zhao, X. Lv, C. Ma, X. Li, Y. Zhao, B. Peng, W. Ying, X. Qian, *Anal Chem* 2014, 86, 6804.
- [22] J. Yu, M. Schorlemer, A. Gomez Toledo, C. Pett, C. Sihlbom, G. Larson, U. Westerlind, J. Nilsson, *Chemistry - A European Journal* 2016, 22, 1114.
- [23] A. Halim, U. Westerlind, C. Pett, M. Schorlemer, U. Rüetschi, G. Brinkmalm, C. Sihlbom, J. Lengqvist, G. Larson, J. Nilsson, *J. Proteome Res.* 2014, 13, 6024.
- [24] A. M. Mayampurath, Y. Wu, Z. M. Segu, Y. Mechref, H. Tang, *Rapid Commun. Mass Spectrom.* 2011, 25, 2007.
- [25] S. Toghi Eshghi, W. Yang, Y. Hu, P. Shah, S. Sun, X. Li, H. Zhang, *Scientific Reports* 2016, 6, 37189.
- [26] B. Domon, C. E. Costello, *Glycoconj. J.* 1988, 5, 397.
- [27] F. G. Hanisch, *Biol. Chem.* 2001, 382, 143.

- [28] A. M. Falick, W. M. Hines, K. F. Medzihradzsky, M. A. Baldwin, B. W. Gibson, *J. Am. Soc. Mass Spectrom.* 1993, 4, 882.
- [29] I. A. Papayannopoulos, *Mass Spectrom. Rev.* 1995, 14, 49.
- [30] M. Wuhrer, M. I. Catalina, A. M. Deelder, C. H. Hokke, *J Chromatogr B Analyt Technol Biomed Life Sci* 2007, 849, 115.
- [31] Q. Dong, X. Yan, Y. Liang, S. E. Stein, *J Proteome Res* 2016, 15, 1472.
- [32] J. Stadlmann, J. Taubenschmid, D. Wenzel, A. Gattinger, G. Durnberger, F. Dusberger, U. Elling, L. Mach, K. Mechtler, J. M. Penninger, *Nature* 2017, 549, 538.
- [33] M. H. Selman, M. Hemayatkar, A. M. Deelder, M. Wuhrer, *Anal Chem* 2011, 83, 2492.
- [34] J. E. Huffman, M. Pučić-Baković, L. Klarić, R. Hennig, M. H. J. Selman, F. Vučković, M. Novokmet, J. Krištić, M. Borowiak, T. Muth, O. Polašek, G. Razdorov, O. Gornik, R. Plomp, E. Theodoratou, A. F. Wright, I. Rudan, C. Hayward, H. Campbell, A. M. Deelder, U. Reichl, Y. S. Aulchenko, E. Rapp, M. Wuhrer, G. Lauc, *Mol. Cell. Proteomics* 2014, 13, 1598.
- [35] F. Clerc, K. R. Reiding, B. C. Jansen, G. S. M. Kammeijer, A. Bondt, M. Wuhrer, *Glycoconj. J.* 2016, 33, 309.
- [36] T. Liu, W. J. Qian, M. A. Gritsenko, D. G. Camp II, M. E. Monroe, R. J. Moore, R. D. Smith, *J Proteome Res* 2005, 4, 2070.
- [37] S. Karav, J. B. German, C. Rouquie, A. Le Parc, D. Barile, *International journal of molecular sciences* 2017, 18, 870.
- [38] P. H. van Berkel, H. A. van Veen, M. E. Geerts, H. A. de Boer, J. H. Nuijens, *Biochem. J.* 1996, 319 (Pt 1), 117.
- [39] M. Wuhrer, A. R. de Boer, A. M. Deelder, *Mass Spectrom. Rev.* 2009, 28, 192.
- [40] K. Jiang, H. Zhu, C. Xiao, D. Liu, G. Edmunds, L. Wen, C. Ma, J. Li, P. G. Wang, *Anal. Chim. Acta* 2017, 962, 32.
- [41] M. Wuhrer, C. A. Koeleman, C. H. Hokke, A. M. Deelder, *Rapid Commun Mass Spectrom* 2006, 20, 1747.
- [42] K. F. Medzihradzsky, K. Kaasik, R. J. Chalkley, *Anal. Chem.* 2015, 87, 3064.