Journal of proteome research

Article

pubs.acs.org/jpr

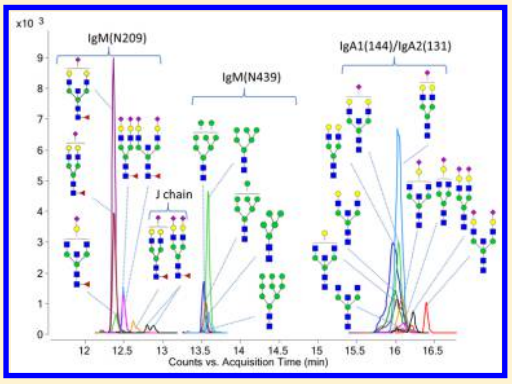
A Method for Comprehensive Glycosite-Mapping and Direct Quantitation of Serum Glycoproteins

Qiuting Hong,[†] L. Renee Ruhaak,[†] Carol Stroble,[†] Evan Parker,[†] Jincui Huang,[†] Emanual Maverakis,[‡] and Carlito B. Lebrilla*[,†]

[†]Department of Chemistry and [‡]Department of Dermatology, School of Medicine, University of California, Davis, California 95616, United States

Supporting Information

ABSTRACT: A comprehensive glycan map was constructed for the top eight abundant glycoproteins in plasma using both specific and nonspecific enzyme digestions followed by nano liquid chromatography (LC)−chip/quadrupole time-of-flight mass spectrometry (MS) analysis. Glycopeptides were identified using an in-house software tool, GPFinder. A sensitive and reproducible multiple reaction monitoring (MRM) technique on a triple quadrupole MS was developed and applied to quantify immunoglobulins G, A, M, and their site-specific glycans simultaneously and directly from human serum/plasma without protein enrichments. A total of 64 glycopeptides and 15 peptides were monitored for IgG, IgA, and IgM in a 20 min ultra high performance (UP)LC gradient. The absolute protein contents were quantified using peptide calibration curves. The glycopeptide ion abundances were normalized to the respective protein abundances to separate protein glycosylation from protein expression. This technique yields higher method reproducibility and less sample loss when compared with the quantitation method that involves protein enrichments. The absolute protein quantitation has a wide linear range (3−4 orders of magnitude) and low limit of quantitation (femtomole level). This rapid and robust quantitation technique, which provides quantitative information for both proteins and glycosylation, will further facilitate disease biomarker discoveries.



KEYWORDS: *immunoglobulin, absolute quantitation, glycopeptide, site specific glycan analysis*

INTRODUCTION

Changes in post-translation modifications such as glycosylation are the hallmark of many common diseases.[1−4] The most extensively studied example is the lack of galactosylation of IgG in several autoimmune diseases such as rheumatoid arthritis, systemic lupus erythematosus, and Crohn's disease.[5,6] Another example is the aberrant O-linked galactosylation and N-linked sialylation of serum IgA1 in primary IgA nephropathy and primary Sjögren's syndrome, respectively.[7−10] Monitoring protein specific glycosylation remains an active goal for monitoring diseases because glycosylation is highly sensitive to local and global biological changes.[11]

However, site-specific glycosylation analysis is a challenging task that involves the determination of both the glycosylation site and the glycan microheterogeneity. One popular approach is to monitor the global changes in glycosylation by releasing glycans from a protein mixture.[12−15] The released glycan analysis however provides no protein-specific information. Additionally, quantitation of glycosylation on the site-specific level is of great interest, but it is currently limited to gross comparisons due to the lack of glycan/glycopeptide standards. For relative comparison, glycan ion abundances are usually normalized to the most abundant glycan or the total glycan ion abundance for relative comparison.[1,16] Because glycan abundances in biological fluids are greatly affected by protein concentration, direct comparison of released glycan abundance may not yield information related to protein expression. Indeed, it has been shown that proteins may decrease while specific glycoforms increase in some diseases states.[17] Thus, it is necessary to monitor protein content and degree of glycosylation simultaneously for disease biomarker discovery.

Site-specific glycosylation of many of the abundant glycoproteins in serum has been explored using different approaches by several groups.[16,18−28] A typical approach is the liquid chromatography coupled with tandem mass spectrometry (LC−MS/MS) analysis of glycopeptides yielded by specific proteases.[18,24] An improvement involves enrichment of specifically glycosylated peptides followed by glycopeptide or released glycan and peptide analysis.[29,30] However, often these studies yield mainly partial glycan heterogeneity or incomplete site-specific information. For this reason, the site specific glycan heterogeneity of even common and abundant immunoglobulins with multiple sites of glycosylation such as IgA and IgM is not well fully described.[29,31,32]

Received: June 1, 2015

DOI: 10.1021/acs.jproteome.5b00756
J. Proteome Res. XXXX, XXX, XXX−XXX

The large interest in glycomic and glycoproteomic biomarkers necessitates the glycan maps of proteins. In this research, we begin this effort by constructing comprehensive glycan maps of the common glycoproteins in serum/plasma where each site is characterized for glycan heterogeneity. Both specific and nonspecific proteases, trypsin and Pronase E, were applied to obtain comprehensive site-specific N-glycan maps of the top abundant serum glycoproteins. An in-house software tool, GPFinder, was used to assign glycopeptides based on the LC−MS/MS analysis.[33] The specificity of trypsin yields well-defined peptides suitable for quantitation but may provide limited glycosite information mainly due to large peptide size or presence of multiple glycosites. Nonspecific proteases usually produce better glycosite coverage but yield variable peptide sequences that complicate the data analysis and limit its application to quantitation. Combining these two approaches produces a more comprehensive glycan map with extensive site-specific heterogeneity. The tryptic map is then used to set the multiple reaction monitoring (MRM) conditions to monitor site-specific glycoforms. Nonspecific glycan map is not used for quantitation due to the variable peptide sequence.

MRM is well-known for its sensitivity and specificity in quantitation. Yet, this powerful technique has been used on limited bases to examine glycopeptides.[34−38] In our previous study, we employed MRM to monitor site-specific IgG glycosylation. To decouple glycan and protein expression, the glycan abundances were normalized to protein concentration that represented by its peptide abundance.[39] In the present application, we significantly expanded this method to IgG, IgA, and IgM. While the three proteins represent a limited and focused application, they still represent the most extensive use of MRM on glycopeptides to date. Furthermore, this MRM method can be readily implemented to a much larger group of serum proteins utilizing the resulting glycan maps. The method for monitoring site-specific glycosylation is attractive because it is especially fast with ultra high-performance (UP)LC separation and convenient with little sample pretreatment. It will find immediate applications in monitoring and diagnosing diseases.

EXPERIMENTAL PROCEDURES

The experimental procedures are briefly described here. A detailed version can be found in the Supporting Information.

Chemicals and Reagents

Immunoglobulin G and M, transferrin, haptoglobin, alpha-2-macroglobulin, alpha-1-antitrypsin, complement C3, and alpha-1-acid glycoprotein purified from human serum/plasma were purchased from Sigma-Aldrich (St. Louis, MO). Immunoglobulin A purified from human plasma was purchased from CalBiochem (La Jolla, CA). Antihuman-IgA (α-chain specific)-agarose and antihuman-IgM (μ-chain specific)-agarose used for protein enrichment were purchased from Sigma-Aldrich (St. Louis, MO). Sequencing grade modified trypsin and dithiothreitol (DTT) were purchased from Promega (Madison, WI). Pronase E from *Streptomyces griseus* and iodoacetamide (IAA) were purchased from Sigma-Aldrich (St. Louis, MO).

Plasma/Serum Sample Collection

Healthy volunteers (n = 13) were recruited during routine outpatient clinic visits or through approved flyers. Following venipuncture and blood sampling, plasma was isolated via use of a Ficoll-Paque (p = 1.077 g/mL) gradient. All research was approved by the UC Davis Institutional Review Board (IRB). One pooled human serum sample was purchased from Sigma-Aldrich (St. Louis, MO) for method development. All serum/plasma samples were stored at −80 °C prior to sample preparation.

Protein Enrichment and Proteolysis

The sample treatment is described briefly here, and more details can be found in the Supporting Information. Briefly, amount of 10 μL of antihuman-IgA (α-chain specific)-agarose and 10 μL of antihuman-IgM (μ-chain specific)-agarose were applied to enrich IgA and IgM proteins from 5 μL of serum/plasma. Both native serum samples and enriched proteins were treated using Pronase E and trypsin for glycan mapping. The tryptic digests were also used for quantitation.

LC−MS Analysis

The glycan mapping was performed on an Agilent 1200 series HPLC-Chip system coupled to an Agilent 6520 quadrupole time-of-flight (Q-TOF) (Agilent Technologies, Santa Clara, CA). Agilent Zorbax C18 stationary phase (300 Å, 5 μm) and porous graphitized carbon (250 Å, 5 μm) stationary phase were used to separate the trypsin digest and Pronase digest, respectively. The microfluidic C18 chip consists of two columns: one for enrichment (4 mm, 40 nL) and one for separation (150 mm × 75 μm). The microfluidic PGC chip also consists of one enrichment column (4 mm, 40 nL) and one separation column (43 mm × 75 μm).

The quantitative analyses were performed on an Agilent 1290 infinity LC system coupled to an Agilent 6490 triple quadrupole (QqQ) mass spectrometer (Agilent Technologies, Santa Clara, CA). An Agilent Eclipse plus C18 (RRHD 1.8 μm, 2.1 × 100 mm^2) coupled with an Agilent Eclipse plus C18 precolumn (RRHD 1.8 μm, 2.1 × 5 mm^2) was used for UPLC separation.

Data Analysis

An in-house software, GPFinder,[33] was applied to identify glycopeptides based on the glycan fragmentation and accurate mass. Briefly, common glycan oxonium fragments, such as m/z 204.08 (HexNAc), 366.14 ($Hex_1HexNAc_1$), 292.09 (Neu5Ac), and 657.24 ($Hex_1HexNAc_1Neu5Ac_1$), were used to identify glycopeptide candidates from the complete peptide ion list. The glycopeptide candidate ion list was matched to the theoretical glycopeptide masses, which were generated from an in-house serum N-glycan library and the protein sequence.[40] A score was generated to evaluate the confidence of each match on the basis of the mass accuracy and glycan fragmentation pattern. A glycopeptide list with known glycan and peptide compositions was thus obtained using a 5% false discovery rate as the cutoff (confidence level = 95%). MRM results were analyzed using Agilent MassHunter Quantitative Analysis B.5.0 software. The limit of detection (LOD) and limit of quantitation (LOQ) were defined as signal-to-noise ratio (S/N) ≥ 3 and 6, respectively.

RESULTS AND DISCUSSION

We present here a comprehensive N-glycan map for plasma glycoproteins utilizing both specific and nonspecific proteases and a targeted MRM method to quantify absolute protein content and the relative quantity of their glycosylation at the site-specific level simultaneously from serum/plasma samples. We select a subset of this group, the antibody proteins, for illustrative purposes. As shown in Supplementary Figure S-1, for transferrin and haptoglobin, this method can be further expanded to quantitate the other proteins described in this

DOI: 10.1021/acs.jproteome.5b00756
J. Proteome Res. XXXX, XXX, XXX−XXX

Table 1. Trypsin Glycan Map for Alpha-2 Macroglobulin

glycopeptide mass (Da)							
exp.	cal.	RT	glycosite	peptide sequence	Hex−HexNAc−Fuc−Neu5Ac	ion abundance[a]	relative[a]
4595.034	4595.086	28.7	869	SLGNVNFTVSAEALESQELCGTEVPSVPEHGRK	4−2−0−0	185585	6.4%
4757.104	4757.138	28.7	869	SLGNVNFTVSAEALESQELCGTEVPSVPEHGRK	5−2−0−0	2694326	93.6%
4075.803	4075.851	44.4	1424	VSNQTLSLFFTVLQDVPVR	5−4−0−1	1700837	16.6%
4221.902	4221.908	44.4	1424	VSNQTLSLFFTVLQDVPVR	5−4−1−1	2272376	22.2%
4366.905	4366.946	45.5	1424	VSNQTLSLFFTVLQDVPVR	5−4−0−2	5617431	54.8%
4512.947	4513.004	45.6	1424	VSNQTLSLFFTVLQDVPVR	5−4−1−2	668015	6.5%

[a]Absolute and relative ion abundance from LC−Q-TOF full scans.

study along with their site-specific glycosylation, although protein enrichment may be required to increase the signal of the lower abundant proteins. The quantitation of glycopeptides from IgG, IgA, and IgM is performed directly from 2 μL of serum/plasma without protein enrichments using a 20 min UPLC gradient. Sigma serum was used for method development, but the method also works well on plasma samples.

Comprehensive Glycan Map of Selected Glycoproteins within Plasma

Both specific protease (trypsin) and nonspecific protease (Pronase E) were employed to obtain a more comprehensive glycan map. The specificity of trypsin yields well-defined peptides suitable for quantitation but may provide limited glycosite information depending on the protein sequence. One reason is that multiple glycosites may be present on a single tryptic peptide making glycan assignment by tandem MS difficult. Possible missed cleavages may yield large peptide not readily characterized by collision-induced dissociation (CID). Conversely, nonspecific proteases usually yield more glycosite information and provide more glycopeptide isoform information.[41] However, the cocktail of nonspecific proteases may produce peptides that are too small and may not designate the site. Employing both allows more complete glycosite mapping.

Take alpha-2 macroglobulin (A2MG) as an example. With trypsin, only two out of eight glycosites were observed, that is, N869 and 1424 (Table 1). In a previous glycan map using another specific protease, chymotrypsin, three out of the eight glycosites, that is, site 70, 396, and 1424, were observed containing similar glycoforms identified in this study.[24] As listed in Table 2, the cocktail of Pronase provides better glycosite coverage yielding full glycan map for all the eight sites. High mannose, hybrid, and complex-type glycans were all observed on site 869, which is consistent with the previous finding.[42] Biantennary sialylated glycans were the most common glycoforms observed on other sites, while bisecting complex type glycans were also observed on site 70 and 991, and tri- and tetra- antennary types were also observed on site 247 and 410. The use of nonspecific proteases generally yielded more sites, but they sometimes yielded ambiguous glycopeptide assignments. When there is ambiguity in glycosite assignments, the results are not used to reduce false positives (Table 2). On occasion, trypsin may provide more site heterogeneity as shown for site 1424 in Table 1. Combining the two types of proteases, specific and nonspecific, provides a more complete glycan map for alpha-2 macroglobulin (Table 3).

Two other examples used for illustrative purposes are human IgA and IgM including their joining chain (Ig J chain). Human IgA has two subclasses: IgA1 and IgA2. IgA1 contains two N-glycosites (N144 and 340) on its Fc region, while IgA2 has five N-glycosites (N47, 92, 131, 205, and 327). IgM does not have subclasses, but it contains multiple glycosites (N46, 209, 272, 279, and 439) on the Fc region. Both immunoglobulins may also include a glycosylated joining (J) chain (18 kDa, UniProtKB ID: P01876), which helps them polymerization. Glycan mapping of these large protein complexes have generally been challenging, and only partial glycan maps were obtained in the previous studies. Table 3 summarizes the results obtained in this study and compared to the previous findings. As with the other proteins, nonspecific proteases yielded better glycosite coverage, while specific proteases provide further complementary information (Figure 1). Our glycan map shown in Figure 1 supports the previous finding that sites N279, 439 of IgM contain high mannose type glycans, while other sites contain complex type glycans (mono-, bi-, triantennary, bisecting) with more specific glycoform information provided using our comprehensive glycan map method.[32] The glycan map for IgA shows site N144 of IgA1 and N131 of IgA2 contains no fucosylation, while glycoforms on N205 of IgA2 and N340 of IgA1 are all fucosylated. Our results do not support plasma IgA2 glycosylation at N47, 92, and 327, as has been reported for secretory IgA2 in milk and saliva, suggesting variations between tissues.[43,44] The component IgA2 is much less abundant in plasma than IgA1 (IgA1:IgA2 = 9:1).[45] The lack of glycosylation in the above sites may be due to the low IgA2 concentrations and low glycosylation expressions on these sites. Furthermore, site N92 contains a proline residue nearby that may decrease the extent of glycosylation as described previously.[11,46]

The glycan map method was further expanded to transferrin, haptoglobin, alpha-2 macroglobulin, alpha-1-antitrypsin, complement C3, and alpha-1 acid glycoprotein from normal serum/plasma. The glycan map for these proteins including results from previous studies is summarized in Table 3. The glycosylation of IgG has been characterized in great detail previously and is not presented here.[16,28,47−49]

Protein and Site-Specific Glycan Quantitation Using MRM of Antibody Proteins

As shown in Table 3, glycoproteins have large N-glycan microheterogeneity. MRM is used to determine absolute protein amounts while monitoring site-specific glycosylation. For the MRM, we focus on the antibody proteins. However, the method can be applied to any or all of the proteins characterized in this study (Figure S-1) with protein enrichment to increase the signal of low abundant proteins and the specificity of quantitation.

Commercially available immunoglobulin standards were used to develop the MRM methods. The MRM transitions for IgG were obtained in a previous study.[39] Those used in this study are listed in Table S-1. The IgA and IgM tryptic glycopeptides were first profiled using nano-LC/QTOF, and their fragmenta-

DOI: 10.1021/acs.jproteome.5b00756
J. Proteome Res. XXXX, XXX, XXX−XXX

Table 2. Pronase E Glycan Map for Alpha-2 Macroglobulin[a]

glycopeptide mass (Da)					
exp.	cal.	RT	glycosite	peptide sequence	Hex−HexNAc−Fuc−Neu5Ac
3316.301	3316.312	37.3	55	SYLNETVTVS	5−4−0−2
2812.052	2812.043	25.8	55	NETV	5−4−1−2
3462.358	3462.370	37.6	55	SYLNETVTVS	5−4−1−2
2601.042	2601.055	23.2	70	SVRGNR	5−4−0−1
2706.063	2706.050	24.2	70	RGNR	5−4−0−2
2892.137	2892.150	24.7	70	SVRGNR (VRGNRS)	5−4−0−2
3134.26	3134.277	26.4	70	LESVRGNR	5−4−0−2
3308.343	3308.341	25.1	70	SLESVRGNRS	5−4−0−2
2747.102	2747.113	25.9	70	SVRGNR (VRGNRS)	5−4−1−1
2989.229	2989.239	26.9	70	LESVRGNR	5−4−1−1
3163.301	3163.303	24.0	70	SLESVRGNRS	5−4−1−1
2852.105	2852.108	27.0	70	RGNR	5−4−1−2
3038.19	3038.208	28.3	70	SVRGNR (VRGNRS)	5−4−1−2
2852.105	2852.096	26.8	70	VRGN	5−5−0−2
2458.896	2458.904	24.3	247	EEEMNVS	5−4−0−0
2749.99	2749.999	26.6	247	EEEMNVS	5−4−0−1
2855.007	2854.994	28.4	247	EEEMN	5−4−0−2
3041.084	3041.095	28.7	247	EEEMNVS	5−4−0−2
3803.496	3803.478	48.9	247	ILEEEMNVSVCG (LEEEMNVSVCGL)	6−5−1−1
2595.938	2595.943	21.4	396	SNAT	5−4−0−2
2922.051	2922.069	35.4	396	YYSNAT	5−4−0−2
3135.133	3135.144	25.0	396	EANYYSNA (SNATTDEHG)	5−4−0−2
3138.151	3138.144	50.2	396	YYSNATTD	5−4−0−2
2450.92	2450.905	19.3	396	SNAT	5−4−1−1
2579.887	2579.948	22.8	396	NATTD	5−4−1−1
2741.998	2742.001	25.2	396	SNAT	5−4−1−2
2870.976	2871.043	26.2	396	NATTD	5−4−1−2
2198.868	2198.857	18.8	410	INTT	4−4−0−1
2285.884	2285.889	20.2	410	SINTT	4−4−0−1
2360.911	2360.910	19.3	410	INTT	5−4−0−1
2447.952	2447.942	21.2	410	SINTT	5−4−0−1
2594.991	2595.010	27.5	410	FSINTT	5−4−0−1
2652.026	2652.005	23.9	410	INTT	5−4−0−2
2506.979	2506.968	23.6	410	INTT	5−4−1−1
4828.829	4828.841	49.9	410	QFSINTTNVMGT	7−6−0−4
1950.746	1950.731	18.4	869	GNVN	4−3−0−1
2153.816	2153.810	18.3	869	GNVN	4−4−0−1
2602.045	2602.043	29.6	869	NVNFTVSA (SLGNVNFT)	4−4−0−1
1618.616	1618.609	16.1	869	GNVN	5−2−0−0
2066.835	2066.841	27.4	869	NVNFTVSA(SLGNVNFT)	5−2−0−0
2112.789	2112.784	17.2	869	GNVN	5−3−0−1
2315.875	2315.863	19.1	869	GNVN	5−4−0−1
2515.972	2515.979	23.0	869	SLGNVN	5−4−0−1
2764.092	2764.096	29.6	869	NVNFTVSA (SLGNVNFT)	5−4−0−1
2483.902	2483.894	20.4	869	NF	5−4−0−2
2606.964	2606.959	23.4	869	GNVN	5−4−0−2
3055.179	3055.191	33.4	869	NVNFTVSA (SLGNVNFT)	5−4−0−2
2274.844	2274.837	18.5	869	GNVN	6−3−0−1
2723.063	2723.069	29.4	869	NVNFTVSA (SLGNVNFT)	6−3−0−1
2531.941	2531.938	22.6	991	NETQQ	5−4−0−1
2694.976	2694.975	23.4	991	NETQ	5−4−0−2
2823.053	2823.033	23.8	991	NETQQ	5−4−0−2
2936.129	2936.117	25.7	991	LNETQQ (NETQQL)	5−4−0−2
2958.068	2958.091	49.5	991	DYLNET	5−4−0−2
3170.236	3170.243	49.9	991	VLDYLNET	5−4−0−2
3895.589	3895.614	36.5	991	DYLNETQQLTPEIK	5−4−0−2
4107.748	4107.723	37.7	991	LFAPNIYVLDYLN	5−5−1−2
2704.068	2704.059	20.0	1424	DKVSNQT	5−4−0−1
2789.141	2789.148	19.8	1424	KVSNQTLS	5−4−0−1

DOI: 10.1021/acs.jproteome.5b00756
J. Proteome Res. XXXX, XXX, XXX−XXX

Table 2. continued

glycopeptide mass (Da)					
exp.	cal.	RT	glycosite	peptide sequence	Hex−HexNAc−Fuc−Neu5Ac
2565.93	2565.932	21.4	1424	NQT	5−4−0−2
2995.151	2995.155	23.3	1424	DKVSNQT	5−4−0−2
3080.232	3080.244	22.4	1424	KVSNQTLS	5−4−0−2
2850.12	2850.117	23.4	1424	DKVSNQT	5−4−1−1
2935.203	2935.206	23.7	1424	KVSNQTLS	5−4−1−1
3141.242	3141.212	24.2	1424	DKVSNQT	5−4−1−2
2144.800	2144.799	17.5	247 (869)	NV	5−4−0−1
2423.867	2423.858	19.4	396 (1424)	SN	5−4−0−2
2520.929	2520.922	20.4	396 (410)	SNATTD (QFSIN)	5−4−0−1
2812.052	2812.017	25.8	396 (410)	SNATTD (QFSIN)	5−4−0−2
2666.993	2666.980	22.2	396 (410)	SNATTD (QFSIN)	5−4−1−1
2958.079	2958.075	27.5	396 (410)	SNATTD (QFSIN)	5−4−1−2
3881.579	3881.562	36.0	396 (410)	NATTDEHGLVQFSI (ATTDEHGLVQFSIN)	5−4−1−2
2361.881	2361.869	17.5	410 (1424)	NTTN (SNQT)	5−4−0−1
2474.958	2474.953	20.1	410 (1424)	INTTN (DKVSN, SNQTL)	5−4−0−1
2652.977	2652.964	21.7	410 (1424)	NTTN (SNQT)	5−4−0−2
2507.939	2507.927	19.6	410 (1424)	NTTN (SNQT)	5−4−1−1
2275.839	2275.821	18.3	55 (991, 247)	NET (NET, MNV)	5−4−0−1
2566.912	2566.916	22.5	55 (991, 247)	NET (NET, MNV)	5−4−0−2
2421.888	2421.879	20.1	55 (991, 247)	NET (NET, MNV)	5−4−1−1
2712.978	2712.974	26.8	55 (991, 247)	NET (NET, MNV)	5−4−1−2
2549.952	2549.949	20.7	70 (869)	RGN (NVN)	5−4−0−2
2695.978	2695.995	25.4	70 (869)	RGN (NVN)	5−4−1−2
3426.349	3426.371	49.3	70(991)	GNRSLFTDLEA(VLDYLNETQQ)	5−4−0−2
2404.899	2404.900	18.4	869 (70)	NVN (RGN; GNR)	5−4−1−1

[a]No glycoform abundance is listed because the same glycoform is spread out on different peptide sequences.

tion behavior was characterized by CID. Shown in Figure 2, panels a and b are the tandem mass spectra of two selected glycopeptides. More examples can be found in Figure S-2. Small glycan fragment "oxonium" ions, such as m/z 204.08 (HexNAc), 366.14 ($Hex_1HexNAc_1$), 292.09 (Neu5Ac), and 657.24 ($Hex_1HexNAc_1Neu5Ac_1$), are the most common and abundant fragment ions of glycopeptides under CID. Because of its strong abundances in nearly all the glycopeptide tandem mass spectra, the fragment ion, m/z 366.1, was selected as the MRM product ion for the complex-type glycans. High mannose-type glycans do not produce the same glycan fragment ions in sufficient abundances (Figure 2c). Instead, they yield abundant peptide−GlcNAc fragment ions under CID. These fragment ions were thus monitored for the high mannose-type glycans. The collision conditions were optimized to yield the highest abundances of this species, and the MRM transitions for IgA and IgM are listed in Table 4.

Peptide MRM Transitions for Absolute Protein Quantitation. Tryptic peptides that are unique for each protein were monitored to quantify the proteins abundance in the protein mixture. Because the IgA standard used is a mixture of subclasses IgA1 and IgA2, common peptides, YLTWASR and WLQGSELPR, which are shared by two subclasses, were monitored to quantify the total IgA content. As listed in Table 4, other peptide candidates monitored for protein quantitation are TPLTATLSK and DASGVTFTWTPSSGK for IgA1, DASGATFTWTPSSGK for IgA2, YAATSQVLLPSK and FTCTVTHTDLPSPLK for IgM, and SSEDPNEDIVER and IIVPLNNR for the joining chain. It was found in our peptide map that IgM glycosite N439 is not fully occupied. Thus, its unglycosylated peptide, STGKPTLYNVSLVMSDTAGTCY, was also monitored to study the glycosylation occupancy for this glycosite. Fragmentation behaviors of these peptides under CID were studied, and some selected tandem spectra are shown in Figure S-2. The most abundant fragments, mostly b or y ions, were monitored to increase the MRM sensitivity.[39] The MRM transitions and their optimized collision energies for peptide quantitation are listed in Table 4.

Ultra High Performance LC (UPLC). An UPLC-QqQ method was developed to monitor IgG, IgA, and IgM simultaneously in serum/plasma without protein enrichments. The C18 stationary phase separates glycopeptides mainly on the basis of their peptide moieties (Figure 3), which provides some identity information, that is, same glycosites elute closely together. In this regard, dynamic MRM whereby transitions were monitored only within specified retention windows was applied. The largest differences in retention times are observed between neutral and acidic glycopeptides. For example, glycan 5410 (Hex:HexNAc:Fuc:Neu5Ac) on glycosite N205 of IgA2 elutes 0.6 min earlier than 5411 and 1.2 min earlier than 5412 (Figure S-3c,d). The elution behavior of each glycopeptide is an important parameter that provides glycosite information. Signals that do not elute accordingly are likely false positives and are not glycoforms of the peptide. The resulting total MRM chromatograms for an IgG, IgA, and IgM standard mixture (protein concentration = 0.300, 0.197, and 0.165 mg/mL, respectively) are depicted in Figure 3, panel a. The annotations for individual glycopeptides from both serum/plasma samples and protein standards are provided in Figure S-3. Although serum/plasma samples are very complicated mixtures, we observe consistent MRM chromatograms for the protein standards and serum/plasma digests.

DOI: 10.1021/acs.jproteome.5b00756
J. Proteome Res. XXXX, XXX, XXX−XXX

Table 3. Site-Specific Glycan Map for Top Eight Abundant Glycoproteins in Normal Serum[a]

protein (Uniprot ID)	glycosite	glycoform observed[b]	literature[c]
Transferrin (P02787)	...NYN^{432}K...	5402	5402 (predominant), 5401, 5412, 6520[18]
	...GSN^{630}V...	5402, 5412, 6503, 6513	5402 (predominant), 5412, 5401, 5422, 6502, 6512[18]
	...IN^{491}HC...	ND	5402[56]
IgA (Fc) IgA1 (P01876), IgA2 (P01877)	...EAN$^{144(IgA1)/131(IgA2)}$L...	4400, 5400, 3500, 4500, 5500, 4401, 5401, 5402, 4501, 5501, 5502	Released analysis[15,29]
	...TAN^{205}I(IgA2)...	5410, 5411, 5412, 4510, 5510, 5511, 5512	
	...HVN^{340}V (IgA1)...	4410, 4411, 5410, 5411, 5412, 5510, 5511, 5512, 6511	
Ig Joining Chain (P01876)	...REN^{71}I...	5401, 5411, 5412	5400, 5410, 5401, 5402, 5412[32]
α2-macroglobulin (P01023)	...YLN^{55}E...	5402, 5412	ND
	...GN^{70}R...	5401, 5411, 5402, 5412, 5502	5401, 5411, 5402, 5412[24]
	...EMN^{247}V...	5400, 5401, 5402, 6511	ND
	...YSN^{396}A...	5401, 5402, 5411, 5412	54015402,, 5411, 5412[24]
	...SIN^{410}T...	4401, 5401, 5411, 5402, 7604	5401, 5411, 5402[24]
	...NVN^{869}F...	4301, 4401, 5401, 5402, 4200, 5200, 5301, 6301	Man, complex type[42]
	...YLN^{991}E...	5401, 5402, 5512	ND
	...VSN^{1424}Q...	5401, 5411, 5402, 5412	ND
IgM (Fc, P01871)	...N^{46}NSD...	4311, 4401, 5401, 5411, 5412, 5501, 5511, 5502	mono-, bi-, triantennary, bisecting, hybrid[32]
	...QQN^{209}A...	5410, 5411, 5412, 4511, 5511, 5512	
	...HTN^{272}I...	5411	
	...HPN^{279}A...	5300, 5200, 6200, 7200, 8200, 9200	Man5→9[32]
	...LYN^{439}V...	5200, 6200, 7200, 8200	
α1-antitrypsin (P01009)	...STN^{70}I...	5402, 5412	bi-, triantennary[19,57]
	...NFN^{107}L...	5402, 5412, 6503, 6513	bi, tri, tetra-antennary[19,57]
	...LGN^{271}A...	4301, 5401, 5402, 5412, 6503, 6513	biantennary[19,57]
Complement C3 (P01024)	...MGN^{85}V...	3200, 5200, 6200, 7200, 6301	Man5,6[22] Man9+Glc[58]
	...MN^{939}K...	8200, 9200	Man8,9[22]
	...KPN^{1617}L...	9200, 10−000	
Haptoglobin (P00738)	...HHN^{184}L...	5401, 5402, 5512, 6503	5401, 5402, 5411, 5412, 6501, 6502, 6503, 6513[21]
	...LFLN^{207}H...	5401, 5402, (5411,6502, 6503)	5401, 5402, 6502, 6503, 6513[3]
	...SEN^{211}A...	5401, 5402, (5411,6502, 6503)	5401, 5402, 6502, 6503, 6513[3]
	...HPN^{241}Y...	5402, 6501, 6502, 6503, 6512, 6513	ND
α1-acid glycoprotein1 (A1AG1, P02763), α1-acid glycoprotein 2 (A1AG2, P19652)	...ITN^{33}A...	4401, 5401, 5402, 5501, 5502, 5512, 6501, 6511, 6502, 6512, 6503, 6513	bi-, tri-, tetra-antennary (+S)/(+F)[25,26]
	...EYN^{56}K...	4301, 4401, 5401, 5402, 5412, 6502, 6512, 6503, 6513	bi-, triantennary (+S)/(+F)[25,26]
	...TPN^{72}K...	6502, 6503, 6513, 7602, 7612, 7603, 7613, 7604, 7614, 8703	bi-, tri-, tetra-antennary (+S)/(+F)[25,26]
	...YN^{93}TT...(A1AGP1)	6503, 6513, 7602, 7612,7603, 7613, 7604, 7614, 8704	tri-, tetra-, penta-, hexa-antennary (+S)/(+F)[25,26]
	...EN^{103}GT...	6503	tri-, tetra-, penta-, hexa-antennary (+S)/(+F)[25,26]

[a]Glycan profile may change in disease states. [b]Glycan composition: Hex−HexNAc−Fuc−Neu5Ac. [c]ND: not determined.

APPLICATIONS

Glycoprotein analysis by MS of biological samples, such as serum and plasma, is hampered by the sample complexity and the low concentrations in protein glycosylations. Protein or glycan enrichment is thus widely employed in glycan analysis.[6,16,28] For example, IgG proteins are usually enriched using protein G or protein A prior to glycosylation analysis.[16] Glycoprotein/glycopeptide enrichments using lectin or hydrophilic interaction LC (HILIC) are also common.[44,50] However, there are several concerns regarding these enrichment procedures. The most prominent ones include sample loss, reproducibility, and potentially enrichment bias for particular types of glycans. To this end, direct analysis of proteins in serum/plasma without enrichment is an attractive technique. Utilizing the power of targeted MRM techniques, we found that IgG, IgA, IgM, and their glycosylation can be monitored directly from serum without any protein enrichment or sample cleanup.

Application of Method to Healthy Human Serum

To determine the IgG, IgA, and IgM abundances with their glycoforms in the general population, the MRM method was applied to plasma samples from 13 healthy individuals and a pooled commercial serum sample (Sigma-Aldrich, St. Louis, MO). Although it was developed on serum samples, the method was found to work well on plasma samples. The peptide calibration curves used for label-free absolute

DOI: 10.1021/acs.jproteome.5b00756
J. Proteome Res. XXXX, XXX, XXX−XXX

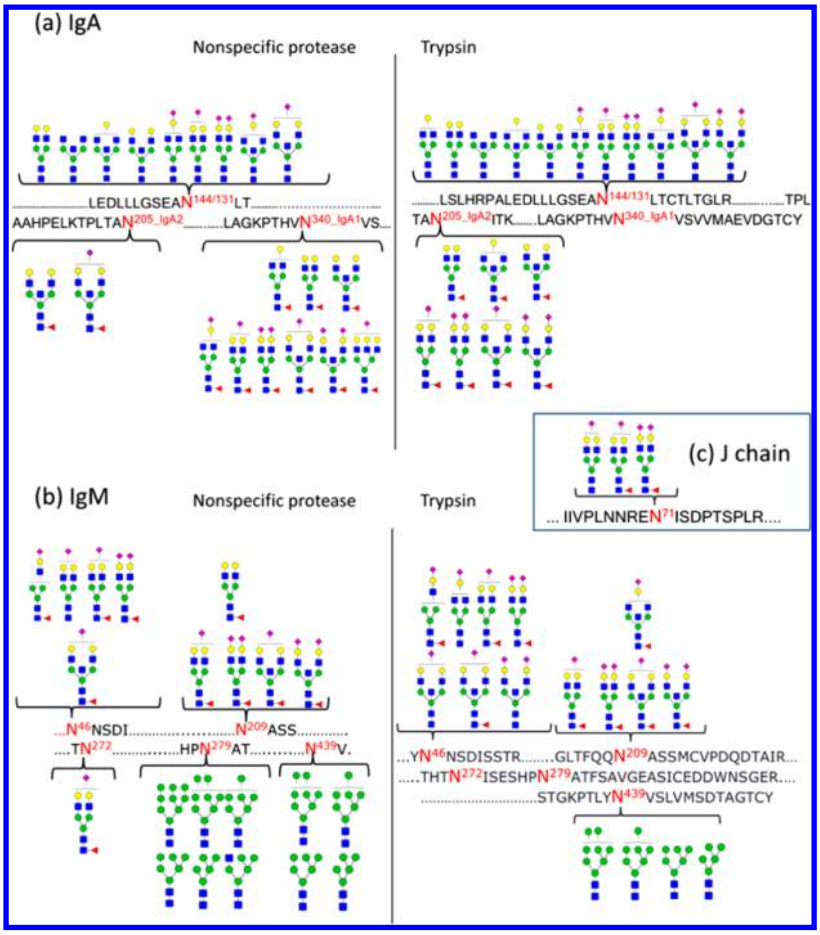


Figure 1. Site-specific glycan mapping for IgA an IgM using specfic and nonspecfic proteases. (a) Glycan map for IgA. IgA1 glycosite N144 and IgA2 glycosite N131 share the same tryptic peptide, thus cannot be distinguished from each other. While no fucosylation was observed for glycosite N144/131, glycans on glycosite N205 (IgA2) and N340 (IgA2) are all fucosylated. (b) Glycan map for IgM. Glycans on site N46, N209, and N272 are all highly sialylated, while mostly high mannose type glycans were observed on site N279 (one hybrid-type glycan identified) and N439. (c) Three glycans were observed for J chain. Symbol key: yellow circles, galactose (Gal); green circles, mannose (Man); blank circles, hexose (Hex); blue squares, N-acetylglucosamine (GlcNAc); yellow squares, N-acetylgalactosamine (GalNAc); red triangles, fucose (Fuc); purple diamonds, N-acetylneuraminic acid (Neu5Ac).

quantitation are shown in Figure S-4. The Ig concentrations in the Sigma serum sample were 14.9 ± 2.6 mg/mL (determined using peptide DTLMISR) for IgG, 2.23 ± 0.05 mg/mL (YLTWASR) for IgA, and 1.17 ± 0.04 mg/mL (FTCTVTHTDLPSPLK) for IgM. The standard deviation is based on triplicate digestions for the serum sample indicating the method repeatability. The Ig concentrations in the 13 healthy plasma samples were 10.4 ± 4.7 mg/mL (DTLMISR) for IgG, 1.72 ± 0.65 mg/mL (YLTWASR) for IgA, 1.13 ± 0.63 mg/mL (FTCTVTHTDLPSPLK) for IgM. The standard deviation here represents the biological variations. The MRM quantitation results well reproduced the previously reported values using nephelometry assay for the general population, that is, 11.2 ± 2.5 mg/mL, 2.62 ± 1.19 mg/mL, 1.47 ± 0.84 mg/mL for IgG, A, and M, respectively.[51] Some discrepancies were observed that may be due to the impurity of the protein standard used. All protein standards (IgG, A, and M) were at least 95% pure. However, the fact that the MRM results are within the experimental range indicates that these impurities do not have significant impact on the glycopeptide quantitation.

It was found here and in the previous study that glycosite N439 of IgM is not fully occupied.[32] To determine the relative unoccupancy, the unoccupied glycosite was monitored (STGKPTLYN^{439}VSLVMSDTAGTCY, m/z 1183.1 → 342.1), and its ion abundance was normalized to the total IgM amount using the following equation:

$$\text{glycosite unoccupancy} = \log_{10} \frac{\text{unoccupied peptide ion counts}}{\text{protein abundance (peptide ion counts)}} \quad (1)$$

where the unoccupied peptide is STGKPTLYN^{439}VSLVMSDTAGTCY without glycosylation and the IgM protein abundance are represented by the ion abundance of FTCTVTHTDLPSPLK. As shown in Figure 4, the unoccupancy for glycosite N439 is very similar across 13 individuals (CV = 10%), while the variation in IgM concentration is relatively large (CV = 50%). Indeed, the biological variation in the protein content is larger than the degree of glycosylation. Although an absolute amount of glycosite occupancy was not obtained due to the lack of glycopeptide standards, the variation across samples can be examined for relative comparisons.

Variation in glycan abundances may be due to either higher protein expression or glycan expression. To remove the contribution of protein expression and focus on the protein

DOI: 10.1021/acs.jproteome.5b00756
J. Proteome Res. XXXX, XXX, XXX−XXX

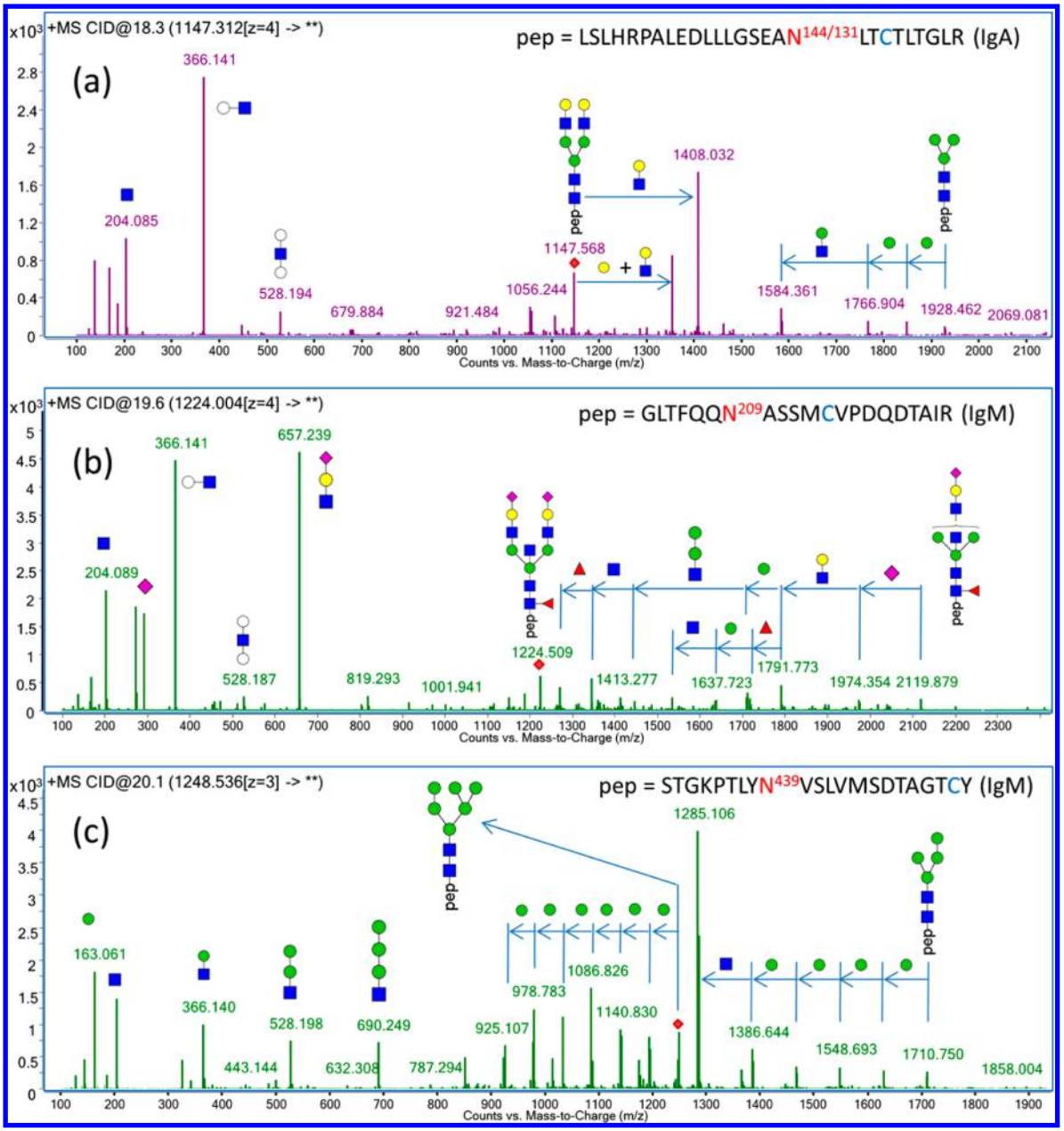


Figure 2. Representative fragmentation spectra for glycopeptides with (a) a neutral glycan, (b) a sialylated glycan, and (c) a high mannose glycan show mainly glycan fragments using CID. Common fragments such as m/z 204.08 (HexNAc), 366.14 ($Hex_1HexNAc_1$), 292.09 (Neu5Ac), and 657.24 ($Hex_1HexNAc_1Neu5Ac_1$) were used to identify glycopeptide candidates from the peptide ion pool. High mannose glycopeptides also produce abundant peptide−GlcNAc fragments. These abundant fragment ions were monitored in the MRM transition and optimized to obtain the maximum sensitivity.

glycosylation, the glycopeptide abundances were normalized to the protein abundances using eq 2:

$$\text{degree of glycosylation} = \frac{\text{glycopeptide ion counts}}{\text{protein abundance (nonglycopeptide ion counts)}} \quad (2)$$

The peptides used to represent protein abundance are DASGATFTWTPSSGK (IgA2), YLTWASR (IgA1,2), FTCTVTHTDLPSPLK (IgM), and SSEDPNEDIVER (J chain). The absolute and normalized abundance of one IgM glycopeptide is shown as representative in Figure 5. The error bar for the pooled serum sample is for three replicate digestions of the same serum sample showing a good method reproducibility (CV < 20%). Figure 5 shows that the trend of absolute ion abundance does not necessarily follow the trend of the degree of protein glycosylation, that is, high absolute ion abundance may have lower degree of glycosylation or vice versa. Indeed, the protein abundances need to be taken into account when protein specific alteration in glycosylation is monitored. Our result also shows a large biological variation in glycan expression even within the general healthy individuals (CV > 50% when using absolute glycopeptide abundance). However, if normalized to the protein content, the biological variation decreases (CV < 20%) indicating a similar degree of glycosylation across samples. Indeed, degree of glycosylation yields information focused on the alternation in glycan profile.

DOI: 10.1021/acs.jproteome.5b00756
J. Proteome Res. XXXX, XXX, XXX−XXX

Table 4. MRM Transition List for IgA and IgM

protein	peptide	glycan[a]	precursor ion (m/z)	product ion (m/z)	RT (min)	collision energy (eV)
IgA2	TPLTAN^{205}ITK	5410	909.8	366.1	6.22	19
		5510	977.5	366.1	6.36	25
		4510	923.5	366.1	6.48	18
		5411	1006.8	366.1	6.9	25
		5511	1074.5	366.1	6.94	25
		5412	1103.8	366.1	7.52	25
		5512	878.9	366.1	7.66	17
IgA1,2	LSLHRPALEDLLLGSEAN$^{144/131}$LT	5501	1016.9	366.1	15.95	25
	CTLTGLR	5401	976.3	366.1	15.96	25
		3500	1117.1	366.1	16	25
		4500	1157.6	366.1	16	25
		5400	1147.3	366.1	16	25
		5500	1198.1	366.1	16	25
		4401	943.9	366.1	16.01	20
		4501	1230.4	366.1	16.01	30
		5402	1292.9	366.1	16.19	30
		5502	1075.1	366.1	16.4	25
IgA1	^{213}TPLTATLSK^{221}	—	466.3	415.8	9.03	15
IgA1	^{154}DASGVTFTWTPSSGK^{168}	—	770.9	475.3	12.73	20
IgA2	^{141}DASGATFTWTPSSGK^{155}	—	756.9	475.3	11.91	25
IgA1,2	^{276}YLTWASR^{282}	—	448.7	620.3	10.78	12
IgA1,2	^{264}WLQGSELPR^{273}	—	607.4	914.5	11.37	18
IgM	YKN^{46}NSDISSTR	5501	851.1	366.1	2.18	14
		4401	1012.8	366.1	2.19	25
		5411	1115.5	366.1	2.19	25
		4311	993.8	366.1	2.21	25
		5511	1183.2	366.1	2.27	25
		5601	901.9	366.1	2.27	14
		5412	909.6	366.1	2.3	14
		5502	923.9	366.1	2.3	19
IgM	GLTFQQN^{209}ASSMCVPDQDTAIR	5411	1100.5	366.1	12.37	19
		5511	1151.3	366.1	12.37	25
		4511	1110.8	366.1	12.41	24
		5412	1173.3	366.1	12.49	23
		5512	1224.1	366.1	12.5	25
IgM	STGKPTLYN^{439}VSLVMSDTAGTCY	9200	1058.3	1284.7	13.51	27
		8200	1356.6	1284.7	13.53	27
		7200	1302.6	1284.7	13.55	27
		5200	1194.5	1284.7	13.58	27
		6200	1248.5	1284.7	13.58	24
IgM	STGKPTLYN^{439}VSLVMSDTAGTCY	—	1183.1	342.1	14.28	38
IgM	^{65}YAATSQVLLPSK^{76}	—	639.4	331.2	11.28	15
IgM	^{301}FTCTVTHTDLPSPLK^{315}	—	573.0	734.9	12.07	15
J Chain	EN^{71}ISDPTSPLR	5401	1048.1	366.1	9.07	25
J Chain	IIVPLNNREN^{71}ISDPTSPLR	5411	1052.8	366.1	12.63	24
J Chain	IIVPLNNREN^{71}ISDPTSPLR	5412	1125.6	366.1	12.8	18
J Chain	^{47}SSEDPNEDIVER^{58}	—	695.3	971.5	5.78	24
J Chain	^{62}IIVPLNNR^{69}	—	469.8	613.3	11.11	12

[a]Glycan composition: Hex−HexNAc−Fuc−Neu5Ac.

The normalized glycopeptide abundances were used for relative comparison of glycoforms. Although individual glycoforms may differ in their ionization and fragmentation efficiencies in electrospray ionization (ESI)−LC−MS, recent results show that the direct comparison of MS signals from the same types of glycoforms may be generally appropriate.[39,52] As shown in Figure 6, the glycosylation profile is relatively similar between the commercial pooled serum and the 13 plasma samples. A glycoform was excluded from the analysis if no significant MRM signal (S/N > 6) was observed for more than three out of 13 individuals. Only the most common glycoforms were illustrated in Figure 6 to better represent the general population. Figure 6, panel a shows the relative abundances of the three glycans identified and quantified on IgM site N46. The glycoform 5411 (Hex:HexNAc:Fuc:Neu5Ac) and 5501 are both abundant at this glycosite, while 5511 is 50% lower. Glycans observed on IgM site N209 are fucosylated and highly sialylated (Figure 6b). The bisecting glycan 5511 is the most abundant, the biantennary glycan 5411 is 50% lower, and the other glycans observed, 5412, 5512, and 4511, are much less

DOI: 10.1021/acs.jproteome.5b00756
J. Proteome Res. XXXX, XXX, XXX−XXX

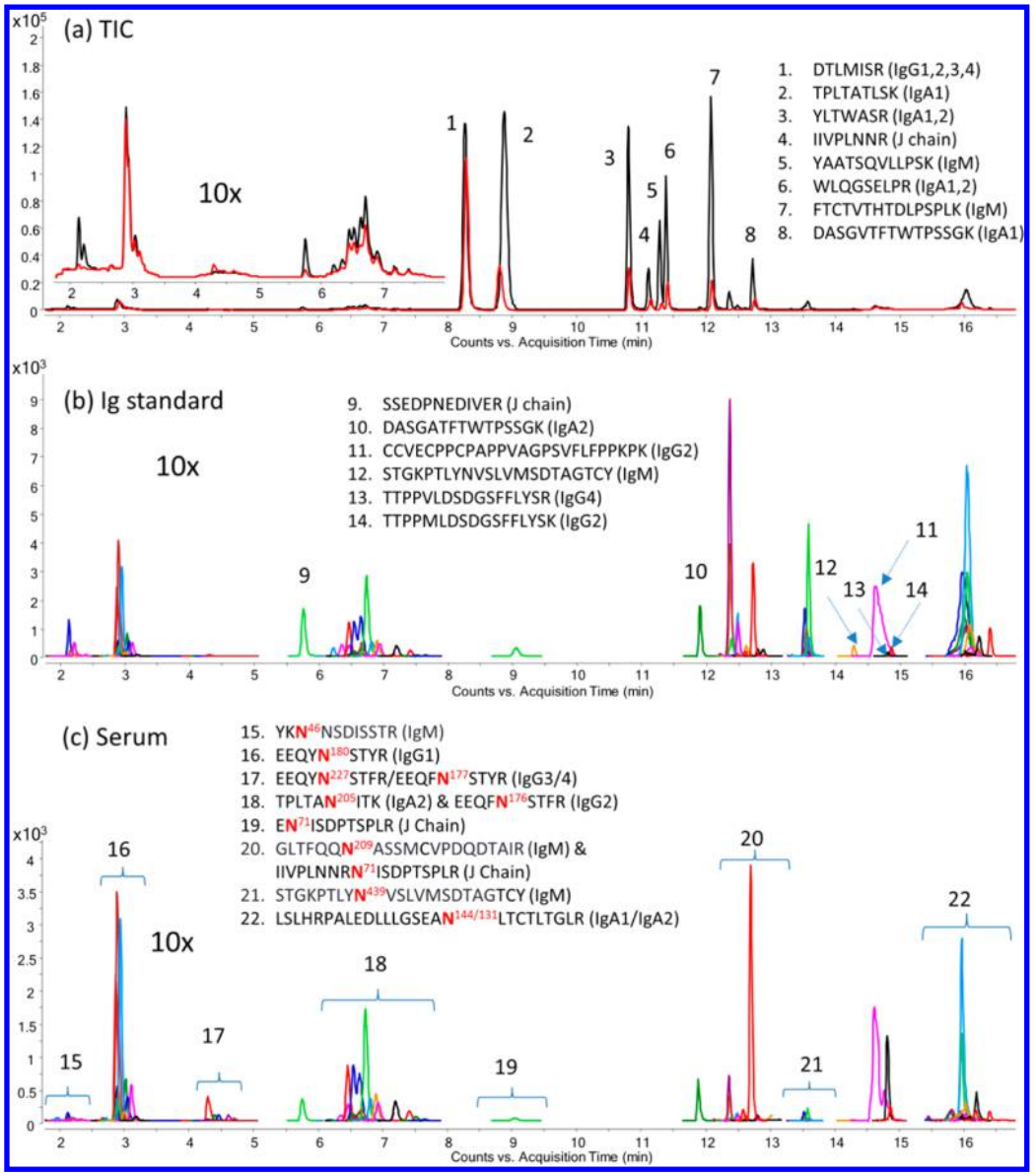


Figure 3. (a) Total MRM chromatogram for the protein standard mixture (black trace) and the pooled serum sample (red trace). Extracted MRM chromatograms of (b) protein standards and (c) the pooled serum sample. The most abundant peaks correspond to peptide signals 1−8 representing the proteins shown. Lower abundance peptides corresponding to peaks 9−14, and glycosites corresponding to signal 15−22. The eight most abundant peptide signals were removed intentionally to magnify the glycopeptide and peptide signals in panels b and c. Glycopeptides were separated on the basis of their peptide backbone using the C18 stationary phase, and the detail glycoform annotation is provided in Figure S-3.

abundant (<20%) at site N209. The results show that the monosialylated and monofucosylated glycans are the major structures at site N46 and N209, which is consistent with the previous findings that monosialylated structures contribute 71% of the total glycosylation from N46, 209, and 272; the bisected structure contributes about 50%; and fucosylated structure accounts for 85%.[32] Only high mannose-type glycans were observed at site N439, with Man6 and Man8 being the most abundant glycoforms. The results are consistent with previous reports.[32]

For IgA, glycoforms 5401 and 5501 are the major components of the 12 glycoforms quantified at site N144/131, while 5411, 5511, and 4510 are the top three abundant glycoforms (seven glycoforms quantified in total) at site N205. The result was consistent with the previous findings that monosialylated bisecting and biantennary glycans are the two major glycan components attached to serum IgA.[53,54] The result shown in Figure 6, panel f is also consistent with the previous findings that the biantennary glycans 5401, 5412, and 5411 are the three major structures at the joining chain.[32,55]

Method Validation

Linearity and Limit of Detection. A serial dilution for a standard protein mixture, which contains 0.30, 0.20, and 0.17 mg/mL of IgG, IgA, and IgM, respectively, was performed to study the linearity and limit of detections of the MRM method. Eight different concentrations as listed in the Experimental section were examined, and linear regressions were performed. As depicted in Figure S-4, the linear range spans three orders of magnitude (R^2 > 0.995) for all peptides used in absolute quantitation. The LOQ is defined as a S/N ratio over 6. The results provided in Figure S-4 show that the LOQ for IgA (MW = 160 kDa) is one femtomole using peptide, YLTWASR, and

DOI: 10.1021/acs.jproteome.5b00756
J. Proteome Res. XXXX, XXX, XXX−XXX

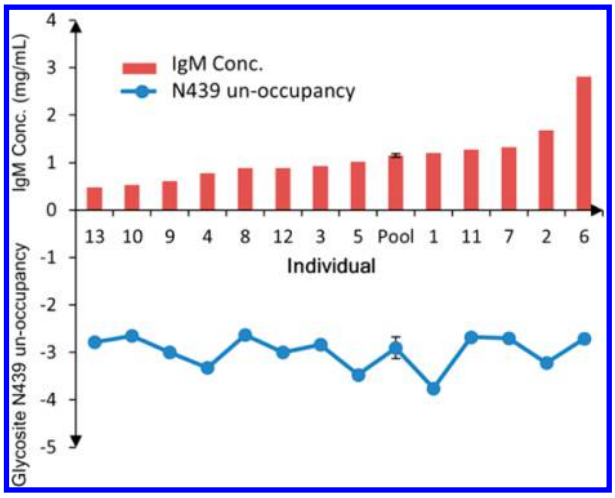


Figure 4. Determination of relative site occupancy for 439N of IgM. The ion abundance of nonglycosylated peptide STGKPTLYNV-SLVMSDTAGTCY was normalized to the quantitating peptide FTCTVTHTDLPSPLK using eq 1. Each bar corresponds to a specific individual (1−13) and a pooled serum sample (Pool). The red bar is the IgM concentration, and the blue circle is the peptide abundance normalized to the total protein abundance as calculated by the quantitating peptide. The pooled serum sample was analyzed in triplicate digestion and LC−MS analysis, and the standard deviation was represented by the error bar. The biological variation in protein abundances was high (CV = 50%); however, the relative glycosylation occupancy for this site was very similar across the 13 individuals (CV = 10%).

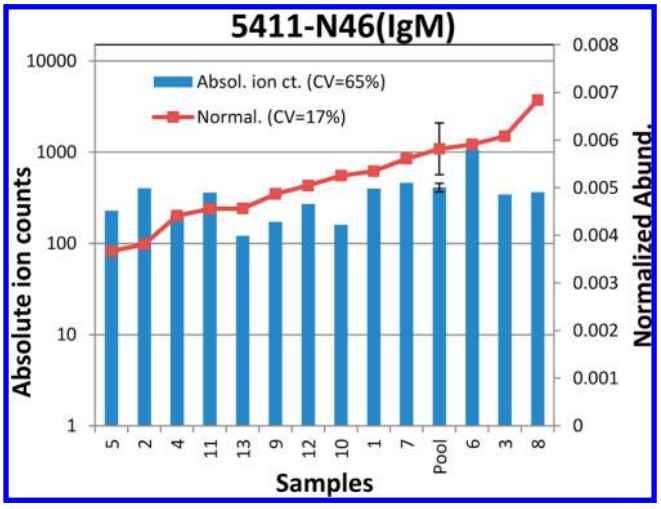


Figure 5. Variations in absolute ion abundances (blue bar) and normalized abundances (read square) of glycopeptide 5411 on N46 of IgM. The absolute ion abundances represent the concentrations of the glycopeptides in serum/plasma. The normalized abundances are representative of the glycan abundances relative to the protein abundance. The error bar for the pooled serum sample is from triplicate digestion experiments. The absolute ion count does not necessarily follow the normalized abundances.

10 femtomole using peptide, WLQGSELPR. The LOQ for IgM (MW = 190 kDa) is one femtomole for both peptides monitored, that is, YAATSQVLLPSK and FTCTVT-HTDLPSPLK. Similar MRM sensitivities were observed for IgG in the previous study.[39] While there are several peptide candidates for quantitation, peptides YLTWASR and FTCTVTHTDLPSPLK were studied for absolute quantitation of IgA and IgM, respectively, because they yielded the highest MRM responses. Although the linear range and LOQ of individual glycopeptides is hard to validate due to the lack of glycopeptide standards, the experiment as discussed further indicates that there is a linear relationship for glycopeptide MRM signal over a range of 0−50 μg/mL added protein standards.

Matrix Effects. Standard addition experiment takes matrix effect into account, producing more accurate quantitation. Five additional concentration levels as explained in the Supporting Information, were examined to evaluate the matrix effects. The resulting calibration curves (Figure S-5) show good linearity, R^2 > 0.95, over the concentration window examined. The protein concentrations obtained from the standard addition experiment are 2.55 ± 0.14 and 1.27 ± 0.09 mg/mL compared to 2.23 ± 0.05 and 1.17 ± 0.04 mg/mL obtained from standard curves for IgA and IgM, respectively. The difference is within 10%, which indicates very limited matrix effects and interference signals. Similar result were obtained with IgG.[39]

Method Repeatability. Method reproducibility plays a key role in quantitation analysis of large sample sets. Good reproducibility generally increases the method sensitivity in distinguishing small changes across samples with a higher confidence level. The instrument repeatability was thus evaluated by injecting the same tryptic serum sample 11 times spanning over a 24 h period. As shown in Figure S-6a, our instrument yielded excellent repeatability throughout the run (CV < 5%). To study the intraday repeatability of trypsin digestions, 10 replicate trypsin digestions were performed for the same serum sample in the same day. The result shown in Figure S-6b indicates a good intraday repeatability for our enzyme digestion (CV ≤ 10%). The interday repeatability generally shows relatively larger variation, which may due to the differences in benchtop conditions such as temperature and pressure. However, only the peptides that show relatively good interday repeatability (CV ≤ 15%) were selected for protein quantitation. These peptides were very likely to have little or no amino acid modifications and or missed-cleavages.

Comparison of Enrichment and Direct Analysis. To illustrate this method of direct analysis, IgA and IgM were quantified both with and without pre-enrichment. The results are summarized in Figure S-6. In the direct analysis, 2 μL of serum/plasma was digested in 100 μL of 50 mM NH_4HCO_3, and the sample was injected directly for MRM quantitation. In the enrichment experiment, IgA and IgM proteins were first immunoprecipitated from 5 μL of serum/plasma using anti-IgA/IgM antibodies and then digested in 50 μL of 50 mM NH_4HCO_3 for MS analysis. The protein abundance is approximately five-fold higher in the enrichment experiment than in the direct analysis. However, as illustrated in Figure S-6c, the two methods yield similar MRM signals indicating a relatively large sample loss in the enrichment method. Our sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) result indicates a near complete depletion of other serum proteins and a significant increase of IgA/M signal after enrichment. The sample loss may be due to inefficient binding of IgA/M to the resin. Increase resin to sample ratio may reduce sample loss, although only the manufacturer suggested protocol was tested in our experiment.

The repeatability of enrichment experiment is illustrated in Figure S-6c. The result shows a relatively larger variation for the protein enrichment experiment (CV of 20% vs 10% for the direct analysis as shown in Figures S-6c and 6b). The same

DOI: 10.1021/acs.jproteome.5b00756
J. Proteome Res. XXXX, XXX, XXX−XXX

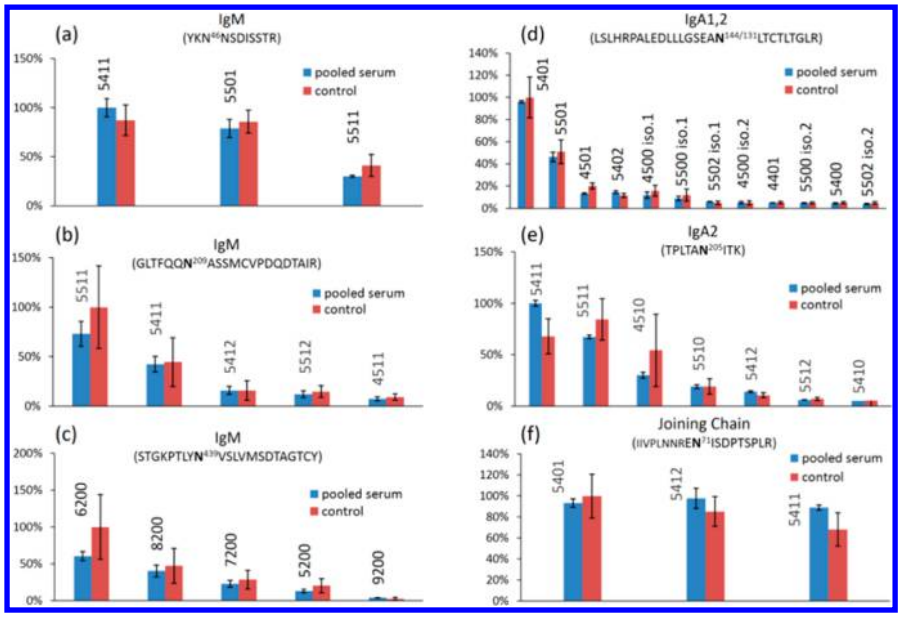


Figure 6. Normalized abundance of IgM and IgA glycopeptides of the pooled serum sample and the 13 plasma samples. IgM glycosite (a) N46, (b) N439, (c) N209; IgA glycosite (d) N205, (e) N144/131; joining chain glycosite (f) N71. The error bar for pooled serum sample is from triplicate digestions of the same serum. The error bar for the 13 plasma samples represents the biological variation. Blue bar, pooled serum sample; red bar, 13 plasma samples. The glycan composition is annotated above the histogram, for example, 5411 corresponds to $Hex_5HexNAc_4Fuc_1Neu5Ac_1$.

peptides were observed in all three sets; however, direct analysis is not only far easier and faster, but also it is less costly and yields better reproducibility.

CONCLUSION

There has been previously a lack of complete glycan map for even the most abundant serum glycoproteins. There is further no general technique for quantitating protein-specific glycosylation. We employed MRM for the quantitation of proteins and their glycosylation directly from human serum without sample enrichment. Although glycan enrichment increases the sensitivity and specificity of glycosylation analysis, sample loss and cost are big concerns. Targeted MRM is attractive because of its capability in detecting low abundant compounds directly in complicated mixtures. Site-specific glycosylation analysis can now be performed directly on serum/plasma samples to quantitate both proteins and glycans simultaneously. The result shows a low femtomole limit of quantitation for IgG, IgA, and IgM protein. The direct analysis using MRM provides higher reproducibility and less sample loss, which makes it a robust technique for biomarker discovery studies. Furthermore, the use of UPLC allows rapid analysis (<20 min per sample) requiring only about 2 μL of serum/plasma. This method will be useful for biomarker discovery and for understanding the role of glycosylation in specific proteins.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jproteome.5b00756.

Experimental procedures; MRM chromatogram for transferrin and haptoglobin; tandem spectrum of peptides and glycopeptides; glycopeptide annotation for Figure 3; peptide calibration curves; standard addition experiment; study of method reproducibility; IgG MRM transitions (PDF)

AUTHOR INFORMATION

Corresponding Author

*E-mail: cblebrilla@ucdavis.edu. Phone: +1-530-752-5504. Fax: +1-530-752-8995.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors are thankful for the funding provided by the National Institutes of Health (RO1 CA136647, R01HD061923, R01GM049077, RO1AT008759, and 1DP2OD008752).

REFERENCES

(1) de Leoz, M. L.; Young, L. J.; An, H. J.; Kronewitter, S. R.; Kim, J.; Miyamoto, S.; Borowsky, A. D.; Chew, H. K.; Lebrilla, C. B. High-mannose glycans are elevated during breast cancer progression. *Mol. Cell. Proteomics* **2011**, *10*, M110.002717.

(2) An, H. J.; Kronewitter, S. R.; de Leoz, M. L. A.; Lebrilla, C. B. Glycomics and disease markers. *Curr. Opin. Chem. Biol.* **2009**, *13*, 601−607.

(3) Fujimura, T.; Shinohara, Y.; Tissot, B.; Pang, P.-C.; Kurogochi, M.; Saito, S.; Arai, Y.; Sadilek, M.; Murayama, K.; Dell, A.; Nishimura, S.-I.; Hakomori, S.-i. Glycosylation status of haptoglobin in sera of patients with prostate cancer vs. benign prostate disease or normal subjects. *Int. J. Cancer* **2008**, *122*, 39−49.

(4) Chen, G.; Wang, Y.; Qiu, L.; Qin, X.; Liu, H.; Wang, X.; Wang, Y.; Song, G.; Li, F.; Guo, Y.; Li, F.; Guo, S.; Li, Z. Human IgG Fc-glycosylation profiling reveals associations with age, sex, female sex hormones and thyroid cancer. *J. Proteomics* **2012**, *75*, 2824−2834.

(5) Axford, J. S. Glycosylation and rheumatic disease. *Biochim. Biophys. Acta, Mol. Basis Dis.* **1999**, *1455*, 219−229.

(6) Bond, A.; Alavi, A.; Axford, J. S.; Bourke, B. E.; Bruckner, F. E.; Kerr, M. A.; Maxwell, J. D.; Tweed, K. J.; Weldon, M. J.; Youinou, P.; Hay, F. C. A Detailed Lectin Analysis of IgG Glycosylation,

DOI: 10.1021/acs.jproteome.5b00756
J. Proteome Res. XXXX, XXX, XXX−XXX

Demonstrating Disease Specific Changes in Terminal Galactose and N-acetylglucosamine. *J. Autoimmun.* **1997**, *10*, 77−85.
(7) Mestecky, J.; Tomana, M.; Crowley-Nowick, P. A.; Moldoveanu, Z.; Julian, B. A.; Jackson, S. Defective galactosylation and clearance of IgA1 molecules as a possible etiopathogenic factor in IgA nephropathy. *Contrib. Nephrol.* **1993**, *104*, 172−82.
(8) Hiki, Y.; Odani, H.; Takahashi, M.; Yasuda, Y.; Nishimoto, A.; Iwase, H.; Shinzato, T.; Kobayashi, Y.; Maeda, K. Mass spectrometry proves under-O-glycosylation of glomerular IgA1 in IgA nephropathy. *Kidney Int.* **2001**, *59*, 1077−85.
(9) Basset; Durand; Jamin; Clément; Pennec; Youinou; Dueymes; Roitt. Increased N-Linked Glycosylation Leading to Oversialylation of Monomeric Immunoglobulin A1 from Patients with Sjögren's Syndrome. *Scand. J. Immunol.* **2000**, *51*, 300−306.
(10) Coppo, R.; Amore, A. Aberrant glycosylation in IgA nephropathy (IgAN). *Kidney Int.* **2004**, *65*, 1544−7.
(11) Bause, E. Structural requirements of N-glycosylation of proteins. Studies with proline peptides as conformational probes. *Biochem. J.* **1983**, *209*, 331−6.
(12) Marino, K.; Bones, J.; Kattla, J. J.; Rudd, P. M. A systematic approach to protein glycosylation analysis: a path through the maze. *Nat. Chem. Biol.* **2010**, *6*, 713−23.
(13) Ruhaak, L. R.; Zauner, G.; Huhn, C.; Bruggink, C.; Deelder, A. M.; Wuhrer, M. Glycan labeling strategies and their use in identification and quantification. *Anal. Bioanal. Chem.* **2010**, *397*, 3457−81.
(14) Viseux, N.; Hronowski, X.; Delaney, J.; Domon, B. Qualitative and Quantitative Analysis of the Glycosylation Pattern of Recombinant Proteins. *Anal. Chem.* **2001**, *73*, 4755−4762.
(15) Aldredge, D.; An, H. J.; Tang, N.; Waddell, K.; Lebrilla, C. B. Annotation of a serum N-glycan library for rapid identification of structures. *J. Proteome Res.* **2012**, *11*, 1958−68.
(16) Wuhrer, M.; Stam, J. C.; van de Geijn, F. E.; Koeleman, C. A.; Verrips, C. T.; Dolhain, R. J.; Hokke, C. H.; Deelder, A. M. Glycosylation profiling of immunoglobulin G (IgG) subclasses from human serum. *Proteomics* **2007**, *7*, 4070−81.
(17) Ahn, J.-M.; Sung, H.-J.; Yoon, Y.-H.; Kim, B.-G.; Yang, W. S.; Lee, C.; Park, H.-M.; Kim, B.-J.; Kim, B.-G.; Lee, S.-Y.; An, H.-J.; Cho, J.-Y. Integrated Glycoproteomics Demonstrates Fucosylated Serum Paraoxonase 1 Alterations in Small Cell Lung Cancer. *Mol. Cell. Proteomics* **2014**, *13*, 30−48.
(18) Satomi, Y.; Shimonishi, Y.; Hase, T.; Takao, T. Site-specific carbohydrate profiling of human transferrin by nano-flow liquid chromatography/electrospray ionization mass spectrometry. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 2983−2988.
(19) Kolarich, D.; Weber, A.; Turecek, P. L.; Schwarz, H.-P.; Altmann, F. Comprehensive glyco-proteomic analysis of human α1-antitrypsin and its charge isoforms. *Proteomics* **2006**, *6*, 3369−3380.
(20) Harazono, A.; Kawasaki, N.; Itoh, S.; Hashii, N.; Matsuishi-Nakajima, Y.; Kawanishi, T.; Yamaguchi, T. Simultaneous glycosylation analysis of human serum glycoproteins by high-performance liquid chromatography/tandem mass spectrometry. *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.* **2008**, *869*, 20−30.
(21) 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*, 1281−1293.
(22) Hirani, S.; Lambris, J. D.; Müller-Eberhard, H. J. Structural analysis of the asparagine-linked oligosaccharides of human complement component C3. *Biochem. J.* **1986**, *233*, 613−616.
(23) Ritchie, G. E.; Moffatt, B. E.; Sim, R. B.; Morgan, B. P.; Dwek, R. A.; Rudd, P. M. Glycosylation and the Complement System. *Chem. Rev. (Washington, DC, U. S.)* **2002**, *102*, 305−320.
(24) Lin, Z.; Lo, A.; Simeone, D. M.; Ruffin, M. T.; Lubman, D. M. An N-glycosylation Analysis of Human Alpha-2-Macroglobulin Using an Integrated Approach. *J. Proteomics Bioinf.* **2012**, *5*, 127−134.
(25) Imre, T.; Schlosser, G.; Pocsfalvi, G.; Siciliano, R.; Molnár-Szöllősi, É.; Kremmer, T.; Malorni, A.; Vékey, K. Glycosylation site analysis of human alpha-1-acid glycoprotein (AGP) by capillary liquid chromatography—electrospray mass spectrometry. *J. Mass Spectrom.* **2005**, *40*, 1472−1483.
(26) Treuheit, M. J.; Costello, C. E.; Halsall, H. B. Analysis of the five glycosylation sites of human alpha 1-acid glycoprotein. *Biochem. J.* **1992**, *283* (Pt 1), 105−12.
(27) Arnold, J. N.; Radcliffe, C. M.; Wormald, M. R.; Royle, L.; Harvey, D. J.; Crispin, M.; Dwek, R. A.; Sim, R. B.; Rudd, P. M. The Glycosylation of Human Serum IgD and IgE and the Accessibility of Identified Oligomannose Structures for Interaction with Mannan-Binding Lectin. *J. Immunol.* **2004**, *173*, 6831−6840.
(28) Bakovic, M. P.; Selman, M. H.; Hoffmann, M.; Rudan, I.; Campbell, H.; Deelder, A. M.; Lauc, G.; Wuhrer, M. High-throughput IgG Fc N-glycosylation profiling by mass spectrometry of glycopeptides. *J. Proteome Res.* **2013**, *12*, 821−31.
(29) Baenziger, J.; Kornfeld, S. Structure of the Carbohydrate Units of IgA1 Immunoglobulin: I. COMPOSITION, GLYCOPEPTIDE ISOLATION, AND STRUCTURE OF THE ASPARAGINE-LINKED OLIGOSACCHARIDE UNITS. *J. Biol. Chem.* **1974**, *249*, 7260−7269.
(30) Zhao, J.; Qiu, W.; Simeone, D. M.; Lubman, D. M. N-linked Glycosylation Profiling of Pancreatic Cancer Serum Using Capillary Liquid Phase Separation Coupled with Mass Spectrometric Analysis. *J. Proteome Res.* **2007**, *6*, 1126−1138.
(31) Mattu, T. S.; Pleass, R. J.; Willis, A. C.; Kilian, M.; Wormald, M. R.; Lellouch, A. C.; Rudd, P. M.; Woof, J. M.; Dwek, R. A. The Glycosylation and Structure of Human Serum IgA1, Fab, and Fc Regions and the Role of N-Glycosylation on Fcα Receptor Interactions. *J. Biol. Chem.* **1998**, *273*, 2260−2272.
(32) Arnold, J. N.; Wormald, M. R.; Suter, D. M.; Radcliffe, C. M.; Harvey, D. J.; Dwek, R. A.; Rudd, P. M.; Sim, R. B. Human Serum IgM Glycosylation: IDENTIFICATION OF GLYCOFORMS THAT CAN BIND TO MANNAN-BINDING LECTIN. *J. Biol. Chem.* **2005**, *280*, 29080−29087.
(33) Strum, J. S.; Nwosu, C. C.; Hua, S.; Kronewitter, S. R.; Seipert, R. R.; Bachelor, R. J.; An, H. J.; Lebrilla, C. B. Automated assignments of N- and O-site specific glycosylation with extensive glycan heterogeneity of glycoprotein mixtures. *Anal. Chem.* **2013**, *85*, 5666−75.
(34) Song, E.; Pyreddy, S.; Mechref, Y. Quantification of glycopeptides by multiple reaction monitoring liquid chromatography/tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **2012**, *26*, 1941−1954.
(35) Kurogochi, M.; Matsushista, T.; Amano, M.; Furukawa, J.-i.; Shinohara, Y.; Aoshima, M.; Nishimura, S.-I. Sialic Acid-focused Quantitative Mouse Serum Glycoproteomics by Multiple Reaction Monitoring Assay. *Mol. Cell. Proteomics* **2010**, *9*, 2354−2368.
(36) Hülsmeier, A. J.; Paesold-Burda, P.; Hennet, T. N-Glycosylation Site Occupancy in Serum Glycoproteins Using Multiple Reaction Monitoring Liquid Chromatography-Mass Spectrometry. *Mol. Cell. Proteomics* **2007**, *6*, 2132−2138.
(37) Sanda, M.; Pompach, P.; Brnakova, Z.; Wu, J.; Makambi, K.; Goldman, R. Quantitative liquid chromatography-mass spectrometry-multiple reaction monitoring (LC-MS-MRM) analysis of site-specific glycoforms of haptoglobin in liver disease. *Mol. Cell. Proteomics* **2013**, *12*, 1294−305.
(38) Sanda, M.; Pompach, P.; Brnakova, Z.; Wu, J.; Makambi, K.; Goldman, R. Quantitative Liquid Chromatography-Mass Spectrometry-Multiple Reaction Monitoring (LC-MS-MRM) Analysis of Site-specific Glycoforms of Haptoglobin in Liver Disease. *Mol. Cell. Proteomics* **2013**, *12*, 1294−1305.
(39) Hong, Q.; Lebrilla, C. B.; Miyamoto, S.; Ruhaak, L. R. Absolute quantitation of immunoglobulin g and its glycoforms using multiple reaction monitoring. *Anal. Chem.* **2013**, *85*, 8585−93.
(40) Kronewitter, S. R.; An, H. J.; de Leoz, M. L.; Lebrilla, C. B.; Miyamoto, S.; Leiserowitz, G. S. The development of retrosynthetic glycan libraries to profile and classify the human serum N-linked glycome. *Proteomics* **2009**, *9*, 2986−2994.
(41) Nwosu, C. C.; Seipert, R. R.; Strum, J. S.; Hua, S. S.; An, H. J.; Zivkovic, A. M.; German, B. J.; Lebrilla, C. B. Simultaneous and

DOI: 10.1021/acs.jproteome.5b00756
J. Proteome Res. XXXX, XXX, XXX−XXX

Extensive Site-specific N- and O-Glycosylation Analysis in Protein Mixtures. *J. Proteome Res.* **2011**, *10*, 2612−2624.
(42) Arnold, J. N.; Wallis, R.; Willis, A. C.; Harvey, D. J.; Royle, L.; Dwek, R. A.; Rudd, P. M.; Sim, R. B. Interaction of mannan binding lectin with alpha2 macroglobulin via exposed oligomannose glycans: a conserved feature of the thiol ester protein family? *J. Biol. Chem.* **2006**, *281*, 6955−63.
(43) Ramachandran, P.; Boontheung, P.; Xie, Y.; Sondej, M.; Wong, D. T.; Loo, J. A. Identification of N-Linked Glycoproteins in Human Saliva by Glycoprotein Capture and Mass Spectrometry. *J. Proteome Res.* **2006**, *5*, 1493−1503.
(44) Picariello, G.; Ferranti, P.; Mamone, G.; Roepstorff, P.; Addeo, F. Identification of N-linked glycoproteins in human milk by hydrophilic interaction liquid chromatography and mass spectrometry. *Proteomics* **2008**, *8*, 3833−3847.
(45) Kerr, M. A. The structure and function of human IgA. *Biochem. J.* **1990**, *271*, 285−96.
(46) Mellquist, J. L.; Kasturi, L.; Spitalnik, S. L.; Shakin-Eshleman, S. H. The Amino Acid Following an Asn-X-Ser/Thr Sequon Is an Important Determinant of N-Linked Core Glycosylation Efficiency. *Biochemistry* **1998**, *37*, 6833−6837.
(47) Huhn, C.; Selman, M. H.; Ruhaak, L. R.; Deelder, A. M.; Wuhrer, M. IgG glycosylation analysis. *Proteomics* **2009**, *9*, 882−913.
(48) Mimura, Y.; Ashton, P. R.; Takahashi, N.; Harvey, D. J.; Jefferis, R. Contrasting glycosylation profiles between Fab and Fc of a human IgG protein studied by electrospray ionization mass spectrometry. *J. Immunol. Methods* **2007**, *326*, 116−126.
(49) Selman, M. H. J.; Derks, R. J. E.; Bondt, A.; Palmblad, M.; Schoenmaker, B.; Koeleman, C. A. M.; van de Geijn, F. E.; Dolhain, R. J. E. M.; Deelder, A. M.; Wuhrer, M. Fc specific IgG glycosylation profiling by robust nano-reverse phase HPLC-MS using a sheath-flow ESI sprayer interface. *J. Proteomics* **2012**, *75*, 1318−1329.
(50) Calvano, C. D.; Zambonin, C. G.; Jensen, O. N. Assessment of lectin and HILIC based enrichment protocols for characterization of serum glycoproteins by mass spectrometry. *J. Proteomics* **2008**, *71*, 304−17.
(51) Gonzalez-Quintela, A.; Alende, R.; Gude, F.; Campos, J.; Rey, J.; Meijide, L. M.; Fernandez-Merino, C.; Vidal, C. Serum levels of immunoglobulins (IgG, IgA, IgM) in a general adult population and their relationship with alcohol consumption, smoking and common metabolic abnormalities. *Clin. Exp. Immunol.* **2008**, *151*, 42−50.
(52) Sinha, S.; Pipes, G.; Topp, E. M.; Bondarenko, P. V.; Treuheit, M. J.; Gadgil, H. S. Comparison of LC and LC/MS Methods for Quantifying N-Glycosylation in Recombinant IgGs. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1643−1654.
(53) Arnold, J. N.; Wormald, M. R.; Sim, R. B.; Rudd, P. M.; Dwek, R. A. The Impact of Glycosylation on the Biological Function and Structure of Human Immunoglobulins. *Annu. Rev. Immunol.* **2007**, *25*, 21−50.
(54) Tanaka, A.; Iwase, H.; Hiki, Y.; Kokubo, T.; Ishii-Karakasa, I.; Toma, K.; Kobayashi, Y.; Hotta, K. Evidence for a site-specific fucosylation of N-linked oligosaccharide of immunoglobulin A1 from normal human serum. *Glycoconjugate J.* **1998**, *15*, 995−1000.
(55) Royle, L.; Roos, A.; Harvey, D. J.; Wormald, M. R.; Van Gijlswijk-Janssen, D.; Redwan, E.-R. M.; Wilson, I. A.; Daha, M. R.; Dwek, R. A.; Rudd, P. M. Secretory IgA N- and O-Glycans Provide a Link between the Innate and Adaptive Immune Systems. *J. Biol. Chem.* **2003**, *278*, 20140−20153.
(56) Satomi, Y.; Shimonishi, Y.; Takao, T. N-glycosylation at Asn491 in the Asn-Xaa-Cys motif of human transferrin. *FEBS Lett.* **2004**, *576*, 51−56.
(57) 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*, e12419.
(58) Crispin, M. D.; Ritchie, G. E.; Critchley, A. J.; Morgan, B. P.; Wilson, I. A.; Dwek, R. A.; Sim, R. B.; Rudd, P. M. Monoglucosylated glycans in the secreted human complement component C3: implications for protein biosynthesis and structure. *FEBS Lett.* **2004**, *566*, 270−4.

DOI: 10.1021/acs.jproteome.5b00756
J. Proteome Res. XXXX, XXX, XXX−XXX