Targeted Mass Spectrometric Approach for Biomarker Discovery and Validation with Nonglycosylated Tryptic Peptides from N-linked Glycoproteins in Human Plasma*⑤

Ju Yeon Lee‡§‡‡, Jin Young Kim‡‡‡, Gun Wook Park‡¶, Mi Hee Cheon‡¶, Kyung-Hoon Kwon‡, Yeong Hee Ahn‡, Myeong Hee Moon§, Hyoung–Joo Lee‖, Young Ki Paik‖, and Jong Shin Yoo‡¶**

A simple mass spectrometric approach for the discovery and validation of biomarkers in human plasma was developed by targeting nonglycosylated tryptic peptides adjacent to glycosylation sites in an *N*-linked glycoprotein, one of the most important biomarkers for early detection, prognoses, and disease therapies. The discovery and validation of novel biomarkers requires complex sample pretreatment steps, such as depletion of highly abundant proteins, enrichment of desired proteins, or the development of new antibodies. The current study exploited the steric hindrance of glycan units in *N*-linked glycoproteins, which significantly affects the efficiency of proteolytic digestion if an enzymatically active amino acid is adjacent to the *N*-linked glycosylation site. Proteolytic digestion then results in quantitatively different peptide products in accordance with the degree of glycosylation. The effect of glycan steric hindrance on tryptic digestion was first demonstrated using alpha-1-acid glycoprotein (AGP) as a model compound *versus* deglycosylated alpha-1-acid glycoprotein. Second, nonglycosylated tryptic peptide biomarkers, which generally show much higher sensitivity in mass spectrometric analyses than their glycosylated counterparts, were quantified in human hepatocellular carcinoma plasma using a label-free method with no need for *N*-linked glycoprotein enrichment. Finally, the method was validated using a multiple reaction monitoring analysis, demonstrating that the newly discovered nonglycosylated tryptic peptide targets were present at different levels in normal and hepatocellular carcinoma plasmas. The area under the receiver operating characteristic curve generated through analyses of nonglycosylated tryptic peptide from vitronectin precursor protein was 0.978, the highest observed in a group of patients with hepatocellular carcinoma. This work provides a targeted means of discovering and validating nonglycosylated tryptic peptides as biomarkers in human plasma, without the need for complex enrichment processes or expensive antibody preparations. *Molecular & Cellular Proteomics 10: 10.1074/mcp.M111.009290, 1–12, 2011.*

Glycosylation is one of the most important post-translational modifications of proteins secreted in serum and is related to protein folding (1), quality control, sorting, degradation, and secretion. The glycan helps to stabilize polypeptide folding and indirectly allows glycoproteins to interact with various lectins, glycosidases, and glycosyltransferases by acting as recognition "tags" (2, 3). The glycoform distribution and degree of glycosylation of a glycoprotein can be significantly altered by diseases such as cancer, and glycoproteins have been reported in association with a variety of abnormal phenomena in patients with cancer (4–10). Thus, quantitative analyses of the glycosylation of proteins may be useful in discovering biomarkers for cancer.

Human plasma is clinically the most important sample, although it contains an extraordinarily large and complex proteome, with its dynamic range spanning more than 10 orders of magnitude in concentration. As conventional proteomic technologies only work within a small subset of this dynamic range (11), glycoprotein analysis in low concentrations typically requires glycoprotein enrichment methods, such as reactions with lectin (12–18), hydrazide chemistry (19, 20), HILIC (21, 22), or other procedures (23, 24) to reduce sample complexity while enriching the glycoproteins. However, such complex pretreatment steps are a cause of poor reproducibility. Also, the identification of glycopeptides and characterization of glycan information from mass spectrometric analysis is difficult because the instrumental sensitivity for glycopeptides is lower than that for general peptides, and a general data-

base for glycopeptides is lacking. These factors make the discovery of a glycoprotein biomarker, and its validation in real clinical samples, extremely difficult.

Two analytical strategies have been used with glycoproteins: glycan and glycopeptide analyses. Many current glycan analyses chemically or enzymatically cleave glycoproteins to generate a pool of glycan moieties prior to mass spectrometry-based analysis (25–27) These analyses provide both quantitative and structural information on glycan moieties. However, because the sample is a pool of glycan moieties from a glycoprotein mixture, these analyses are unsuitable for the detection of specific glycoforms released from an interesting glycoprotein and for the identification of glycosylation sites.

For glycopeptide analysis, deglycosylated peptides released by endoglycosidase are generally analyzed by MS (18, 19, 25, 28–31). This method is useful for identifying glycosylation sites; however, it usually requires glycopeptide enrichment. Enrichment methods such as lectin affinity chromatography or reaction with hydrazide beads have been widely applied for the analysis of glycoproteins at low concentrations.

However, lectin affinity methods have shown less satisfactory in enrichment efficiencies. For example, in a membrane protein study of breast tumor cells, only 25 of 88 proteins (28.4%) and 34 of 152 proteins (22.3%) were identified as glycoproteins using a concanavalin A (Con A) lectin affinity column, and only 22 of 87 proteins (26.4%) and 27 of 146 proteins (18.5%) were identified as glycoproteins using a wheat germ agglutinin lectin affinity column (31). The more general approach of the hydrazide method may isolate a higher number of glycoproteins or glycopeptides compared with lectin affinity methods designed to enrich them with specific glycan moiety. Although McDonald *et al.* (28) reported considerable nonspecific binding of abundant proteins to hydrazide resin in their analysis of cell surface glycoproteins, they identified 589 proteins in HeLa cell plasma membranes using the hydrazide method, and 191 (32%) corresponded to glycoproteins. Lee *et al.* (32) demonstrated binding of nonglycosylated proteins with both the lectin affinity and hydrazide methods of glycosylation enrichment. These low enrichment efficiencies are attributable to weak binding of glycoproteins or glycopeptides and nonspecific binding of abundant proteins, and can result in data with low reproducibility. It is not adequate for quantitatively analyzing a large number of samples. Thus, a simpler and more practical MS-based proteomics technology without the need for any complex and irreproducible glycoprotein enrichment steps or ambiguous glycan structural interpretations is desirable for the discovery of new glycoprotein biomarkers.

Tryptic digestions tend to be incomplete when applied to glycosylated proteins, because of glycan steric hindrance (25, 33) However, the ways in which post-translational modifications affect trypsin digestion can provide crucial clues for the discovery of new glycoprotein biomarkers. For example, if there are differences in the degree of glycosylation of a protein between normal and diseased states, it would result in quantitatively different peptide products in accordance with the degree of glycosylation.

In particular, because peptides adjacent to *N*-linked glycosylation sites are sterically affected by bulky glycan groups during proteolytic digestion, the relative abundance of nonglycosylated tryptic peptides adjacent to *N*-linked glycosylation sites is altered post-digestion. Many studies (18, 34, 35) and a review article (36) have reported that the degree of glycosylation can be aberrantly high in diseased samples. Thus, we hypothesized that these nonglycosylated tryptic peptides adjacent to *N*-linked glycosylation sites may be a valuable biomarker, revealing quantitative differences in the glycosylation states of patients with cancer. Nonglycosylated tryptic peptides are not only detected with higher sensitively by MS, but also are readily identified and quantified using existing proteomic technologies, because of the absence of heterogeneous glycans.

In the current study, targeted nonglycopeptide biomarkers were identified and validated among glycoproteins in human plasma. First, the effect of glycan steric hindrance in glycoproteins was demonstrated via a tryptic digestion of glycosylated alpha-1-acid glycoprotein (AGP-T)[1] and deglycosylated AGP treated with PNGase F (AGP-PT). AGP-T and AGP-PT indirectly indicated highly glycosylated disease plasmas and less glycosylated normal plasma, according to the degree of glycosylation. Second, one pooled normal and ten hepatocellular carcinoma (HCC) human plasma samples (sample information in supplemental Table S1) were analyzed using a label-free method consisting of nano-ultrahigh-performance liquid chromatography (UPLC) followed by quadrupole time-of-flight (Q-TOF) MS to identify candidate nonglycopeptide biomarkers. Nonglycosylated tryptic peptides adjacent to *N*-linked glycosylation sites from glycoproteins were targeted as candidate biomarkers because of observed, quantitative differences between normal and HCC plasma samples. Third, newly discovered nonglycosylated tryptic peptides were quantitatively validated by multiple reaction monitoring (MRM) using a nano-UPLC/triple quadrupole mass spectrometer from ten normal and eighteen HCC human plasma samples (sample information in supplemental Table S2).

Nonglycosylated tryptic peptide biomarkers were discovered without the need for complex glycoprotein enrichment methods using this conventional MS-based proteomics technology, where many of the limitations inherent in currently accepted glycoprotein biomarker discovery methods were avoided.

EXPERIMENTAL PROCEDURES

Materials—AGP standard protein (source: human), glucose-6-phosphate dehydrogenase (G6PD) standard protein (source: yeast), [Glu[1]]-fibrinopeptide B (GFP), dithiothreitol, iodoacetamide, and formic acid were purchased from Sigma-Aldrich (St. Louis, MO). Trypsin for protein digestion was obtained from Promega (Madison, WI). Peptide-N-Glycosidase F (PNGase F) for deglycosylation was supplied by New England Biolabs (NEB, Beverly, MA). Acetonitrile was of HPLC grade and obtained from J.T. Baker (Phillipsburg, NJ). Water was deionized through a Millipore system prior to use (Millipore, Eugene, OR). An HP1100LC system (Agilent, Santa Clara, CA) was used to deplete the levels of the six most abundant proteins using a multiple affinity removal system (MARS) consisting of an LC column (Agilent, 5185–5984), solutions of Buffer A for sample loading, washing, and equilibration (Agilent, 5185–5987), and Buffer B for elution (Agilent, 5185–5988). The stable isotope-labeled peptides TEDTIFL*R, TINPAV*DHCCK, ENFLFLTP*DCK and GQYCYEL*DEK (isotopically labeled [^{13}C and ^{15}N] at amino acid sites marked with asterisks) were obtained from Anygen Co. (Kwangju, Korea). Carbamidomethylation of cysteine residues was performed in our laboratory.

Alpha-1-Acid Glycoprotein Preparation—Two AGP standard solutions were prepared at concentrations of 100 μg/10 μl. AGP solutions were reduced by adding with 1 μl of 1 M dithiothreitol at 60 °C for 30 min. Proteins were alkylated by adding 5 μl of 500 mM iodoacetamide in the dark at room temperature and allowing the solution to react for 30 min. One of the 100-μg AGP solutions was deglycosylated by reacting with 1 μl of PNGase F (500,000 u/mg) at 37 °C for 6 h followed by incubation and digestion in a solution of 10:1 trypsin (total protein:trypsin, by weight) at 37 °C overnight (AGP-PT). The other AGP solution was incubated in 1 μl of deionized water, instead of PNGase F, at 37 °C for 6 h. This solution was digested as described above (AGP-T). The two AGP digests were dried in a SpeedVac. The dried samples were diluted in mobile phase A and, as needed, spiked with isotope-labeled TEDTIFL*R prior to LC/MS/MS analyses.

Plasma Sample Preparation—Plasma samples were obtained with informed consent and in accordance with IRB guidelines from Yonsei University College of Medicine (Seoul, Korea).

Plasma samples of blood from healthy donors and HCC cancer patients were divided into four equal volume bags with an appropriate concentration of K_2EDTA. Each aliquot was frozen and stored at −80 °C until use.

Depletion of Major Abundant Proteins from Human Plasma by Immunoaffinity—The six most abundant proteins in human plasma: albumin, transferrin, IgG, IgA, haptoglobin, and α_1-antitrypsin were depleted using a HP1100LC system (Agilent) equipped with a multiple affinity removal column (MARC; Agilent). Crude human plasma samples were diluted by a factor of five with Buffer A (for example: 80 μl of Buffer A was added to 20 μl of human plasma) containing the proper amount of protease inhibitor (Complete Protease Inhibitor Mixture tablet (Roche, Indianapolis, IN)) and filtered through 0.22-μm filters by centrifugation (16,000 × g, room temperature, 1–2 min). Diluted crude plasma (70–100 μl) was injected at a flow rate of 0.25 ml/min; flow-through fractions were collected and stored at −20 °C. Depleted plasma samples were desalted by centrifugal filtration using 5000-Da MWCO (molecular weight cutoff) VIVASPIN filters (product VS061; Sartorius, Göttingen, Germany).

Digestion of Plasma Proteins and Internal Standards—Aliquots (5–10 μg) of human plasma samples that had been quantitatively analyzed by Bradford protein assays were diluted with 100 mM Tris-HCl buffer (pH 8.00). An internal standard of G6PD was prepared in 100 mM Tris-HCl buffer (pH 8.00). The plasma and internal standard solutions were digested using the same AGP digestion process without the deglycosylation step. Briefly, human plasma samples and the

internal standard were reduced, alkylated, and incubated as above. Digested samples were dried using a SpeedVac. Dried samples were frozen and stored at −20 °C until use. For label-free quantification, digested internal standard solutions were spiked equally prior to liquid chromatography tandem MS (LC/MS/MS) analyses. For MRM quantification, stable isotope-labeled peptide mixtures in each plasma sample were spiked equally prior to LC/MS/MS analyses.

Nano-LC-ESI-MS/MS for Proof (DDA) of Glycan Steric Hindrance and Selection (DDA, MS^E) of Candidate Biomarkers—Digested AGP samples were dissolved in mobile phase A, and stable isotope-labeled peptide standards of AGP were spiked equally in each AGP sample prior to nano-LC/electrospray ionization (ESI)-MS/MS analyses. Digested plasma samples were also dissolved in mobile phase A, and an internal standard of digested G6PD was added equally to each plasma sample prior to nano-LC/ESI-MS/MS analyses. MS/MS experiments for identification and quantification of proteins from peptide mixtures were performed using a nano-LC/MS system consisting of a nano ACQUITY UPLC system (Waters Corp., Milford, MA) and a Q-TOF mass spectrometer (Premier; Waters Corp.) equipped with a nano-ESI source. An autosampler was used to load 5-μl aliquots of the peptide solutions onto a C_{18} trap column (I.D. 180 μm, length 20 mm, and particle size 5 μm; Waters Corp.). The peptides were desalted and concentrated on the trap column for 5 min at a flow rate of 20 μl/min. The peptides were then separated at a flow of 450–500 nL/min on an analytical microcapillary column (C_{18}, I.D. 75 μm, length 150 mm, particle size 1.7 μm; Waters Corp.). Mobile phase A consisted of water with 0.1% formic acid. Mobile phase B contained acetonitrile with 0.1% formic acid. Gradient elution began with 1% B for 0.33 min, ramped to 35% B over 84.67 min, then to 50% B over 0.5 min, retained at 50% B for 6.5 min, then ramped to 95% B over 0.5 min, retaining for 7.5 min and then to 1% B for another 0.5 min. The column was equilibrated with 1% B for 19.5 min before each run. For AGP, the total run time was 75 min with ramping to 40% B over 39.67 min. All analyses were performed using positive and V mode ESI using a nano-spray source. In label-free analyses, for accurate MS measurements, a lock mass of [Glu1]-fibrinopeptide B(GFP) at 300 fmol/μl was injected once every 30 s during a single LC/MS/MS run at a flow rate of 0.5 μl/min with NanoLock spray source. The voltage applied to produce the electrospray was 2.2 kV and the cone voltage was 35 eV. Argon was introduced as a collision gas at a pressure of 20 psi. DDA mode was used for AGP samples and to focus the database of human plasma samples. The MS scan range was 400–1600 m/z and the spectral acquisition time in each mode was 1.0 s with a 0.1-s interscan delay for DDA mode. Data-dependent peak selection of the three most abundant MS ions was applied and the collision energy was used as a means of charge state recognition. Elevated energy (MS^E) mode was used for label-free quantification of plasma samples and all samples were analyzed in triplicate.

MS^E is used in LC/MS studies in order to obtain fragment information about small molecules in complex mixtures, based on acquisition of exact mass data at alternating high and low collision energies. In low-energy MS mode, data were collected at constant collision energy of 5 eV to yield peptide molecular weight information. In elevated MS^E mode, the collision energy was increased from 20 to 35 eV for peptide fragmentation. All MS data generated during the elution of each peptide from LC were collected for the identification and quantitative analysis of the peptide. In both low energy (MS) and MS^E modes of acquisition, the mass scan range was 50–2000 m/z with spectral acquisition times of 1.5 s and 1.0 s, respectively, for MS and MS/MS modes with a 0.1-s interscan delay in MS^E mode.

Database Search-Identification for DDA Data—Peak lists were generated and processed using the MassLynx software (ver. 4.1; Waters Corp.). MS spectra were smoothed once using a 5-point Savitzky-Golay method, centered on the top 50% of each peak. The resulting

raw files from each analysis were automatically processed into a single *.pkl peak list file. The resulting *.pkl peak list files were processed against the International Protein Index (IPI) Human database (version 3.22; 57,846 sequences; 26,015,783 residues) using the Mascot search engine (version 2.1; Matrix Science, London, UK). Mascot was used with monoisotopic mass selected, a precursor mass tolerance of ± 1.5 Da, and a fragment mass tolerance of ± 0.8 Da. Trypsin was selected as the digestion enzyme, with one potential missed cleavage. ESI-Q-TOF was selected as the instrument type. Oxidized methionine, carbamidomethylated cysteine, propionamide cysteine, and pyroglutamate (N-term E, Q) were chosen as variable modifications for identifying nano-UPLC/ESI-Q-TOF results in DDA mode. Deglycosylated AGP analyses also included deamidation (N). In a Mascot search, AGP came up as the first hit with high-scoring peptides above the search threshold (expected $p < 0.05$, peptide score >42). All of the Mascot search results from normal human plasma, and individual HCC plasma samples were summed. The focused database was composed of sequences from all of the identified proteins in the pooled normal plasma sample and the ten HCC individual plasma samples plus the sequence of the G6PD internal standard.

Database Search and Quantification for MS^E Data—A continuum of LCMS^E data was generated and processed using ProteinLynx Global Server software version 2.2 (PLGS2.2; Waters Corp.). Protein identifications were performed by using the forced database composed of the Mascot search results from the previous step. The ion detection, clustering of peptide components by mass and retention time, and data normalization were performed using PLGS, as described by Silva and *et al.* (37, 38). Briefly, all raw data were corrected for mass accuracy by m/z 785.8426, corresponding to the GFP ([Glu^1]-fibrinopeptide B) standard that was used as the lock mass. MS spectra were smoothed with a 7-point Savitzky-Golay method, centered on the top 80% of each peak. Trypsin was used as the digestion enzyme with one potential missed cleavage. ESI-QTOF was selected as the instrument type. Oxidized methionine and carbamidomethylated cysteine were chosen as modifications. For quantification, exact mass was selected with a mass tolerance of 20 ppm and the fine retention time window was 1-min. The intensity measurements for the entire data set were normalized to the internal standard, G6PD. From these criteria, the peak, which was detected at least twice, was used to quantify both normal and HCC individual plasma samples.

Nano-LC-ESI-MS/MS for Validation of Selected Biomarker Candidate Peptides: MRM Quantification by Online Nano-LC-MS—MRM experiments were performed with a nanoACQUITY UPLC system (Waters Corp.) and a TSQ quantum ultra EMR triple-quadrupole mass spectrometer (Thermo Finnigan, San Jose, CA) equipped with a nanospray source. With the exception of the gradient, LC conditions were the same as those described above for the selection of candidate biomarkers. The gradient began with 5% B for 0.33 min, then ramped to 40% B over 39.67 min, to 50% over 10 min, to 95% over 0.5 min, retained at 95% B over 9.5 min, and then ramped to 5% B over 0.5 min. The column was equilibrated with 5% B for 14.5 min before the next run. Optimized conditions were used for MRM quantification: quadrupoles Q1 and Q3 were held at 0.7 m/z FWHM with a scan time of 50 ms per peptide and 2.0 kV spray voltage. Optimum transitions and collision energy parameters were determined for each peptide by infusion of 10 pmol/μl peptide solution (Table III). Three transitions for each peptide were selected and monitored. First, precursor ions with a specific mass were transmitted from Q1 to the collision cell for fragmentation. Three fragment ions were then transmitted through Q3, yielding the signals used for quantification (39).

Statistical Analyses—For quantitative comparisons of the label-free plasma samples, the results of exact mass and retention time (EMRTs) from the PLGS analysis were exported to an Excel spreadsheet (Microsoft, Redmond, WA), and all data were normalized using a linear regression analysis. Validated MRM results of target nonglycopeptides from the normal and HCC cancer groups were compared statistically using MedCalc (version 10.1.8.0, demo version). The diagnostic accuracy of each peptide from the candidate biomarkers was evaluated using receiver operating characteristic (ROC) curve analyses. The area under the curve (AUC) is reported at the 95% confidence interval (CI). The related sensitivity and specificity were determined.

RESULTS AND DISCUSSION

Proof-Of-Concept for the Effects of Glycan Steric Hindrance on Proteolytic Digestion Using AGP—Successful enzyme digestion results in a solution of peptides that are completely and accurately cleaved at expected sites. This only occurs when the enzymes can efficiently access the reactive cleavage sites. Thus, steric hindrance in a protein that interferes with enzyme access affects proteolytic digestion (25, 33). This effect is more pronounced in proteins such as *N*-linked glycoproteins that have been highly modified post-translationally. Steric hindrance from the glycan group(s) in *N*-linked glycoproteins has been thought to significantly affect the efficiency of proteolytic digestion if the enzymatically active amino acid is adjacent to the *N*-linked glycosylation site, resulting in quantitatively different peptide products in accordance with the degree of glycosylation.

This concept is shown in Fig. 1. The illustrations depicting *N*-linked glycoproteins are based on an article by Bunkenborg (40). Both normal (left) and aberrant (right) glycosylated *N*-linked glycoproteins are shown. After tryptic digestion, most of the peptides that are not related to the *N*-linked glycosylation sites will be present at the same concentration, with a relative abundance close to unity. However, concentrations of peptides that were adjacent to *N*-linked glycosylation sites, such as peptide 7, will be lower due to steric hindrance by the glycan in aberrantly glycosylated *N*-linked glycoproteins. Because steric hindrance increases with the size of the glycan unit, the abundance of peptide 7, which boasts a relatively large glycan moiety, will be significantly decreased. In diseased plasma samples, aberrant glycosylation of glycoproteins, such as increased levels of fucose and sialic acid, the addition of polylactosamine units, or higher-order branching of *N*-linked glycans, are known (33). The peak area ratios (aberrant/normal) of specific nonglycosylated tryptic peptides that originate from sites adjacent to *N*-linked glycosylation sites are thus expected to be lower than those of other peptides. These peptides were the target of this study: nonglycosylated tryptic peptides from *N*-linked glycoproteins.

First, the effect of glycan steric hindrance on trypsin digestion was demonstrated using AGP as a model compound. Two AGP digest solutions were prepared and analyzed by LC/MS/MS. One consisted of a tryptic digest of AGP (AGP-T). The other consisted of a tryptic digest of deglycosylated AGP that had already been treated with PNGase F (AGP-PT).

Fig. 2*A* shows a quantitative comparison of peptides from AGP-T and AGP-PT samples. The peptide sequences are
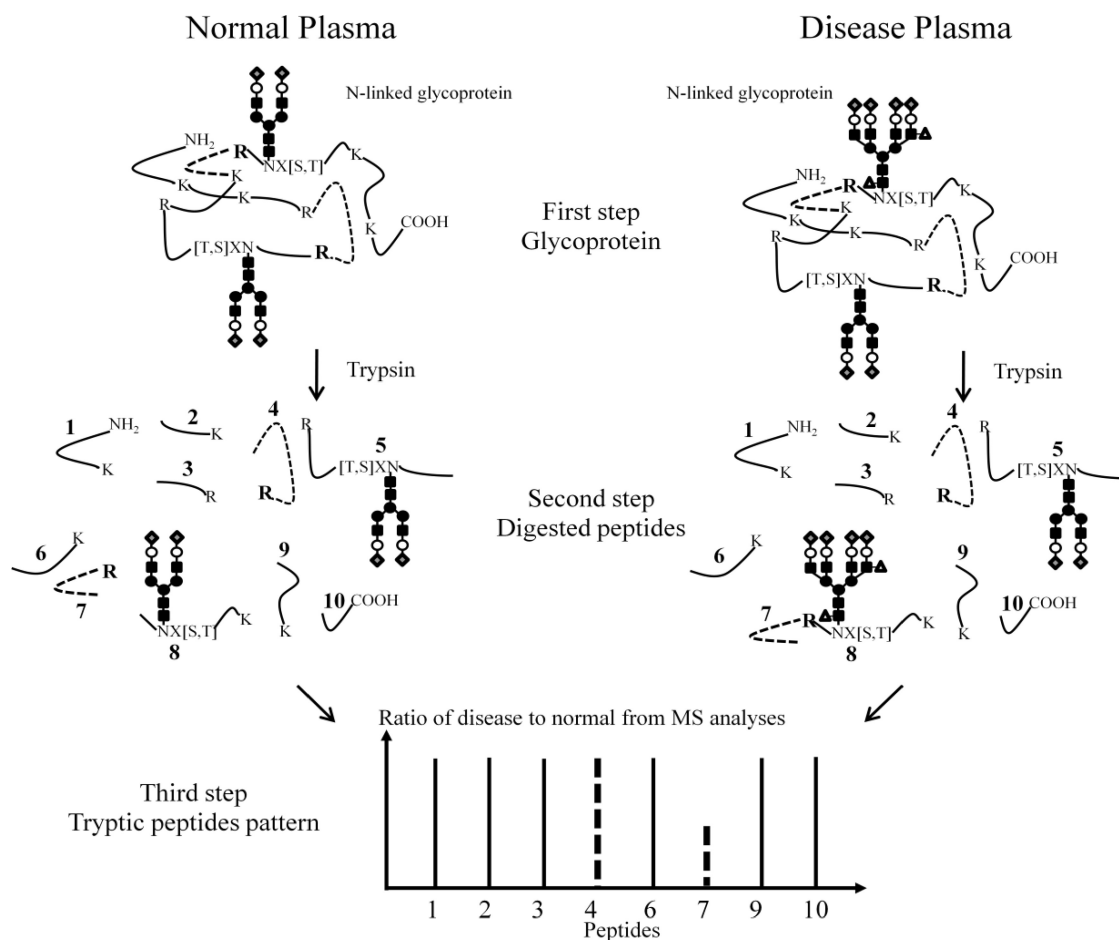
FIG. 1. **Effects of glycan steric hindrance and the degree of glycosylation on the efficiency of proteolytic digestion are illustrated.** If the degree of glycosylation or size of the glycan unit(s) differs between normal and diseased samples, then these differences will be reflected in the relative abundance of nonglycosylated tryptic peptides that were adjacent to N-linked glycosylation sites following digestion. In the diseased sample shown here, nonglycosylated tryptic peptides (4 and 7) and N-linked glycosylated peptides (5 and 8) will be cleaved with a different efficiency compared with those in a normal plasma sample. Furthermore, the relative abundance of nonglycosylated tryptic peptide 7 will differ in the mass spectra of normal and diseased samples. Peptide 4 will not be affected by glycan steric hindrance because "R" (the active site for trypsin) is distant from the N-linked glycosylation site. Thus, peptide 4 will be detected at similar levels in both the normal and diseased samples. Other peptides, except nonglycosylated tryptic peptide 7 and peptide 4, will be present at similar levels. Full lines represent general peptides and dotted lines indicate nonglycosylated tryptic peptides. Glycopeptides 5 and 8 could not be identified in the MS analyses and are not displayed in the third step.

given in Table I, where the underlined sequences indicate the quantified peptides. The nearest N-linked glycosylation sites in each identified peptide are indicated as 72N and 103N (in parentheses). Methionine oxidation generally occurs when a methionine residue is exposed to oxygen during digestion, and cysteine carbamidomethylation occurs during treatment with iodoacetic acid. The peak intensity ratio was much lower for the nonglycosylated tryptic peptide TEDTIFLR than for the other peptides (peptide 2, 3, and 4), which had peak intensity ratios of approximately one.

This indicates that the production of enzyme-digested peptides that had been adjacent to the N-linked glycosylation site in AGP was affected by the steric hindrance of the glycan unit. Peptide TEDTIFLR was produced by tryptic cleavage at both sites 73K and 81R. However, because of the steric hindrance by the glycan linked at 72N, trypsin's access to the 73K cleavage site was inhibited. Indeed, it caused the peptide TEDTIFLR to be present at low levels in the AGP-T sample. Notably, the same peptide exhibited a much higher abundance in the AGP-PT sample where the glycan had already been removed by PNGase F treatment prior to tryptic digestion. Furthermore, the degree of this effect would increase with the size of the glycan and decrease with the length between the enzyme active site and the N-linked glycosylation site.

Stable isotope-labeled peptide, TEDTIFL*R, was synthesized and spiked equally into the AGP-T and AGP-PT samples prior to LC/MS/MS analysis. The relative abundance of peptide TEDTIFLR (light) and peptide TEDTIFL*R (heavy) in the AGP-T and AGP-PT samples is shown in Fig. 2B. The level of TEDTIFLR peptide in the AGP-T sample was less
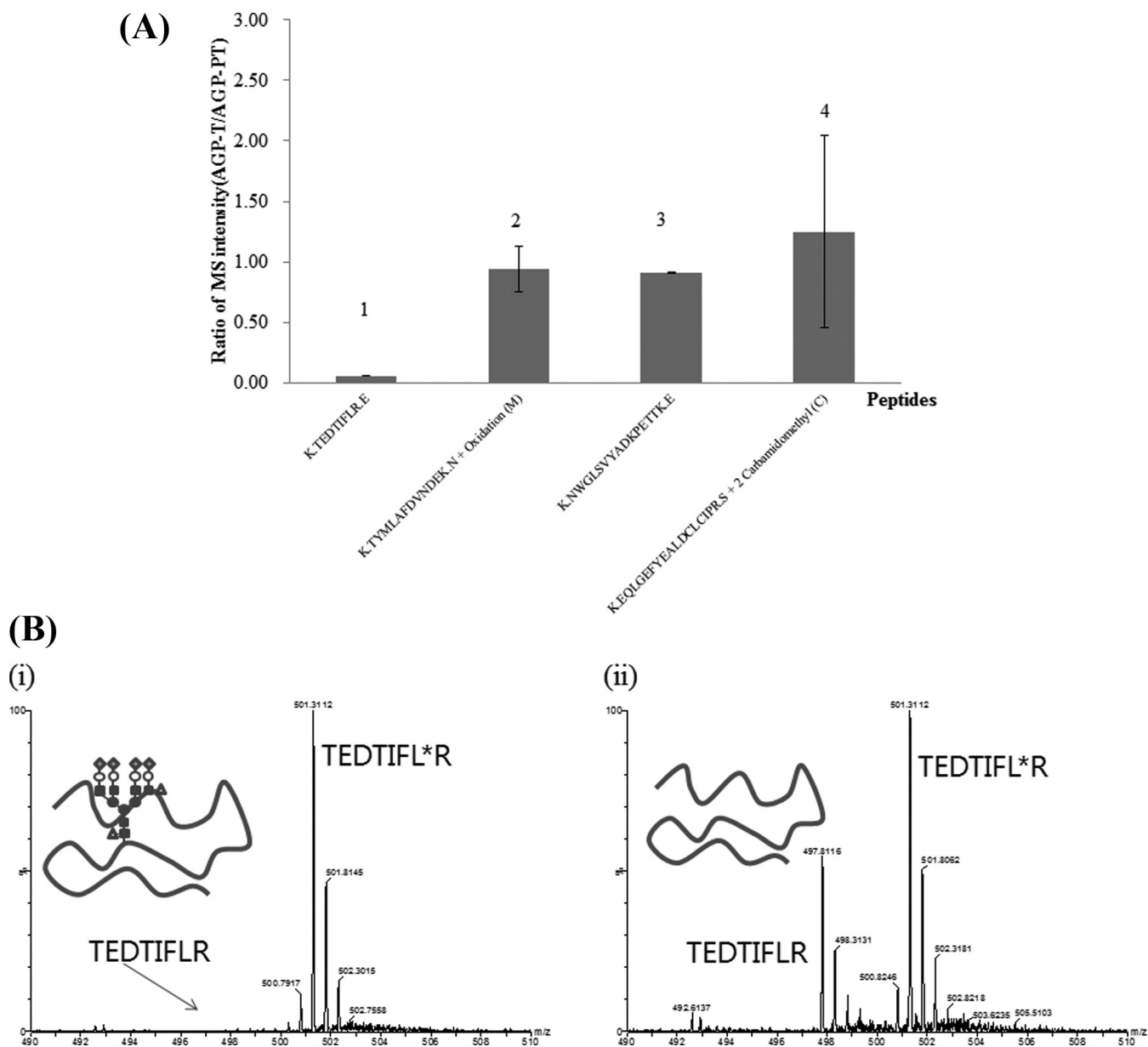
(A)

(B)

(i)

(ii)

FIG. 2. **Quantitative comparison of tryptic peptide MS intensities from glycosylated AGP (AGP-T) and deglycosylated AGP (AGP-PT) samples.** AGP-T; tryptic digests of glycosylated AGP. AGP-PT; tryptic digests of deglycosylated AGP with PNGase F. *A*, The ratio of peptide MS intensities. Peptide 1: TEDTIFLR, nonglycosylated tryptic peptide adjacent to an *N*-linked glycosylation site. Peptide 2: TY*M*LAFDVNDEK. Peptide 3: NWGLSVYADKPETTK. Peptide 4: EQLGEFYEALD**C**L**C**IPR. Italic bold cysteine residues and methionine residues were carbamidomethylated and oxidized, respectively. *B*, MS spectra of TEDTIFLR of AGP-T and AGP-PT. (I) AGP-T, (II) AGP-PT. AGP-T and AGP-PT samples were spiked equally with the stable isotope-labeled peptide TEDTIFL*R (of TEDTIFLR) prior to quantitative LC/MS/MS analyses. TEDTIFLR was adjacent to an N-glycosylated site and exhibited a lower abundance in AGP-T relative to AGP-PT.

than 10% of that in the AGP-PT sample. This experiment again demonstrates that the amount of tryptic peptide, TEDTIFLR, was quantitatively different in AGP-T and AGP-PT samples.

Discovery of Candidate Nonglycopeptide Biomarkers by Label-Free Analyses—A label-free quantitative LC/MS/MS method was applied to identify nonglycosylated tryptic peptide biomarker from HCC human plasma. Many of the labeling approaches in quantitative proteomics have potential limitations, including complex and expensive sample preparation and incomplete labeling. Conversely, label-free quantification approaches do not have these limitations and, above all, are not limited by the number of samples. Label-free methods are limited only by the reproducibility of LC and the accuracy of mass measurements in mass spectrometry.

TABLE I

Quantified peptide sequences of AGP, determined by LC/MS/MS. Underlined sequences are quantified peptides. Bold letters indicate an N-linked glycosylated motif. 72 N and 103 N are the nearest N-linked glycosylation sites of each peptide. Italic bold cysteine and methionine were carbamidomethylated and oxidated, respectively

No.	Peptide sequences from AGP	m/z (charge)	Peptide score (mascot)
1	P**NKT**EDTIFLRE (71–82, 72 N)	497.6892(+2)	60
2	KTY***M***LAFDVNDEKN (126–139, 103 N)	731.2680(+2)	64
3	KNWGLSVYADKPETTKE (138–154, 103 N)	570.2266(+3)	76
4	KEQLGEFYEALD***C***LC***I***PRS (153–171, 103 N)	1056.8601(+2)	88

In this study, all 11 plasma samples (one pooled normal and ten HCC plasma samples) exhibited highly reproducible retention times and accurate mass measurements for separated peptides. The standard deviations in the retention time and in the mass accuracy of the eight tryptic peptides from the internal standards (G6PD) were 0.19–0.30 min and 0.53–3.45 ppm, respectively (supplemental Table S3).

All of the plasma samples were analyzed in DDA mode for identification prior to label-free quantification. All identified proteins from 11 analyses (one normal and ten HCC human plasma samples) were combined to form a focused database. This database consisted of 94 proteins including single peptide identifications. The small size of this database, relative to the total human database, increased the reliability of label-free MS^E results by the PLGS algorithm. Each sample was analyzed in triplicate, in MS^E mode by LC/MS/MS. Peptides were identified and quantified by PLGS, developed by Sliva *et al.* (37, 38). Quantitative results were statistically normalized by linear regression, although the retention times of the label-free results were normalized by the internal standard peptides in PLGS.

Between the two groups, 133 peptides were quantified; 17 nonglycosylated tryptic peptides related to *N*-linked glycosylation sites are listed in Table II. In terms of the average ratio from all quantified peptides for each protein, the normal plasmas were similar to those of the HCC plasmas (Table II). In general, it ranged within ±30%. At the peptide level, however, some peptides changed over 30% between the two groups or showed quite different ratios in comparison with the averaged ratios of the protein levels.

We selected four targets of peptides, 1, 2, 4, and 7, as nonglycosylated tryptic peptide biomarkers that were not only adjacent to *N*-linked glycosylation sites, but also exhibited quantitative differences between normal and HCC human plasma samples (marked with an asterisk in Table II). Despite their proximity to *N*-linked glycosylation sites, peptides 8, 9, 13, 14, 15, and 17 showed small quantitative differences of less than 20% and were excluded from consideration of nonglycosylated tryptic peptide biomarkers. Nonglycosylated tryptic peptides, such as 3, 5, 12, and 16, were also not selected because they were distant from the *N*-linked glycosylation sites. Also, peptides 10 and 11 were not selected based on their poor MS/MS spectra by MS^E (supple-mental Fig. S1 presents MS/MS spectra of all nonglycosylated tryptic peptides in Table II except for peptides 10 and 11; the peptide probabilities from the PLGS search algorithm were over 95 for all peptides). Isoform HWM of Kininogen 1 precursor was identified by peptides 3, 4, 5, and 6. Among these peptides, peptide 4 was selected based on its proximity to *N*-linked glycosylation site.

Fig. 3*A* shows the extracted ion chromatogram of peptide TEDTIFLR, a nonglycosylated tryptic peptide biomarker (No. 1 in Table II) from AGP. Compared with those of other peptides (Figs. 3*B*, 3*C*, 3*D*), TEDTIFLR from HCC plasma showed a relatively low intensity in peak abundance. Such a result is consistent with our previous experimental data, obtained as a proof of principle. As we expected, this demonstrated that the difference in a protein's glycosylation state between a normal and diseased sample can be determined by specific nonglycosylated tryptic peptides. Other MS-based label-free results showing only quantitative differences in specific nonglycosylated tryptic peptide biomarkers are presented in supplemental Fig. S2.

Although the label-free method showed quantitative differences in the nonglycosylated tryptic peptide TEDTIFLR of AGP, the Western blot analysis of AGP was very similar between the normal and HCC plasma samples (supplemental Fig. S3). Western blot analysis without specific glycoprotein enrichment did not reveal any distinguishing information between the normal and disease groups, despite differences in the glycan groups or degree of glycosylation reported previously (41–43). Western blot analysis shows total protein levels and cannot accurately detect changes in protein glycosylation.

At the peptide level, peptide TEDTIFLR from AGP, peptide TINPAVD***CC***K from afamin precursor, peptide ENFLFLTP-D***C***K from isoform HMW of kininogen 1 precursor, and peptide GQY***C***YELDEK from vitronectin precursor protein could be identified as nonglycosylated tryptic peptide biomarkers for HCC, without glycoprotein enrichment.

Verification of Candidate Nonglycopeptide Biomarkers by MRM Analyses—Triple quadrupole mass spectrometry yielded very specific and sensitive results, because all targeted peptides were monitored by two mass filters (MS1: parent ion, and MS2: specific fragment ions of the parent ion) during the MRM. MS-based MRM assays (44–49) have been

TABLE II

List of peptides related to N-linked glycosylation sites. Bold letters are N-linked glycosylation motifs and underlined sequences are quantified peptides among normal and HCC plasma. Asterisks indicate peptides that were selected based on the proof-of-concept demonstrated here and candidate nonglycopeptides. Italic bold cysteine residues and methionine residues were carbamidomethylated and oxidized, respectively

No.	Peptide sequences with N-linked glycosylated motifs	IPI accession no. protein name	MH+	Targeted peptide		Each protein		
				Average of HCC/normal ratio	Std. dev. of HCC/normal ratio	Average of HCC/normal ratio range	Std. dev. of HCC/normal ratio range	Concentration mg/L[b]
1[a]	NKTEDTIFLR	IPI00022429 Alpha-1-acid Glycoprotein 1 precursor	994.5197	0.76	±0.24	1.40	±0.46	350–840
2[a]	NRTINPAVDHCCK	IPI00019943 Afamin precursor	1314.5969	0.38	±0.31	1.21	±0.18	50–120
3	YNSQNQSNNQFVLYRITEATKTVGSDTFYSFK	IPI00032328 Isoform HMW of Kininogen 1 precursor	1251.5918	0.63	±0.09	0.74	±0.15	180
4[a]	ITYSIVQTNCSKENFLFLTPDCK		1383.6638	0.64	±0.10	0.74	±0.15	180
5	ICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFK		861.4065	0.26	±0.18	0.74	±0.15	180
6	ICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFK		2139.0798	0.64	±0.11	0.74	±0.15	180
7[a]	NGSLFAFRGQYCYELDEK	IPI00298971 Vitronectin precursor	1304.5535	0.58	±0.34	1.17	±0.24	60–170
8	FEDGVLDPDYPRNISDGFDGIPDNVDAALALPAHSYSGRER		1422.6573	0.94	±0.20	1.17	±0.24	60–170
9	NVTLTCQARLETPDFQLFK	IPI00022895 Alpha 1B glycoprotein precursor	1237.6484	1.04	±0.22	1.16	±0.24	150–300
10	LGACNDTLQQLMEVFK	IPI00032179 Antithrombin III variant	1882.8894	2.08	±2.51	1.46	±0.88	170~300
11	IADAHLDRVENTTVYYLVLDVQESDCSVLSR	IPI00022371 Histidine-rich glycoprotein precursor	910.4670	0.71	±0.26	0.81	±0.14	50–150
12	VIDFNCTTSSVSSALANTKDSPVLIDFFEDTER		1682.7937	0.73	±0.44	0.81	±0.14	50–150
13	SWPAVGNCSSALRWLGRYYCFGGNQFLR	IPI00022488 Hemopexin precursor	1495.6808	1.18	±0.19	1.09	±0.37	500–1150
14	NDFTWFKLNDTLDYECHDGY	IPI00029739 Isoform 1 of Complement factor H precursor	957.4464	0.95	±0.28	0.97	±0.13	240–760
15	QCVDGEWTTLPVCIVEESTCGDIPELEHGWAQLSSPPYYYGDSVEFNCSESFTMIGHRSITCIHGVWTQLPQCVAIDK		2326.1636	1.13	±0.31	0.97	±0.13	240–760
16	DPEVNCSMAQIQLCPPPPQIPNSHNMTTTLNYRDGEKVSVLCQENYLIQEGEEITCK		2412.1501	0.78	±0.31	0.97	±0.13	240–760
17	LSYTCEGGFRISEENETTCYMGK		1189.5320	0.96	±0.17	0.97	±0.13	240–760

[a] Selected nonglycosylated tryptic peptides by label free quantification.
[b] Approximate molar abundance values of polypeptides in plasma from healthy subjects from Hortin *et al.* (58).
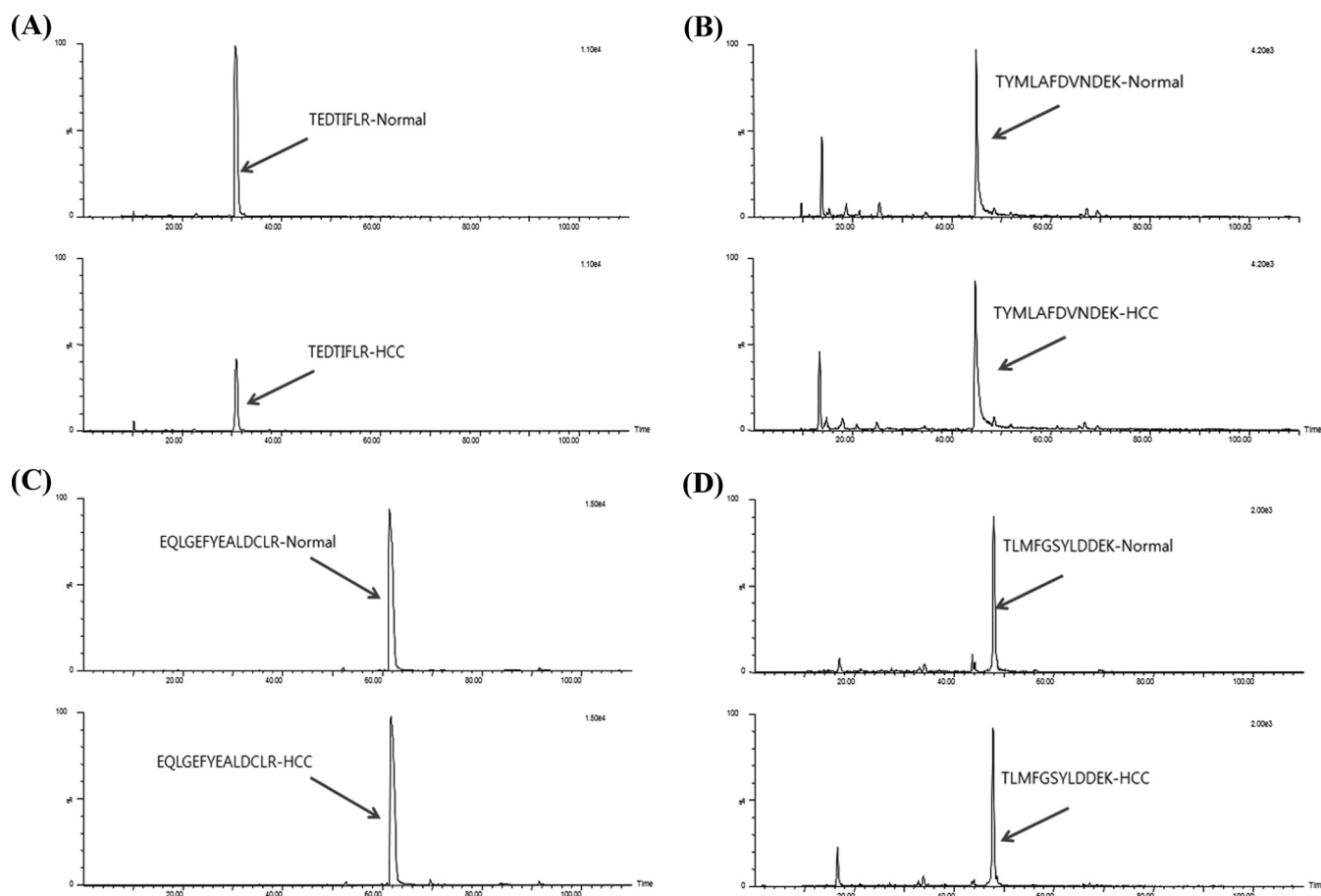
FIG. 3. **Chromatograms of nonglycopeptide biomarker candidates and general peptides from AGP targeted proteins are shown as acquired by the MS-based label-free method (*A–D*).** Nonglycosylated tryptic peptides and general tryptic peptides were differentially quantified between the normal and HCC groups. Italic bold cysteine residue and methionine residue were carbamidomethylated and oxidated, respectively. *A*, AGP - TEDTIFLR, nonglycosylated tryptic peptide biomarker candidate. *B*, AGP - TY*M*LAFDVNDEK, general peptide. *C*, AGP - EQLGEFYEALD*C*LR, general peptide. *D*, AGP - TLMFGSYLDDEK, general peptide.

proposed as an alternative to antibody-based biomarker verification due to the high throughput, selectivity, and sensitivity of using two mass filter stages. This approach was used to verify selected nonglycosylated tryptic peptide biomarkers in ten normal and eighteen HCC human plasma samples.

All of the plasma samples were prepared as in the previous identification step. Prior to the LC/MS/MS run, all of the plasma samples were spiked equally with four stable isotope-labeled standard peptides (heavy) of selected native nonglycosylated tryptic peptides (light) and analyzed by triple quadrupole mass spectrometry combined with the same UPLC system used for the label-free analyses. The optimum transition ions and collision energy conditions for each peptide were determined through direct infusion nanoflow-ESI experiments with stable isotope labeled standards (Table III). Calibration curves for each peptide also showed strong linearity in the range of concentrations from 10 to 500 fmol for the human plasma sample (2 μg; supplemental Fig. S4). The MS and MS/MS tolerances were $\Delta 1$ Da and $\Delta 0.7$ Da, respectively.

Quantitative MRM analyses were run in triplicate for each sample. Data that resulted in detection events in two or more of the three trials, within CV (coefficient of variation) $<$ 30%, were accepted. supplemental Table S4 shows the CV% of light/heavy ratios of targeted nonglycosylated tryptic peptides from normal and HCC plasma samples by MRM. Data that yielded results under the limit of quantification for each peptide were removed. MRM chromatograms of the targeted nonglycosylated tryptic peptide and those of stable isotope labeled standard peptides are shown in supplemental Fig. S5.

The results of the MRM analyses were analyzed statistically using MedCalc (version 10.1.8.0). ROC curves and quantitative scatter plots showing the MRM results in both the normal and HCC groups are shown in Fig. 4. The target nonglycosylated tryptic peptides were present at quantitatively lower levels in the HCC samples relative to the normal group. This was consistent with the results from the label-free quantitative analysis (Table II). This corroborates the results of the label-free analyses described above. The area under the ROC curve (AUC) values of each peptide are shown in the Table IV.

TABLE III

*MRM conditions used for target nonglycopeptides and stable isotope-labeled standard peptides. Italic bold cysteine residues were carbamidomethylated. *isotopically labeled sites*

| Sequences	Q1 (*m/z*, +2)	Q3 (*m/z*, +1)	CE	Protein	Isotope (ΔDa)
GQY**C**YELDEK	652.777	796.372(y6), 956.402(y7), 1119.466(y8)	22	Vitronectin precursor	
GQY**C**YEL*DEK	656.285	803.389(y6), 963.420(y7), 1126.483(y8)	22	Isotope labeled standard	L:$^{13}C_6$,^{15}N (7.017)
TEDTIFLR	497.764	435.271(y3), 649.403(y5), 764.430(y6)	20	Alpha-1 acid glycoprotein 1 precursor	
TEDTIFL*R	501.273	442.288(y3), 656.420(y5), 771.447(y6)	20	Isotope labeled standard	L:$^{13}C_6$,^{15}N (7.017)
TINPAVDH**CC**K	657.800	818.328(y6), 986.418(y8), 1100.461(y9)	28	Afamin precursor	
TINPAV*DH**CC**K	660.807	824.342(y6), 992.432(y8), 1106.474(y9)	28	Isotope labeled standard	V:$^{13}C_5$,^{15}N (6.014)
ENFLFLTPD**C**K	692.334	519.223(y4), 620.270(y5), 880.423(y7)	22	Isoform HMW of Kininogen-1 precursor	
ENFLFLTP*D**C**K	695.341	525.236(y4), 626.284(y5), 886.437(y7)	22	Isotope labeled standard	P:$^{13}C_5$,^{15}N (6.014)
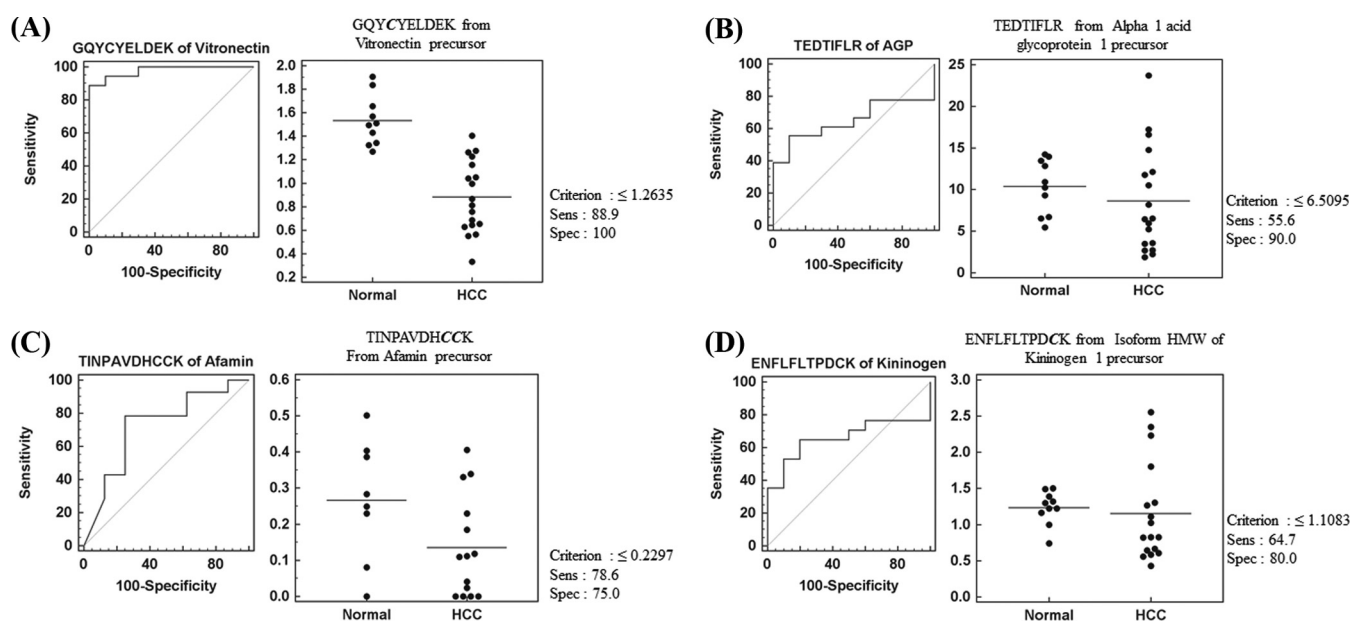
FIG. 4. **ROC curves and scatter plots of target nonglycopeptides from normal and HCC plasma, as determined by MRM quantification using triple-quadrupole mass spectrometry.** Italic bold cysteine residues were carbamidomethylated. *A*, GQY**C**YELDEK- Vitronectin precursor. *B*, TEDTIFLR – AGP. *C*, TINPAVDH**CC**K- Afamin precursor. *D*, ENFLFLTPD**C**K- Isoform HMW of kininogen 1 precursor.

Sensitivity for the four targets ranged from 55.6 to 88.9%. Specificity was between 75 and 100% at the 95% confidence interval (CI). The AUC of peptide GQY**C**YELDEK, from a vitronectin precursor protein, was particularly high (0.978).

In fact, the mother proteins of four nonglycosylated tryptic peptides, AGP (41–43, 50), afamin precursor (51, 52), isoform HMW of kininogen 1 precursor (53), and a vitronectin precursor (35, 54–57), are well-characterized proteins related to HCC. Most studies have suggested that changes in the glycosylation of proteins are associated with hepatic diseases.

Recently, an antibody-lectin sandwich assay that recognizes asialo-alpha-1-acid glycoprotein (AsAGP) was developed and applied to 610 serum specimens from patients with hepatic diseases, 41 healthy donors, and 155 patients with

TABLE IV

Area under the ROC curves (AUC) of selected and monitored nonglycopeptides from normal and HCC plasma samples. Italic bold cysteine residues were carbamidomethylated

Peptide	AUC	Sensitivity (%)	Specificity (%)
GQY**C**YELDEK	0.978	88.9	100.0
TEDTIFLR	0.650	55.6	90.0
TINPAVDH**CC**K	0.723	78.6	75.0
ENFLFLTPD**C**K	0.659	64.7	80.0

nonhepatic diseases. AsAGP was validated as a marker for hepatic disease, with an AUC value of 0.919 (42). Pei *et al.* reported that afamin expression was significantly reduced in HCC tumor tissue compared with expression in normal hepa-

tocytes (51). For kininogen, the level of the fucosylated kininogen (Fc-Kin) was significantly higher in patients with HCC (53). The AUC from the combination of Fc-Kin, Golgi protein 73 (GP73), and alpha-fetoprotein (AFP) as multipanel biomarkers was 0.93, which was higher than the individual AUC values of 0.79 for Fc-Kin, 0.88 for GP73, and 0.75 for AFP. Lee *et al.* found higher expression of the *N*-linked glycopeptides from vitronectin in HCC tumor tissue and HCC patient plasma (35).

CONCLUSIONS

In this study, we developed a novel mass spectrometric approach for the discovery and validation of biomarkers in human plasma targeting specific nonglycosylated tryptic peptides in an *N*-linked glycoprotein. Based on our hypothesis, steric hindrance by glycan groups in *N*-linked glycoproteins would significantly affect the efficiency of tryptic digestion if enzymatically active amino acids were adjacent to an *N*-linked glycosylation site. This results in quantitatively different peptide products in accordance with the degree of glycosylation. We identified four nonglycosylated tryptic peptide targets using one pooled normal and ten HCC plasma samples, and verified them using ten normal and eighteen HCC plasma samples.

The current approach has significant benefits for the discovery and validation of biomarkers. In terms of sample preparation, complex, time-consuming procedures and those with low reproducibility, such as glycoprotein/glycopeptide enrichment, are not necessary, whereas they have previously been required in conventional glycoprotein research. In terms of sensitivity, because we do not target glycopeptides, but instead nonglycosylated tryptic peptides, they can readily be detected and identified by MS without PNGase F treatment. This approach also reveals differences in the degree of glycosylation in glycoproteins, which are not seen by Western blot analyses. Above all, this approach is simple and practically applicable to large numbers of complex clinical samples, such as plasma.

Ⓢ This article contains supplemental Figs. S1 to S5 and Tables S1 to S4.

** To whom correspondence should be addressed: Division of Mass Spectrometry, Korea Basic Science Institute, 804-1 Yangcheong-Ri, Ochang-Myun, Cheongwon-Gun 363-883, Republic of Korea. Tel.: +82-43-240-5150; Fax:+82-43-240-5159; E-mail: jongshin@kbsi.re.kr.

‡‡ These authors are equally contributed.

REFERENCES

1. Parodi, A. J. (2000) Protein glucosylation and its role in protein folding. *Annu. Rev. Biochem.* **69,** 69–93
2. Helenius, A., and Aebi, M. (2004) Roles of N-linked glycans in the endoplasmic reticulum. *Annu. Rev. Biochem.* **73,** 1019–1049
3. Kim, Y. J., and Varki, A. (1997) Perspectives on the significance of altered glycosylation of glycoproteins in cancer. *Glycoconj. J.* **14,** 569–576
4. Furmanek, A., and Hofsteenge, J. (2000) Protein C-mannosylation: facts and questions. *Acta Biochim. Pol.* **47,** 781–789
5. Durand, G., and Seta, N. (2000) Protein glycosylation and diseases: blood and urinary oligosaccharides as markers for diagnosis and therapeutic monitoring. *Clin. Chem.* **46,** 795–805
6. Spiro, R. G. (2002) Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology* **12,** 43R–56R
7. Hwang, H., Zhang, J., Chung, K. A., Leverenz, J. B., Zabetian, C. P., Peskind, E. R., Jankovic, J., Su, Z., Hancock, A. M., Pan, C., Montine, T. J., Pan, S., Nutt, J., Albin, R., Gearing, M., Beyer, R. P., Shi, M., and Zhang, J. (2010) Glycoproteomics in neurodegenerative diseases. *Mass Spectrom. Rev.* **29,** 79–125
8. Kim, Y. S., Kang, H. Y., Kim, J. Y., Oh, S., Kim, C. H., Ryu, C. J., Miyoshi, E., Taniguchi, N., and Ko, J. H. (2006) Identification of target proteins of N-acetylglucosaminyl transferase V in human colon cancer and implications of protein tyrosine phosphatase kappa in enhanced cancer cell migration. *Proteomics* **6,** 1187–1191
9. Kim, Y. S., Hwang, S. Y., Kang, H. Y., Sohn, H., Oh, S., Kim, J. Y., Yoo, J. S., Kim, Y. H., Kim, C. H., Jeon, J. H., Lee, J. M., Kang, H. A., Miyoshi, E., Taniguchi, N., Yoo, H. S., and Ko, J. H. (2008) 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* **7,** 1–14
10. Anderson, N. L., and Anderson, N. G. (2002) The human plasma proteome: history, character, and diagnostic prospects. *Mol. Cell. Proteomics* **1,** 845–867
11. Yang, Z., and Hancock, W. S. (2004) Approach to the comprehensive analysis of glycoproteins isolated from human serum using a multi-lectin affinity column. *J. Chromatogr. A* **1053,** 79–88
12. Yamamoto, K., Ito, S., Yasukawa, F., Konami, Y., and Matsumoto, N. (2005) Measurement of the carbohydrate-binding specificity of lectins by a multiplexed bead-based flow cytometric assay. *Anal. Biochem.* **336,** 28–38
13. Wang, Y., Wu, S. L., and Hancock, W. S. (2006) Approaches to the study of N-linked glycoproteins in human plasma using lectin affinity chromatography and nano-HPLC coupled to electrospray linear ion trap–Fourier transform mass spectrometry. *Glycobiology* **16,** 514–523
14. Zhao, J., Simeone, D. M., Heidt, D., Anderson, M. A., and Lubman, D. M. (2006) Comparative serum glycoproteomics using lectin selected sialic acid glycoproteins with mass spectrometric analysis: application to pancreatic cancer serum. *J. Proteome Res.* **5,** 1792–1802
15. Heo, S. H., Lee, S. J., Ryoo, H. M., Park, J. Y., and Cho, J. Y. (2007) Identification of putative serum glycoprotein biomarkers for human lung adenocarcinoma by multilectin affinity chromatography and LC-MS/MS. *Proteomics* **7,** 4292–4302
16. Kim, Y. S., Son, O. L., Lee, J. Y., Kim, S. H., Oh, S., Lee, Y. S., Kim, C. H., Yoo, J. S., Lee, J. H., Miyoshi, E., Taniguchi, N., Hanash, S. M., Yoo, H. S., and Ko, J. H. (2008) Lectin precipitation using phytohemagglutinin-L(4) coupled to avidin-agarose for serological biomarker discovery in colorectal cancer. *Proteomics* **8,** 3229–3235
17. Abbott, K. L., Aoki, K., Lim, J. M., Porterfield, M., Johnson, R., O'Regan, R. M., Wells, L., Tiemeyer, M., and Pierce, M. (2008) Targeted glycoproteomic identification of biomarkers for human breast carcinoma. *J. Proteome Res.* **7,** 1470–1480
18. Ueda, K., Takami, S., Saichi, N., Daigo, Y., Ishikawa, N., Kohno, N., Katsumata, M., Yamane, A., Ota, M., Sato, T. A., Nakamura, Y., and Nakagawa, H. (2010) Development of serum glycoproteomic profiling technique; simultaneous identification of glycosylation sites and site-specific quantification of glycan structure changes. *Mol. Cell. Proteomics* **9,** 1819–1828
19. Zhang, H., Li, X. J., Martin, D. B., and Aebersold, R. (2003) Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat. Biotechnol.* **21,** 660–666
20. Liu, T., Qian, W. J., Gritsenko, M. A., Camp, D. G., 2nd, Monroe, M. E., Moore, R. J., and Smith, R. D. (2005) Human plasma N-glycoproteome analysis by immunoaffinity subtraction, hydrazide chemistry, and mass spectrometry. *J. Proteome Res.* **4,** 2070–2080

21. Hägglund, P., Bunkenborg, J., Elortza, F., Jensen, O. N., and Roepstorff, P. (2004) 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.* **3,** 556–566

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

23. Zou, Z., Ibisate, M., Zhou, Y., Aebersold, R., Xia, Y., and Zhang, H. (2008) Synthesis and evaluation of superparamagnetic silica particles for extraction of glycopeptides in the microtiter plate format. *Anal. Chem.* **80,** 1228–1234

24. Larsen, M. R., Jensen, S. S., Jakobsen, L. A., and Heegaard, N. H. (2007) Exploring the sialiome using titanium dioxide chromatography and mass spectrometry. *Mol. Cell. Proteomics* **6,** 1778–1787

25. An, H. J., Peavy, T. R., Hedrick, J. L., and Lebrilla, C. B. (2003) Determination of N-glycosylation sites and site heterogeneity in glycoproteins. *Anal. Chem.* **75,** 5628–5637

26. Wuhrer, M., Koeleman, C. A., and Deelder, A. M. (2009) Two-dimensional HPLC separation with reverse-phase-nano-LC-MS/MS for the characterization of glycan pools after labeling with 2-aminobenzamide. *Methods Mol. Biol.* **534,** 79–91

27. North, S. J., Hitchen, P. G., Haslam, S. M., and Dell, A. (2009) Mass spectrometry in the analysis of N-linked and O-linked glycans. *Curr. Opin. Struct. Biol.* **19,** 498–506

28. McDonald, C. A., Yang, J. Y., Marathe, V., Yen, T. Y., and Macher, B. A. (2009) Combining results from lectin affinity chromatography and glycocapture approaches substantially improves the coverage of the glycoproteome. *Mol. Cell. Proteomics* **8,** 287–301

29. Lewandrowski, U., Moebius, J., Walter, U., and Sickmann, A. (2006) Elucidation of N-glycosylation sites on human platelet proteins: a glycoproteomic approach. *Mol. Cell. Proteomics* **5,** 226–233

30. Wollscheid, B., Bausch-Fluck, D., Henderson, C., O'Brien, R., Bibel, M., Schiess, R., Aebersold, R., and Watts, J. D. (2009) Mass-spectrometric identification and relative quantification of N-linked cell surface glycoproteins. *Nat. Biotechnol.* **27,** 378–386

31. Wang, Y., Ao, X., Vuong, H., Konanur, M., Miller, F. R., Goodison, S., and Lubman, D. M. (2008) Membrane glycoproteins associated with breast tumor cell progression identified by a lectin affinity approach. *J. Proteome Res.* **7,** 4313–4325

32. Lee, A., Kolarich, D., Haynes, P. A., Jensen, P. H., Baker, M. S., and Packer, N. H. (2009) Rat liver membrane glycoproteome: enrichment by phase partitioning and glycoprotein capture. *J. Proteome Res.* **8,** 770–781

33. Whiteaker, J. R., Zhao, L., Anderson, L., and Paulovich, A. G. (2010) An automated and multiplexed method for high throughput peptide immunoaffinity enrichment and multiple reaction monitoring mass spectrometry-based quantification of protein biomarkers. *Mol. Cell. Proteomics* **9,** 184–196

34. de Leoz, M. L., Y. L. J., An, H. J., Kronewitter, S. R., Kim, J., Miyamoto, S., Borowsky, A. D., Chew, H. K., and Lebrilla, C. B. (2010) High-mannose glycans are elevated during breast cancer progression. *Mol. Cell. Proteomics*, DOI:10.1074/mcp.M110.002717, in press

35. Lee, H. J., Na, K., Choi, E. Y., Kim, K. S., Kim, H., and Paik, Y. K. (2010) Simple method for quantitative analysis of N-linked glycoproteins in hepatocellular carcinoma specimens. *J. Proteome Res.* **9,** 308–318

36. Drake, P. M., Cho, W., Li, B., Prakobphol, A., Johansen, E., Anderson, N. L., Regnier, F. E., Gibson, B. W., and Fisher, S. J. (2010) Sweetening the pot: adding glycosylation to the biomarker discovery equation. *Clin. Chem.* **56,** 223–236

37. Silva, J. C., Denny, R., Dorschel, C. A., Gorenstein, M., Kass, I. J., Li, G. Z., McKenna, T., Nold, M. J., Richardson, K., Young, P., and Geromanos, S. (2005) Quantitative proteomic analysis by accurate mass retention time pairs. *Anal. Chem.* **77,** 2187–2200

38. Silva, J. C., Gorenstein, M. V., Li, G. Z., Vissers, J. P., and Geromanos, S. J. (2006) Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition. *Mol. Cell. Proteomics* **5,** 144–156

39. Barnidge, D. R., Dratz, E. A., Martin, T., Bonilla, L. E., Moran, L. B., and Lindall, A. (2003) Absolute quantification of the G protein-coupled receptor rhodopsin by LC/MS/MS using proteolysis product peptides and synthetic peptide standards. *Anal. Chem.* **75,** 445–451

40. Bunkenborg, J., Pilch, B. J., Podtelejnikov, A. V., and Wiśniewski, J. R. (2004) Screening for N-glycosylated proteins by liquid chromatography mass spectrometry. *Proteomics* **4,** 454–465

41. Lee, E. Y., Kang, J. H., Kim, K. A., Chung, T. W., Kim, H. J., Yoon, D. Y., Lee, H. G., Kwon, D. H., Kim, J. W., Kim, C. H., and Song, E. Y. (2006) Development of a rapid, immunochromatographic strip test for serum asialo alpha1-acid glycoprotein in patients with hepatic disease. *J. Immunol. Methods* **308,** 116–123

42. Kim, K. A., Lee, E. Y., Kang, J. H., Lee, H. G., Kim, J. W., Kwon, D. H., Jang, Y. J., Yeom, Y. I., Chung, T. W., Kim, Y. D., Yoon do, Y., and Song, E. Y. (2006) Diagnostic accuracy of serum asialo-alpha1-acid glycoprotein concentration for the differential diagnosis of liver cirrhosis and hepatocellular carcinoma. *Clin. Chim. Acta* **369,** 46–51

43. Fabris, C., Pirisi, M., Soardo, G., Toniutto, P., Falleti, E., Vitulli, D., Pezzetta, F., Gonano, F., and Bartoli, E. (1996) Diagnostic usefulness of acute-phase protein measurement in hepatocellular carcinoma. *Cancer Invest.* **14,** 103–108

44. Anderson, L., and Hunter, C. L. (2006) Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. *Mol. Cell. Proteomics* **5,** 573–588

45. Stahl-Zeng, J., Lange, V., Ossola, R., Eckhardt, K., Krek, W., Aebersold, R., and Domon, B. (2007) High sensitivity detection of plasma proteins by multiple reaction monitoring of N-glycosites. *Mol. Cell. Proteomics* **6,** 1809–1817

46. Shin, B., Jung, H. J., Hyung, S. W., Kim, H., Lee, D., Lee, C., Yu, M. H., and Lee, S. W. (2008) Postexperiment monoisotopic mass filtering and refinement (PE-MMR) of tandem mass spectrometric data increases accuracy of peptide identification in LC/MS/MS. *Mol. Cell. Proteomics* **7,** 1124–1134

47. Fortin, T., Salvador, A., Charrier, J. P., Lenz, C., Lacoux, X., Morla, A., Choquet-Kastylevsky, G., and Lemoine, J. (2009) Clinical quantitation of prostate-specific antigen biomarker in the low nanogram/milliliter range by conventional bore liquid chromatography-tandem mass spectrometry (multiple reaction monitoring) coupling and correlation with ELISA tests. *Mol. Cell. Proteomics* **8,** 1006–1015

48. Kurogochi, M., Matsushista, T., Amano, M., Furukawa, J., Shinohara, Y., Aoshima, M., and Nishimura, S. (2010) Sialic Acid-focused Quantitative Mouse Serum Glycoproteomics by Multiple Reaction Monitoring Assay. *Mol. Cell. Proteomics* **9,** 2354–2368

49. Kim, K., Kim, S. J., Yu, H. G., Yu, J., Park, K. S., Jang, I. J., and Kim, Y. (2010) Verification of biomarkers for diabetic retinopathy by multiple reaction monitoring. *J. Proteome Res.* **9,** 689–699

50. Bachtiar, I., Santoso, J. M., Atmanegara, B., Gani, R. A., Hasan, I., Lesmana, L. A., Sulaiman, A., Gu, J., and Tai, S. (2009) Combination of alpha-1-acid glycoprotein and alpha-fetoprotein as an improved diagnostic tool for hepatocellular carcinoma. *Clin. Chim. Acta.* **399,** 97–101

51. Wu, G. X., Lin, Y. M., Zhou, T. H., Gao, H., and Pei, G. (2000) Significant down-regulation of alpha-albumin in human hepatoma and its implication. *Cancer Lett.* **160,** 229–236

52. Noel, E. S., Reis, M. D., Arain, Z., and Ober, E. A. (2010) Analysis of the Albumin/alpha-Fetoprotein/Afamin/Group specific component gene family in the context of zebrafish liver differentiation. *Gene Expr. Patterns* **10,** 237–243

53. Wang, M., Long, R. E., Comunale, M. A., Junaidi, O., Marrero, J., Di Bisceglie, A. M., Block, T. M., and Mehta, A. S. (2009) Novel fucosylated biomarkers for the early detection of hepatocellular carcinoma. *Cancer Epidemiol. Biomarkers Prev.* **18,** 1914–1921

54. Ogawa, H., Yoneda, A., Seno, N., Hayashi, M., Ishizuka, I., Hase, S., and Matsumoto, I. (1995) Structures of the N-linked oligosaccharides on human plasma vitronectin. *Eur. J. Biochem.* **230,** 994–1000

55. Edwards, S., Lalor, P. F., Tuncer, C., and Adams, D. H. (2006) Vitronectin in human hepatic tumours contributes to the recruitment of lymphocytes in an alpha v beta3-independent manner. *Br. J. Cancer* **95,** 1545–1554

56. Zhao, Y., Jia, W., Sun, W., Jin, W., Guo, L., Wei, J., Ying, W., Zhang, Y., Xie, Y., Jiang, Y., He, F., and Qian, X. (2010) Combination of improved (18)O incorporation and multiple reaction monitoring: a universal strategy for absolute quantitative verification of serum candidate biomarkers of liver cancer. *J. Proteome Res.* **9,** 3319–3327

57. Paradis, V., Degos, F., Dargère, D., Pham, N., Belghiti, J., Degott, C., Janeau, J. L., Bezeaud, A., Delforge, D., Cubizolles, M., Laurendeau, I., and Bedossa, P. (2005) Identification of a new marker of hepatocellular carcinoma by serum protein profiling of patients with chronic liver diseases. *Hepatology* **41,** 40–47

58. Hortin, G. L., Sviridov, D., and Anderson, N. L. (2008) High-abundance polypeptides of the human plasma proteome comprising the top 4 logs of polypeptide abundance. *Clin. Chem.* **54,** 1608–1616