

Characterization of Site-Specific *N*-Glycopeptide Isoforms of α -1-Acid Glycoprotein from an Interlaboratory Study Using LC–MS/MS

Ju Yeon Lee,^{†,•} Hyun Kyoung Lee,^{†,‡,•} Gun Wook Park,^{†,‡} Heeyoun Hwang,[†] Hoi Keun Jeong,^{†,‡} Ki Na Yun,^{†,§} Eun Sun Ji,^{†,||} Kwang Hoe Kim,^{†,‡} Jun Seok Kim,[⊥] Jong Won Kim,[#] Sung Ho Yun,[▽] Chi-Won Choi,[▽] Seung Il Kim,[▽] Jong-Sun Lim,[○] Seul-Ki Jeong,[○] Young-Ki Paik,[○] Soo-Youn Lee,^{◆,¶} Jisook Park,⁺ Su Yeon Kim,[□] Young-Jin Choi,[▲] Yong-In Kim,[▲] Jawon Seo,[▲] Je-Yoel Cho,[▲] Myoung Jin Oh,[‡] Nari Seo,[‡] Hyun Joo An,[‡] Jin Young Kim,^{*,†} and Jong Shin Yoo^{*,†,‡}

[†]Biomedical Omics Group, Korea Basic Science Institute, Ochang 28119, Republic of Korea

[‡]Graduate School of Analytical Science and Technology, Chungnam National University, Daejeon 34134, Republic of Korea

[§]Department of Chemistry, Sogang University, Seoul 04107, Republic of Korea

^{||}Department of Chemistry, Hannam University, Daejeon 34430, Republic of Korea

[⊥]Department of Biomedical Systems Engineering, Korea Polytechnics, Gyeonggi 13590, Republic of Korea

[#]New Drug Development Center, Osong Medical Innovation Foundation, Cheongju 28160, Republic of Korea

[▽]Drug & Disease Target Group, Korea Basic Science Institute, Daejeon 34133, Republic of Korea

[○]Yonsei Proteome Research Center, Yonsei University, Seoul 03722, Republic of Korea

[◆]Department of Laboratory & Genetics, Samsung Medical Center, Sungkyunkwan University of Medicine, Seoul 06351, Republic of Korea

[¶]Department of Clinical Pharmacology and Therapeutics, Samsung Medical Center, Seoul 06351, Republic of Korea

⁺Samsung Biomedical Research Institute, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul 06351, Republic of Korea

[□]Department of Clinical Research Supporting Team, Clinical Research Institute, Samsung Medical Center, Seoul 06351, Republic of Korea

[▲]Department of Biochemistry, BK21 PLUS Program for Creative Veterinary Science Research and Research Institute for Veterinary Science, College of Veterinary Medicine, Seoul National University, Seoul 08826, Republic of Korea

Supporting Information

Samples numbers	Experimental workflow	Test goals
Test sample 1	AGP protein standard → Protein reduction and alkylation → Trypsin digestion → LC-MS/MS analysis ←----- Laboratory 1 -----> ←----- Laboratory 1-7 ----->	To check performance of mass spectrometers
Test sample 2	AGP protein standard → Protein reduction and alkylation → Trypsin digestion → Glycopeptides enrichment → LC-MS/MS analysis ←----- Laboratory 1 -----> ←----- Laboratory 1-7 ----->	To check performance of sample enrichment
Test sample 3	AGP protein standard → Protein reduction and alkylation → Trypsin digestion → Glycopeptides enrichment → LC-MS/MS analysis ←----- Laboratory 1-7 ----->	To check performance of sample digestion/enrichment, and To check performance of mass spectrometer for standard sample
Blind sample	Crude sample → Protein reduction and alkylation → Trypsin digestion → Glycopeptides enrichment → LC-MS/MS analysis ←----- Laboratory 1-7 ----->	To validate performance of sample digestion/enrichment, and To check performance of mass spectrometer for real sample

ABSTRACT: Glycoprotein conformations are complex and heterogeneous. Currently, site-specific characterization of glycopeptides is a challenge. We sought to establish an efficient method of *N*-glycoprotein characterization using mass spectrometry (MS). Using alpha-1-acid glycoprotein (AGP) as a model *N*-glycoprotein, we identified its tryptic *N*-glycopeptides and examined the data reproducibility in seven laboratories running different LC–MS/MS platforms. We used three test samples and one blind sample to evaluate instrument performance with entire sample preparation workflow. 165 site-specific *N*-glycopeptides representative
continued...

Received: December 24, 2015

Published: October 19, 2016

of all *N*-glycosylation sites were identified from AGP 1 and AGP 2 isoforms. The glycopeptide fragmentations by collision-induced dissociation or higher-energy collisional dissociation (HCD) varied based on the MS analyzer. Orbitrap Elite identified the greatest number of AGP *N*-glycopeptides, followed by Triple TOF and Q-Exactive Plus. Reproducible generation of oxonium ions, glycan-cleaved glycopeptide fragment ions, and peptide backbone fragment ions was essential for successful identification. Laboratory proficiency affected the number of identified *N*-glycopeptides. The relative quantities of the 10 major *N*-glycopeptide isoforms of AGP detected in four laboratories were compared to assess reproducibility. Quantitative analysis showed that the coefficient of variation was <25% for all test samples. Our analytical protocol yielded identification and quantification of site-specific *N*-glycopeptide isoforms of AGP from control and disease plasma sample.

KEYWORDS: *interlaboratory study, site-specific N-glycopeptide, isoforms, PTM (post translational modification)*

■ INTRODUCTION

Comprehensive characterization of post-translational modifications (PTMs), such as glycosylation, phosphorylation, and acetylation, is one of the major tasks in understanding biological processes. More specifically, glycosylation is directly involved in numerous biological processes, such as embryonic development, cell adhesion, signal transduction, and immune responses.^{1–5} High-performance liquid chromatography (HPLC)–mass spectrometer (MS)^{6–9} is a widespread and effective technology used to identify glycosylation site and glycan structure, site occupancy, and glycan isoform distribution. This technique has been used to discover candidate glycoprotein biomarkers and yield information on aberrant glycan structure.¹⁰

Glycoproteins are modified by the attachment of glycans on a glycosylation site. *N*-glycosylation occurs mainly at the carboxamido–nitrogen ligands of the asparagine (Asn or N) residues where the consensus tripeptide sequence NXS/T is located (where X is any amino acid except P¹¹). At least one *N*-glycosylation^{7,8} motif is present in two-thirds of all human proteins, and 50% of all human proteins are *N*-glycosylated.¹² O-Glycosylation occurs in the hydroxyl side chains of nucleophilic residues such as serine and threonine, and C-glycosylation bonds on the indole rings of the tryptophan residues.¹³

Changes in glycoprotein abundance, glycosylation occupancy, and glycoform distribution are closely associated with various diseases.^{14–17} For example, it can affect protein functions at the cellular level in cancer, which differs from those in normal cells. Therefore, many studies on diseases have focused on monitoring changes in glycoprotein features.^{18,19} In glycoproteomics and glycomics, HPLC–MS is a common technique used to identify glycosylation sites^{20–23} and the structure of glycans.^{24,25} However, glycoprotein analysis is difficult and challenging because of the microheterogeneity of glycans attached to proteins and the diversity of glycosylation site. MS analysis of glycopeptides in complex samples remains more difficult due to the low proportions and poor ionization of individual glycopeptides.^{26,27}

To improve glycopeptide analysis, selective enrichment and separation of glycopeptides from peptide mixtures such as lectin affinity^{28–31} and hydrazide chemistry^{32,33} are generally coupled to HPLC–MS. The successful application of hydrophilic interaction chromatography (HILIC) for unbiased glycopeptide enrichment such as zwitterionic-HILIC (ZIC-HILIC),^{34–36} Sepharose-CL-4B^{37,38} (or -6B³⁹), and click maltose,^{40–42} has recently improved the relatively narrow specificity of lectin affinity and complemented the lack of glycan structural information from hydrazide chemical analysis.

In this study, ZIC-HILIC was used for glycopeptides enrichment. HILIC has been employed to purify various polar compounds in complex samples; the retention patterns are the opposite of those of reverse-phase liquid chromatography (RPLC).^{36,43} Hydrophilic partitioning of analytes builds up the

hydrogen bonding with the water-enriched layer of the polar stationary phase; electrostatic interactions also contribute in this context. To effectively enrich glycopeptides from peptide mixtures, trifluoroacetic acid (TFA) was used as the hydrophobic ion-pairing reagent,³⁶ where it reduces the hydrogen-bonding potential of peptides and the electrostatic interactions between charged peptides and the stationary phase. Glycans increase the hydrophilic character of glycopeptides.

From enriched *N*-glycopeptides, glycans can be released using chemicals and enzymes such as *N*-glycanase (PNGase F). The resulted deglycopeptides^{16,44–46} and the glycans^{47,48} can be analyzed by LC–MS/MS. The analysis of deglycopeptides leads to the identification of glycosylation sites, although information on glycan composition is lost. The analysis of the released glycans provides compositional information but no site-specific information. Recently, many researchers have interested in the analyses of intact *N*-glycopeptides^{12,38,49–51} to gain comprehensive information on glycosylation sites and site-specific glycan composition. The combination of multiple dissociation techniques (CID/HCD,²¹ HCD/ETD,⁵² or CID/ETD/HCD⁵³) and the relevant search algorithms such as GP Finder,⁵⁴ GPS,⁵⁵ Boynic,⁵⁶ GlycoFragwork,⁵⁷ Sweet-Heart,⁵⁸ MAGIC,⁵⁹ and I-GPA²¹ have improved the identification of intact *N*-glycopeptides.

To obtain useful information on glycoproteins, all analyses must be reliable and reproducible. In some studies, the suitability of analytical protocols or standardization methods from different laboratories was evaluated in same previous studies.^{60–62}

Glycoprotein characterization^{51,63,64} was also explored in the interlaboratory analyses for the *N*-glycopeptides of human prostate-specific antigen⁵¹ and O-glycomics of IgA1.⁶⁴ MS has been used to analyze *N*-linked glycans and their relative abundances.⁶³ Different methods including top-down and bottom-up approaches with or without PNGase F treatment were used by the participating laboratories, which employed MALDI-MS or ESI-MS in the analyses. They discussed differences among laboratories and the reproducibility of analytical methods. Leymarie et al.⁵¹ explored consistency among sample preparation, separation techniques, and analytical methods. It was concluded that bottom-up methods employing proteolytic enzymes were likely to be required for reliable glycoprotein analysis.

The goal of the present study was to establish a standard analytical protocol for identification and quantification of the site-specific *N*-glycopeptide isoforms from glycoproteins with multiple *N*-glycosylation sites. We explored differences among the instrument in glycopeptide fragmentation/identification/quantification. AGP standard and human serum samples were analyzed by different nano LC–ESI–MS/MS instruments in seven laboratories. Four samples (three test samples of AGP standard and one blind sample of human serum) were given to each laboratory, along with the sample information on the three test samples and the analytical protocols including protein digestion,

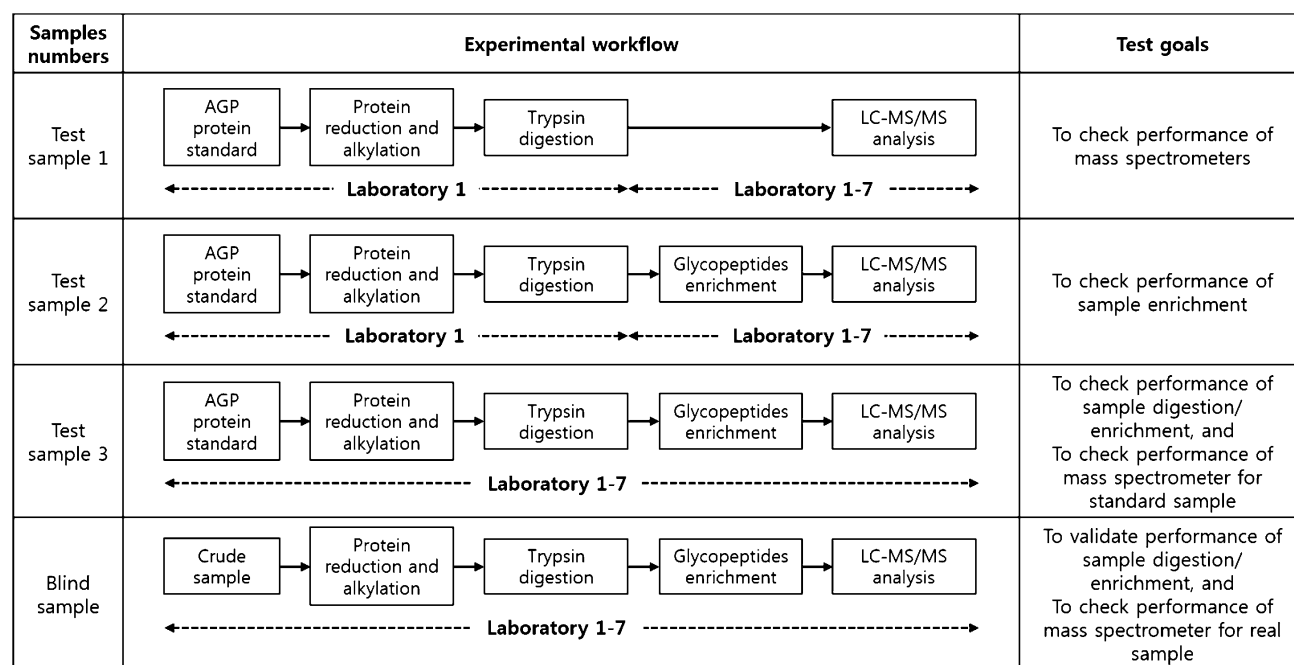


Figure 1. Experimental workflow for three test and one blind samples in an interlaboratory study.

glycopeptide enrichment by HILIC, and LC-ESI-MS/MS conditions. Samples (Figure 1) consisted of (1) a digested AGP standard to check each MS performance (test sample 1), (2) a digested AGP standard to check *N*-glycopeptides enrichment performance by HILIC (test sample 2), (3) an AGP standard protein (test sample 3), and (4) a human serum (a blind sample) to validate the performance of entire experimental processes including digestion, glycopeptide enrichment, and LC-MS/MS analysis. The analytical results for the identification and quantification of the *N*-glycopeptides of AGP from standard and human serum samples were compared among laboratories. Finally, the *N*-glycopeptides of AGP were quantitatively analyzed from pooled normal and hepatocellular carcinoma (HCC) human plasma samples

EXPERIMENTAL METHODS AND MATERIALS

Materials

AGP standard protein (human, cat. no. G9885) was purchased from Sigma-Aldrich (St. Louis, MO). The ZIC-HILIC kit (ProteoExtract Glycopeptide Enrichment Kit) was purchased from EMD Millipore (cat. no. 72103-3, Billerica, MA). 1,4-Dithiothreitol (DTT), iodoacetamide (IAA), formic acid (FA), and standard human serum (cat. no. S7023) were purchased from Sigma-Aldrich. Trypsin was obtained from Promega (Madison, WI). HPLC-grade acetonitrile (ACN) was obtained from J. T. Baker (Phillipsburg, NJ). Water was deionized using the Milli-Q Advantage A10 System (Millipore) prior to use.

Samples

Test and Blind Samples. 10 μg of the digested AGP (test sample 1), 200 μg of the digested AGP (test sample 2), 1 mg of the AGP standard protein (test sample 3), and 200 μg of serum (blind sample) in 50 mM ammonium bicarbonate (ABC) buffer were distributed to seven laboratories with trypsin and glycopeptide enrichment kits. For LC-MS/MS analysis, 0.15 μg of digested AGP (test sample 1), 1.5 μg of digested/enriched AGP (test samples 2 and 3), and 9 μg of digested/enriched serum

(blind sample) were injected onto the column in each laboratory. The amounts of protein injected were quantified using the Bradford protein assay in protein level.

Plasma Samples. Plasma samples in appropriate concentrations of K_2EDTA were obtained from Yonsei University College of Medicine (Seoul, Korea); we conformed to the IRB guidelines on informed consent and patient approval. Human plasma samples were frozen and stored at -80°C until use. Normal and HCC human plasma samples were pooled from 10 individuals for analysis. Information on each sample is given in Supplementary Table 1. For LC-MS/MS analysis, plasma samples were injected in the column as much as serum sample.

Depletion of Serum and Plasma Samples

The six most abundant proteins (albumin, transferrin, IgG, IgA, haptoglobin, and α_1 -antitrypsin) of serum and plasma samples were depleted using a HP1100LC system (Agilent Technologies, Santa Clara, CA) equipped with a multiple-affinity removal column (MARC; Agilent) according to the manufacturer's instructions. Flow-through fractions were collected and stored at -20°C .

Digestion of Samples

Standard AGP solution was prepared at a concentration of 1 $\mu\text{g}/\mu\text{L}$ in 50 mM ABC buffer. The AGP solution was denatured with 8 M urea at room temperature ($<25^\circ\text{C}$) for 10 min. The protein solution was reduced by the addition of 2 μL of 500 mM DTT at room temperature ($<25^\circ\text{C}$) for 1 h and was alkylated by the addition of 5 μL of 500 mM IAA in the dark at room temperature during a 1 h reaction time. Surplus reagents were removed by centrifugal filtration using a 10000 Da membrane cut off filter (Millipore, product code UFC501096). Aliquots (100 μg) of AGP that had been quantitatively analyzed by the Bradford protein assay were diluted with 50 mM ABC buffer. Diluted samples were digested with trypsin at 37°C overnight (16 h). The digestion was quenched in a 3% final concentration of FA. For LC-MS/MS analysis, the solution was diluted by mobile phase A (99.9% water with 0.1% FA) liquid chromatography.

For glycopeptide enrichment, the digested samples were dried in a SpeedVac.

Human sera and plasma samples were digested by the same as above protocol for AGP standard.

Glycopeptide Enrichment

Glycopeptide enrichment was performed using the ProteoExtract Glycopeptide Enrichment Kit (EMD Millipore) according to the manufacturer's instructions. The kit contains ZIC Glycocalculation Resin, ZIC Binding Buffer, ZIC Wash Buffer, and ZIC Elution Buffer. Then, 10 μL samples of AGP standard digest (30 μg), serum digest (20 μg), or plasma digest (30 μg) were prepared. The 10 μL digested samples were diluted with 50 μL of ZIC Binding Buffer. The ZIC Glycocalculation Resin was mixed, and 50 μL of homogeneous ZIC Glycocalculation Resin was transferred to a new microcentrifuge tube. The tube was centrifuged for 1 to 2 min at 2000–2500g, and the supernatant was completely removed and discarded. The diluted sample was added to the ZIC Glycocalculation Resin, mixed by pipetting, and incubated at 1200 rpm for 10–20 min. The tube was centrifuged and the supernatant was removed same as above. Next, 150 μL of ZIC Wash Buffer was added to the ZIC Glycocalculation Resin. It was mixed, incubated, and centrifuged, and the supernatant was removed. Total three repeats were performed. Then, 75–100 μL of ZIC Elution Buffer was added in it for the elution of glycopeptides. The tube was mixed, incubated, and centrifuged. The supernatant was transferred to a new microcentrifuge tube and centrifuged for 2 min at 10 000g, then transferred to a new microcentrifuge tube, avoiding the transfer of any resin particles. The supernatant contained the eluted glycopeptides. The elutions were dried in a SpeedVac and redissolved in 0.1% FA/99.9% H_2O in preparation for LC–MS/MS analysis.

Nano LC–ESI–MS/MS Analysis

General Nano LC Conditions for the Interlaboratory Study. Digested or enriched samples were dissolved in mobile phase A and analyzed using an LC–MS/MS system consisting of a nano LC and a mass spectrometer equipped with a nano electrospray source in laboratories 1–7 (Table 1). An auto-sampler was used to load the sample solutions into a C_{18} trap column (except in laboratory 5). The samples were desalted and concentrated on the trap column for 5–10 min at a flow rate of several $\mu\text{L}/\text{min}$. The trapped samples were then separated on a C_{18} analytical column. The mobile phases were composed of 99.9% water (A) and 99.9% ACN (B), and each contained 0.1% FA. The LC gradient started with 5% of B for 15 min and was ramped to 15% of B for 5 min, 50% of B for 75 min, and 95% of B for 1 min; it was held at 95% of B for 13 min and then at 5% of B for another 1 min. The column was re-equilibrated with 5% of B

for 10 min before the next run. For the serum/plasma sample analysis, the LC gradient time was extended to 180 min. Two laboratories utilized LTQ Orbitrap Elites (Thermo Scientific) differing in analytical dimensions. The other laboratories employed different mass spectrometers including a Triple TOF 5600+ (ABSciex), a Q-Exactive Plus (Thermo Scientific), an LTQ Orbitrap XL (Thermo Scientific), an LTQ Orbitrap (Thermo Scientific), and a Q-TOF (Agilent). All LC–MS/MS conditions including analytical and trapping column dimensions, flow rates, the MS instruments used (vendors and models), resolutions, and MS/MS modes are summarized in Table 1; more detailed information follows.

LTQ Orbitrap Elite (Laboratories 1 and 2) and LTQ Orbitrap XL (Laboratory 5). The voltage applied to produce the electrospray was 2.2 kV. During the chromatographic separation, the mass spectrometers were operated in a data-dependent acquisition (DDA) mode. The MS data were collected using the following parameters: Full scans were acquired in the Orbitrap at a resolution of 60 000 for each sample; five collision-induced dissociation (CID) and five higher-energy collisional dissociations (HCD) scans per full scan; CID scans were acquired in linear trap quadrupole (LTQ) with a 30 ms activation time performed for each sample with 35% normalized collision energy (NCE) and a ± 5.0 Da isolation window; and HCD scans were acquired in the Orbitrap at a resolution of 15 000 with 20 ms activation for each sample with 35% NCE and ± 5.0 Da isolation window. Previously fragmented ions were excluded for 180 s (300 s in laboratory 1).

LTQ Orbitrap (Laboratory 6). The voltage applied to produce the electrospray was 2.0 kV. During the chromatographic separation, the LTQ Orbitrap was operated in a DDA mode. The MS data were acquired using the following parameters: Full scans were acquired in Orbitrap at a resolution of 60 000 for each sample; five CID scans per full scan; CID scans were acquired in LTQ with a 30 ms activation time performed for each sample with 35% NCE and a ± 2.5 Da isolation window. Previously fragmented ions were excluded for 180 s.

Triple TOF (Laboratory 3). The voltage of the nano electrospray was 2.0 kV. During the chromatographic separation, the triple TOF was performed in an information-dependent acquisition (IDA) mode. Full scans were acquired in TOF at a resolution of 35 000 for each sample. The 10 CID scans with a 100 ms acquisition time were performed at a resolution of 15 000 for each sample and rolling energy for CID was set. Previously fragmented ions were excluded for 45 s after two occurrences.

Q-Exactive Plus (Laboratory 4). Conditions for mass spectrometry were similar to those of Orbitrap Elite. The resolution was set to 70 000 for MS and 17 500 for MS/MS. The top 10

Table 1. LC–MS/MS Conditions Used in Participating Laboratories

laboratory	analytical column	trap column	flow rate (nL/min)	MS instrument	MS resolution	MS/MS fragmentation
1	C_{18} 3 μm , 100 μm \times 200 mm	C_{18} 3 μm , 75 μm \times 20 mm	300	Thermo, LTQ Orbitrap Elite	60 000	CID(s)/HCD(s)
2	C_{18} 3 μm , 75 μm \times 150 mm	C_{18} 3 μm , 75 μm \times 20 mm	300	Thermo, LTQ Orbitrap Elite	60 000	CID(s)/HCD(s)
3	C_{18} 3 μm , 100 μm \times 200 mm	C_{18} 5 μm , 180 μm \times 20 mm	300	AB SCIEX, Triple TOF 5600+	35 000	CID(10)
4	C_{18} 2 μm , 75 μm \times 150 mm	C_{18} 5 μm , 100 μm \times 20 mm	300	Thermo, Q-Exactive Plus	70 000	HCD(10)
5	C_{18} 3 μm , 100 μm \times 150 mm	–	500	Thermo, LTQ Orbitrap XL	60 000	CID(s)/HCD(s)
6	C_{18} 3 μm , 100 μm \times 200 mm	C_{18} 5 μm , 180 μm \times 20 mm	300	Thermo, LTQ Orbitrap	60 000	CID(s)
7	C_{18} 5 μm , 75 μm \times 150 mm, chip	C_{18} 5 μm , 9 mm, 160 nL, chip	200	Agilent, Q-TOF G6520B	25 000	CID(3)

HCD scans per full scan were acquired in Orbitrap with a 27% NCE and ± 2.0 Da isolation window. Previously fragmented ions were excluded for 20 s.

Q-TOF (Laboratory 7). The voltage of the nano-electrospray was 2.0 kV. During the chromatographic separation, the Q-TOF was performed in the DDA mode. Full scans were acquired in TOF at a resolution of 25 000 for each sample, and the rolling energy for three CID scans was set. Previously fragmented ions were excluded after two occurrences.

Identification and Quantification

Peptide Identification. MS/MS spectra were analyzed using the following software analysis protocols. For tryptic peptide (nonglycopeptide) identification, MASCOT (version 2.4.0, Matrix Science, London, U.K.), operated on a local server, was used to search the AGP (P02763 and P19652). The AGP database of the Uniprot human 20141029 database release contained two relevant protein sequences (402 residues). MASCOT was used with monoisotopic mass selected, a precursor mass tolerance, and a fragment mass tolerance was set differently according to the mass spectrometer (Supplementary Table 2). The enzyme selected was trypsin, with one potentially missed cleavage. Oxidized methionine and carbamidomethylated cysteine were chosen as the variable modifications and fixed modifications, respectively. Peptides with a *p* value <0.05 (as automatically estimated by MASCOT) were considered authentic.

N-Glycopeptide Identification. Site-specific AGP N-glycopeptides were automatically identified by Integrated GlycoProteome Analyzer (I-GPA);²¹ our group previously developed software facilitating such N-glycopeptide analysis. In brief, N-glycopeptide tandem MS/MS spectra were first selected by reference to 15 glycan-specific oxonium ions. Second, N-glycopeptide candidates were selected by matching their experimental MS isotope patterns to the theoretical patterns of N-glycopeptides in the GAP-DB, by combining possible tryptic peptides including N-glycosites of AGP and 351 N-glycans (331 retrosynthetic glycans from Kronewitter et al.⁶⁵ and 20 glycans of penta and hexa polyactosamine series from Ozohanic et al.⁶⁶). Finally, N-glycopeptides of AGP were identified by their Y-score upon matching of experimental and theoretical fragment ions. To calculate Y-scores, B/Y ions from glycosidic bond cleavage and b/y ions from peptide bond cleavage were considered.

All data sets from seven laboratories were automatically evaluated in host laboratory using I-GPA. The common I-GPA search parameters were: missed cleavages (= 0) and fixed modification of carbamidomethyl cysteine and mammalian N-glycan. The mass tolerances for precursor and fragment ions were set according to MS analyzer used (Supplementary Table 2). To count the number of identified N-glycopeptides, the Y-score was set above 40. For the laboratory 7, the Y-score was set to 20 because there was no result with above 40. Finally, we extracted uniquely identified N-glycopeptides with the highest Y-score and manually checked.

N-glycopeptide nomenclature reflects the relevant peptide sequence and the numbers of hexose (Hex), N-acetylglucosamine (GlcNAc), fucose (Fuc), and N-acetyl neuraminic acid (NeuAc) moieties (#Hex_#HexNac_#Fuc_#NeuAc). The nomenclature of glycan is based on the database reference data⁶⁵ on 331 retrosynthetic glycans. For example, the LVPVPITNATLDR_6_5_0_3 N-glycopeptide is composed of the LVPVPITNATLDR peptide and six hexose, five N-acetylglucosamine, no fucose, and three N-acetyl neuraminic acid moieties. The italicized and underlined N (asparagine) is N-glycosylation site.

All data sets from seven laboratories have been uploaded to the ProteomeXchange Consortium via MassIVE repository (the data set identifiers PXD004677 and MSV000080010, respectively). The raw files are available for download at <ftp://massive.ucsd.edu/MSV000080010>.

Quantification. For quantitative analysis of N-glycopeptides, the extracted ion chromatograms (XICs) of the 10 major abundant N-glycopeptides from AGP were extracted within the various *m/z* intervals described in Supplementary Table 2. Only the monoisotopic signal of the precursor was considered, and the areas derived from different charge states were summed.

RESULTS AND DISCUSSION

Experimental Workflow for the Interlaboratory Study

AGP is a representative N-glycoprotein found in human serum. AGP consists of two protein isoforms, AGP 1 and AGP 2 (Figure 2a), and contains a glycan content of 45% (w/w) in terms of total molecular weight. There are five N-glycosylation sites with a molecular weight of 41–43 kDa.⁶⁷ AGP is composed of 18 amino acids as a signal moiety and 183 amino acids as the mature peptide chain,⁶⁸ where 21 amino acid sites are genetic variants. (The genetic homology between the two protein isoforms is ~90%.) Although each AGP isoform has five N-glycosites, two of the five are shared between AGP 1 and AGP 2 (Figure 2b).

For the interlaboratory study, three test samples and one blind sample, trypsin, and a glycopeptide enrichment kit were provided by host laboratory with the preparation process as well as a guide for the general nano LC–MS/MS conditions based on a bottom-up approach. The sample preparation process of each sample is shown in Figure 1. Test sample 1 of the digested AGP standard was prepared in the host laboratory for an MS performance test in all participating laboratories. In each laboratory, test sample 1 was diluted to the same concentration before LC–MS/MS analysis. Test sample 2 was prepared to validate the reproducibility of N-glycopeptide enrichment by HILIC. The N-glycopeptides of test sample 2 (digested AGP standard) were individually enriched and analyzed in each laboratory. Test sample 3 was used to confirm entire experimental workflow from digestion to LC–MS/MS analysis of a standard sample. The three test samples consisted of AGP standard protein, and the blind sample consisted of human serum depleted of the six most abundant proteins. The entire experiment workflow was additionally evaluated with the blind sample.

The LC–MS/MS systems used in the seven participating laboratories (three universities, three government institutes, and one medical hospital) are summarized in Table 1. Different C_{18} analytical columns with flow rates of 200–500 nL/min were used in all laboratories. Different C_{18} trapping columns were used in six laboratories. Only laboratory 7 used chip-based trapping and analytical columns. Four linear ion trap–Orbitrap analyzers, one quadrupole–Orbitrap analyzer, and two Q-TOF analyzers were used. Except for LC–MS/MS platforms, the entire protocol for the characterization of N-glycopeptides was the same in this study. Unlike in our study, Leymarie et al.⁵¹ allowed different processes, such as native glycoprotein (top-down), N-glycopeptide (bottom-up), and released N-glycan analysis (PNGase F digestion) used in 24 participating laboratories. They only provided information on glycan composition because targeted human prostate specific antigen (PSA) isoforms have one N-glycosite. Here we provided information on site-specific N-glycopeptide isoforms including peptides and glycans.

(a) Amino acid sequences of alpha-1-acid glycoprotein isoforms (AGP 1 and 2).

Alpha 1-acid glycoprotein 1 (AGP1) : P02763

1 **MALS**WVLTVL **SL**LPLEAQI PLCANLVPVP **IT****NATL**DQIT GKWFYIASAF RNEEY**M**KSVQ EIQTFFYFT **P****M**KTEDTIFL
 81 REYQTRQ**D**QC **IY****MT**TYLNVQ **RE****AG**TISRIV GGQEHFAHLL ILRDTKTYML **AF**DVNDEKNW GLSVYADKPE TTKEQLGEFY
 161 EALDCL**R**IPK SDV**V**YTDWKK DKCEPLEKQH EKERKQEEGE S

Alpha 1-acid glycoprotein 2 (AGP2) : P19652

1 **MALS**WVLTVL **SL**LPLEAQI PLCANLVPVP **IT****NATL**DRIT GKWFYIASAF RNEEY**M**KSVQ EIQTFFYFT **P****M**KTEDTIFL
 81 REYQTRQ**N**QC **FY****MS**SYLNVQ **E****AG**TISRIV GGREHVAHLL FLRDTKTYML **GS**YLDDEKNW GLSFYADKPE TTKEQLGEFY
 161 EALDCL**C**IPR SDV**M**YTDWKK DKCEPLEKQH EKERKQEEGE S

(b) List of N-glycopeptides from AGP 1 and 2.

N-glycosite	Alpha-1-acid glycoprotein 1 (AGP 1) Accession No. P02763	Alpha-1-acid glycoprotein 2 (AGP 2) Accession No. P19652
Site 1	LVPVPIT <u>N</u> ATLDQITGK (8-24, Asn ¹⁵)	LVPVPIT <u>N</u> ATLDR (8-20, Asn ¹⁵)
Site 2	NEEY <u>N</u> K(34-39, Asn ³⁸)	
Site 3	SVQEIQATFFYFT <u>P</u> <u>M</u> K(40-55, Asn ⁵⁴)	
Site 4	QDQCIY <u>N</u> TTYLNVQR (69-83, Asn ⁷⁵)	QNQCFY <u>N</u> SSYLNQR (69-83, Asn ⁷⁵)
Site 5	E <u>N</u> GTISR (84-90, Asn ⁸⁵)	E <u>N</u> GTVSR (84-90, Asn ⁸⁵)

Figure 2. Alpha-1-acid glycoprotein sequence information. (a) Amino acid sequences of alpha-1-acid glycoprotein isoforms (AGP 1 and 2). Bold black letters: 18 N-terminal amino acids represent the signal peptide sequence; italic and bold red letters: N-glycosylation sites; blue letters: sites of genetic variance; underlined letters: tryptic N-glycopeptides. (b) List of N-glycopeptides from AGP 1 and 2. Italicized and bold black letters: N-glycosylation sites.

Identification of N-Glycopeptides

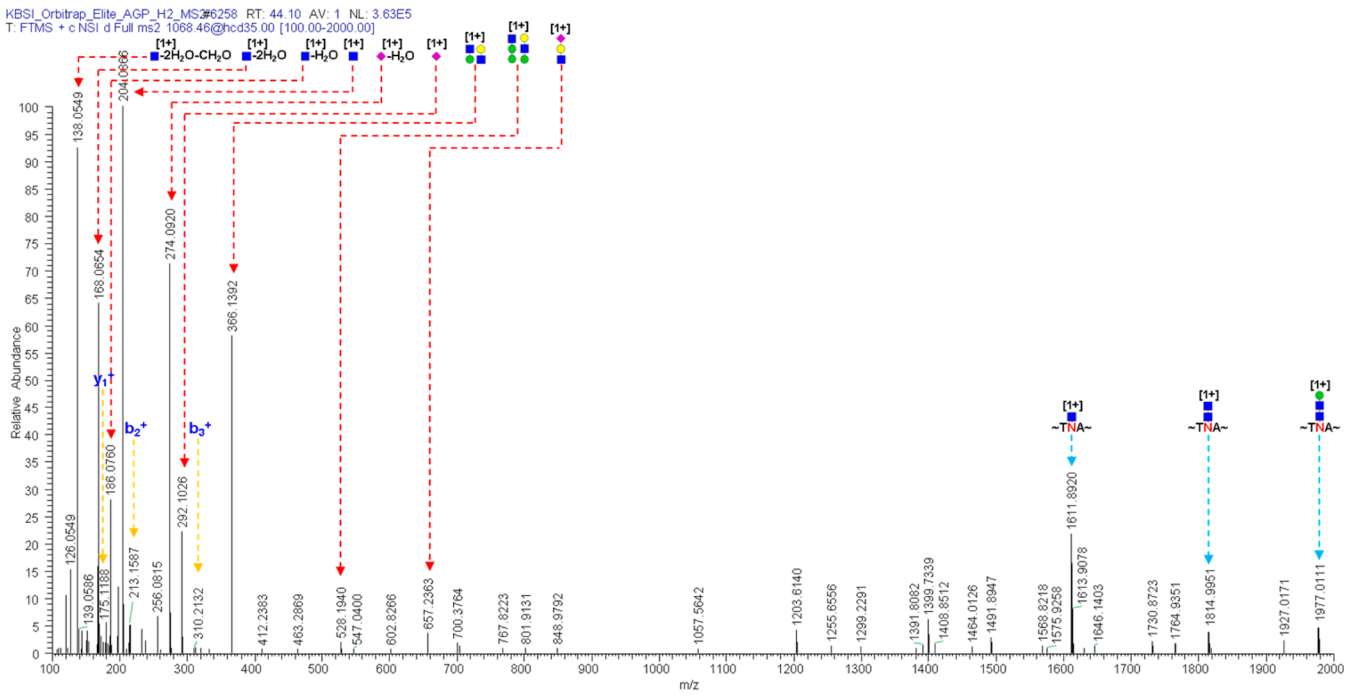
All N-glycopeptides from AGP were automatically identified by I-GPA software as described in the N-glycopeptide identification section of “experimental method and materials.” Figure 3 shows the MS/MS spectra of the representative N-glycopeptide, LVPVPIT**N**ATLDR_6_5_0_3, identified from test sample 2 in all laboratories. They commonly showed oxonium ions, glycan-cleaved glycopeptide fragment ions (B/Y ions), and peptide backbone fragment ions (b/y ions); these are typical fragment ions of an N-glycopeptide. N-Glycopeptides were identified by matching of B/Y ions and b/y ions from HCD spectra (Figure 3a-1,b-1,e-1) as well as from CID spectra (Figure 3a-2,b-2,e-2) in LTQ-Orbitrap Elite (laboratories 1 and 2) and XL (laboratory 5) analyses. In LTQ-Orbitrap (laboratory 6), only the CID spectrum (Figure 3f) was used for the identification of N-glycopeptides (the HCD was not available). In Q-Exactive Plus (Figure 3d, laboratory 4), only the HCD mode is available for the fragmentation of precursor ions. Triple TOF (Figure 3c, laboratory 3) and Q-TOF (Figure 3g, laboratory 7) are possible only in CID mode. The patterns of N-glycopeptide fragment ions from MS/MS spectra were very similar when the results from same type of MS analyzers were compared (the ion trap-Orbitrap CID spectra from LTQ-Orbitrap Elite (laboratory 1 and 2), the LTQ-Orbitrap XL (laboratory 5), and the LTQ-Orbitrap (laboratory 6) or the Q-TOF CID spectra from Triple TOF (laboratory 4) and Q-TOF (laboratory 7)). For the Q-Exactive Plus, the MS/MS patterns were intermediate between those from two types of MS

analyzers. The intensity of oxonium ions was higher than that of the other ions in HCD spectra from the Orbitrap Elite (Figure 3a-1,b-1), Orbitrap XL (Figure 3e-1), and Q-Exactive Plus (Figure 3d) and in CID spectra from the Triple TOF (Figure 3c) and Q-TOF (Figure 3g).

The change of dominant fragment ions from N-glycopeptides was investigated from three representative mass spectrometers (Orbitrap Elite, Q-Exactive Plus, and Triple TOF) according to the type of analyzer (Figure 4). Five N-glycopeptides of different N-glycosites from AGP (LVPVPIT**N**ATLDQITGK, LVPVPIT**N**ATLDR (in Figure 4), E**N**GTISR, E**N**GTVSR, and NEEY**N**K (in Supplementary Figure 1)) were compared. The number of oxonium ions from glycan, B/Y ions from glycan-cleaved glycopeptide, and b/y ions from peptide backbone fragmentation were counted by assigned ions from CID/HCD spectra for Orbitrap Elite, HCD spectra for Q-Exactive Plus, and CID spectra for Triple TOF. The number was averaged from data obtained from test sample 3. According to the increased complexity of the glycan structure attached to the N-glycosite, B/Y ions were more frequently detected than b/y ions in all three instruments because the collision energy was not sufficient to cleave the peptide bond. We also examined trends of peptide dependence in the fragmentation of five different N-glycopeptides. The overall patterns were similar, with the exception of b/y ions, which were slightly peptide-dependent.

Fundamentally, instruments can be characterized by the following fragmentation method, IT-CID (ion trap–CID,

(a-1) HCD spectrum from laboratory 1 (Orbitrap Elite).



(a-2) CID spectrum from laboratory 1 (Orbitrap Elite).

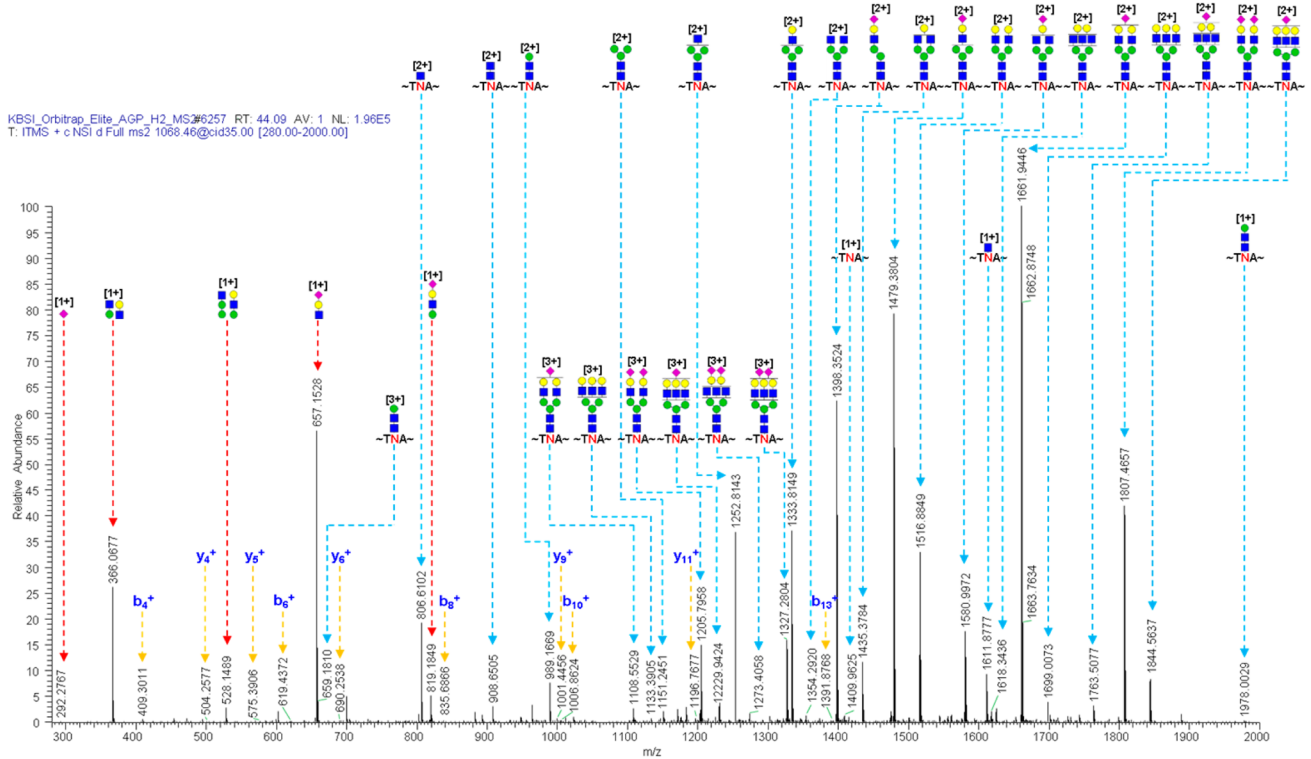
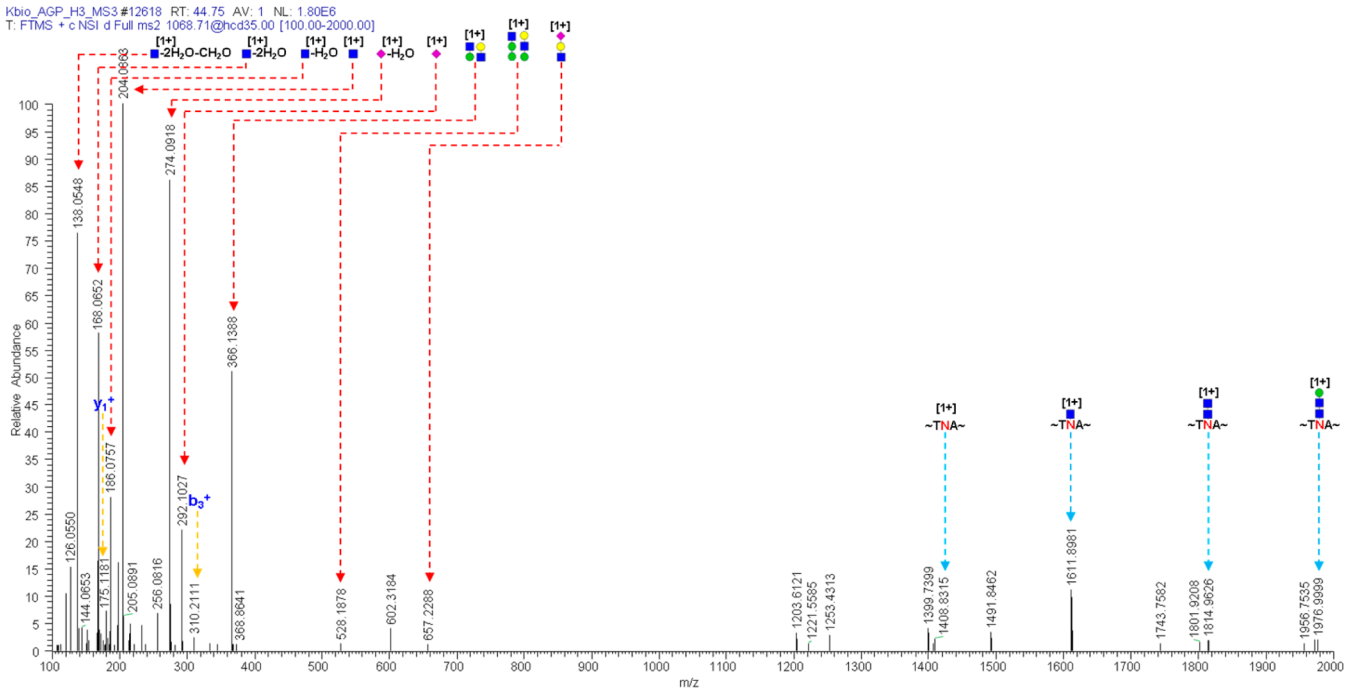


Figure 3. continued

(b-1) HCD spectrum from laboratory 2 (Orbitrap Elite).



(b-2) CID spectrum from laboratory 2 (Orbitrap Elite).

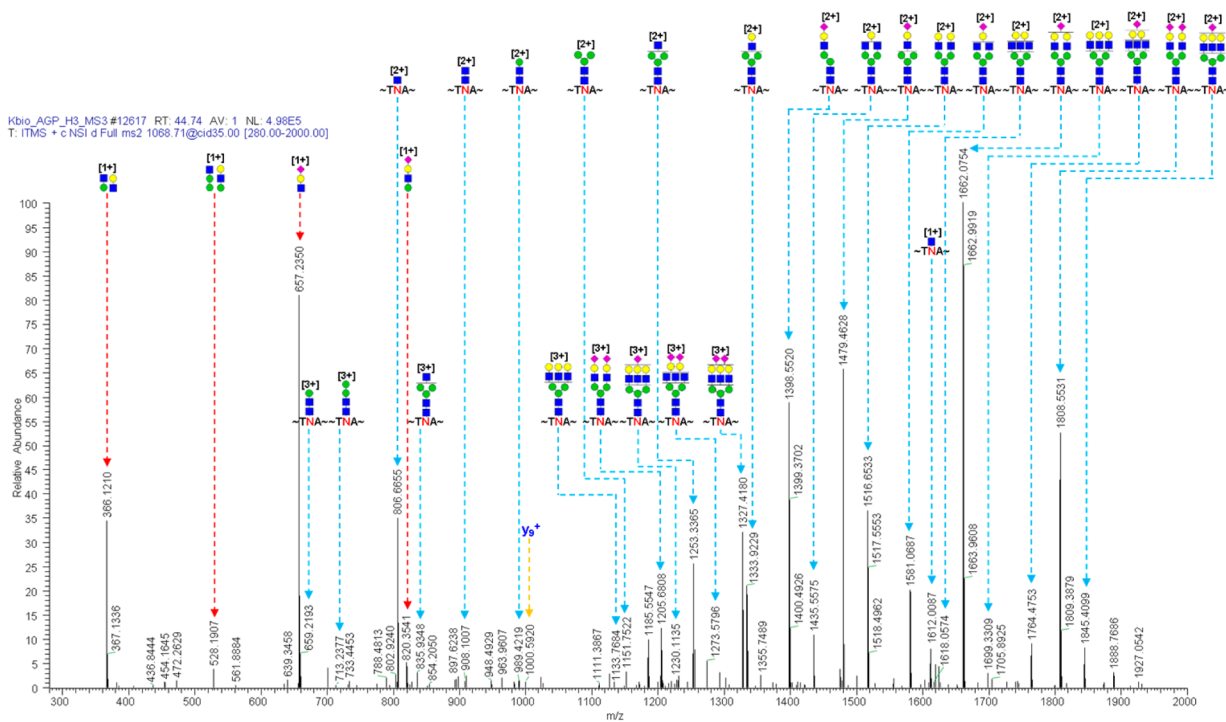
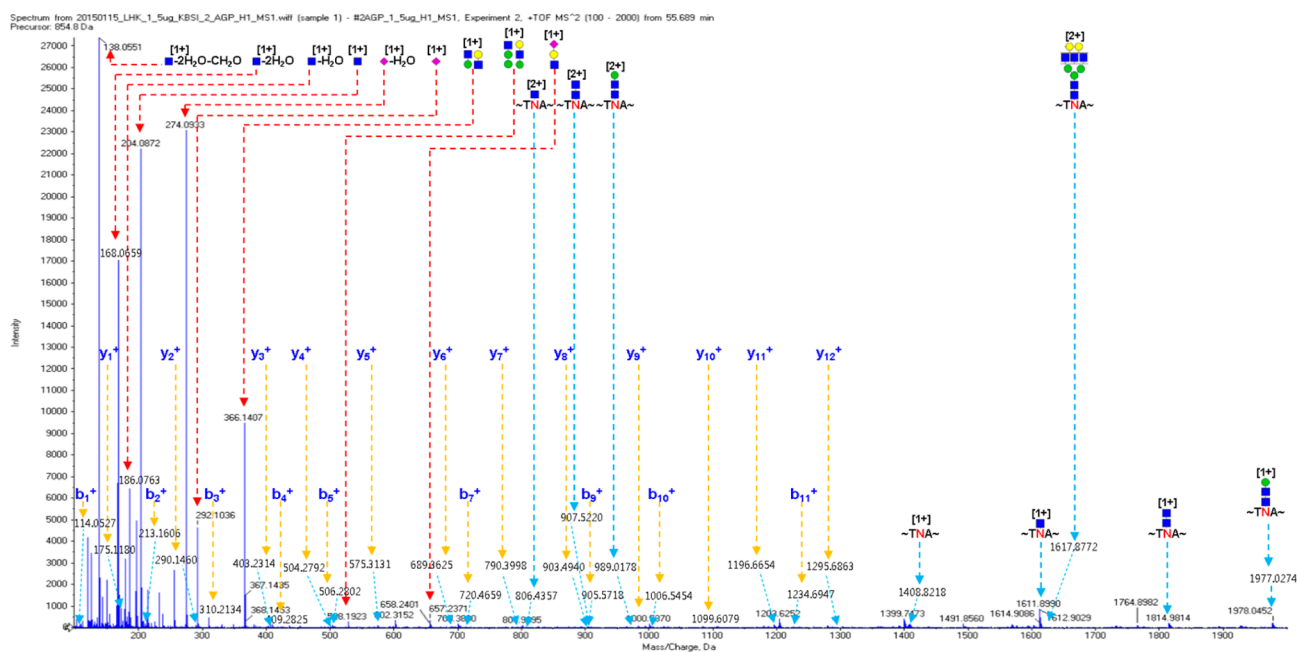


Figure 3. continued

(c) CID spectrum from laboratory 3 (Triple TOF).



(d) HCD spectrum from laboratory 4 (Q-Exactive Plus).

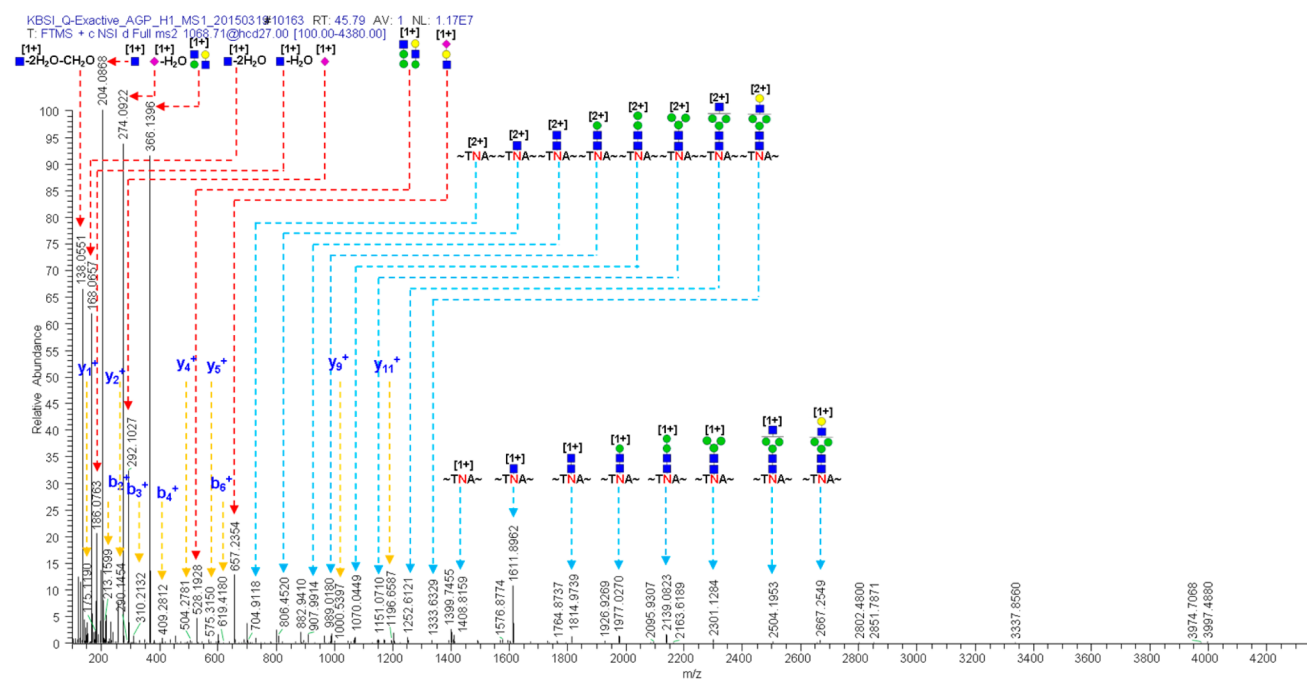
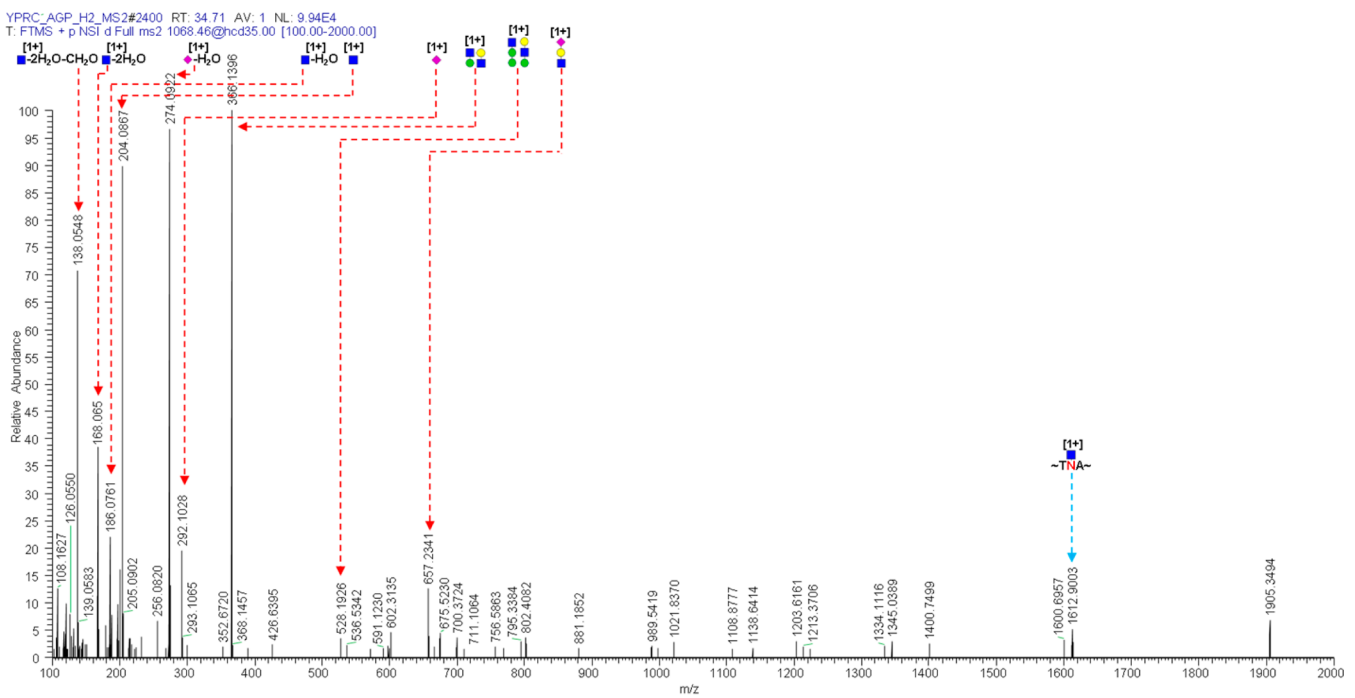


Figure 3. continued

(e-1) HCD spectrum from laboratory 5 (Orbitrap XL).



(e-2) CID spectrum from laboratory 5 (Orbitrap XL).

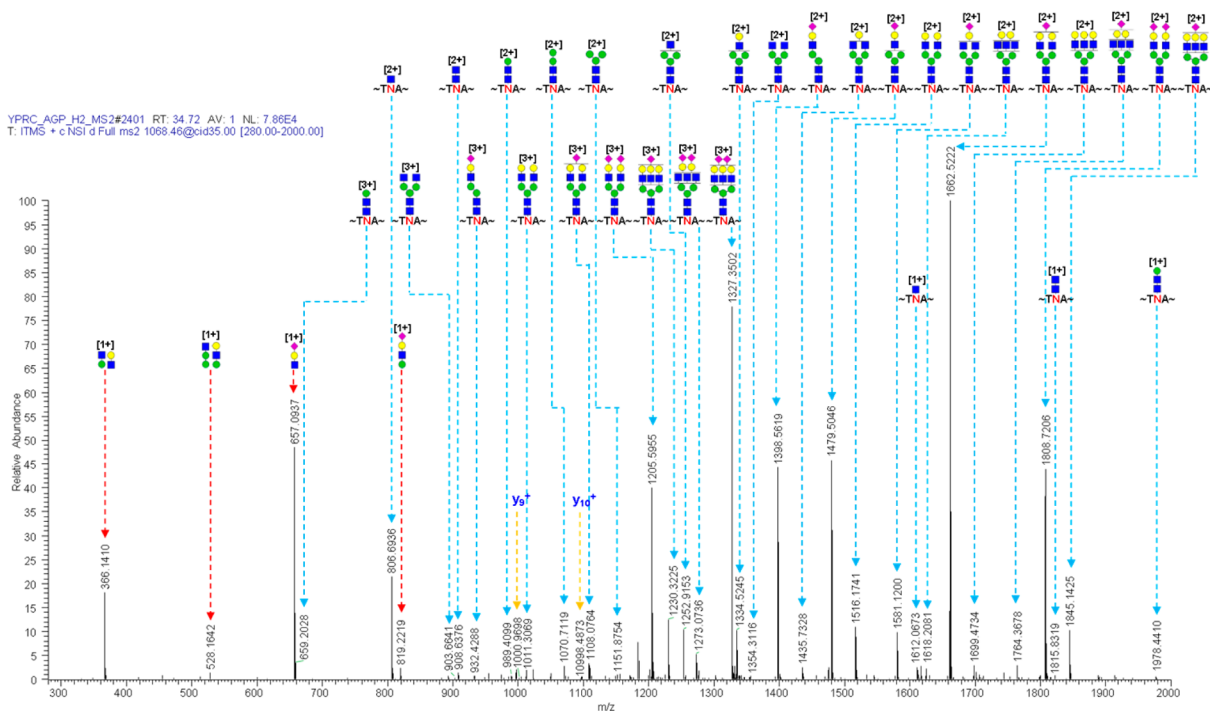
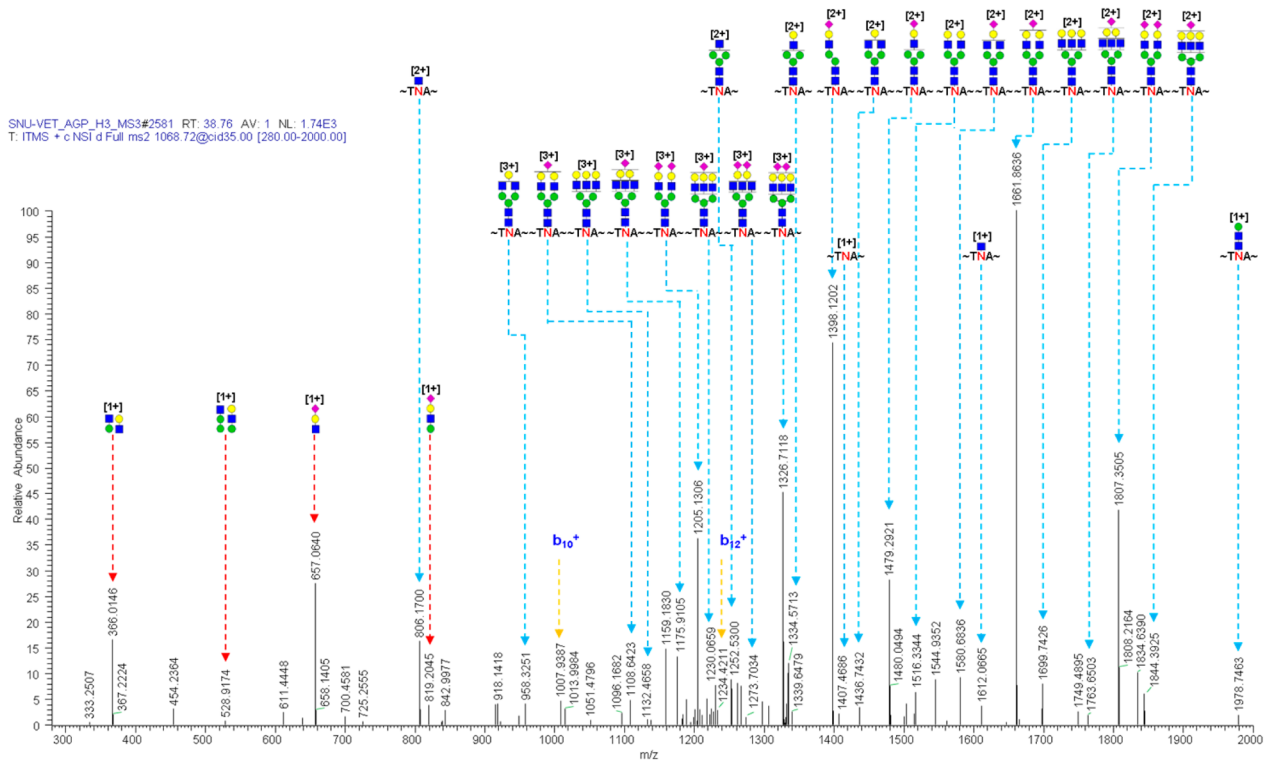


Figure 3. continued

(f) CID spectrum from laboratory 6 (Orbitrap XL).



(g) CID spectrum from laboratory 7 (Q-TOF).

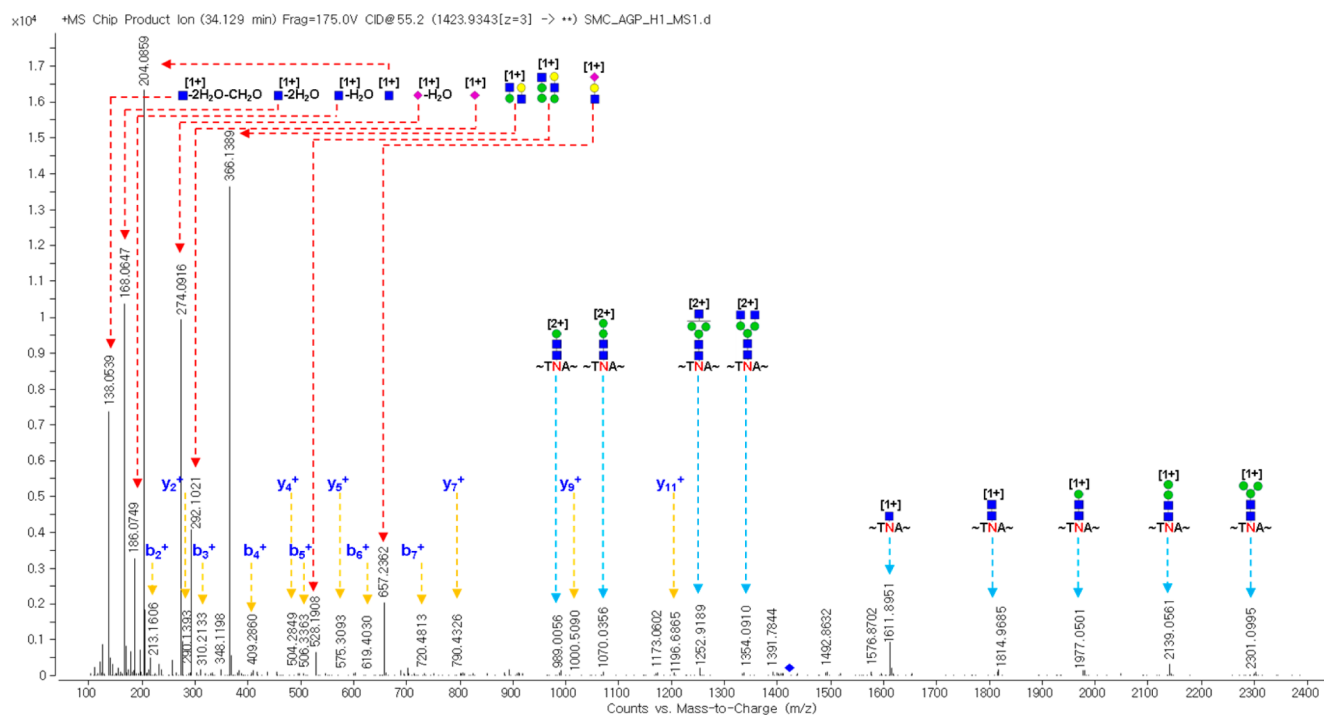


Figure 3. MS/MS spectra of a representative *N*-glycopeptide (LVPVPITNATLDR_6_5_0_3) of AGP produced by each laboratory. The nomenclature of the *N*-glycopeptide is LVPVPITNATLDR_6_5_0_3 (the digits following the peptide sequence represent the number of hexose, *N*-acetylglucosamine (GlcNAc), fucose (Fuc), and *N*-acetyl neuraminic acid (NeuAc) residues, respectively, N: *N*-glycosylation site). (a-1) HCD and (a-2) CID spectrum from laboratory 1 (Orbitrap Elite), (b-1) HCD and (b-2) CID spectrum from laboratory 2 (Orbitrap Elite), (c) CID spectrum from laboratory 3 (Triple TOF), (d) HCD spectrum from laboratory 4 (Q-Exactive Plus), (e-1) HCD and (e-2) CID spectrum from laboratory 5 (Orbitrap XL), (f) CID spectrum from laboratory 6 (Orbitrap), and (g) CID spectrum from laboratory 7 (Q-TOF). The *m/z* values of CID spectra derived from low resolution using ion trap (panels a-2, b-2, and e-2) are effective to one decimal place.

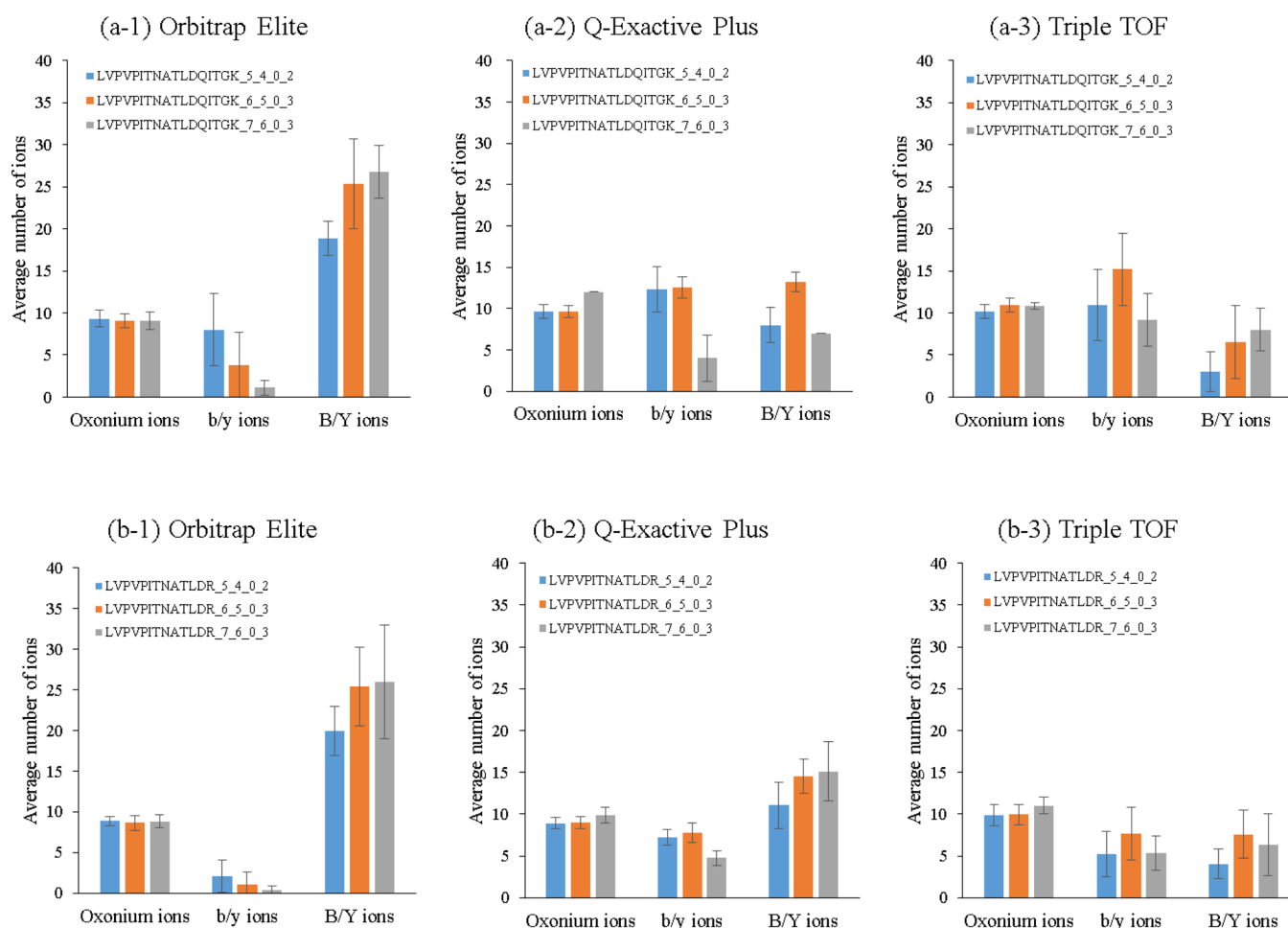


Figure 4. Average numbers of oxonium ions, glycan-cleaved glycopeptide fragment ions (B/Y ions), and peptide backbone fragment ions (b/y ions) matched from the spectra of *N*-glycopeptides. The results from test sample 3 were used. (a) LVPVPITNATLDQITGK from AGP 1. (b) LVPVPITNATLDR from AGP 2. Results from Orbitrap Elite, Q-Exactive Plus, and Triple TOF were from laboratories 1, 4, and 3, respectively. * Average from two MS/MS spectra.

Table 2. Numbers of *N*-Glycopeptides Identified from Three Test Samples and One Blind Sample in Participating Laboratories 1–4^a

sample	number of identified <i>N</i> -glycopeptides			
	laboratory 1 (Orbitrap Elite)	laboratory 2 (Orbitrap Elite)	laboratory 3 (Triple TOF)	laboratory 4 (Q-Exactive Plus)
test sample 1	116 ± 5	105 ± 4	108 ± 9	73 ± 2
test sample 2	133 ± 3	120 ± 9	85 ± 12	41 ± 3
test sample 3	124 ± 9	147 ± 5	76 ± 11	52 ± 1
blind sample	91 ± 1	72 ± 3	53 ± 3	28 ± 3

^aOnly *N*-glycopeptides from AGP were considered.

10^{-1} to 1 eV with 10–100 ms activation time) with low collision energy and HCD (2–10 eV with 0.1 to 0.01 ms activation time) or Q-CID (quadrupole-CID, 10 to 10^2 eV with 0.1 to 1 ms activation time) with high collision energy.⁶⁹ Low-energy fragmentation provided more B/Y ions, which are glycan-cleaved glycopeptide fragment ions resulting from cleavage of glycosidic bond remaining in the peptide bond due to the lability of the glycan group. In contrast, b/y ions that are formed through cleavage of peptide bond and provide peptide sequence information more frequently occurred using high-energy fragmentation methods such as HCD and Q-CID. IT-CID in LTQ-Orbitrap Elite was advantageous for the analysis of glycan composition because it produced more Y ions. In the HCD spectra from LTQ-Orbitrap Elite, oxonium ions were major, and the number of B/Y ions was smaller than those in IT-CID. The distributions of oxonium,

b/y, and B/Y ions were similar between HCD spectra from Q-Exactive Plus and CID spectra from the Triple TOF because they use fragmentation methods with high collision energy. The total numbers of assigned ions were smaller than those from LTQ-Orbitrap Elite due to the small number of Y ions. The average numbers of oxonium ions were almost the same in three mass spectrometers. They were not affected by the complexity of the glycan structures attached to peptides.

In our study, the even generation of oxonium, b/y, and B/Y ions were essential for successful identification. Given this, a combination of IT-CID and HCD fragmentation in Orbitrap-Elite gave the best results in the identification of *N*-glycopeptides.

Table 2 shows the numbers of identified *N*-glycopeptides in each sample. These numbers resulted from triplicate LC–MS/MS runs with or without HILIC enrichment process.

Test sample 1 was provided to check the performance of the mass spectrometer used in each laboratory. More *N*-glycopeptides (73–116) were reported by laboratories 1–4 (Table 2) than laboratories 5–7 (3–25). (The results from laboratories 5–7 were shown in Supplementary Table 3 because the total number of identified *N*-glycopeptides was <30.) The mass spectrometers used in laboratories 1–4 were more sensitive than those used in laboratories 5–7. To validate the sensitivities of the MS instruments, the data for test sample 1 from all the laboratories were compared using peptide search by MASCOT (version 2.4.0) (Supplementary Figure 2). The numbers of identified peptide spectra in laboratory 5–7 were much lower (2–8-fold) than those in laboratories 1–4. On the basis of these results, we decided to discuss the results from laboratories 1–4.

Test sample 2 was digested in the host laboratory and then distributed to each laboratory. To evaluate the performance of *N*-glycopeptides enrichment, the *N*-glycopeptides were individually enriched in each laboratory and analyzed under the same LC–MS/MS conditions as those used for test sample 1. Slightly more *N*-glycopeptides (41–133) were identified from the AGP of test sample 2 than from the AGP of test sample 1 without enrichment. In laboratories 3 and 4, the number of identified *N*-glycopeptides from AGP was somewhat reduced due to a lack of skill in glycopeptide enrichment. The coefficient of variation (CV) percentages from three enrichment experiments (Supplementary Table 4) were 2, 7, 14, and 6% for laboratories 1, 2, 3, and 4, respectively, in terms of the number of identified *N*-glycopeptides. We found that the enrichment process was generally reproducible in each laboratory but required the proficiency.

Test sample 3 (AGP standard sample) was digested, enriched by HILIC, and analyzed by LC–MS/MS in each laboratory. The number of identified *N*-glycopeptides (52–147 for

laboratories 1–4) in test sample 3 was similar to that (41–133 for laboratories 1–4) in test sample 2. Protein digestion with test sample 3 was repeated three times. *N*-Glycopeptide enrichment for one digested sample was repeated three times (Supplementary Table 4). The number of *N*-glycopeptides identified from nine LC–MS/MS runs with triplicated digestion and triplicated HILIC enrichment were averaged. In each laboratory, the entire workflow from digestion to LC–MS/MS analysis was reproducible; the CV percentages were 9, 4, 16, and 10% for laboratories 1, 2, 3, and 4, respectively.

The blind sample was considered as a model of complex sample using human serum. It was prepared and analyzed using the same process as test sample 3. Only AGP *N*-glycopeptides in the blind sample were considered. The numbers of AGP *N*-glycopeptides identified in human serum were less than those in three test samples in all laboratories (Table 2) because of the complexity of the sample. The CV percentages were within ranges 5, 10, 11, and 13% in laboratories 1, 2, 3, and 4, respectively.

Ideally, in an interlaboratory study, the same sample and LC–MS/MS systems should be implemented. Laboratories 1 and 2 used the same Easy-nLC 1000 device (Thermo Scientific, USA) and an LTQ Orbitrap Elite mass spectrometer (Thermo Scientific) equipped with a nanoelectrospray source; they differed only in terms of the dimensions of the column used for separation (Table 1). The data from these laboratories were the most similar in all analyses of test samples 1, 2, and 3 and one blind sample. The combination of multiple fragmentation method such as CID/HCD²¹ or CID/ETD/HCD⁵³ or HCD/HCD with different NCE percentages⁷⁰ is very useful for the comprehensive identification of *N*-glycopeptides, providing information on both glycan structures and peptide sequence.⁷¹ In our study, laboratories 1 and 2 identified the greatest number of *N*-glycopeptides by the combination of CID and HCD fragmentation.

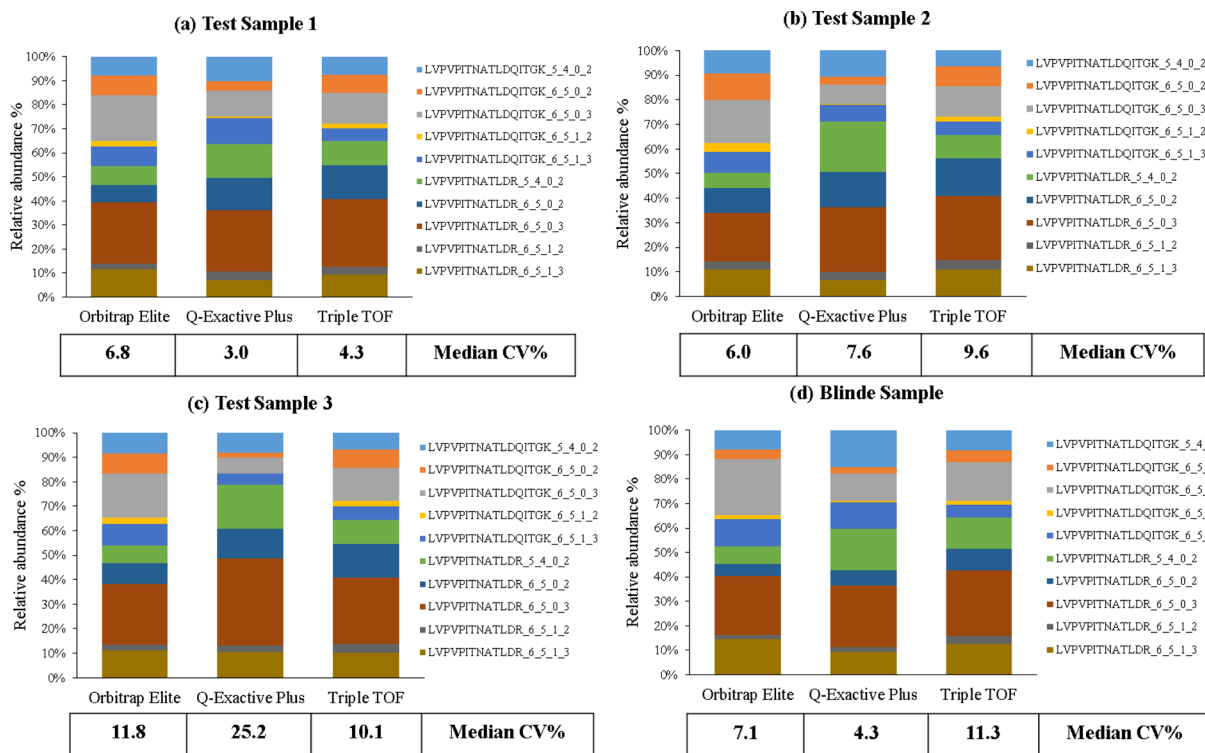


Figure 5. Relative abundance percentages of the 10 major *N*-glycopeptides in three test samples and one blind sample (a–d) from Orbitrap Elite (laboratory 1), Q-Exactive Plus (laboratory 4), and Triple TOF (laboratory 3) analyses.

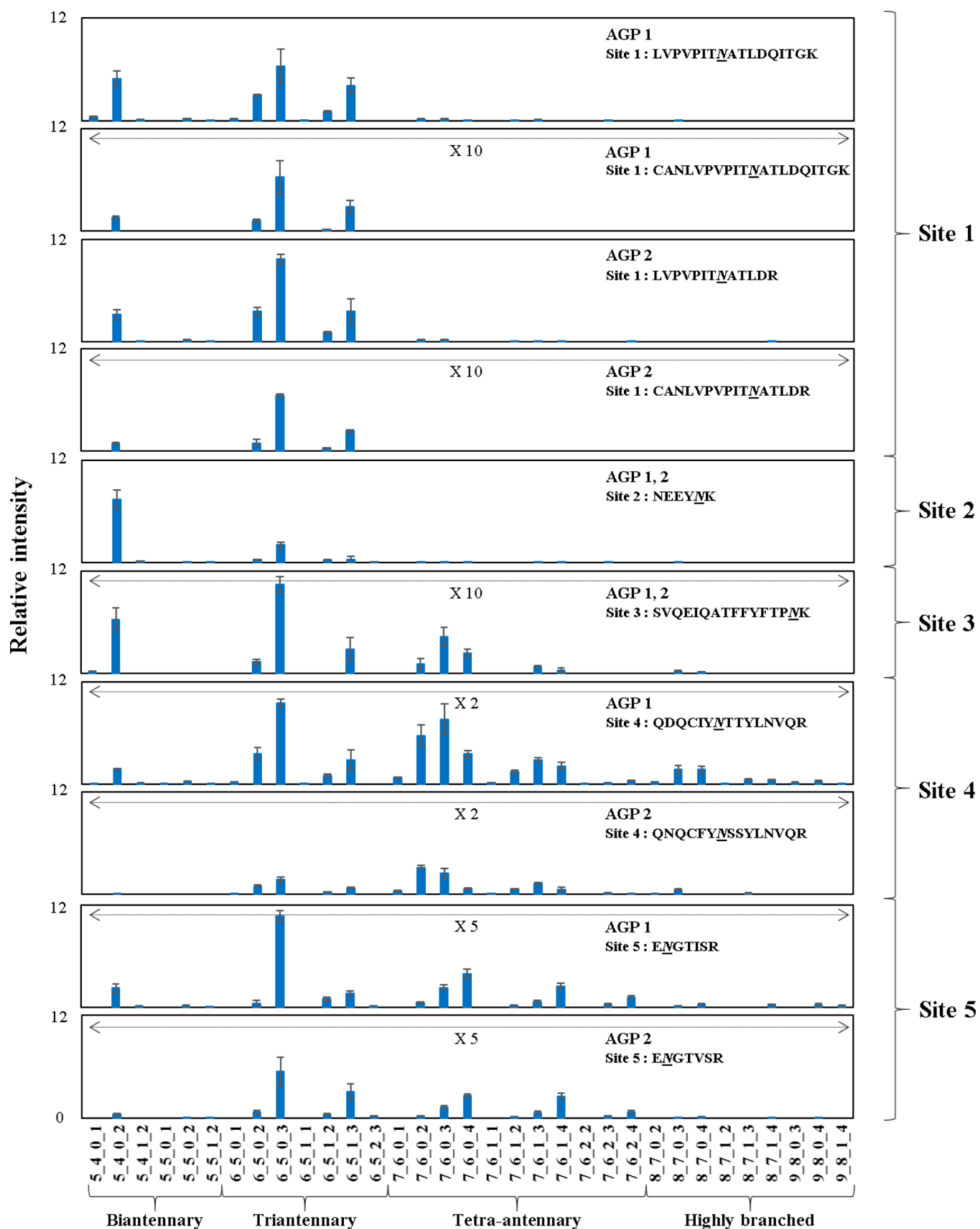


Figure 6. Identification and quantification of 165 site-specific N-glycopeptides of AGP from test sample 3 in laboratory 2. Italicized and underlined N (asparagine) is N-glycosylation site.

In a previous report, Imre et al.⁷² identified 80 AGP N-glycopeptides by the exact mass matching approach and

confirmed 43 site-specific N-glycopeptides using MS/MS spectra and GlycoPattern software⁶⁶ from an AGP standard digest.

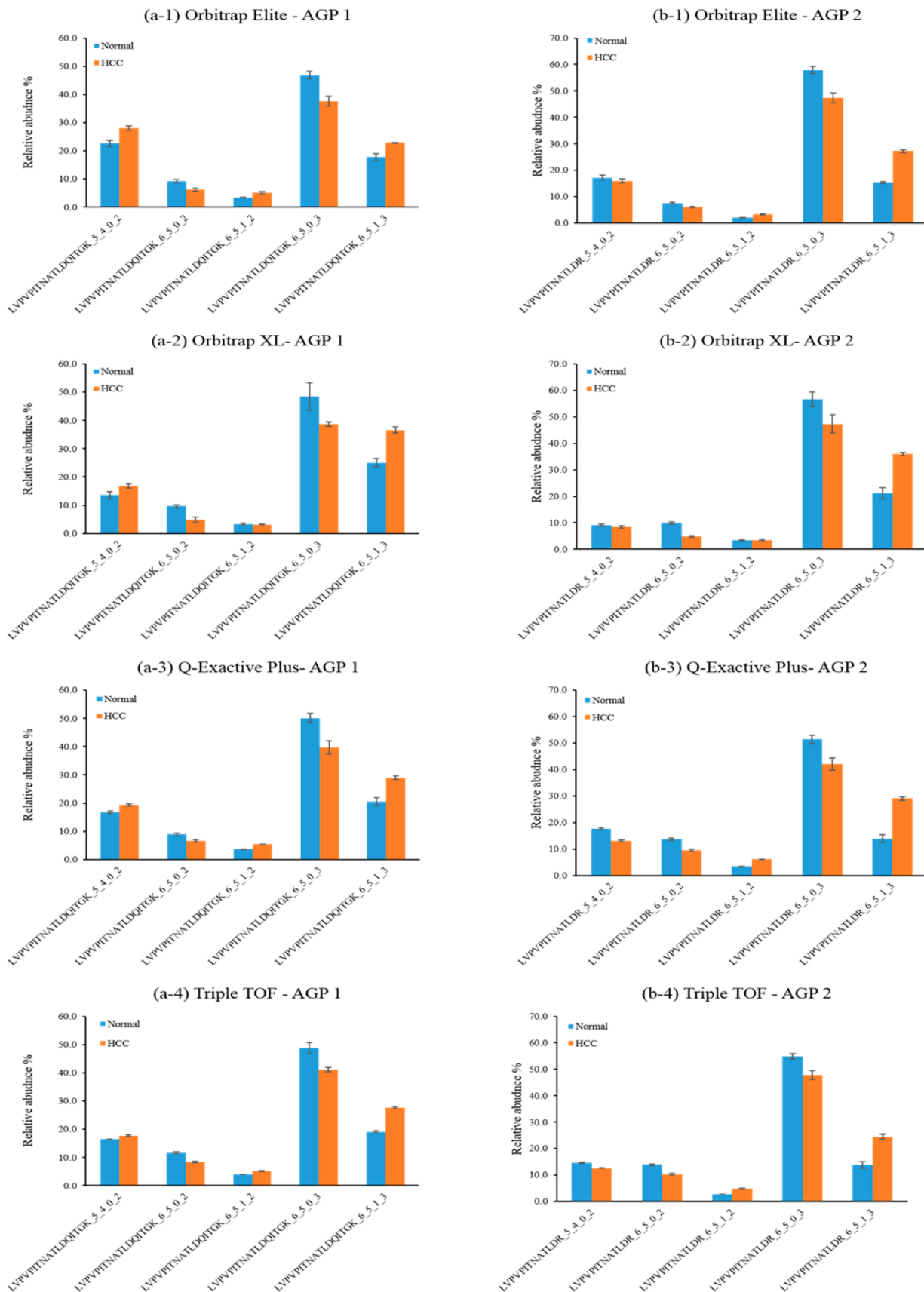


Figure 7. Relative abundance percentages of the 10 major site-specific N-glycopeptides of AGP from pooled normal and hepatocellular carcinoma (HCC) human plasma samples. The results were obtained from (a) LVPVPITNATLDQITGK of AGP 1 and (b) LVPVPITNATLDR of AGP 2. Results from Orbitrap Elite, Triple TOF, Q-Exactive Plus, and Orbitrap XL were from laboratories 1, 3, 4, and 5, respectively. The relative abundances were measured by dividing the sum of representative N-glycopeptides peak areas with the each peak area of site-specific N-glycopeptides.

Quantification of *N*-Glycopeptides

The 10 major *N*-glycopeptides of AGP from three test samples and one blind sample were used for the quantitative analyses because they were commonly identified by three different instruments (Orbitrap Elite-laboratory 1, Q-Exactive Plus-laboratory 4, and Triple TOF-laboratory 3) (Figure 5). On the basis of the small number of identified *N*-glycopeptides, the results from laboratories 5–7 were excluded for the quantitative analyses.

N-Glycopeptides were quantified by summing the peak areas of multiply charged precursor ions because they were generally detected on multiply charged precursor ions of large molecular weight (Supplementary Figure 3). The percentage of each *N*-glycopeptide was calculated from the total peak areas of the 10 major *N*-glycopeptides. The patterns of relative quantification from the 10 major *N*-glycopeptides were similar among three MS analyzers in individual sample and among samples in individual MS analyzer. Pearson's correlation coefficient (*r* values) was calculated to confirm the similarity of their quantitative patterns (Supplementary Table 5a,b). The *r* values were 0.87 to 1.00 when samples were compared using individual MS analyzer (Supplementary Table 5a) and 0.73 to 0.92 (except for one: *r* = 0.53) when samples were compared using different MS analyzers (Supplementary Table 5b). The quantitative similarities among samples (Supplementary Table 5a) in each MS analyzer were slightly better than those among MS analyzers in each sample (Supplementary Table 5b). Therefore, the quantitative data on major *N*-glycopeptides were more affected by instrument than sample specificity.

The CV percentages of the relative abundance of the 10 major *N*-glycopeptides were calculated for three test samples and one blind sample (Figure 6). The quantitative data were reproducible with median CVs of 3–7% for test sample 1, 6–10% for test sample 2, 10–25% for test sample 3, and 4–11% for blind sample in all three instruments. Each value is the average of three (for test samples 1 and 2) or nine (for test sample 3 and blind sample) LC–MS/MS analyses.

Characterization of Site-Specific *N*-Glycopeptide Isoforms of AGP. In total, 165 site-specific *N*-glycopeptide isoforms were identified from all glycosylation sites including *N*-terminal truncated isoforms⁷² of AGP. These types of truncated *N*-glycopeptides might be significant clues to elucidate the biological functions of glycoproteins. The relative abundances of 165 *N*-glycopeptides identified in test sample 3 from laboratory 2 are presented in Figure 6. The order of abundance of site-specific *N*-glycopeptides from AGP 1 and AGP 2 was as follows: AGP 1 - site 4 (Asn,⁷⁵ 25%) > AGP 2 - site 1 (Asn,¹⁵ 23%) > AGP 1 - site 1 (Asn,¹⁵ 22%) > AGP 1 and 2 - site 2 (Asn,³⁸ 11%) > AGP 2-site 4 (Asn,⁷⁵ 7%) > AGP 1 - site 5 (Asn⁸⁵, 6%) > AGP 2 - site 5 (Asn⁸⁵, 4%) > AGP 1 and 2 - site 3 (Asn,⁵⁴ 3%). In general, the pattern of relative abundance between AGP 1 and AGP 2 isoforms is very similar. *N*-Glycopeptides containing a triantennary glycan were common at site 1. Biantennary *N*-glycopeptides were more abundant than triantennary *N*-glycopeptides at site 2. Tetra-antennary *N*-glycopeptides were increased at sites 3, 4, and 5. More highly branched *N*-glycopeptides were detected at sites 4 and 5. Imre et al.⁷² reported the relative abundance of glycan at five different glycosylation sites instead of site-specific *N*-glycopeptides. The main antennary structure at each glycosylation site is consistent with our results. The quantification of site-specific *N*-glycopeptides could be used to investigate the differences between normal and disease samples.

Plasma Sample Analysis

AGP is a well-characterized glycoprotein related to disease.^{73–76} In some cases, a change in glycan composition between control and disease samples was reported.^{74,75} We quantified the *N*-glycopeptide of AGP in normal and HCC human plasma samples using different MS analyzers (Figure 7). We intended to determine whether the relative quantities were similar when various MS analyzers were used. Orbitrap Elite (laboratory 1), Triple TOF (laboratory 3), Q-Exactive Plus (laboratory 4), and Orbitrap XL (laboratory 5) were considered as same or different MS analyzer.

N-Glycopeptides from plasma samples were prepared using a process employed in a previous study on blind sample using human serum, and then distributed equally to each laboratory. The LC–MS/MS analyses were repeated five times under the same conditions. Site-specific *N*-glycopeptides of AGP in pooled normal and HCC plasma sample were identified and quantified from four laboratories. The number of identified *N*-glycopeptides was $66 \pm 5/74 \pm 4$ (normal/HCC) in the Orbitrap Elite, $69 \pm 4/73 \pm 8$ in Q-Exactive Plus, $37 \pm 3/39 \pm 7$ in Triple TOF, and $22 \pm 3/17 \pm 1$ in Orbitrap XL. The total number of quantified *N*-glycopeptides was 57/60 (normal/HCC) in the Orbitrap Elite, 48/50 in Q-Exactive Plus, 40/34 in Triple TOF, and 17/16 in Orbitrap XL, as shown in Supplementary Figure 4. Among them, 10 major *N*-glycopeptides (red boxed) were commonly identified and quantified by four LC–MS/MS systems. Their relative abundances were similar between same type of MS analyzers (Orbitrap Elite (Figure 7a-1,b-1) versus Orbitrap XL (Figure 7a-2,b-2)) and among different type of MS analyzers (Orbitrap Elite and Orbitrap XL vs Q-Exactive Plus (Figure 7a-3,b-3) versus Triple TOF (Figure 7a-4,b-4)). However, the *N*-glycopeptides showed large variations in the number of identification (from 17 to 74) and quantification (from 16 to 60) in four laboratories due to different sensitivities from different LC–MS systems. Therefore, in our analysis, minor *N*-glycopeptides except 10 major ones render quantitative analysis difficult.

The Pearson's correlation coefficients were calculated to evaluate the quantitative similarities among diverse MS analyzers from normal and HCC plasma samples. The *r* values (*r* > 0.89) among the various MS analyzers (Supplementary Table 6) showed that the 10 major AGP *N*-glycopeptides were quantitatively similar in both normal and HCC plasma samples. We found that the quantitative patterns yielded by the Triple TOF and Q-Exactive Plus were the most similar (the highest *r* value in Supplementary Table 6). Despite being the same type of MS analyzer, the data from Orbitrap Elite and Orbitrap XL were less correlated due to a difference in the performance of the LC–MS systems.

In particular, fucosylated peptides with the glycan formula 6_5_1_3 were consistently increased in HCC plasma sample, as revealed by all types of MS analyzers. This shows that differences between normal and HCC samples are detectable using any of the MS analyzers. The quantitative changes in fucosylated peptide levels are consistent with a previous report on fucosylated AGP levels in inflammatory⁷⁴ and hepatic diseases.^{20–22}

CONCLUSIONS

We studied the identification and quantification of the site-specific *N*-glycopeptides from AGP in standard and human serum samples tested in seven laboratories and its application to normal and HCC plasma samples. To identify *N*-glycopeptides, we found that the following criteria were critical: MS/MS spectra of good quality representing oxonium ions by glycans, glycan-

cleaved glycopeptide fragment ions (B/Y ions), and peptide backbone fragment ions (b/y ions) and use of a mass spectrometer of high sensitivity to detect microheterogeneous N-glycopeptides at low abundance. In individual laboratory, the AGP analysis gave reproducible identification and quantification results in all samples. However, the number of identified N-glycopeptides was different among laboratories due to differences of LC–MS performance and the laboratory proficiency. The greatest numbers of AGP N-glycopeptides were identified by the Orbitrap Elite; this is attributable to the sensitivity of the MS analyzer and the use of combinations of multiple tandem techniques (CID and HCD). The quantitative patterns of the 10 most abundant N-glycopeptide isoforms of AGP were similar among the laboratories. Therefore, this analytical method could be used as the standard protocol to characterize site-specific N-glycopeptide isoforms of glycoproteins. The protocol was successfully applied for the relative quantification of the 10 major N-glycopeptides from AGP in normal and HCC human plasma samples using various types of MS analyzers.

■ ASSOCIATED CONTENT

📄 Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jproteome.5b01159.

Supplementary Table 1. The information of normal and HCC plasma samples. Supplementary Table 2. Mass tolerances of precursor and fragment ions for the identification of peptides using MASCOT search for tryptic peptides and I-GPA search for N-glycopeptides from AGP according to mass spectrometers. Supplementary Table 3. The number of N-glycopeptides identified more than two times from triplicate analyses of three test samples and one blind sample in three participating laboratories. Supplementary Table 4. The experimental processing steps and replicates. Supplementary Table 5. Pearson's correlation coefficient value (r) for three test samples and one blind sample. Supplementary Table 6. Pearson's correlation coefficient value (r) for normal and HCC plasma samples. Supplementary Figure 1. Average number of oxonium ions, glycan-cleaved glycopeptide fragment ions (B/Y ions), and peptide backbone fragment ions (b/y ions) from total matched fragment ions of each peptide. Supplementary Figure 2. The number of spectra identified as general tryptic peptides of AGP from a digested AGP (test sample 1) in participated laboratories. Supplementary Figure 3. Charge-state distribution of N-glycopeptides from AGP (test sample 1) in Orbitrap Elite (laboratory 1), in Q-Exactive Plus (laboratory 4), and in Triple TOF (laboratory 3). Supplementary Figure 4. The relative abundance of identified site-specific N-glycopeptides from AGP in pooled normal and HCC (hepatocellular carcinoma) human plasma samples. (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*J.S.Y.: Tel: +82-43-240-5150. Fax: +82-43-240-5159. E-mail: jongshin@kbsi.re.kr.

*J.Y.K.: Tel: +82-43-240-5142. Fax: +82-43-240-5159. E-mail: jinyoung@kbsi.re.kr.

Author Contributions

•J.Y.L. and H.K.L. contributed equally.

Notes

The authors declare no competing financial interest.

All data sets from seven laboratories have been uploaded to the ProteomeXchange Consortium via MassIVE repository (the data set identifier PXD004677 and MSV000080010, respectively). The raw files are available for download at <ftp://massive.ucsd.edu/MSV000080010>.

■ ACKNOWLEDGMENTS

This work was supported by the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HI13C2098), by the Creative Allied Project (CAP) through the National Research Council of Science and Technology (grant number: NTM2371511), by the research program through the Korea Basic Science Institute (grant numbers: K36110), by the Bio-Synergy Research Project (grant number: NRF-2014M3A9C4066461) of the Ministry of Science, ICT and Future Planning through the National Research Foundation, and by the National Research Foundation of Korea Grant funded by the Korea Government (MSIP) (2016, R&D Equipment Engineer Education Program, 2014R1A6A9064166).

■ ABBREVIATIONS

AGP, alpha-1-acid glycoprotein; PNGase F, peptide-N-glycosidase F; DTT, 1,4-dithiothreitol; IAA, iodoacetamide; FA, formic acid; ABC, ammonium bicarbonate; HCC, hepatocellular carcinoma; PTM, post-translational modification; HILIC, hydrophilic interaction liquid chromatography; HPLC, high-performance liquid chromatography; MS, mass spectrometry; ESI, electrospray ionization; CID, collision-induced dissociation; HCD, higher-energy collisional dissociation; LTQ, linear trap quadrupole; Q-TOF, quadrupole-time-of-flight; NCE, normalized collision energy; I-GPA, Integrated GlycoProteome Analyzer; GlcNAc, N-acetylglucosamine; NeuAc, N-acetylneuraminic acid; Fuc, fucose

■ REFERENCES

- (1) Ohtsubo, K.; Marth, J. D. Glycosylation in cellular mechanisms of health and disease. *Cell* **2006**, *126*, 855–867.
- (2) Kim, Y. S.; et al. Functional proteomics study reveals that N-Acetylglucosaminyltransferase V reinforces the invasive/metastatic potential of colon cancer through aberrant glycosylation on tissue inhibitor of metalloproteinase-1. *Mol. Cell. Proteomics* **2008**, *7*, 1–14.
- (3) Andre, G.; Buleon, A.; Haser, R.; Tran, V. Amylose chain behavior in an interacting context. III. Complete occupancy of the AMY2 barley alpha-amylase cleft and comparison with biochemical data. *Biopolymers* **1999**, *50*, 751–762.
- (4) Drake, P. M.; et al. Sweetening the pot: adding glycosylation to the biomarker discovery equation. *Clin. Chem.* **2010**, *56*, 223–236.
- (5) Hwang, H.; et al. Glycoproteomics in neurodegenerative diseases. *Mass Spectrom. Rev.* **2010**, *29*, 79–125.
- (6) Hua, S.; et al. Comprehensive native glycan profiling with isomer separation and quantitation for the discovery of cancer biomarkers. *Analyst* **2011**, *136*, 3663–3671.
- (7) Froehlich, J. W.; et al. Nano-LC-MS/MS of glycopeptides produced by nonspecific proteolysis enables rapid and extensive site-specific glycosylation determination. *Anal. Chem.* **2011**, *83*, 5541–5547.
- (8) Ni, W.; Bones, J.; Karger, B. L. In-depth characterization of N-linked oligosaccharides using fluoride-mediated negative ion microfluidic chip LC-MS. *Anal. Chem.* **2013**, *85*, 3127–3135.

- (9) Ruhaak, L. R.; Miyamoto, S.; Kelly, K.; Lebrilla, C. B. N-Glycan profiling of dried blood spots. *Anal. Chem.* **2012**, *84*, 396–402.
- (10) Paik, Y. K.; et al. The Chromosome-Centric Human Proteome Project for cataloging proteins encoded in the genome. *Nat. Biotechnol.* **2012**, *30*, 221–223.
- (11) An, H. J.; Peavy, T. R.; Hedrick, J. L.; Lebrilla, C. B. Determination of N-glycosylation sites and site heterogeneity in glycoproteins. *Anal. Chem.* **2003**, *75*, 5628–5637.
- (12) Hwang, H.; et al. In-depth analysis of site-specific N-glycosylation in vitronectin from human plasma by tandem mass spectrometry with immunoprecipitation. *Anal. Bioanal. Chem.* **2014**, *406*, 7999–8011.
- (13) Furmanek, A.; Hofsteenge, J. Protein C-mannosylation: facts and questions. *Acta Biochim. Pol.* **2000**, *47*, 781–789.
- (14) Zhang, Y.; Yin, H.; Lu, H. Recent progress in quantitative glycoproteomics. *Glycoconjugate J.* **2012**, *29*, 249–258.
- (15) Pompach, P.; et al. Site-specific glycoforms of haptoglobin in liver cirrhosis and hepatocellular carcinoma. *Mol. Cell. Proteomics* **2013**, *12*, 1281–1293.
- (16) Liu, Y.; et al. Glycoproteomic analysis of prostate cancer tissues by SWATH mass spectrometry discovers N-acyl ethanolamine acid amidase and protein tyrosine kinase 7 as signatures for tumor aggressiveness. *Mol. Cell. Proteomics* **2014**, *13*, 1753–1768.
- (17) Sanda, M.; et al. Quantitative liquid chromatography-mass spectrometry-multiple reaction monitoring (LC-MS-MRM) analysis of site-specific glycoforms of haptoglobin in liver disease. *Mol. Cell. Proteomics* **2013**, *12*, 1294–1305.
- (18) Wei, X.; Li, L. Comparative glycoproteomics: approaches and applications. *Briefings Funct. Genomics Proteomics* **2009**, *8*, 104–113.
- (19) Cerciello, F.; et al. Identification of a seven glycopeptide signature for malignant pleural mesothelioma in human serum by selected reaction monitoring. *Clinical proteomics* **2013**, *10*, 16.
- (20) Kawahara, R.; Saad, J.; Angeli, C. B.; Palmisano, G. Site-specific characterization of N-linked glycosylation in human urinary glycoproteins and endogenous glycopeptides. *Glycoconjugate J.* **2016**, DOI: 10.1007/s10719-016-9677-z.
- (21) Park, G. W.; et al. Integrated GlycoProteome Analyzer (I-GPA) for Automated Identification and Quantitation of Site-Specific N-Glycosylation. *Sci. Rep.* **2016**, *6*, 21175.
- (22) Lih, T. M.; et al. MAGIC-web: a platform for untargeted and targeted N-linked glycoprotein identification. *Nucleic Acids Res.* **2016**, *44*, W575.
- (23) Moh, E. S.; Lin, C. H.; Thaysen-Andersen, M.; Packer, N. H. Site-Specific N-Glycosylation of Recombinant Pentameric and Hexameric Human IgM. *J. Am. Soc. Mass Spectrom.* **2016**, *27*, 1143–1155.
- (24) Ruhaak, L. R.; et al. Differential N-Glycosylation Patterns in Lung Adenocarcinoma Tissue. *J. Proteome Res.* **2015**, *14*, 4538–4549.
- (25) Reiding, K. R.; Blank, D.; Kuijper, D. M.; Deelder, A. M.; Wuhrer, M. High-throughput profiling of protein N-glycosylation by MALDI-TOF-MS employing linkage-specific sialic acid esterification. *Anal. Chem.* **2014**, *86*, 5784–5793.
- (26) Wohlgenuth, J.; Karas, M.; Eichhorn, T.; Hendriks, R.; Andrecht, S. Quantitative site-specific analysis of protein glycosylation by LC-MS using different glycopeptide-enrichment strategies. *Anal. Biochem.* **2009**, *395*, 178–188.
- (27) Geyer, H.; Geyer, R. Strategies for analysis of glycoprotein glycosylation. *Biochim. Biophys. Acta, Proteins Proteomics* **2006**, *1764*, 1853–1869.
- (28) Zielinska, D. F.; Gnad, F.; Wisniewski, J. R.; Mann, M. Precision mapping of an in vivo N-glycoproteome reveals rigid topological and sequence constraints. *Cell* **2010**, *141*, 897–907.
- (29) Zielinska, D. F.; Gnad, F.; Schropp, K.; Wisniewski, J. R.; Mann, M. Mapping N-glycosylation sites across seven evolutionarily distant species reveals a divergent substrate proteome despite a common core machinery. *Mol. Cell* **2012**, *46*, 542–548.
- (30) Drake, P. M.; et al. A lectin affinity workflow targeting glycosite-specific, cancer-related carbohydrate structures in trypsin-digested human plasma. *Anal. Biochem.* **2011**, *408*, 71–85.
- (31) Zhao, J.; Qiu, W.; Simeone, D. M.; Lubman, D. M. 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.
- (32) Wang, L.; et al. Mapping N-linked glycosylation sites in the secretome and whole cells of *Aspergillus niger* using hydrazine chemistry and mass spectrometry. *J. Proteome Res.* **2012**, *11*, 143–156.
- (33) Sun, S.; et al. Isolation of N-linked glycopeptides by hydrazine-functionalized magnetic particles. *Anal. Bioanal. Chem.* **2010**, *396*, 3071–3078.
- (34) Takegawa, Y.; et al. Simple separation of isomeric sialylated N-glycopeptides by a zwitterionic type of hydrophilic interaction chromatography. *J. Sep. Sci.* **2006**, *29*, 2533–2540.
- (35) Hagglund, P.; Bunkenborg, J.; Elortza, F.; Jensen, O. N.; 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.
- (36) Myslins, 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.
- (37) Wada, Y.; Tajiri, M.; Yoshida, S. Hydrophilic affinity isolation and MALDI multiple-stage tandem mass spectrometry of glycopeptides for glycoproteomics. *Anal. Chem.* **2004**, *76*, 6560–6565.
- (38) Reusch, D.; et al. High-throughput work flow for IgG Fc-glycosylation analysis of biotechnological samples. *Anal. Biochem.* **2013**, *432*, 82–89.
- (39) Ding, W.; Hill, J. J.; Kelly, J. Selective enrichment of glycopeptides from glycoprotein digests using ion-pairing normal-phase liquid chromatography. *Anal. Chem.* **2007**, *79*, 8891–8899.
- (40) Yu, L.; Li, X.; Guo, Z.; Zhang, X.; Liang, X. Hydrophilic interaction chromatography based enrichment of glycopeptides by using click maltose: a matrix with high selectivity and glycosylation heterogeneity coverage. *Chem. - Eur. J.* **2009**, *15*, 12618–12626.
- (41) Qu, Y.; et al. Integrated sample pretreatment system for N-linked glycosylation site profiling with combination of hydrophilic interaction chromatography and PNGase F immobilized enzymatic reactor via a strong cation exchange precolumn. *Anal. Chem.* **2011**, *83*, 7457–7463.
- (42) Zhu, J.; et al. Centrifugation assisted microreactor enables facile integration of trypsin digestion, hydrophilic interaction chromatography enrichment, and on-column deglycosylation for rapid and sensitive N-glycoproteome analysis. *Anal. Chem.* **2012**, *84*, 5146–5153.
- (43) Buszewski, B.; Noga, S. Hydrophilic interaction liquid chromatography (HILIC)—a powerful separation technique. *Anal. Bioanal. Chem.* **2012**, *402*, 231–247.
- (44) Liu, Y.; et al. Quantitative measurements of N-linked glycoproteins in human plasma by SWATH-MS. *Proteomics* **2013**, *13*, 1247–1256.
- (45) Bock, T.; et al. Proteomic analysis reveals drug accessible cell surface N-glycoproteins of primary and established glioblastoma cell lines. *J. Proteome Res.* **2012**, *11*, 4885–4893.
- (46) Mayampurath, A. M.; Wu, Y.; Segu, Z. M.; Mechref, Y.; Tang, H. Improving confidence in detection and characterization of protein N-glycosylation sites and microheterogeneity. *Rapid Commun. Mass Spectrom.* **2011**, *25*, 2007–2019.
- (47) Ruhaak, L. R.; Lebrilla, C. B. Applications of Multiple Reaction Monitoring to Clinical Glycomics. *Chromatographia* **2015**, *78*, 335–342.
- (48) Totten, S. M.; et al. Rapid-throughput glycomics applied to human milk oligosaccharide profiling for large human studies. *Anal. Bioanal. Chem.* **2014**, *406*, 7925–7935.
- (49) Thaysen-Andersen, M.; et al. Human neutrophils secrete bioactive paucimannosidic proteins from azurophilic granules into pathogen-infected sputum. *J. Biol. Chem.* **2015**, *290*, 8789–8802.
- (50) Zhu, Z.; Su, X.; Go, E. P.; Desaire, H. New glycoproteomics software, GlycoPep Evaluator, generates decoy glycopeptides de novo and enables accurate false discovery rate analysis for small data sets. *Anal. Chem.* **2014**, *86*, 9212–9219.
- (51) Leymarie, N.; et al. Interlaboratory study on differential analysis of protein glycosylation by mass spectrometry: the ABRF glycoprotein

research multi-institutional study 2012. *Mol. Cell. Proteomics* **2013**, *12*, 2935–2951.

(52) Singh, C.; Zampronio, C. G.; Creese, A. J.; Cooper, H. J. 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.

(53) Ye, H.; Boyne, M. T., 2nd; Buhse, L. F.; 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.

(54) Strum, J. S.; et al. Automated assignments of N- and O-site specific glycosylation with extensive glycan heterogeneity of glycoprotein mixtures. *Anal. Chem.* **2013**, *85*, 5666–5675.

(55) Chandler, K. B.; 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.

(56) Bern, M.; Kil, Y. J.; Becker, C. Byonic: advanced peptide and protein identification software. *Current Protocols in Bioinformatics* **2012**, DOI: 10.1002/0471250953.bi1320s40.

(57) Mayampurath, A.; et al. Computational framework for identification of intact glycopeptides in complex samples. *Anal. Chem.* **2014**, *86*, 453–463.

(58) Wu, S. W.; Liang, S. Y.; Pu, T. H.; Chang, F. Y.; Khoo, K. H. Sweet-Heart - an integrated suite of enabling computational tools for automated MS2/MS3 sequencing and identification of glycopeptides. *J. Proteomics* **2013**, *84*, 1–16.

(59) Lynn, K. S.; et al. MAGIC: an automated N-linked glycoprotein identification tool using a Y1-ion pattern matching algorithm and in silico MS(2) approach. *Anal. Chem.* **2015**, *87*, 2466–2473.

(60) Percy, A. J.; et al. Method and platform standardization in MRM-based quantitative plasma proteomics. *J. Proteomics* **2013**, *95*, 66–76.

(61) Abbatiello, S. E.; et al. Design, implementation and multisite evaluation of a system suitability protocol for the quantitative assessment of instrument performance in liquid chromatography-multiple reaction monitoring-MS (LC-MRM-MS). *Mol. Cell. Proteomics* **2013**, *12*, 2623–2639.

(62) Percy, A. J.; Chambers, A. G.; Smith, D. S.; Borchers, C. H. Standardized protocols for quality control of MRM-based plasma proteomic workflows. *J. Proteome Res.* **2013**, *12*, 222–233.

(63) Wada, Y.; et al. Comparison of the methods for profiling glycoprotein glycans—HUPO Human Disease Glycomics/Proteome Initiative multi-institutional study. *Glycobiology* **2007**, *17*, 411–422.

(64) Wada, Y.; et al. Comparison of methods for profiling O-glycosylation: Human Proteome Organisation Human Disease Glycomics/Proteome Initiative multi-institutional study of IgA1. *Mol. Cell. Proteomics* **2010**, *9*, 719–727.

(65) Kronewitter, S. R.; et al. The development of retrosynthetic glycan libraries to profile and classify the human serum N-linked glycome. *Proteomics* **2009**, *9*, 2986–2994.

(66) Ozohanics, O.; Turiak, L.; Puerta, A.; Vekey, K.; Drahos, L. High-performance liquid chromatography coupled to mass spectrometry methodology for analyzing site-specific N-glycosylation patterns. *Journal of chromatography. A* **2012**, *1259*, 200–212.

(67) Fournier, T.; Medjoubi-N, N.; Porquet, D. Alpha-1-acid glycoprotein. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* **2000**, *1482*, 157–171.

(68) Dente, L.; Ruther, U.; Tripodi, M.; Wagner, E. F.; Cortese, R. Expression of human alpha 1-acid glycoprotein genes in cultured cells and in transgenic mice. *Genes Dev.* **1988**, *2*, 259–266.

(69) Mitchell Wells, J.; McLuckey, S. A. Collision-induced dissociation (CID) of peptides and proteins. *Methods Enzymol.* **2005**, *402*, 148–185.

(70) Cao, L.; et al. Characterization of intact N- and O-linked glycopeptides using higher energy collisional dissociation. *Anal. Biochem.* **2014**, *452*, 96–102.

(71) Liu, H.; et al. Mass spectrometry-based analysis of glycoproteins and its clinical applications in cancer biomarker discovery. *Clinical proteomics* **2014**, *11*, 14.

(72) Imre, T.; et al. Glycosylation site analysis of human alpha-1-acid glycoprotein (AGP) by capillary liquid chromatography-electrospray mass spectrometry. *J. Mass Spectrom.* **2005**, *40*, 1472–1483.

(73) Ayyub, A.; et al. Glycosylated Alpha-1-acid glycoprotein 1 as a potential lung cancer serum biomarker. *Int. J. Biochem. Cell Biol.* **2016**, *70*, 68–75.

(74) Clerc, F.; et al. Human plasma protein N-glycosylation. *Glycoconjugate J.* **2016**, *33*, 309.

(75) Gimenez, E.; et al. Quantitative analysis of N-glycans from human alpha-acid-glycoprotein using stable isotope labeling and zwitterionic hydrophilic interaction capillary liquid chromatography electrospray mass spectrometry as tool for pancreatic disease diagnosis. *Anal. Chim. Acta* **2015**, *866*, 59–68.

(76) Luo, Z.; Lei, H.; Sun, Y.; Liu, X.; Su, D. F. Orosomucoid, an acute response protein with multiple modulating activities. *J. Physiol. Biochem.* **2015**, *71*, 329–340.