ORIGINAL PAPER

A lectin-coupled, multiple reaction monitoring based quantitative analysis of human plasma glycoproteins by mass spectrometry

Yeong Hee Ahn · Park Min Shin · Eun Sun Ji · Hoguen Kim & Jong Shin Yoo

Received: 23 September 2011 /Revised: 1 December 2011 /Accepted: 7 December 2012 / Published online: 13 January 2012 \circ Springer-Verlag 2012

Abstract Aberrant protein glycosylation may be closely associated with cancer pathology. To measure the abundance of protein glycoforms with a specific glycan structure in plasma samples, we developed a lectin-coupled multiple reaction monitoring (MRM)-based mass spectrometric method. It was confirmed that the method could provide reproducible results with precision sufficient to distinguish differences in the abundance of protein glycoforms between individuals. Plasma samples prepared from hepatocellular carcinoma (HCC) patients without immuno-depletion of highly abundant plasma proteins were fractionated by use of fucose-specific aleuria aurantia lectin (AAL) immobilized on magnetic beads by use of a biotin–streptavidin conjugate. The lectin-captured fractions were digested by trypsin and profiled by tandem mass spectrometry. From the proteomic profiling data, target glycoproteins were selected and analyzed quantitatively by MRM-based analysis. The reproducibility of MRM-based quantification of the selected target proteins was reliable, with precision (CV; \leq 14% for batch-to-batch replicates and \leq 19% for replicates over three days) sufficient to distinguish differences in the abundance of AAL-captured glycoforms between individual plasma samples. This lectin-coupled, MRM-based method, measuring only lectin-captured glycoforms of a target

Electronic supplementary material The online version of this article (doi:[10.1007/s00216-011-5646-3\)](http://dx.doi.org/10.1007/s00216-011-5646-3) contains supplementary material, which is available to authorized users.

Y. H. Ahn \cdot P. M. Shin \cdot E. S. Ji \cdot J. S. Yoo (\boxtimes) Division of Mass Spectrometry, Korea Basic Science Institute, 804-1 Yangcheong-Ri, Ochang-Eup, Cheongwon-Gun 363-883, Republic of Korea e-mail: jongshin@kbsi.re.kr

H. Kim

Department of Pathology, Yonsei Medical Center, Yonsei University, Seoul 120-752, Republic of Korea

protein rather than total target protein, is a tool for monitoring differences between individuals by measuring the abundance of aberrant glycoforms of a target protein related to a disease. This method may be further applied to rapid verification of biomarker candidates involved in aberrant protein glycosylation in human plasma.

Keywords Plasma . Glycoprotein . Lectin . Multiple reaction monitoring . Mass spectrometry . Cancer

Abbreviations

Introduction

Protein glycosylation has been reported to be important in adhesion, metastasis, and signaling through cell-to-cell interactions. Furthermore, abnormal glycosylation patterns, for example, increased glycan size and extra branching of glycan chains with over-sialylation and fucosylation, are closely associated with cancer progression [\[1](#page-10-0)–[5\]](#page-10-0). Thus, efficient differentiation of these aberrant protein glycosylation patterns and of quantitative levels between healthy and individuals with malignant cells may be useful for understanding the pathological mechanism of cancer and for identifying

specific cancer biomarkers [[6](#page-10-0)–[9](#page-10-0)]. Neoglycoforms aberrantly glycosylated in the malignant transformation of cells can be secreted from the cells, via conventional secretary pathways, or shed from the cell membrane as a result of enhanced protease activity, and reach the bloodstream, reflecting abnormal states of the malignant cells.

Neoglycoforms present in the bloodstream are attractive as potent candidates for serological cancer biomarkers [[10](#page-10-0)–[13](#page-10-0)]; thus, an efficient quantitative assay for neoglycoforms would be valuable. To measure the abundance of these neoglycoforms, many obstacles must be overcome, for example severe masking by highly abundant blood proteins, low substoichiometric abundance of each aberrant glycoform of a glycoprotein, because of glycan microheterogeneity, and difficulties in identifying and quantifying aberrant glycoforms because of the complex nature of the glycan structure.

Recent advances in proteomics, including glycoproteinenrichment techniques, using various lectin or hydrazide beads, and mass spectrometric techniques for efficient analysis of complex tryptic digests, have enabled rapid progress in analytical and functional studies of glycoproteins [\[14](#page-10-0)–[20](#page-11-0)]. A variety of lectins have been used to capture a selective lectin-active glycoform from a complex glycoproteome before subsequent analysis, for example by immunoassay and mass spectrometry [\[14](#page-10-0)–[18](#page-11-0)]. Because a lectin with binding affinity for a glycoform of a specific glycan structure, in general, captures a relatively low portion of the total mass of the blood proteome, a negative masking effect in subsequent mass spectrometry, because of highly abundant proteins that do not have affinity for the lectin, can be removed. A variety of lectins with narrow specificity for a specific glycan structure can also act as good selectors for enriching cancer-related aberrant glycoforms that may be further developed as cancer biomarkers [[17,](#page-11-0) [21](#page-11-0)].

A typical example is fucose-specific lectin, which has been commonly used in the discovery of cancer biomarkers, especially those targeting hepatocellular carcinoma (HCC). Among many glycoproteins known as potential serological biomarkers for HCC, alpha-fetoprotein (AFP) is known as a potent biomarker for HCC [\[9](#page-10-0)]. In particular, AFP-L3, a glycoform glycosylated aberrantly by fucosylation that has increased affinity for fucose-specific lectins such as lens culinaris agglutinin (LCA), has been reported to have improved specificity as a biomarker for HCC [[22](#page-11-0)–[24\]](#page-11-0).

Quantification of these lectin-enriched aberrant glycoforms related to cancer has primarily been achieved by use of an immunoassay approach, for example the lectin–antibody sandwich array technique [\[25](#page-11-0)–[27\]](#page-11-0). In this immunoassay approach to quantification, development of an antibody to target the glycoprotein to be quantified is required as a prerequisite; this is a costly and tedious process. Furthermore, the number of antibody reagents that can be developed for an immunoassay is limited compared with the number of new glycoproteins

found in a variety of biomarker discovery experiments, and development of some antibodies for a target antigen may be impossible. To overcome these limitations and to conduct rapid preliminary verification of potential biomarker candidates, a more general analytical method featuring a simple, rapid, cost-effective, and multiplexing analytical system would be an attractive alternative.

Mass spectrometry is an alternative tool capable of identifying new target proteins and measuring the abundance of the newly identified serological proteins that need to be validated as viable biomarker candidates [\[28](#page-11-0)–[34](#page-11-0)]. We have previously reported that the high sensitivity and precision of multiple reaction monitoring (MRM)-based techniques are useful for measuring the abundance of aberrant glycoforms of target glycoproteins between subproteome samples obtained by comparative lectin capturing. However, in the study, the lectincoupled, MRM-based analysis was conducted with lectinenriched secretomes obtained from culture media of colorectal cancer cell lines rather than from plasma samples [\[17\]](#page-11-0).

In addition to MRM quantification for cell line samples that were lectin-enriched under well-restricted conditions, MRM quantification for samples that are lectin-enriched from human plasma, with high sample complexity, a wide dynamic range in protein abundance, and dramatic variation between individuals, must be developed. Also necessary are robust and welldesigned sample manipulation and quantification processes capable of lectin-capture of target glycoforms with a specific glycan structure and suitable for MRM mass analysis of tryptic digests of the lectin-captured subproteome. Reproducible quantification of protein glycoforms aberrantly glycosylated because of a variety of pathological processes can be applied to the diagnosis and monitoring of diseases that involve those biological processes. Here, we introduce a lectincoupled, MRM-based quantification method using lectin fractionation of human plasma glycoproteins by aleuria aurantia lectin (AAL) [\[35,](#page-11-0) [36\]](#page-11-0) and MRM mass analysis of in-solution tryptic digests of the fractionated glycoforms. In contrast with some earlier approaches, for example lectin-coupled proteomic profiling [\[5](#page-10-0), [14](#page-10-0), [18](#page-11-0), [21,](#page-11-0) [37\]](#page-11-0) and isotope-labeling quantification [[21](#page-11-0)], often used for biomarker discovery, this lectincoupled, MRM-based quantification method can be especially useful for rapid verification of biomarker candidates involved in aberrant protein glycosylation. We demonstrate a reproducible analytical method for measuring the abundance of the AAL-captured target plasma glycoproteins.

Experimental

Plasma sample preparation and lectin-capture

Plasma samples were obtained from normal individuals and HCC patients at Severance Medical Center at Yonsei University

(Seoul, Korea), with agreement from patients for the research use of the blood samples. K₂EDTA-treated plasma (7.5 μ L) was diluted fivefold with 50 mmol L^{-1} ammonium bicarbonate, and the protein concentration was quantified separately by use of Quant-iT protein assay kits (Invitrogen). The diluted plasma samples were spiked with $CaCl₂$ and $MgCl₂$ solutions at concentrations of 5 mmol L^{-1} , and the plasma solutions were then mixed with lectin-immobilized magnetic beads (400 μg) and incubated overnight at 10 °C with gentle shaking. The immobilized lectin was prepared in advance by mixing biotinylated AAL (Vector Lab) with streptavidin–magnetic beads (Invitrogen) for 1 h at room temperature, in accordance with the manufacturer's instructions; just before protein-binding, the mixture was washed extensively with phosphate-buffered saline (PBS). The lectin-bound proteins were washed three times with PBS and eluted from the bead complexes by adding an elution buffer (2 mol L^{-1} urea, 0.5 mmol L^{-1} DTT, 50 mmol L^{-1} ammonium biocarbonate). After cysteine blocking with iodoacetamide for 30 min in the dark, the eluted proteins diluted with 50 mmol L^{-1} ammonium bicarbonate were digested overnight at 37 °C (1 mmol L^{-1} CaCl₂, 1.2 μmol L^{-1} trypsin). The trypsin reaction was stopped by addition of 1% formic acid. For MRM quantification experiments, stable isotope-coded internal standards of the target peptides to be quantified were spiked. Tryptic digests were desalted using a microspin column (Harvard apparatus), lyophilized in a SpeedVac system, and reconstituted with 0.1% formic acid for mass analysis.

LC–MS–MS analysis

The prepared tryptic peptides were analyzed, to profile the lectin-captured glycoproteins, by use of a Surveyor highperformance liquid chromatography (HPLC) system (Thermo Finnigan, Austin, TX, USA) and an LTQ-FT mass spectrometer (Thermo Finnigan) equipped with a nano-electrospray ionization (ESI) source. The reconstituted tryptic digests were loaded on to a C₁₈ trap column (5 μm, 300 μm i.d. \times 5 mm; LC Packings, Amsterdam, The Netherlands) by use of an autosampler and washed with an aqueous solution of 0.5% acetic acid and 0.02% formic acid (mobile phase A; see below) for 10 min at 20 μ L min⁻¹. The peptides were transferred on-line on to a homemade C_{18} column (Aqua, 3 μ m, 75 μm i.d. \times 150 mm; Phenomenex, Torrance, CA, USA), eluted with a gradient, and directly electrosprayed into an LTQ-FT mass spectrometer controlled by Xcalibur software (Thermo Electron, Waltham, MA, USA). Mobile phases A and B were water and water containing 80% acetonitrile, respectively, each containing 0.5% acetic acid and 0.02% formic acid. The gradient started with 5% B for 15 min and the amount of B was then increased to 20% for 3 min, 50% for 47 min, and 95% for 2 min; it was held at 95% for 5 min and then at 5% for 2 min. Before each run, the column was equilibrated with 5% B for 8 min.

By use of a nano-ESI LTQ-FT mass spectrometer, one full-scan mass spectrum was acquired in the m/z range 400–2000, and three ion-trap MS–MS spectra were acquired per data-dependent cycle. The mass analysis method used an ion-spray potential of 2.2 kV, a capillary temperature of 220 °C, 1 FTMS full micro scan with 200 ms FTMS full max ion time, and three ion-trap MS– MS micro scans with 50 ms ion-trap MS–MS max ion time. Dynamic exclusion settings for MS–MS were also used. The collision-induced dissociation (CID) MS–MS was conducted with 35% normalized collision energy, 3000 counts minimum signal threshold, an isolation

2 Springer

width of 2, 0.25 activation Q, and 30 ms activation time for MS–MS acquisitions.

MS–MS data analysis and target peptide selection

Protein profiling from tandem mass data for lectin-captured plasma samples was conducted using the Mascot search engine (ver. 2.2.04, Matrix Science, Boston, MA, USA) against the Swiss-Prot 51.6 database confined to Homo sapiens taxonomy, with a decoy database. Raw LTQ-FT tandem mass spectra were converted into mzXML (.XML) format by use of Bioworks software (ThermoElectron). A Mascot search was performed with options of monoisotopic mass, a significance threshold $p<0.05$, a peptide mass tolerance of ± 0.2 Da, and a fragment mass tolerance ± 0.4 Da. Trypsin was selected as the enzyme, considering three potentially missed cleavages and a possible cleavage at the K/R-P amide bond. Oxidized methionine was selected for variable modifications. For selection of the identified proteins, we considered only the top-scoring peptides for each peptide query.

Target glycoproteins for MRM mass quantification were selected among the lectin-captured proteins profiled by the Mascot search, with an option noted as the N-linked glycoprotein in the Swiss-Prot 51.6 database. Also, although not identified in this proteomic profiling experiment, alphafetoprotein, known to be fucosylated in the plasma of HCC patients, was also included [[9](#page-10-0), [22](#page-11-0)–[24\]](#page-11-0). A group of tryptic target peptides, surrogates of the selected target glycoproteins for MRM mass quantification, was constructed according to the following elimination criteria:

- 1. The peptide should not contain Met, Cys, or modified residues (noted in the Swiss-Prot 51.6 database), for example glycosylation, phosphorylation, acetylation, and natural variants within the peptide sequence. The mass of a peptide modified at these residues is different to that of the unmodified peptide. For example, methionine oxidation of a peptide containing the methionine residue results in a mass increase (+16 Da) for the peptide oxidized. Therefore the using this peptide as a target peptide may not reflect correctly the real abundance of the lectin-captured target protein.
- 2. Peptides containing a peptide bond K/R-P were not considered, because partial tryptic hydrolysis at the peptide bond is often observed in tandem mass spectrometry.

The following acceptance criteria were considered:

- 1. based on the proteomic profiling experiments, peptides identified with the best signal intensity and symmetrical chromatographic peak with a mass range greater than 900 Da and less than 2000 Da were selected; and
- 2. in the absence of a properly identified peptide that fit the above criteria, frequently detected peptides registered in

the Peptide Atlas Database or the PRIDE database were selected.

MRM mass analysis

Using MS–MS spectra of the synthetic target peptides and their stable isotope-coded counterparts (Anygene, Korea), three transition channels for each of the target peptides were set for MRM mass quantification. The prepared tryptic digests were analyzed with a nano ultra performance liquid chromatography (nano UPLC) system (Waters, UK) coupled to a TSQ Quantum Ultra (Thermo Finnigan) mass spectrometer equipped with a nano-ESI source. The tryptic digests were loaded on to a trap column (5 μm symmetry C₁₈, 180 μm i.d. × 20 mm; Waters, UK) by use of an autosampler and washed with water for 5 min at 10 μL min−¹ . Peptides were transferred on-line to an

a) From the total plasma proteome

run 1 run 2 run 3 total

b) From the AAL-captured subproteome of plasma

run 1 run 2 run 3 total

Fig. 2 Venn diagrams for peptides identified by 1D LC–ESI–MS–MS analysis of tryptic digests obtained from the total and AAL-captured plasma proteome. Each 1 μg of tryptic digest was injected in triplicate for tandem mass analysis. From the total plasma samples, 33 peptides of 60 identified peptides were repeatedly identified at least twice, and from AAL-captured samples, 61 peptides of 101 identified peptides were identified at least twice

Table 1 Glycoproteins identified at least twice by each triplicate tandem mass analysis for the total and AAL-captured plasma proteome

Swiss-Prot accession no.	gene name	protein name	peptide sequence ^a	monoisotopic mass	Q ₁ $(2+, m/z)$	$C.E^b$	$Q3^c$ (m/z)	$Q3^c$ (m/z)	$Q3^c$ (m/z)	LOQ ^d (fmol)
PO2763	ORM1	Alpha-acid glycoprotein 1	TEDTIFLR TEDTIFL*R	993.5 1000.5	497.7 501.2	19.0 19.0	548.4 555.4	649.4 656.4	764.4 3 771.4	
PO1011	SERPINA3	Alpha-1-antichymotrypsin	NLAVSOVVHK NLAVSOVV*HK	1093.6 1099.6	547.8 550.8	21.0 21.0	697.4 703.4	796.5 802.5	$867.5 \quad 3$ 873.5	
P01009	SERPINA1	Alpha-1-antitrypsin	SVLGQLGITK SVLGOL*GITK	1014.6 1021.6	508.3 511.8	17.0 17.0	531.4	716.5 538.4 723.5	829.5 30 836.5	
P10909	CLU	Clusterin	ELDESLOVAER ELDESLQV*AER	1287.6 1293.6	644.8 647.8	22.0 22.0	715.4 721.4	802.4 808.4	1046.5 30 1052.5	
P00738	HP	Haptoglobin	VGYVSGWGR VGYV*SGWGR	979.5 985.5	490.8 493.8	18.0 18.0	661.3 667.3	824.4 830.4	881.4 3 887.4	
P01876	1GHA1	Ig alpha-1 chain C region	TPLTATLSK TPLTATL*SK	930.5 937.5	466.3 469.8	19.0 19.0	519.3 526.3	620.4 627.4	733.5 3 740.5	
P01871	IGHM	Ig mu chain C region	NVPLPVIAELPPK NVPLPVIAEL*PPK	1385.8 1392.8	693.9 697.4	22.0 22.0	767.5 774.5	963.6 970.6	1173.7 1180.7	-30
P02771	AFP	Alpha-fetoprotein ^{e)}	DFNOFSSGEK DFNOF*SSGEK	1157.5 1167.5	579.7 584.8	20.0 20.0	654.3 664.3	782.4 792.4	896.4 30 906.4	

Table 2 Target glycoproteins selected from the AAL-captured protein list obtained by the proteomic profiling experiment to quantify fucosylated glycoforms

a Asterisk indicates stable isotope-coded sites

^b CE, collision energy

^c Bold letters show a transition selected for quantifying each target peptide

 d Based on concentrations measured with \langle CV 15% for triplicate MRM quantification

^e Alpha-fetoprotein (AFP), known to be fucosylated in plasma of HCC patients, was also included, although not identified in this proteomic profiling experiment

analytical column (1.7 μm BEH130 C₁₈ UPLC column, 75 μm i.d. \times 150 mm; Waters, UK), eluted under a gradient, and

Fig. 3 Linear concentration ranges and limits of quantification obtained by MRM mass analysis of the target peptides. Each MRM analysis was conducted with an equal amount of stable isotope-coded internal standard (heavy, 300 fmol) and a different amount of synthetic target peptide (light, from 1.0 fmol to 3,000 fmol). Analyses at concentrations with CV <15% in triplicate MRM mass measurements for each target peptide are displayed. Each spot represents the average value obtained by triplicate MRM mass analysis. The *dotted line* is a theoretical line

directly electrosprayed into a TSQ Quantum ultra mass spectrometer. Mobile phases A and B were 100% water and 100% acetonitrile, respectively, each containing 0.1% formic acid. The gradient started at 3% B for 0.33 min and was then increased to 35% B for 40 min, 50% B for 50 min, and 95% B for 50.5 min; it was held at 95% B for 60 min and then reduced to 3% B for 60.5 min. The column was equilibrated with 3% B for 14.5 min between each run. MRM analysis was conducted at an ion-spray potential of 1.7 kV, a capillary temperature of 270 °C, and an inclusion window of ± 1 Da for each precursor ion and for the fragment ions. MRM analyses were conducted in triplicate for each sample.

Results and discussion

Proteomic profiling for the AAL-fractionated and the total plasma sample

AAL was used to capture fucosylated glycoforms of plasma glycoproteins from human plasma from an HCC patient. AAL has binding affinity for fucose moieties linked to GlcNAc on asparagine residues of proteins. Fucosylated glycoproteins captured by magnetic bead (MB)-immobilized AAL were denatured, alkylated, and digested insolution with trypsin (Fig. [1\)](#page-2-0). To immobilize biotinylated

^a Represents averaged value calculated from three quantification values obtained in three batch experiments, independently conducted from sample preparation by lectin to MRM-based analysis in a day

^bCalculated on the basis of three quantification values obtained from three batch experiments in a day

AAL on beads we used streptavidin-conjugated MB. Thus, MB-immobilized AAL–biotin–streptavidin conjugates were used to capture fucosylated glycoproteins from plasma samples. Profiling patterns of a proteome sample obtained by ESI– MS–MS analysis can be affected by several analytical processes, for example sample preparation, tryptic digestion, chromatographic separation, and mass acquisition. Furthermore, a wide dynamic range in the abundance of plasma proteins can lead to substantial variability in the proteomic profiling results. Thus, tryptic digests (1 μg) of the total plasma proteome and the AAL-captured subproteome of an HCC patient were analyzed in triplicate (Electronic Supplementary Material Table S1). Among the proteins identified, proteins that were not annotated in the Swiss-Prot database to undergo N or Oglycosylation were also identified; these might have originated from proteins not yet included in the Swiss-Prot database, nonspecific binding to the lectin, or protein–protein interactions with a lectin-binding protein. Non-specific binding or nonglycoprotein capture via protein–protein interactions during lectin-affinity enrichment have been reported by several groups [\[21,](#page-11-0) [38](#page-11-0), [39](#page-11-0)]. Nonetheless, as shown in Electronic Supplementary Material Table S1, the proteomic profiling for each sample, the total plasma proteome, and the lectin-captured subproteome were highly reproducible in triplicate MS–MS runs. Thus the variation between the AAL-captured and the total proteome sample in the identified protein list is much more dramatic than that between each MS–MS run of each sample. Figure [2](#page-3-0) (summarized from the profiling results of Electronic Supplementary Material Table S1) shows the number of peptides identified in each run for the total plasma proteome samples and subproteome samples that were lectin-captured. More than half of all peptides identified from triplicate runs (61% for the lectincaptured plasma and 54% for total plasma) were repeatedly identified at least twice in both samples. Because AAL-capture from the plasma proteome has been conducted without depleting highly abundant serological proteins, it is assumed that the lectin-captured sample still has a wide dynamic range for the abundance of captured proteins. Additionally, because we analyzed the tryptic digests by one-dimensional (1D) LC– MS–MS analysis, the number of profiled proteins was low. Nonetheless, the number of glycoproteins (101 proteins) identified from triplicate profiling experiments of the lectincaptured subproteome was greater than that (61 proteins) of the total proteome. The masking effect of highly abundant plasma proteins in 1D LC–MS–MS analysis was removed by lectin-capture; thus, many peptides from less abundant proteins were identified.

Also, the benefits of lectin-capture were obvious because of the difference in the profiled protein lists between the two samples (Table [1](#page-4-0)). Among the proteins identified by the Mascot search for raw mass data obtained by the 1D LC– MS–MS analysis, glycoproteins noted on the Swiss-Prot DB with N or O-glycosylation sites and identified at least twice through the tandem mass analysis are summarized in Table [1.](#page-4-0) Among the 40 glycoproteins identified in the AAL-captured sample, 16 glycoproteins (40%) were newly identified after lectin-capture, whereas 24 glycoproteins (60%) were identified in both samples. In contrast, four glycoproteins observed in the total plasma sample, including Ig gamma-4 and serotransferrin, were not identified in the lectin-captured sample. This dramatic change in the lists of profiled glycoproteins after lectin-capture indicates that lectin-capture worked efficiently and is a powerful tool for extracting a subproteome with a specific glycan moiety from a total proteome sample.

Target peptide selection for quantitative analysis and calibration

Target glycoproteins used to quantify AAL-captured protein glycoforms were selected from the protein list obtained from

Table 4 MRM-based analysis of fucosylated target proteins AAL-captured from eight individual plasma samples⁹⁾

the proteomic profiling experiment, with an option noted as the N-linked glycoprotein in the Swiss-Prot database (Table [2\)](#page-5-0). On the other hand, AFP has been reported to be fucosylated in the plasma of HCC patients and to be captured by AAL [\[22](#page-11-0)–[24\]](#page-11-0). Unfortunately AFP was not identified in our proteomic profiling by 1D LC–MS–MS analysis, possibly because the abundance of the protein was too low and because of the high complexity of the AAL-captured fraction. Nonetheless, AFP was also selected as a target protein to evaluate the sensitivity of the lectin-coupled, MRM-based method of quantification. For MRM-based quantification of the selected target proteins, a mixture of tryptic target peptides, surrogates for the selected target proteins, was constructed by following the restricted criteria discussed in the "Experimental" section. Linear concentration ranges for the selected target peptides were investigated for five concentrations from 1.0 fmol to 3 pmol, prepared by adding the target peptides (light) at different concentrations to tryptic peptide matrix solutions with internal standards (heavy) of each target peptide, each at a concentration of 300 fmol. The peptide matrix solution was prepared by tryptic digestion of a standard protein mixture containing serum albumin (bovine), α-casein (bovine), β casein (bovine), and ovalbumin (chicken). Concentrations of each target peptide with a signal intensity 10 times greater than the noise in the chromatograms extracted for selected transitions and with acceptable CV values ($CV \leq 15\%$) during triplicate MRM quantifications were accepted tentatively as limits of quantification (LOQ) for the target peptides (Table [2\)](#page-5-0). The extracted chromatograms for each target peptide at the LOQ concentration are summarized in Electronic Supplementary Material Fig. S1. For most target peptides good concentration linearity was obtained in the concentration range above the LOQ level (Fig. [3\)](#page-5-0).

Reproducibility in the lectin-coupled, MRM-based analysis

The objectives of the study were to develop an analytical method for reproducibly quantifying lectin-captured protein glycoforms with a specific glycan structure and for reliably distinguishing differences in the abundance of the lectincaptured glycoforms between individuals. For this, we prepared a pooled liver-disease plasma sample by mixing 10 plasma samples from each of HCC, cholangiocarcinoma, and hepatitis-B patients, to obtain the amount of control human plasma sample necessary to optimize the MRMbased method for quantification of lectin-captured target proteins and to achieve a consistent result without a bias for each plasma sample of individuals during this optimizing study. By MRM-based analysis for the selected target proteins using the AAL-captured subproteome of the pooled liver-disease human plasma, reproducibility of the developed lectin-coupled, MRM-based method was investigated. Because it is impossible practically to obtain standard protein glycoforms with

fucosylated glycan microheterogeneity and structural conformations identical with those of the proteins in plasma, we evaluated the developed method by investigating the overall reproducibility throughout lectin-capture, enzymatic digestion, and MRM-based quantification.

The overall process from lectin-capture to MRM quantification was run with intra-day triplicate replication and with inter-day triplicate replication. Table [3](#page-6-0) shows that all selected target proteins except AFP were quantified at concentrations from tens of fmol to hundreds of fmol. As expected, AFP, known to be fucosylated in HCC patients and present in low abundance in plasma samples, was not quantified, although a dramatic reduction in sample complexity was achieved by the lectin-fractionation step by removing major serological proteins, non-glycosylated proteins, and proteins with no binding affinity for AAL. For detection of a low-abundance target protein, for example AFP, an additional enrichment technique using an antipeptide antibody may be useful, as exemplified in our previous paper [[16\]](#page-11-0). Table [3](#page-6-0) shows that the variation of measured values in intra-day triplicate replication was calculated within a CV of 14.3%. The variation in inter-day replication reached a CV of 20.5%, which was larger than the intra-day replication, indicating that some unwanted biases might have been present in the overall process of lectin-capture, trypsin digestion, internal standard spiking, and MRM quantification. Nonetheless, the variation of each target protein in the inter-day analysis was comparable with that for the overall analyses calculated for nine replicate assays (CV 11.0–19.1%). Thus it was confirmed that this MRM-based analytical method enabled highly reproducible quantitative analysis of lectin-captured glycoforms of target proteins. This analysis strategy was further applied to measurement of the abundance of AAL-captured glycoforms of target proteins in each individual plasma sample.

Precision of the lectin-coupled, MRM-based analysis

To validate this method of quantification and confirm that its precision was sufficient to distinguish differences between the abundance of protein glycoforms of target proteins in individuals, we used the method to analyze eight individual plasma samples. From a plasma sample group consisting of four healthy individuals and four HCC patients, each individual plasma sample was picked up randomly and analyzed by use of the developed lectin-coupled, MRM-based method. Table [4](#page-7-0) shows that each of the lectin-captured target proteins was quantified with a very wide range of concentrations for each plasma sample. Also, variation of each target protein among the eight individual cases was always larger than that of triplicate MRM measurements for a plasma sample. For example, the variation of alpha-1-acid glycoprotein 1 (AGP) among the eight individual cases was CV 66.1%, which is dramatically larger than those of the triplicate MRM measurements for each plasma sample, for which CVs were within 10.5%. These results showed that it was possible to distinguish differences in the abundance of the target proteins between plasma samples by using the quantification method developed. The ability of the quantification method to distinguish individual samples was also evaluated by comparing variations observed between individual samples with variations in the nine replicate measurements for the pooled liverdisease plasma sample used for the reproducibility validation. Figure 4 compares the variations in the measurements for eight individual cases with the variations in the nine replicate measurements for a pooled liver-disease plasma sample. For

Fig. 4 Comparison between variations (diamonds) in the measured values for eight individual plasma cases, variations (squares) in the nine replicate analysis for the pooled liver-disease plasma sample, and the variation value (CV 13.8%, dotted line) of AAL released from the immobilized lectin conjugates measured together during MRM analysis of the eight individual cases

all target proteins, the variations between individual cases were significantly larger than those between the replicated measurements of the pooled liver-disease control plasma.

The reproducibility and precision of the quantification method was further evaluated by observing the abundance of each AAL partially released from the AAL–streptavidin– magnetic bead conjugates during the analysis for the eight individual cases. The biotinylated AAL was used in an identical amount for lectin-captured sample preparation for all individual plasma samples and linked to magnetic beads by biotin–streptavidin binding. Although the biotin–streptavidin binding was assumed to be much stronger than the binding between AAL and AAL-active glycoforms of the target proteins, a small amount of AAL was continuously released from the conjugates during the elution of AALcaptured glycoproteins and quantified by this method of analysis, using the target peptide PTEFLYTSK as a surrogate for AAL and its stable isotope-coded counterpart, PTE-FL*YTSK (L*, isotope-labeled site) as an internal standard. The measured values were calculated using the chromatographic peak-area ratios of the transition pair, the m/z 543.3→887.5 transition of the endogenous target peptide (light), the m/z 546.8 \rightarrow 894.5 transition of the internal standard (heavy), and the molecular weight, 34 kDa, of AAL. As each MRM analysis of the eight individual plasma cases was conducted, measurements of the released AAL were conducted together. The amounts of AAL quantified in each analysis of eight individual samples corresponded to approximately 1% of the AAL used for lectin-fractionation in each analysis; the variation (CV 13.8%) in the quantified amounts of lectin is shown in Fig. [4](#page-8-0) with a dotted line. This quantification of the released AAL can be used as control data to evaluate the reproducibility of the analytical method for the steps elution of the lectin-bound glycoprotein, tryptic digestion, and MRM quantification. Figure [4](#page-8-0) shows that the variation of the eluted AAL was comparable with that of nine replicate measurements of each lectin-captured target protein using the pooled liver-disease control sample, whereas the variation of each lectin-captured target protein among eight individual plasmas was notably greater than those of the eluted AAL and of nine replicate measurements using the pooled control sample. This indicates that the abundance of each AAL-captured target protein in each individual case could be monitored quantitatively by use of the lectin-coupled, MRM-based method.

Quantitative monitoring of AAL-captured glycoproteins

The results from analysis of AAL-captured glycoforms of some target proteins, obtained from eight plasma samples consisting of four healthy individuals and four HCC patients, are presented in Fig. 5. The biological variation in the abundance of each lectin-captured target protein

between individual cases was quite remarkable for some target proteins. For example, the abundance of AGP was measured with wide ranges, from 5 fmol to 438 fmol, for the eight individual cases, with precision in the MRM quantification sufficient to distinguish the abundance of each individual case. The abundance of haptoglobin was also measured similarly with wide ranges, from 5 fmol to 989 fmol, for the eight individual cases. In contrast, variation in the abundance of the Ig alpha-1 chain C region over the eight individual cases was relatively small; the difference in the measured values was only a factor of two at most. In this study, we focused on measuring the abundance of AALcaptured glycoforms of the selected target proteins rather than on measuring the abundance of total target proteins,

Fig. 5 Plot of results from analysis of AAL-captured glycoforms of some target proteins obtained from eight individual plasma samples from four healthy individuals and four HCC patients. Each individual plasma sample was selected randomly, irrespective of its sample information, and analyzed sequentially by use of the developed lectincoupled, MRM-based quantification method. Normal and cancer patients are denoted N and C, respectively. Each point is an individual measurement of MRM-based analysis

and confirmed that the precision and reproducibility of the lectin-coupled, MRM-based analytical method was sufficient for measurement of the abundance of AAL-captured target proteins.

Different strategies using lectin have been introduced for discovery of serum glycoprotein biomarkers [10, 12, 14–[17\]](#page-11-0). Proteome samples active with a variety of lectins have been prepared by gel-based technology or immunoaffinity-based technology before subsequent mass analytical process for identification of target glycoproteins. This lectin-coupled MS-based strategy for identification of biomarker candidates has further evolved into immunoassay technology, for example antibody–lectin sandwich array, for identification and verification of biomarker candidates [\[15](#page-11-0), [25,](#page-11-0) [26,](#page-11-0) [37](#page-11-0)]. In this study, as an alternative to the immunoassay method inevitably requiring use of antibody against biomarker candidates to be verified, we focused on developing an MRM-based mass spectrometric method for verification of the biomarker candidate to be identified by lectin fractionation. Finally it was confirmed that the lectin-coupled, MRM-based method could be used to monitor differences between individuals by measuring the abundance of lectin-captured glycoforms of target proteins.

Conclusions

A quantitative method for measurement of protein glycoforms with a specific glycan structure, rather than quantifying total glycoforms, from the human plasma proteome was developed and evaluated to confirm that its reproducibility and precision were sufficient to distinguish differences between the abundance of protein glycoforms in individuals. In this study, we focused on quantitative analysis of AALcaptured glycoforms of some target proteins selected from profiling experiments using a subproteome fractionated by AAL with specific affinity for fucose on N-linked glycans. The lectin-coupled, MRM-based method for substoichiometrically fucosylated glycoforms of the target proteins was developed without highly abundant protein depletion or affinity enrichment of the target protein. The precision and reproducibility of the method were sufficient to distinguish differences between individuals by measuring the abundance of AAL-captured target proteins. This method can be extended to analysis of a glycan-specific subproteome captured by a variety of lectins with binding affinity for a specific glycan structure. Progression of some diseases may involve aberrant glycosylation in a glycoprotein; thus, the ability to analyze quantitatively aberrant glycoforms of glycoproteins provides a tool for monitoring differences between individual cases in the abundance of aberrant glycoforms of target proteins related to the diseases. Thus the lectin-coupled, MRM-based method is expected to be

further applied to differentiation of diseased patients from healthy individuals and to validate biomarker candidates involved in aberrant protein glycosylation.

Acknowledgments This work was supported by the Convergence Research Center Program (grant number 2011 K000884) from the Korean Ministry of Education, Science, and Technology.

References

- 1. Orntoft TF, Vestergaard EM (1999) Clinical aspects of altered glycosylation of glycoproteins in cancer. Electrophoresis 20:362– 371
- 2. Pierce M, Arango J (1986) Rous sarcoma virus-transformed baby hamster kidney cells express higher levels of asparagine-linked triand tetraantennary glycopeptides containing [GlcNAc-beta (1,6) Man-alpha (1,6)Man] and poly-N-acetyllactosamine sequences than baby hamster kidney cells. J Biol Chem 261:10772–10777
- 3. Dennis JW, Laferte S, Waghorne C, Breitman ML, Kerbel RS (1987) Beta 1-6 branching of Asn-linked oligosaccharides is directly associated with metastasis. Science 236:582–585
- 4. Barrabés S, Pagès-Pons L, Radcliffe CM, Tabarés G (2007) Glycosylation of serum ribonuclease 1 indicates a major endothelial origin and reveals an increase in core fucosylation in pancreatic cancer. Glycobiology 17:388–400
- 5. Zhao J, Qiu W, Simeone DM, Lubman DM (2007) N-linked glycosylation profiling of pancreatic cancer serum using capillary liquid phase separation coupled with mass spectrometric analysis. J Proteome Res 6:1126–1138
- 6. Mizuochi T, Nishimura R, Derappe C, Taniguchi T, Hamamoto T, Mochizuki M, Kobata A (1983) Structures of the asparaginelinked sugar chains of human chorionic gonadotropin produced in choriocarcinoma: appearance of triantennary sugar chains and unique biantennary sugar chains. J Biol Chem 258:14126–14129
- 7. Peracaula R, Tabares G, Royle L, Harvey DJ, Dwek RA, Rudd PM, de Llorens R (2003) Altered glycosylation pattern allows the distinction between prostatespecific antigen (PSA) from normal and tumor origins. Glycobiology 13:457–470
- 8. Ohyama C, Hosono M, Nitta K, Oh-eda M, Yoshikawa K, Habuchi T, Arai Y, Fukuda M (2004) Carbohydrate structure and differential binding of prostate specific antigen to Maackia amurensis lectin between prostate cancer and benign prostate hypertrophy. Glycobiology 14:671–679
- 9. Gomaa AI, Khan SA, Leen EL, Waked I, Taylor-Robinson SD (2009) Diagnosis of hepatocellular carcinoma. World J Gastroenterol 15:1301–1314
- 10. Drake RR, Schwegler EE, Malik G, Diaz J, Block T, Mehta A, Semmes OJ (2006) Lectin capture strategies combined with mass spectrometry for the discovery of serum glycoprotein biomarkers. Mol Cell Proteomics 5:1957–1967
- 11. Drake PM, Cho W, Li B, Prakobphol A, Johansen E, Anderson NL, Regnier FE, Gibson BW, Fisher SJ (2010) Sweetening the pot: adding glycosylation to the biomarker discovery equation. Clin Chem 56:223–236
- 12. Dai Z, Zhou J, Qiu SJ, Liu YK, Fan J (2009) Lectin-based glycoproteomics to explore and analyze hepatocellular carcinomarelated glycoprotein markers. Electrophoresis 30:2957–2966
- 13. Polanski M, Anderson NL (2007) A list of candidate cancer biomarkers for targeted proteomics. Biomarker Insights 7:1–48
- 14. Ito H, Kuno A, Sawaki H, Sogabe M, Ozaki H, Tanaka Y, Mizokami M, Shoda J, Angata T, Sato T, Hirabayashi J, Ikehara Y, Narimatsu H (2009) Strategy for glycoproteomics: identification of glyco-

alteration using multiple glycan profiling tools. J Proteome Res 8:1358–1367

- 15. Qiu Y, Patwa TH, Xu L, Shedden K, Misek DE, Tuck M, Jin G, Ruffin MT, Turgeon DK, Synal S, Bresalier R, Marcon N, Brenner DE, Lubman DM (2008) Plasma glycoprotein profiling for colorectal cancer biomarker identification by lectin glycoarray and lectin blot. J Proteome Res 7:1693–1703
- 16. Ahn YH, Lee JY, Lee JY, Kim YS, Ko JH, Yoo JS (2009) Quantitative analysis of an aberrant glycoform of TIMP1 from colon cancer serum by L-PHA-enrichment and SISCAPA with MRM mass spectrometry. J Proteome Res 8:4216–4224
- 17. Ahn YH, Kim YS, Ji ES, Lee JY, Jung JA, Ko JH, Yoo JS (2010) Comparative quantification of aberrant glycoforms by lectin-based glycoprotein enrichment coupled with multiple-reaction monitoring mass spectrometry. Anal Chem 82:4441–4447
- 18. Heo SH, Lee SJ, Ryoo HM, Park JY, Cho JY (2007) Identification of putative serum glycoprotein biomarkers for human lung adenocarcinoma by multilectin affinity chromatography and LC–MS– MS. Proteomics 7:4292–4302
- 19. Stahl-Zeng J, Lange V, Ossola R, Eckhardt K, Krek W, Aebersold R, Domon B (2007) High sensitivity detection of plasma proteins by multiple reaction monitoring of N-glycosites. Mol Cell Proteomics 6:1809–1817
- 20. Hua S, Nwosu CC, Strum JS, Seipert RR, An HJ, Zivkovic AM, German JB, Lebrilla CB (2011) Site-specific protein glycosylation analysis with glycan isomer differentiation. Anal Bioanal Chem. doi[:10.1007/s00216-011-5109-x](http://dx.doi.org/10.1007/s00216-011-5109-x)
- 21. Jung K, Cho W, Regnier FE (2009) Glycoproteomics of plasma based on narrow selectivity lectin affinity chromatography. J Proteome Res 8:643–650
- 22. Kumada T, Nakano S, Takeda I, Kiriyama S, Sone Y, Hayashi K, Katoh H, Endoh T, Sassa T, Satomura S (1999) Clinical utility of Lens culinaris agglutinin-reactive alphafetoprotein in small hepatocellular carcinoma: special reference to imaging diagnosis. J Hepatol 30:125–130
- 23. Khien VV, Mao HV, Chinh TT, Ha PT, Bang MH, Lac BV, Hop TV, Tuan NA, Don LV, Taketa K, Satomura S (2001) Clinical evaluation of lentil lectin-reactive alpha-fetoprotein-L3 in histology-proven hepatocellular carcinoma. Int J Biol Markers 16:105–111
- 24. Miyaaki H, Nakashima O, Kurogi M, Eguchi K, Kojiro M (2007) Lens culinaris agglutinin-reactive alpha-fetoprotein and protein induced by vitamin K absence II are potential indicators of a poor prognosis: a histopathological study of surgically resected hepatocellular carcinoma. J Gastroenterol 42:962–968
- 25. Chen S, LaRoche T, Hamelinck D, Bergsma D, Brenner D, Simeone D, Brand RE, Haab BB (2007) Multiplexed analysis of glycan variation on native proteins captured by antibody microarrays. Nat Methods 4:437–444
- 26. Kuno A, Uchiyama N, Koseki-Kuno S, Ebe Y, Takashima S, Yamada M, Hirabayashi J (2005) Evanescent-field fluorescenceassisted lectin microarray: a new strategy for glycan profiling. Nat Methods 2:851–856
- 27. Pilobello KT, Krishnamoorthy L, Slawek D, Mahal LK (2005) Development of a lectin microarray for the rapid analysis of protein glycopatterns. Chembiochem 6:985–989
- 28. Kuhn E, Wu J, Karl J, Liao H, Zolg W, Guild B (2004) Quantification of C-reactive protein in the serum of patients with rheumatoid arthritis using multiple reaction monitoring mass spectrometry and 13 C-labeled peptide standards. Proteomics 4:1175–1186
- 29. Anderson L, Hunter CL (2006) Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. Mol Cell Proteomics 5:573–588
- 30. Martínez-Villalba A, Moyano E, Martins CPB, Galceran MT (2010) Fast liquid chromatography/tandem mass spectrometry (highly selective selected reaction monitoring) for the determination of toltrazuril and its metabolites in food. Anal Bioanal Chem 397:2893–2901
- 31. McKay MJ, Sherman J, Laver MT, Baker MS, Clarke SJ, Molloy MP (2007) The development of multiple reaction monitoring assays for liver-derived plasma proteins. Proteomics Clin Appl 1:1570–1581
- 32. Kuzyk MA, Smith D, Yang J, Cross TJ, Jackson AM, Hardie DB, Anderson NL, Borchers CH (2009) Multiple reaction monitoringbased, multiplexed, absolute quantification of 45 proteins in human plasma. Mol Cell Proteomics 8:1860–1877
- 33. Addona TA, Abbatiello SE, Schilling B (2009) Multi-site assessment of the precision and reproducibility of multiple reaction monitoring-based measurements of proteins in plasma. Nat Biotechnol 27:633–641
- 34. Zhao Y, Jia W, Sun W, Jin W, Guo L, Wei J, Ying W, Zhang Y, Xie Y, Jiang Y, He F, 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
- 35. Kochibe N, Furukawa K (1980) Purification and properties of a novel fucose-specific hemagglutinin of Aleuria aurantia. Biochemistry 19:2841–2846
- 36. Matsumura K, Higashida K, Ishida H, Hata Y, Yamamoto K, Shigeta M, Mizuno-Horikawa Y, Wang X, Miyoshi E, Gu J, Taniguchi N (2007) Carbohydrate binding specificity of a fucosespecific lectin from Aspergillus oryzae: a novel probe for core fucose. J Biol Chem 282:15700–15708
- 37. Comunale MA, Wang M, Hafner J, Krakover J, Rodemich L, Kopenhaver B, Long RE, Junaidi O, Bisceglie AM, Block TM, Mehta AS (2009) Identification and development of fucosylated glycoproteins as biomarkers of primary hepatocellular carcinoma. J Proteome Res 8:595–602
- 38. Wang Y, Ao X, Vuong H, Konanur M, Miller FR, Goodison S, Lubman DM (2008) Membrane glycoproteins associated with breast tumor cell progression identified by a lectin affinity approach. J Proteome Res 7:4313–4325
- 39. Kullolli M, Hancock WS, Hincapie M (2008) Preparation of a high-performance multi-lectin affinity chromatography (HP-M-LAC) adsorbent for the analysis of human plasma glycoproteins. J Sep Sci 31:2733–2739