

Protein and Site Specificity of Fucosylation in Liver-Secreted Glycoproteins

Petr Pompach,^{†,‡,§} David J. Ashline,^{||} Zuzana Brnakova,[§] Julius Benicky,[§] Miloslav Sanda,[§] and Radoslav Goldman^{*,§,⊥}

[†]Institute of Microbiology v.v.i., Czech Academy of Sciences, Videnska 1083, Prague 142 20, Czech Republic

[‡]Department of Biochemistry, Faculty of Sciences, Charles University, Hlavova 8, Prague 128 43, Czech Republic

[§]Department of Oncology, Lombardi Comprehensive Cancer Center, Georgetown University, 3970 Reservoir Road NW, Washington, DC 20007, United States

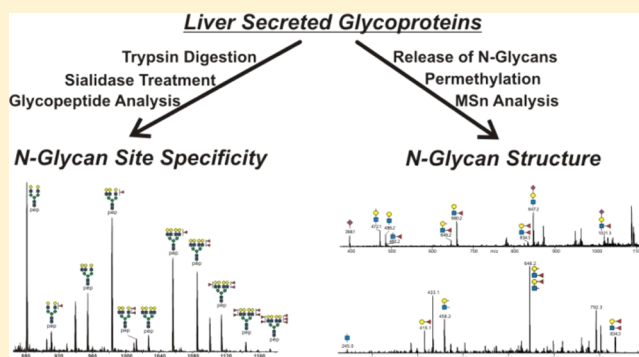
^{||}The Glycomics Center, University of New Hampshire, Durham, New Hampshire 03824, United States

[⊥]Department of Biochemistry and Molecular & Cellular Biology, Georgetown University, Washington, DC 20007, United States

Supporting Information

ABSTRACT: Chronic liver diseases are a serious health problem worldwide. One of the frequently reported glycan alterations in liver disease is aberrant fucosylation, which was suggested as a marker for noninvasive serologic monitoring. We present a case study that compares site specific glycoforms of four proteins including haptoglobin, complement factor H, kininogen-1, and hemopexin isolated from the same patient. Our exoglycosidase-assisted LC–MS/MS analysis confirms the high degree of fucosylation of some of the proteins but shows that microheterogeneity is protein- and site-specific. MSn analysis of permethylated detached glycans confirms the presence of LeY glycoforms on haptoglobin, which cannot be detected in hemopexin or complement factor H; all three proteins carry Lewis and H epitopes. Core fucosylation is detectable in only trace amounts in haptoglobin but with confidence on hemopexin and complement factor H, where core fucosylation of the bi-antennary glycans on select glycopeptides reaches 15–20% intensity. These protein-specific differences in fucosylation, observed in proteins isolated from the same patient source, suggest that factors other than up-regulation of enzymatic activity regulate the microheterogeneity of glycoforms. This has implications for selection of candidate proteins for disease monitoring and suggests that site-specific glycoforms have structural determinants, which could lead to functional consequences for specific subsets of proteins or their domains.

KEYWORDS: mass spectrometry, exoglycosidase treatment, microheterogeneity, N-glycans, permethylation, MSn structural analysis, glycopeptides, hepatocellular carcinoma, fucosylation, site specificity



INTRODUCTION

N-Glycosylation sequons of proteins (Asn-Xaa-Ser/Thr, where Xaa is any amino acid except proline) typically carry a diverse repertoire of structures, called glycan microheterogeneity.¹ The diversity of protein glycosylation is influenced by the activity of many enzymes involved in the N-glycan biosynthetic machinery.^{2,3} The source and the location of the glycoproteins also strongly influence the final glycan structure and site specificity.⁴ Microheterogeneity of protein N-glycoforms changes significantly in cancer diseases often in connection with changes in the activity of enzymes involved in N-glycoprotein synthesis.^{5–7} One of the best documented examples of aberrant glycosylation is the progression of liver disease to hepatocellular carcinoma (HCC).^{8,9} In the case of liver disease, changes in protein N-glycosylation can be analyzed efficiently by serologic monitoring of liver-secreted glycoproteins.^{10–12} Reported changes of N-glycan repertoire in cancer include the formation of bisecting

glycans with one molecule of N-acetylglucosamine attached via β 1–4 bond to the branched core mannose, asialo and agalacto glycans, truncated structures with terminal galactoses or N-acetylglucosamines, (poly)-N-acetylglucosamines with beta 1–6 branching, and fucosylated glycans with fucose attached mainly at the C-3 position of outer arm N-acetylglucosamine or less frequently at the C-6 N-acetylglucosamine of the N-glycan core.^{13–20}

We have previously detected multiply outer arm fucosylated glycoforms of haptoglobin (Hp) in the context of HCC.¹⁰ We have now isolated three additional glycoproteins from serum of a patient that carries hyper-fucosylated Hp (up to six fucoses per glycan). These liver-secreted glycoproteins copurify on hemin agarose and include hemopexin (Hpx), kininogen-1 (Kng-1),

Received: June 4, 2014

Published: September 29, 2014

and complement factor H (CFH). Our assumption is that the same glycosylation machinery that hyper-fucosylates Hp would lead to hyper-fucosylation of additional proteins. We report the analysis of the distribution of N-glycans between the four proteins and on specific glycopeptides using two complementary analytical methods. We have used sialidase-assisted tandem mass spectrometry of glycopeptides to study the site-specific protein glycoforms and further determined by ESI-ITMSn the analysis of permethylated glycans detached from each of the purified proteins. This is to our knowledge the first case report of detailed structural characterization of site-specific glycoforms of proteins isolated directly from patient serum.

■ EXPERIMENTAL SECTION

Patient Samples

Proteins were isolated from serum of a patient previously shown to carry multiply fucosylated glycoforms of Hp.¹⁰ This patient was enrolled under protocols approved by the Georgetown University Institutional Review Board in collaboration with the Georgetown University Hospital, Department of Hepatology and Liver Transplantation, Washington, DC. The patient is a 61 year old Caucasian male, blood group O, diagnosed by the attending physician as a cirrhotic patient of HCV etiology with HCC stage T2N0M0, according to the seventh edition of the American Joint Committee on Cancer Staging manual.

Isolation of Glycoproteins from Human Plasma

Hp was isolated as described elsewhere.¹⁰ Hpx, Kng-1, and CFH were isolated by incubation of 500 μ L of human plasma with 400 μ L of heme-agarose beads (Sigma Aldrich, St. Louis, MO) at 4 °C overnight. After the incubation, the resin was washed with 10 mL of PBS buffer, and bound glycoproteins were released by the addition of 500 μ L of 0.2 M citric acid, pH 2 and immediately neutralized by 270 μ L of 1 M Tris/Cl buffer, pH 9. Proteins were further separated by reversed-phase chromatography using C18 macroporous mRP Hi-recovery protein column (4.6 \times 50 mm, Agilent Technologies, Santa Clara, CA) under the following conditions: 0 min 1% B, 5 min 35% B, 30 min 45% B, 31 min 100% B, 32 min 100% B, 33 min 1% B, 45 min 1% B (A: 2% ACN (Fisher Scientific, Fair Lawn, NJ), 0.1% TFA in water (Fisher Scientific, Fair Lawn, NJ), B: 98% ACN, 0.08% TFA in water), flow rate 0.75 mL/min, and column was heated to 40 °C. Eluted proteins were monitored by UV detector set at 214 nm. The CFH fraction was collected in the time range 9.7 to 10 min, Kng-1 fraction in the time range 13.5–15 min, and Hpx fraction in the time range 17.5–18.5 min. Collected fractions were dried in SpeedVac concentrator for further processing.

Protein Digest and Sialidase Treatment

Isolated glycoproteins were resuspended in 20 μ L of 50 mM NH_4HCO_3 (Sigma-Aldrich) at pH 7.8 with 0.05% RapiGest (Waters, Milford, MA). Proteins were reduced with 5 mM DTT (Sigma-Aldrich) and alkylated with 15 mM iodoacetamide (Sigma-Aldrich). Trypsin (Promega, Madison, WI) digestion (2.5 ng/ μ L) was carried out at 37 °C in Barocycler NEP2320 (Pressure BioSciences, South Easton, MA) for 1 h. Protein digests were loaded on a Micro Trap peptide cartridge (Michrom Bioresources, Auburn, CA) and washed three times with 250 μ L of 0.1% aqueous TFA. Peptides were eluted with 100 μ L of 60% ACN with 0.1% TFA and the eluate was dried using a SpeedVac concentrator. For exoglycosidase treatment, peptides were resuspended in 20 μ L of reaction buffer containing 50 mM sodium citrate, pH 6.0. Two microliters (100 units) of

neuraminidase (New England BioLabs, Ipswich, MA) was added to peptide digests derived from 2 μ g of each glycoprotein and incubated at 37 °C for 20 h. Peptides were desalted as previously described.¹⁰

Nano LC–MS/MS

Mixture of tryptic peptides was resuspended in 40 μ L of 2% ACN, 0.1% formic acid and analyzed by nanoLC–ES–MS/MS using a C18 column (1.7 μ m particles, 75 μ m ID \times 150 mm, Waters) coupled to a QStar Elite mass spectrometer (Applied Biosystems, Foster City, CA). Two picomoles of each glycoprotein digest were separated at a flow rate of 0.4 μ L/min, using the following gradient: 0 min 1% B, 35 min 35% B, 37 min 90% B, 42 min 90% B, 44 min 1% B, 55 min 1% B (Solvent A: 0.1% formic acid in 2% ACN; Solvent B: 0.1% formic acid in 98% ACN). The mass spectrometer was operated in data-dependent mode; after the full-scan (m/z 400 to m/z 1800) survey, the three most intense precursor ions were selected for collision-induced dissociation. Collision energy and MS/MS accumulation time were set automatically, and the MS/MS spectra were recorded from m/z 150 to m/z 2000. Dynamic exclusion was set at 15 s and five counts for two repeated precursors. Fragment intensity multiplier was set at 16 with maximum accumulation 2 s, which resulted in a total cycle time of 6.5 s.

Structural Characterization of Detached Glycans

Hp, Hpx, and CFH purified as previously described (50 μ g each) were enzymatically deglycosylated with N-glycanase at 37 °C overnight (Prozyme, Hayward, CA). Released N-glycans were purified by C18 and porous graphitized carbon solid-phase extraction steps as described previously.²¹ Briefly, N-glycans were reduced using borane-ammonia and permethylated using spin columns, as previously described.^{22,23}

MALDI-ToF analysis was performed on a Shimadzu Kratos Axima-CFR (Columbia, MD) using DHB as a matrix. This step provided a survey of the compositions present in each sample. Detailed mass spectrometric analysis was performed on a Thermo LTQ (San Jose, CA) mass analyzer equipped with an Advion Triversa Nanomate (Ithaca, NY). Samples were dissolved in 1:1 methanol/water for direct infusion into the LTQ. All ions were sodium adducts, and select fucosylated compositions were disassembled to determine detailed structure as described.²¹

Data Interpretation

LC–MS/MS of glycopeptides: Analyst QS 2.0 software (Applied Biosystems, Foster City, CA) was used to process MS data sets, and glycoproteins were identified by Mascot search engine. Protein sequences were downloaded from UniProt database. MS/MS spectra were converted to mzXML format using msconvert from the ProteoWizard project. GlycoPeptideSearch (GPS) engine was used for finding of glycopeptides forms for each glycoprotein.^{24,25} Identified glycopeptides were further confirmed by manual data inspection. We adopt the N-glycan nomenclature from the NIBRT GlycoBase databases.²⁶ Peak area counts from extracted ion chromatogram (XIC) of the precursor ion of each identified glycoform were used for calculation of percentage distribution of glycopeptides and also for calculation of ratio between fucosylated and nonfucosylated glycoforms. All samples were analyzed in triplicate.

■ RESULTS AND DISCUSSION

Our recent analysis of Hp isolated from serum of a patient with HCC detected multiply fucosylated glycans with up to six

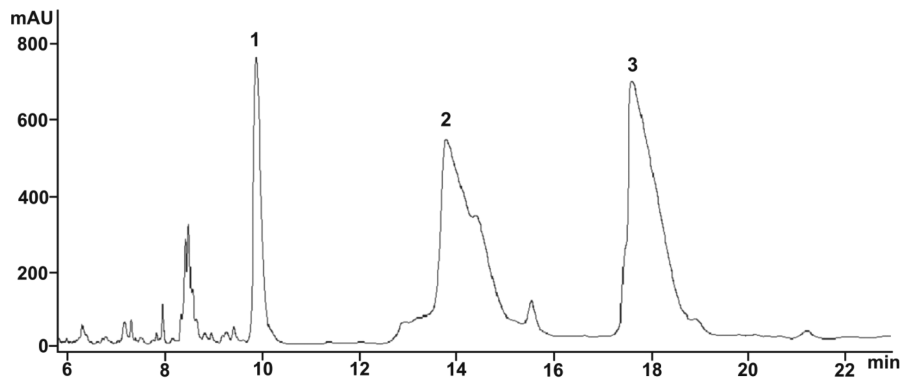


Figure 1. UV detection (214 nm) of a glycoprotein complex isolated by heme-agarose affinity chromatography and separated by reversed-phase chromatography. Peaks containing the proteins of interest are numbered as follows: 1-CFH, 2-Kng-1, and 3-Hpx.

Table 1. List of Identified Glycoforms of Hp, Kng-1, Hpx, and CFH and Their Intensity Distribution at All Identified Glycopeptides Determined after Neuraminidase Treatment^a

Glycopeptide (AA position)	Glycoform	%	Observed m/z (Charge)	Glycopeptide (AA position)	Glycoform	%	Observed m/z (Charge)
Kininogen-1				Hemopexin			
ITY.....CSK (197-208)	A2G2	11	1012.7 (3+)	SWP.....ALR (158-170)	A2G2	84	1009.7 (3+)
	A3G3	39	851.1 (4+)		A2G2F1	4	1058.5 (3+)
	A3G3F1	50	887.6 (4+)		A3G3	3	1131.4 (3+)
LNA.....YFK (289-300)	A2G2	22	1018.7 (3+)	A3G3F1	6	1180.2 (3+)	
	A2G2F1	2	1067.5 (3+)	A3G3F2	1	1228.8 (3+)	
	A3G3	23	855.6 (4+)	A4G4	<1	940.4 (4+)	
	A3G3F1	42	892.1 (4+)	A4G4F1	1	976.9 (4+)	
YNS.....LYR (44-58)	A3G3F2	11	928.6 (4+)	A4G4F2	1	1013.2 (4+)	
	A2G2	100	875.1 (4+)	A4G4F3	<1	1049.7 (4+)	
	A2G2	28	880.8 (5+)	ALP.....CTH (424-439)	A2G2	94	1120.2 (3+)
HGI.....VKR (162-184)	A3G3	9	953.8 (5+)	A2G2F1	4	1168.9 (3+)	
	A3G3F1	22	983.1 (5+)	A3G3	1	1241.9 (3+)	
	A3G3F2	2	1012.3 (5+)	A3G3F1	1	1290.5 (3+)	
	A4G4	2	1026.9 (5+)	A3G3F2	<1	1339.2 (3+)	
	A4G4F1	15	1053.1 (5+)	CSD.....FFK (27-48)	A2G2	85	1038.4 (4+)
	A4G4F2	13	1085.3 (5+)		A2G2F1	7	1074.9 (4+)
	A4G4F3	6	1114.5 (5+)		A3G3	2	1129.7 (4+)
	A4G4F4	2	1143.9 (5+)	A3G3F1	5	1166.2 (4+)	
A4G4F5	<1	1173.1 (5+)	A3G3F2	1	1202.7 (4+)		
Haptoglobin				Complement factor H			
VVL.....LIK (218-233) – T3	A2G2	38	855.1 (4+)	MDG.....NSR (1024-1036)	A2G2	83	1016.4 (3+)
	A2G2F1	<1	891.6 (4+)	A2G2F1	17	1065.1 (3+)	
	A3G3	15	946.4 (4+)	A2G2	52	1062.1 (3+)	
	A3G3F1	21	982.9 (4+)	A2G2F1	13	1110.7 (3+)	
	A3G3F2	4	1019.4 (4+)	A3G3	13	888.1 (4+)	
	A3G3F3	<1	1055.9 (4+)	A3G3F1	21	924.6 (4+)	
	A3G3F4	<1	1092.4 (4+)	A3G3F2	2	961.1 (4+)	
	A4G4	3	1037.6 (4+)	IPC.....SSR (868-885)	A2G2	76	911.7 (4+)
	A4G4F1	7	1074.2 (4+)	A2G2F1	17	948.2 (4+)	
	A4G4F2	8	1110.7 (4+)	A3G3	1	1002.9 (4+)	
	A4G4F3	3	1147.2 (4+)	A3G3F1	5	1039.4 (4+)	
	A4G4F4	1	1183.7 (4+)	A3G3F2	1	1076.0 (4+)	
	A4G4F5	<1	1220.2 (4+)	WDP.....NYR (797-830)	A2G2+A2G2	30	1210.0 (6+)
A4G4F6	<1	1256.8 (4+)	A2G2+A3G3		39	1271.0 (6+)	
MVS.....TAK (161-184) – T1	A2G2	74	1076.2 (4+)		A2G2+A3G3F1	21	1295.4 (6+)
A2G2F1	3	1112.7 (4+)	A2G2+A3G3F2	10	1319.0 (6+)		
A3G3	7	1167.5 (4+)	SPY.....QCK (1079-1103)	A2G2	79	1161.0 (4+)	
A3G3F1	15	1204.0 (4+)	A2G2F1	16	1197.4 (4+)		
A3G3F2	1	1240.5 (4+)	A3G3	3	1252.5 (4+)		
A4G4	<1	1258.7 (4+)	A3G3F1	2	1288.8 (4+)		
A4G4F1	<1	1295.2 (4+)	LND.....YER (528-567)	A2G2	25	1249.5 (5+)	
A4G4F2	<1	1331.7 (4+)		A2G2F1	2	1278.8 (5+)	
A4G4F3	<1	1368.2 (4+)		A3G3	34	1322.2 (5+)	
NLF.....TAK (185-197) – T2	A2G2+A2G2	41	1176.7 (4+)	A3G3F1	39	1351.4 (5+)	
	A2G2+A2G2F1	6	1213.2 (4+)	A2G2	51	1413.6 (6+)	
	A3G3+A2G2	26	1268.0 (4+)	A2G2F1	18	1437.9 (6+)	
A3G3+A2G2F1	27	1304.5 (4+)	A3G3	19	1474.5 (6+)		
				A3G3F1	13	1498.8 (6+)	

^aGlycan abbreviations are adopted from the NIBRT GlycoBase databases (A2G2, bi-antennary glycan terminated with two galactoses; A3G3, tri-antennary glycan terminated with three galactoses, etc.; F, fucose; we do not resolve core and outer arm).

fucoses.^{10,12} We have used glycosidase assisted LC-MS/MS to show the presence of Lewis Y type glycoforms at one glycopeptide of Hp. In this study, we use the recently optimized glycosidase-assisted LC-MS/MS methods and MSn structural

analysis of detached N-glycans to compare glycoforms of Hp with glycoforms of Hpx, Kng-1, and CFH isolated from plasma of the same patient. Our goal is to compare site-specific microheterogeneity of proteins synthesized by the same

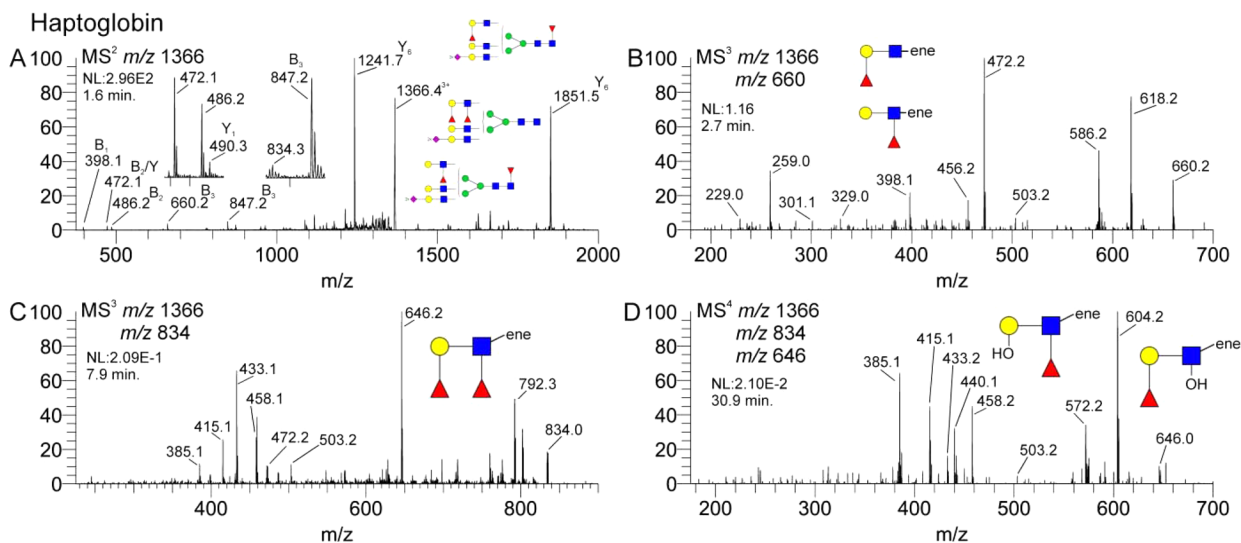


Figure 2. Structural analysis of permethylated tetra-antennary doubly fucosylated glycan (m/z 1366 (3+)) from Hpt by MSn fragmentation. Spectrum A shows the MS² spectrum with zoomed-in views of the core fuc-GlcNAc (Y1, m/z 490) and Lewis Y (m/z 834) fragments. Panel B shows the MS³ spectrum of m/z 660, showing a mixture of Lewis X and H2 structures. Panels C and D show MS³ and MS⁴ spectra, respectively, of the Lewis Y epitope disassembly. Antenna localization was not empirically determined.

enzymatic machinery and under the same pathophysiological conditions. The proteins were purified by reversed-phase chromatography after incubation of the patient's plasma with heme-agarose beads. Isolation of these proteins as a complex on heme-agarose is expected because all three proteins have high affinity for either heme or each other.²⁷ Because N-glycans have minimal interaction with the reversed-phase resin, all glycoforms of the protein coelute and their microheterogeneity can be analyzed efficiently in the isolated protein. Hpx and CFH were both purified to apparent homogeneity, but Kng-1 coeluted from the reversed-phase column with histidine-rich glycoprotein and other minor contaminants (Figure 1). We were able to analyze glycopeptides of Kng-1 efficiently in the simple coeluting mixture of proteins, but we were not able to isolate sufficient amounts of purified Kng-1 for MSn structural analysis; MSn analysis was therefore completed on Hp, CFH, and Hpx. We have focused on LC-MS analysis of tryptic digests because trypsin provides sufficient coverage of each protein's glycosites to allow comparison of microheterogeneity and because the tryptic digest is standardized sufficiently to allow semiquantitative comparison of precursor ion intensities of the microheterogeneous glycopeptides. We describe the first results of analysis of each protein and follow the description by comparison of the results across the proteins.

Haptoglobin: Hp is a liver-secreted glycoprotein with four N-glycosylation sites. Its major function is capture of released hemoglobin during intravascular hemolysis. Our recent data revealed that Hp isolated from serum of HCC patients carries up to six fucoses per glycan.¹⁰ In the present study, we have used neuraminidase assisted LC-MS/MS to confirm that all glycosites of Hp carry complex glycans with variable proportion of bi-, tri-, and tetra-antennary structures and variable degree of fucosylation (Table 1). Tri-antennary structures on peptide V(218)-K(233) are as abundant as bi-antennary structures.

To further refine structure assignments, we have used MSn analysis of detached permethylated N-glycans. Our previous study used glycosidase-assisted LC-MS/MS to show that hyper-fucosylated glycoforms of Hp contain mostly Lewis-X- and Lewis-Y-type glycoforms.¹⁰ Detailed MSn structural analysis of

tetra-antennary doubly fucosylated glycan with monosaccharide composition (Hex)₄ (HexNAc)₄ (Deoxyhexose)₂ (NeuAc)₂ + (Man)₃(GlcNAc)₂ supports the presence of Lewis X, H2, and Lewis Y glycoforms. Dissociation (MS²) of this doubly fucosylated permethylated glycan at m/z 1366 (3+) gives Lewis type fragments at m/z 834 and 660. Both fragments were further dissociated (MS³) and product ion at m/z 646 (derived from 834 fragment) was further collided (MS⁴). MSn analysis shows at least three possible isomers at this composition, with both core and antennal fucosylation. Figure 2 shows selected MSn spectra for the reduced and permethylated tetra-antennary, disialylated, and doubly fucosylated composition (m/z 1366 (3+)). Figure 2A shows the MS² spectrum of this composition, with insets displaying zoomed-in views of the terminal (B2) and internal (B3/Y) lactosamine fragments as well as the Y1 fragment (m/z 490) indicative of the fucosylated reducing-end GlcNAc; the second inset shows the terminal doubly fucosylated lactosamine (m/z 834) and the sialylated lactosamine (m/z 847) fragments. The m/z 490 fragment indicates traces of core fucosylated glycans. The MS² spectrum also clearly indicates two different antennal fucosylation fragments: the terminal (B-type) monofucosylated fragment at m/z 660 and the terminal (B-type) difucosylated fragment at m/z 834. Because there are several structural isomeric possibilities for these fragments, these ions were subjected to further MSn interrogation. The MS³ fragmentation of the m/z 660 ion (Figure 2B) is indicative of a mixture of Lewis X and H2 epitopes (Figure 2). These isomeric glycans differ in fucose position and linkage and are particularly difficult to separate by LC-MS analysis; MSn fragmentation is a complementary method that resolves the linkage efficiently. The specific diagnostic fragments are the m/z 259, a C-type terminal hexose that can arise from a Lewis X structure but not an H2 structure, m/z 329, a ^{3,5}A cross-ring cleavage fragment that can only arise from Lewis X, not H2 or Lewis A, and the m/z 503, another ^{3,5}A cross-ring cleavage fragment which can only arise from H2, not H1 or Lewis A or X. Detailed structures with fragment identifications are in the Supporting Information. The low-intensity fragment at m/z 834 (terminal difucosylated lactosamine) in the MS/MS spectrum is

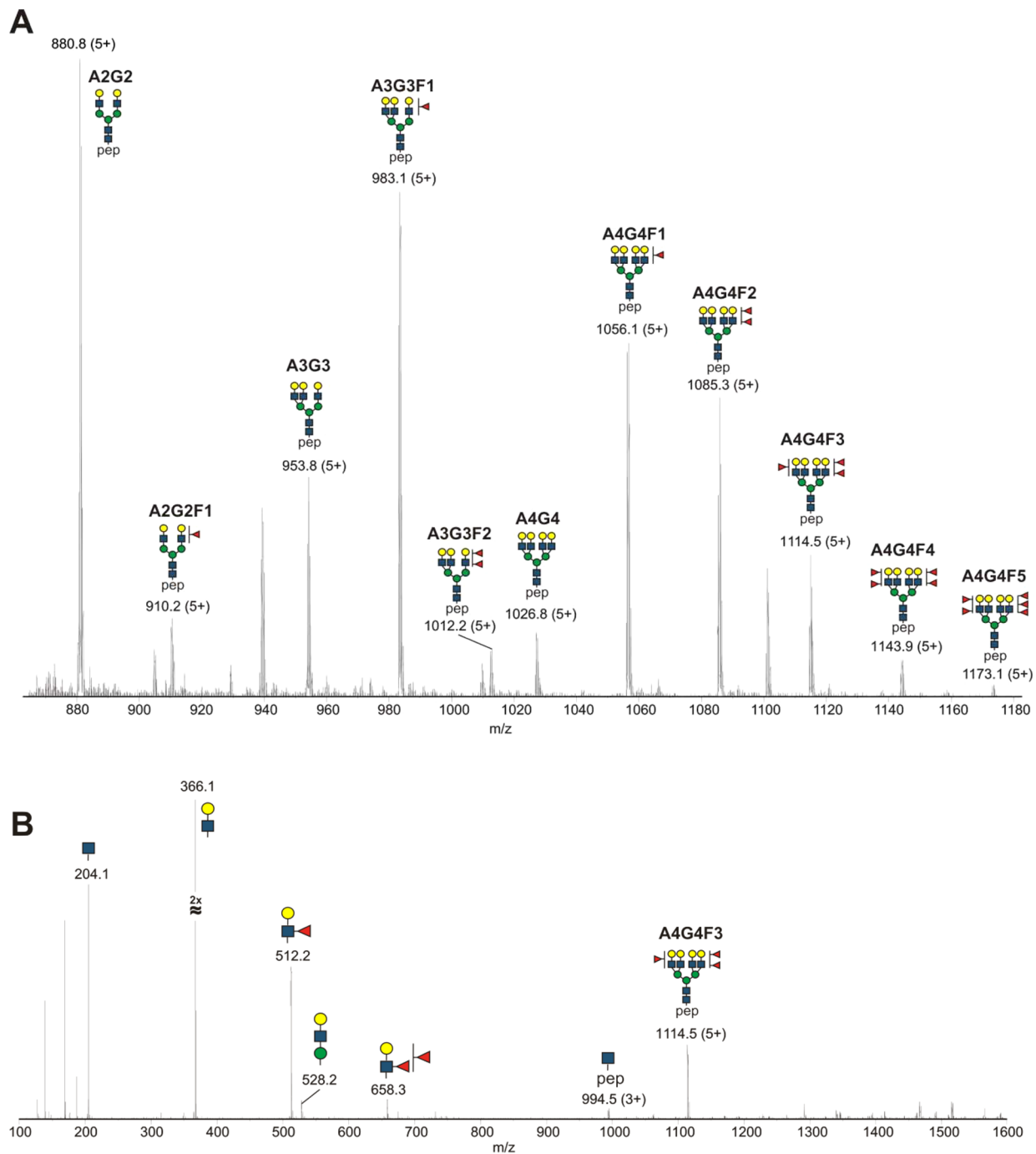


Figure 3. (A) MS profile of Kng-1 peptide HGIQYFNNNTQHSSLFMLNEVKR bearing tetra-antennary glycans with up to five fucoses. (B) CID spectrum of tetra-antennary triply fucosylated HGIQYFNNNTQHSSLFMLNEVKR glycopeptide with oxonium ions corresponding to GlcNAc and GlcNAc-Hex fragments (m/z 204 and 366) and GlcNAc-Gal-Fuc fragment (m/z 512) consistent with Lewis X structure.

consistent with the possible Lewis Y epitope. Disassembly of this ion through MS3, m/z 834 (Figure 2C) and MS4, m/z 834 \rightarrow 646 (Figure 2D) are consistent with the Lewis Y epitope.²¹ The MS3 and MS4 spectra both contain the m/z 503 fragment, a ^{3,5}A cross-ring cleavage ion, of the same structure as that found in H2, which distinguishes this from Lewis B.²¹ Because this fragment is of the same structure as that from H2, the fragmentation pathway, through the m/z 834 ion, is of great importance in assigning this structure.

We were not able to detect any core-fucosylated glycoforms in Hp by LC-MS/MS. The low abundant core-fucosylated

glycoform detected by MSn (see above) is either a minor component below the detection limit of our LC-MS/MS analysis or derives from a trace-contaminating protein. We cannot exclude that low percentage of core fucose exists especially on the doubly fucosylated site, but it is clear that core-fucosylated glycoforms represent at most a very minor component of the Hp glycoforms.

Hemopexin: Hpx is a heme-binding 60 kDa plasma glycoprotein containing five N-glycosylation sites and one N-terminal O-glycosylation site. Its major biological role is defense against hemoglobin-mediated oxidative damage during intra-

Table 2. Ratios of Ion Area Counts of Fucosylated versus Non-Fucosylated Glycoforms for Each Glycopeptide

Glycopeptide	Glycoform	Ratio of Fuc/nonfuc	SD	Glycopeptide	Glycoform	Ratio of Fuc/nonfuc	SD
Haptoglobin				Hemopexin			
VV.....LIK (218-233) – T3	A2G2F1	0.01	0.00	CSD.....FFK (27-48)	A3G3F1	2.21	0.08
	A3G3F1	1.41	0.19		A3G3F2	0.38	0.04
	A3G3F2	0.30	0.03	Kininogen-1			
	A3G3F3	0.02	0.01	ITY.....CSK(197-208)	A3G3F1	1.30	0.22
	A3G3F4	0.01	0.00		A2G2F1	0.09	0.02
	A4G4F1	2.81	0.46		A3G3F1	1.83	0.04
	A4G4F2	3.36	0.62		A3G3F2	0.52	0.31
	A4G4F3	1.25	0.07	HGI.....VKR (162-184)	A3G3F1	2.49	0.12
	A4G4F4	0.26	0.01		A3G3F2	0.19	0.03
	A4G4F5	0.09	0.01		A4G4F1	6.51	0.18
	A4G4F6	0.02	0.00		A4G4F2	5.75	1.38
MV.....TAK (161-184) – T1	A2G2F1	0.04	0.01		A4G4F3	2.42	0.16
	A3G3F1	2.18	0.22		A4G4F4	0.66	0.09
	A3G3F2	0.10	0.01		A4G4F5	0.19	0.02
	A4G4F1	3.12	1.50	Complement factor H			
	A4G4F2	2.13	0.86	MDG.....NSR (1024-1036)	A2G2F1	0.20	0.01
	A4G4F3	0.73	0.14	ISE.....MGK (907-919)	A2G2F1	0.25	0.03
NLF.....TAK (185-197) – T2	A2G2+A2G2F1	0.16	0.08		A3G3F1	1.71	0.25
	A3G3+A2G2F1	1.20	0.54		A3G3F2	0.13	0.01
Hemopexin				IPC.....SSR (868-885)	A2G2F1	0.23	0.01
SWP.....ALR (158-170)	A2G2F1	0.05	0.01		A3G3F1	3.23	0.01
	A3G3F1	1.81	0.13		A3G3F2	0.42	0.03
	A3G3F2	0.18	0.01	WD.....NYR (797-830)	A2G2+A3G3F1	0.53	0.01
	A4G4F1	5.91	1.30		A2G2+A3G3F2	0.25	0.01
	A4G4F2	7.21	2.58	SPY.....QCK (1079-1103)	A2G2F1	0.20	0.01
	A4G4F3	4.45	1.58		A3G3F1	0.86	0.49
ALP.....CTH (424-439)	A2G2F1	0.05	0.02	LND.....YER (528-567)	A2G2F1	0.09	0.06
	A3G3F1	1.62	0.09		A3G3F1	1.19	0.14
	A3G3F2	0.11	0.04	IQC.....GHR (671-729)	A2G2F1	0.34	0.05
CSD.....FFK (27-48)	A2G2F1	0.08	0.01		A3G3F1	0.63	0.13

vascular hemolysis.²⁸ We have analyzed microheterogeneity of three tryptic N-glycopeptides of Hpx. Each glycopeptide is occupied by complex glycans, primarily bi- and tri-antennary glycoforms with up to two fucoses. Bi-antennary glycans dominate, and their fucosylated counterparts represent a minor (<10%) contribution (Table 1). One tetra-antennary glycoform with three fucoses was observed at SWPAVGNCSALR peptide, but multiply fucosylated and highly branched structures are rare on Hpx, contrary to Hp. Although the ion intensities of most singly fucosylated bi-antennary glycopeptides were not sufficient for fragmentation, the measurable CID spectra contain low-intensity peptide-HexNAc-Fuc fragment consistent with core fucosylation. The presence of oxonium ions at m/z 512.2 indicates the presence of outer-arm fucose. The intensity of fragments, which could distinguish fucose linkage in doubly and triply fucosylated glycoforms, was under the detection limit of our instrument. Fragmentation spectra of SWPAVGNCSALR glycopeptide bearing A3G3F1 and A3G3F2 glycans contain both oxonium ion m/z 512.2 consistent with Lewis X type linkage. The presence of oxonium ion at m/z 658.2 in fragmentation spectrum of the A3G3F2 glycopeptide could suggest the presence of Lewis-Y-type structure, but rearrangement of fucose at outer arms is a likely event;²⁹ independent methods would have to confirm such linkage assignment. We were not able to confirm Lewis Y linkage by MSn analysis and conclude that rearrangement of the Lewis X structures is more likely in this case. This fucose rearrangement, however, was not described to happen between core and outer arms in tri- and tetra-antennary glycans, and we can conclude that core fucosylation is minor and below the limit of our detection in the tri-antennary structures.²⁹ These assignments are quite consistent with analysis of N-glycans detached from Hpx in a study of cirrhotic and HCC patients.¹⁴ Debruyne et al. observed intense bi-antennary glycan with small contribution of fucosylation and minor tri- and tetra-antennary glycoforms compared with the dominant bi-antennary structure. The authors found that outer-arm fucosylated tri- and tetra-antennary glycans are increased in HCC compared with bi-

antennary glycan without fucose and suggest that their ratio can be used as a diagnostic test for detection of HCC.

MALDI-ToF spectrum of glycans released from isolated Hpx is dominated by bi-antennary structures (Supplemental Figure 1 in the Supporting Information). Disassembly of a bi-antennary, monosialylated, monofucosylated composition shows a mixture of core fucosylation, Lewis X, and sialylated Lewis X-containing isomers (Supplemental Figure 4 in the Supporting Information).

Kininogen-1: Kng-1, a plasma glycoprotein of 644 amino acids, contains a heavy chain with four N-glycosylation sites.^{30,31} Serine proteases cleave from Kng-1 bradykinin, a nine amino acid peptide that regulates many biological processes including inflammation, angiogenesis, and cell migration.³² We have identified bi- and tri-antennary glycans and their fucosylated forms at three of four N-glycosylation sites of Kng-1. The glycopeptide YNSQNQSNQFVLYR was observed only with single glycoform A2G2. However, the intensity of this peptide is very low, and we cannot exclude the presence of other glycoforms below the detection limit of our method. Two glycopeptides carry tri-antennary glycoforms with up to two fucoses per glycan; moreover, glycopeptide HGIQYFNNNTQHSSLFMLNEVKR carries tetra-antennary glycans with up to five fucoses (Figure 3A). This is similar to the multiply fucosylated glycoforms of Hp and in contrast with the limited fucosylation of glycoforms associated with Hpx and CFH. Figure 3B shows a representative CID fragmentation spectrum of triply fucosylated tetra-antennary glycoform of HGIQYFNNNTQHSSLFMLNEVKR peptide with glycan fragments consistent with Lewis X structure at m/z 512.2. There is no evidence of core fucosylation in Kng-1 glycopeptides based on their CID fragmentation. Unfortunately, we were not able to complete MSn structural analysis of Kng-1's glycans because the protein was not available in sufficient amount and purity.

Complement factor H: CFH is a 140 kDa (without carbohydrate moiety) plasma glycoprotein that regulates enzymatic activity of complement C3.³³ CFH has nine N-glycosylation sequons and is primarily secreted to plasma by the liver even though a membrane bound form exists as well.³⁴ We

have detected seven glycopeptides in the tryptic digest of CFH which cover eight sequons (one peptide is doubly glycosylated). All peptides carry a limited variety of glycoforms of the complex type; high mannose or hybrid glycoforms were not detected. We have observed dominant bi-antennary glycoforms on all sequons except peptide L(528)-R(567), which carries more intense tri-antennary glycoforms. The bi-antennary glycoforms are singly fucosylated, but the intensity of the fucosylated glycoforms is somewhat higher (up to 20%) than for Hpx (<10%). Similar to Hpx, the singly fucosylated tri-antennary glycoforms tend to have higher intensity than their nonfucosylated counterparts. We have detected the minimal amount of doubly fucosylated glycoforms; tetra-antennary glycoforms and glycoforms with more than two fucoses were not detected (Table 2). Structural analysis of detached permethylated glycans uncovered the presence of core-fucosylated glycans represented by Fuc-GlcNAc fragment at m/z 490 (Supplemental Figure 3 in the Supporting Information). The intensity of this fragment is roughly similar to the fragment at m/z 660 corresponding to LeX epitope. This is in contrast with other analyzed singly fucosylated glycoforms from Hpx and Hp, where the 660 fragment was dominant, and the 490 fragment was small but detectable. These findings demonstrate that the process of glycan branching and fucosylation is site-specific and differs significantly between different glycoproteins isolated from the same patient.

CONCLUSIONS

Our aim was to examine whether hyper-fucosylation of Hp is associated with hyper-fucosylation of other proteins isolated from serum of the same HCC patient. Our analysis focused on fucosylation because we observe only complex N-glycans on all isolated proteins examined in this study and because we have previously observed hyper-fucosylated structures on Hp.^{10,12} Comparison of the protein and site-specific heterogeneity of the glycoforms shows interesting general trends. Distribution of fucosylated glycoforms is protein- and site-specific (Table 1). Bi-antennary glycan without fucose is typically the major glycoform observed. Singly fucosylated bi-antennary glycoforms are absent or low-intensity compared with the nonfucosylated bi-antennary forms in almost all studied glycoproteins. The highest intensity of bi-antennary glycoforms was found on CFH, where fucosylated glycoforms reach up to 20% ion intensity of the nonfucosylated glycoforms (Table 1). Singly fucosylated tri- and tetra-antennary glycoforms have higher intensity than their nonfucosylated forms (Table 2). Tri-antennary singly fucosylated glycans are the most intense glycoforms across virtually all identified tri-antennary glycopeptides. Tetra-antennary glycans are, in general, less abundant and observed on select glycopeptides only.

The distribution of multifucosylated tri- and tetra-antennary glycoforms differs by protein and peptide. Hyper-fucosylated glycoforms with more fucoses than arms were observed only in Hp (up to six fucoses per glycopeptide) and Kng-1 (five fucoses per glycopeptide).¹⁰ In contrast with Hp and Kng-1, we have detected at most two fucoses on three of eight glycopeptides of CFH and only one glycopeptide with three fucoses in the case of Hpx (A4G4F3 on SWPAVGNCSALR peptide). Structural analysis of detached permethylated glycans shows differences in level of core fucosylation between the glycoproteins. For example, in both Hp and Hpx, core fucosylation is less abundant compared with CFH. Furthermore, it was observed that doubly fucosylated glycoforms of Hp contain both LewisX and LewisY epitopes. This shows that factors other than up-regulated glycosyltransferases define how and which glycosite is fucosy-

lated. The determinants of site-specific microheterogeneity are not entirely known to our knowledge, but it is expected that the types of mature glycoforms detectable in disease context depend not only on the activity of glycosyltransferases but also on sequence of their action, distribution of the synthesized glycoproteins, their stability and turnover, and also on protein structure.

N-Glycosylation is a cotranslational process that originates in the ER, but the N-glycan structures continue to mature on folded proteins in Golgi apparatus.³⁵ The protein's secondary and tertiary structure determines the interaction with binding partners and accessibility of glycosyltransferases responsible for fucosylation, sialylation, and branching.³⁶ A recent study correlated structural attributes of the protein substrates with glycan microheterogeneity.³⁷ The authors suggest that protein structure (e.g., solvent accessibility, protein size, and hydrophobicity index) affects glycan microheterogeneity. This is in agreement with our observation that glycoproteins isolated from serum of the same patient are differentially fucosylated, although we assume that the repertoire of glycan biosynthetic machinery is the same for all of the glycoproteins being characterized. We cannot exclude the possibility that some glycoforms are selectively removed from circulation, but it is hard to believe that all tri- and tetra-antennary fucosylated glycoforms of CFH or Hpx would be so efficiently eliminated. In addition, differences in site-specific distribution of glycoform on each protein support influence of other factors (Table 1). For example, Kng-1 glycopeptide Y(44)-R(58) carries only one bi-antennary glycan compared with the tetra-antennary multiply fucosylated glycans on the neighboring peptide. It is more likely that the selection process is defined by local structure-recognition, which regulates the activity of glycosyltransferases at specific sequons.

ASSOCIATED CONTENT

Supporting Information

Supplemental Figure 1: Reduced and permethylated N-glycans from haptoglobin, hemopexin, and complement factor H were profiled using MALDI-ToF. Supplemental Table 1: List of glycan compositions observed in MALDI-ToF analysis of reduced and permethylated N-glycans from haptoglobin, hemopexin, and complement factor H with their relative distribution. Supplemental Figure 2: Detailed structure of haptoglobin N-glycans m/z 1246⁽³⁺⁾, showing a mixture of isomers, including Lewis X, H2, and Lewis Y epitopes. Supplemental Figure 3: Detailed analysis of biantennary monosialylated, monofucosylated N-glycan from complement factor H. Supplemental Figure 4: Detailed analysis of biantennary monosialylated, monofucosylated N-glycan from hemopexin. Supplemental Figure 5: Structures of the B-type fragment ions for Lewis X, H2, and Lewis Y with diagnostically important fragment ions assigned. Peak lists for MSⁿ mass spectra. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*Tel: 202-687-9868. Fax: 202-687-1988. E-mail: rg26@georgetown.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by NCI's U01 CA168926 and RO1 CA135069 awarded to R.G. and CCSG Grant P30 CA51008 (to Lombardi Comprehensive Cancer Center supporting the Proteomics and Metabolomics Shared Resource). P.P. was supported by AMVIS Czech Republic-U.S. exchange program LH13051 and by Grant Agency of the Czech Republic (Grant P206/12/0503), Charles University (ProjectUNCE_204025/2012). D.J.A. was supported by a partnership arrangement between the The UNH Glycomics Center and Glycan Connections, LLC, Lee, NH.

REFERENCES

- (1) Kornfeld, R.; Kornfeld, S. Assembly of asparagine-linked oligosaccharides. *Annu. Rev. Biochem.* **1985**, *54*, 631–664.
- (2) Roth, J.; Taatjes, D. J.; Lucocq, J. M.; Weinstein, J.; Paulson, J. C. Demonstration of an extensive trans-tubular network continuous with the Golgi apparatus stack that may function in glycosylation. *Cell* **1985**, *43* (1), 287–295.
- (3) Lairson, L. L.; Henrissat, B.; Davies, G. J.; Withers, S. G. Glycosyltransferases: structures, functions, and mechanisms. *Annu. Rev. Biochem.* **2008**, *77*, 521–555.
- (4) West, M. B.; Segu, Z. M.; Feasley, C. L.; Kang, P.; Klouckova, I.; Li, C.; Novotny, M. V.; West, C. M.; Mechref, Y.; Hanigan, M. H. Analysis of site-specific glycosylation of renal and hepatic γ -glutamyl transpeptidase from normal human tissue. *J. Biol. Chem.* **2010**, *285* (38), 29511–29524.
- (5) Takeda, Y.; Shinzaki, S.; Okudo, K.; Moriwaki, K.; Murata, K.; Miyoshi, E. Fucosylated haptoglobin is a novel type of cancer biomarker linked to the prognosis after an operation in colorectal cancer. *Cancer* **2012**, *118* (12), 3036–3043.
- (6) Mehta, A. S.; Norton, P.; Liang, H.; Comunale, M. A.; Wang, M.; Rodemich-Betesh, L.; Koszycki, A.; Noda, K.; Miyoshi, E.; Block, T. Increased levels of tetra-antennary N-linked glycan but not core fucosylation are associated with hepatocellular carcinoma tissue. *Cancer Epidemiol., Biomarkers Prev.* **2012**, *21* (6), 925–933.
- (7) Miyoshi, E.; Nakano, M. Fucosylated haptoglobin is a novel marker for pancreatic cancer: detailed analyses of oligosaccharide structures. *Proteomics* **2008**, *8* (16), 3257–3262.
- (8) Anderson, N.; Pollacchi, A.; Hayes, P.; Therapondos, G.; Newsome, P.; Boyter, A.; Smith, K. A preliminary evaluation of the differences in the glycosylation of alpha-1-acid glycoprotein between individual liver diseases. *Biomed. Chromatogr.* **2002**, *16* (6), 365–372.
- (9) Fang, M.; Dewaele, S.; Zhao, Y. P.; Starkel, P.; Vanhooren, V.; Chen, Y. M.; Ji, X.; Luo, M.; Sun, B. M.; Horsmans, Y.; Dell, A.; Haslam, S. M.; Grassi, P.; Libert, C.; Gao, C. F.; Chen, C. C. Serum N-glycome biomarker for monitoring development of DENA-induced hepatocellular carcinoma in rat. *Mol. Cancer* **2010**, *9*, 215.
- (10) Pompach, P.; Brnakova, Z.; Sanda, M.; Wu, J.; Edwards, N.; Goldman, R. Site specific glycoforms of haptoglobin in liver cirrhosis and hepatocellular carcinoma. *Mol. Cell. Proteomics* **2013**, *12* (5), 1281–1293.
- (11) Kaji, H.; Ocho, M.; Togayachi, A.; Kuno, A.; Sogabe, M.; Ohkura, T.; Nozaki, H.; Angata, T.; Chiba, Y.; Ozaki, H.; Hirabayashi, J.; Tanaka, Y.; Mizokami, M.; Ikehara, Y.; Narimatsu, H. Glycoproteomic discovery of serological biomarker candidates for HCV/HBV infection-associated liver fibrosis and hepatocellular carcinoma. *J. Proteome. Res.* **2013**, *12* (6), 2630–2640.
- (12) Sanda, M.; Pompach, P.; Brnakova, Z.; Wu, J.; Makambi, K.; Goldman, R. Quantitative LC-MS-MRM analysis of site-specific glycoforms of haptoglobin in liver disease. *Mol. Cell. Proteomics* **2013**, *12* (5), 1294–1305.
- (13) Comunale, M. A.; Rodemich-Betesh, L.; Hafner, J.; Wang, M.; Norton, P.; Di Bisceglie, A. M.; Block, T.; Mehta, A. Linkage specific fucosylation of alpha-1-antitrypsin in liver cirrhosis and cancer patients: implications for a biomarker of hepatocellular carcinoma. *PLoS One* **2010**, *5* (8), e12419.
- (14) Debruyne, E. N.; Vanderschaeghe, D.; Van, V. H.; Vanhecke, A.; Callewaert, N.; Delanghe, J. R. Diagnostic value of the hemopexin N-glycan profile in hepatocellular carcinoma patients. *Physiol. Rev.* **2010**, *56* (5), 823–831.
- (15) Mori, S.; Aoyagi, Y.; Yanagi, M.; Suzuki, Y.; Asakura, H. Serum N-acetylglucosaminyltransferase III activities in hepatocellular carcinoma. *J. Gastroenterol. Hepatol.* **1998**, *13* (6), 610–619.
- (16) Tsumumi, M.; Wang, J. S.; Takada, A. Microheterogeneity of serum glycoproteins in alcoholics: is desialo-transferrin the marker of chronic alcohol drinking or alcoholic liver injury? *Alcohol. Clin. Exp. Res.* **1994**, *18* (2), 392–397.
- (17) Campion, B.; Leger, D.; Wieruszkeski, J. M.; Montreuil, J.; Spik, G. Presence of fucosylated triantennary, tetraantennary and pentaantennary glycans in transferrin synthesized by the human hepatocarcinoma cell line Hep G2. *Eur. J. Biochem.* **1989**, *184* (2), 405–413.
- (18) Comunale, M. A.; Wang, M.; Hafner, J.; Krakover, J.; Rodemich, L.; Kopenhaver, B.; Long, R. E.; Junaidi, O.; Bisceglie, A. M.; Block, T. M.; Mehta, A. S. Identification and Development of Fucosylated Glycoproteins as Biomarkers of Primary Hepatocellular Carcinoma. *J. Proteome Res.* **2009**, *8* (2), 595–602.
- (19) Saldova, R.; Royle, L.; Radcliffe, C. M.; Abd Hamid, U. M.; Evans, R.; Arnold, J. N.; Banks, R. E.; Hutson, R.; Harvey, D. J.; Antrobus, R.; Petrescu, S. M.; Dwek, R. A.; Rudd, P. M. Ovarian cancer is associated with changes in glycosylation in both acute-phase proteins and IgG. *Glycobiology* **2007**, *17* (12), 1344–1356.
- (20) Bones, J.; Byrne, J. C.; O'Donoghue, N.; McManus, C.; Scaife, C.; Boissin, H.; Nastase, A.; Rudd, P. M. Glycomic and glycoproteomic analysis of serum from patients with stomach cancer reveals potential markers arising from host defense response mechanisms. *J. Proteome. Res.* **2011**, *10* (3), 1246–1265.
- (21) Ashline, D. J.; Hanneman, A. J.; Zhang, H.; Reinhold, V. N. Structural documentation of glycan epitopes: sequential mass spectrometry and spectral matching. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (3), 444–453.
- (22) Alley, W. R., Jr.; Madera, M.; Mechref, Y.; Novotny, M. V. Chip-based reversed-phase liquid chromatography-mass spectrometry of permethylated N-linked glycans: a potential methodology for cancer biomarker discovery. *Anal. Chem.* **2010**, *82* (12), S095–S106.
- (23) Kang, P.; Mechref, Y.; Klouckova, I.; Novotny, M. V. Solid-phase permethylation of glycans for mass spectrometric analysis. *Rapid Commun. Mass Spectrom.* **2005**, *19* (23), 3421–3428.
- (24) Pompach, P.; Chandler, K. B.; Lan, R.; Edwards, N.; Goldman, R. Semi-automated identification of N-Glycopeptides by hydrophilic interaction chromatography, nano-reverse-phase LC-MS/MS, and glycan database search. *J. Proteome. Res.* **2012**, *11* (3), 1728–1740.
- (25) Chandler, K. B.; Pompach, P.; Goldman, R.; Edwards, N. Exploring Site-Specific N-Glycosylation Microheterogeneity of Haptoglobin Using Glycopeptide CID Tandem Mass Spectra and Glycan Database Search. *J. Proteome. Res.* **2013**, *12* (8), 3652–3666.
- (26) Royle, L.; Campbell, M. P.; Radcliffe, C. M.; White, D. M.; Harvey, D. J.; Abrahams, J. L.; Kim, Y. G.; Henry, G. W.; Shadick, N. A.; Weinblatt, M. E.; Lee, D. M.; Rudd, P. M.; Dwek, R. A. HPLC-based analysis of serum N-glycans on a 96-well plate platform with dedicated database software. *Anal. Biochem.* **2008**, *376* (1), 1–12.
- (27) Koide, T.; Odani, S. Histidine-rich glycoprotein is evolutionarily related to the cystatin superfamily. Presence of two cystatin domains in the N-terminal region. *FEBS Lett.* **1987**, *216* (1), 17–21.
- (28) Delanghe, J. R.; Langlois, M. R. Hemopexin: a review of biological aspects and the role in laboratory medicine. *Clin. Chim. Acta* **2001**, *312* (1–2), 13–23.
- (29) Wuhler, M.; Koeleman, C. A.; Hokke, C. H.; Deelder, A. M. Mass spectrometry of proton adducts of fucosylated N-glycans: fucose transfer between antennae gives rise to misleading fragments. *Rapid Commun. Mass Spectrom.* **2006**, *20* (11), 1747–1754.
- (30) Kellermann, J.; Lottspeich, F.; Henschen, A.; Muller-Esterl, W. Completion of the primary structure of human high-molecular-mass kininogen. The amino acid sequence of the entire heavy chain and evidence for its evolution by gene triplication. *Eur. J. Biochem.* **1986**, *154* (2), 471–478.

- (31) Lottspeich, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W. The amino acid sequence of the light chain of human high-molecular-mass kininogen. *Eur. J. Biochem.* **1985**, *152* (2), 307–314.
- (32) Dempsey, E.; Rudd, P. M. Acute phase glycoproteins: bystanders or participants in carcinogenesis? *Ann. N.Y. Acad. Sci.* **2012**, *1253*, 122–132.
- (33) Whaley, K.; Ruddy, S. Modulation of the alternative complement pathways by beta 1 H globulin. *J. Exp. Med.* **1976**, *144* (5), 1147–1163.
- (34) Ripoché, J.; Day, A. J.; Harris, T. J.; Sim, R. B. The complete amino acid sequence of human complement factor H. *Biochem. J.* **1988**, *249* (2), 593–602.
- (35) Aebi, M. N-linked protein glycosylation in the ER. *Biochim. Biophys. Acta* **2013**, *1833* (11), 2430–2437.
- (36) Parodi, A. J. Protein glucosylation and its role in protein folding. *Annu. Rev. Biochem.* **2000**, *69*, 69–93.
- (37) Thaysen-Andersen, M.; Packer, N. H. Site-specific glycoproteomics confirms that protein structure dictates formation of N-glycan type, core fucosylation and branching. *Glycobiology* **2012**, *22* (11), 1440–1452.