

Isomer-Specific LC/MS and LC/MS/MS Profiling of the Mouse Serum N-Glycome Revealing a Number of Novel Sialylated N-Glycans

Serenus Hua,^{†,‡} Ha Neul Jeong,[†] Lauren M. Dimapasoc,[§] Inae Kang,[†] Chanyoung Han,[†] Jong-Soon Choi,^{†,⊥} Carlito B. Lebrilla,[§] and Hyun Joo An^{*,†,‡}

[†]Graduate School of Analytical Science and Technology, Chungnam National University, Daejeon, Korea

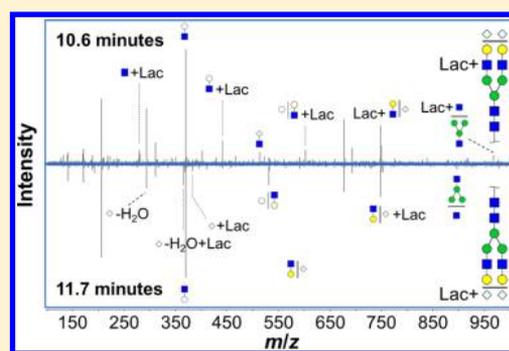
[‡]Cancer Research Institute, Chungnam National University, Daejeon, Korea

[§]Department of Chemistry, University of California, Davis, California, United States

[⊥]Division of Life Science, Korea Basic Science Institute, Daejeon, Korea

S Supporting Information

ABSTRACT: Mice are the premier mammalian models for studies of human physiology and disease, bearing extensive biological similarity to humans with far fewer ethical, economic, or logistic complications. To facilitate glycomic studies based on the mouse model, we comprehensively profiled the mouse serum N-glycome using isomer-specific nano-LC/MS and -LC/MS/MS. N-Glycans were identified by accurate mass MS and structurally elucidated by MS/MS. Porous graphitized carbon nano-LC was able to separate out nearly 300 N-linked glycan compounds (including isomers) from just over 100 distinct N-linked glycan compositions. Additional MS/MS structural analysis was performed on a number of novel N-glycans, revealing the structural characteristics of modifications such as dehydration, O-acetylation, and lactylation. Experimental findings were combined with known glycobiology to generate a theoretical library of all biologically possible mouse serum N-glycan compositions. The library may be used for automated identification of complex mixtures of mouse N-glycans, with possible applications to a wide range of mouse-related research endeavors, including pharmaceutical drug development and biomarker discovery.



In vivo medical research relies primarily on the common laboratory mouse, who exhibits an inherent similarity to its fellow mammal, the human. Mouse models exist for a vast multitude of human diseases, including cancer, diabetes, atherosclerosis, obesity, and even attention deficit hyperactivity disorder.^{1–3} Combined with the small size and rapid life cycle of the mouse, these disease models provide a relatively inexpensive and high-throughput path toward understanding the biology behind human afflictions.

Numerous studies have elucidated the genome, proteome, and metabolome of the various mouse models available, but only a few have paid any attention to the mouse glycome.^{4–6} However, glycosylation is one of the most common eukaryotic post-translational modifications, and aberrant glycosylation has been linked to cancer and other important human diseases by over fifty years of glycobiology.⁷ Additionally, recent studies have shown that changes in mouse glycosylation mirror those in human glycosylation for diseases as diverse as breast cancer,⁸ diabetes,⁹ and radiation sickness.¹⁰ Elucidation of the mouse glycome would aid glycomic and glycoproteomic biomarker studies based on the mouse model and advance our understanding of mice as models for humans.

Unfortunately, glycans possess significant structural complexity, and characterization presents a significant analytical challenge. Whereas sequence is sufficient to determine the

primary structures of linear biomolecules such as nucleic acids and proteins, glycans are branched, with a plethora of stereo- and regioisomers for each composition. Furthermore, since each glycan isomer is produced by a different set of glycosyltransferases, different isomers with the same composition cannot be taken as biologically equivalent. As a result, isomer separation and differentiation is a crucial component of glycan analysis.

Recently, LC/MS has emerged as the premier platform for isomer-specific glycomics.¹¹ Complex glycan mixtures are separated online by isomer-sensitive chromatographic methods such as porous graphitized carbon (PGC) or hydrophilic interaction liquid chromatography (HILIC) and then differentiated by accurate mass analysis with high-resolution Fourier transform (FT-) or time-of-flight (TOF)-MS detectors.^{12,13} Accompanying retention time and/or MS/MS spectral libraries enable rapid identification of specific glycan structures.^{14–16} These methods have already yielded promising results in numerous applications ranging from glycan biomarker discovery to glycosylated biopharmaceutical analysis.^{17–23}

Received: January 28, 2013

Accepted: March 27, 2013

Published: March 27, 2013

Glycomics studies based on the mouse model have great research potential; however, many glycomics tools must first be adapted and optimized for mice. To aid this process, we have comprehensively characterized the mouse serum N-glycome using chip-based nano-LC/MS and -LC/MS/MS. Isomer-specific PGC nano-LC was used to separate native, underivatized mouse serum N-glycans. Accurate mass MS was used to compositionally annotate and profile the N-glycans, while MS/MS was used to confirm N-glycan identities and structurally characterize them. Detailed structural analysis by MS/MS additionally revealed a number of novel N-glycan modifications, including dehydration, O-acetylation, and lactylation. Combining these empirical findings with our knowledge of the N-glycan biosynthesis pathway,²⁴ we created a library of all biologically plausible mouse serum N-glycan compositions.

■ EXPERIMENTAL SECTION

Acquisition of Mouse Sera. Female inbred FVB mice weighing 12 to 14 g were purchased directly from Charles River Laboratories (Wilmington, MA) and maintained in University of California vivariums (Davis, CA) according to NIH guidelines. Sera were collected via orbital eye bleeding, once per day, for five consecutive days.

Enzymatic Release of N-Glycans. N-Glycan release and associated processing steps were performed according to optimized procedures developed for human serum N-glycan analysis.²⁵ Briefly, serum proteins were thermally denatured in an aqueous solution of ammonium bicarbonate and dithiothreitol prior to digestion by peptide N-glycosidase F (New England Biolabs, Ipswich, MA) in a microwave reactor (CEM Corporation, Matthews, NC). Deglycosylated proteins were precipitated out with chilled ethanol, leaving a glycan-rich supernatant that was dried in vacuo.

N-Glycan Enrichment with Graphitized Carbon SPE. Released N-glycans were purified by graphitized carbon solid-phase extraction according to optimized procedures developed for human serum N-glycan analysis.²⁵ Briefly, an automated liquid handler (Gilson, Middleton, WI) conditioned graphitized carbon cartridges (Grace Davison, Deerfield, IL) with water; eluted aqueous N-glycan solutions; washed with water; then eluted serum N-glycans with 40% acetonitrile and 0.05% trifluoroacetic acid (v/v) in water. Samples were dried in vacuo.

Chromatographic Separation and MS Analysis of the Serum N-Glycome. Samples were analyzed using an HPLC-Chip Quadrupole Time-of-Flight (Chip/Q-TOF) MS system (Agilent Technologies, Santa Clara, CA) comprising an autosampler (maintained at 6 °C), capillary pump, nano pump, HPLC-Chip/MS interface, and a 6520 Q-TOF MS detector. The chip used consisted of a 9 × 0.075 mm i.d. enrichment column and a 43 × 0.075 mm i.d. analytical column, both packed with 5 μm porous graphitized carbon as the stationary phase, with an integrated nano-ESI spray tip. Chromatographic separation was performed with slight modifications from optimized procedures developed for human serum N-glycan analysis.²² For each sample, 1.0 μL (corresponding to 200 nL of serum) was loaded onto the enrichment column and washed with a solution of 3.0% acetonitrile and 0.5% formic acid (v/v) in water at 4.0 μL/min. A rapid glycan elution gradient was delivered at 0.4 μL/min using solutions of (A) 3.0% acetonitrile and 0.5% formic acid (v/v) in water and (B) 90.0% acetonitrile and 0.5% formic acid (v/v) in water, at the following proportions and time points:

5% to 32.8% B, 0 to 13.3 min; 32.8% to 35.9% B, 13.3 to 16.5 min. Remaining nonglycan compounds were flushed out with 100% B at 0.8 μL/min for 5 min. Finally, the analytical column was re-equilibrated with 5% B at 0.8 μL/min for 10 min, while the enrichment column was re-equilibrated with 0% B at 8 μL/min for 10 min. The drying gas temperature was set at 325 °C with a flow rate of 4 L/min (2 L of filtered nitrogen gas and 2 L of filtered dry compressed air).

MS spectra were acquired in positive ionization mode over a mass range of m/z 500–2000 with an acquisition time of 1.5 s per spectrum. MS/MS spectra were acquired in positive ionization mode over a mass range of m/z 100–3000 with an acquisition time of 1.5 s per spectrum. Following an MS scan, precursor compounds were automatically selected for MS/MS analysis by the acquisition software based on ion abundance and charge state ($z = 2, 3, \text{ or } 4$) and isolated in the quadrupole with a mass bandpass fwhm (full width at half-maximum) of 1.3 m/z . In general, collision energies for CID fragmentation were calculated for each precursor compound based on the following formula:

$$V_{\text{collision}} = 3.6 V \left(\frac{m/z}{100 \text{ Da}} \right) - 4.8 V$$

Here, $V_{\text{collision}}$ is the potential difference across the collision cell. The slope and offset values of the energy- m/z ramp could be changed as needed to produce more or less fragmentation.²⁶

LC/MS Data Processing and N-Glycan Identification by Accurate Mass. Following data acquisition, raw LC/MS data was processed using the Molecular Feature Extractor algorithm included in the MassHunter Qualitative Analysis software (version B.04.00 SP2, Agilent Technologies). MS peaks were filtered with a signal-to-noise ratio of 5.0 and parsed into individual ion species. Using expected isotopic distribution, charge state information, and retention time, all ion species associated with a single compound (e.g., the doubly protonated ion, the triply protonated ion, and all associated isotopologues) were summed together, and the neutral monoisotopic mass of the compound was calculated. Using this information, a list of all compound peaks in the sample was generated, with abundances represented by chromatographic peak areas.

Computerized algorithms were used to identify N-glycan compositions by accurate mass.²⁷ Deconvoluted experimental masses were compared against theoretical glycan masses using a mass error tolerance of 5 ppm. On the basis of known mouse serum N-glycosylation patterns, glycan compositions containing hexose (Hex), N-acetylhexosamine (HexNAc), fucose (Fuc), N-acetylneuraminic acid (NeuAc), N-glycolylneuraminic acid (NeuGc), O-acetylation (OAc), and dehydration ($- \text{H}_2\text{O}$) were considered.

■ RESULTS AND DISCUSSION

Separation of Mouse Serum N-Glycans by Nano-LC. Chip-based porous graphitized carbon (PGC) nano-LC provided isomer-specific separation of mouse serum N-glycans. Previously, PGC nano-LC has been used with great success for isomer-specific separation of free oligosaccharides, glycopeptides, and human serum N-glycans;^{19–22,26} however, this study represents the first application of PGC toward the separation of mouse serum N-glycans. The isomeric heterogeneity of mouse serum N-glycans presents a considerable analytical challenge that is significantly addressed by the use of an isomer-specific stationary phase such as PGC.

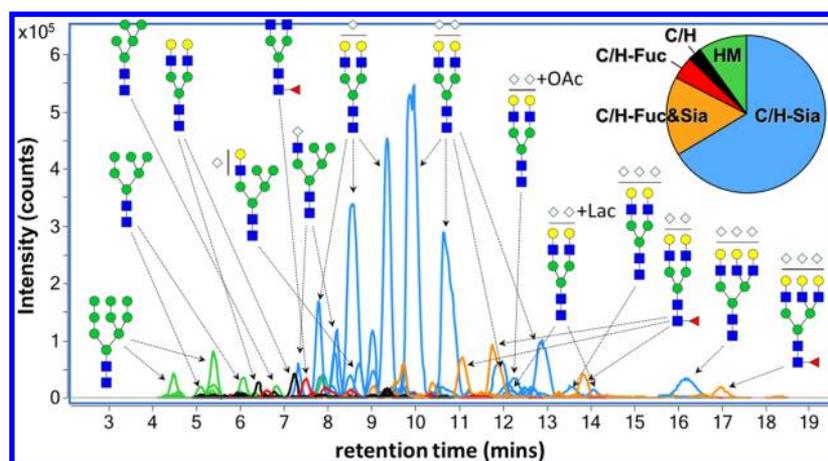


Figure 1. Overlaid chromatographic traces of mouse serum N-glycans identified by PGC nano-LC/MS. Inset, relative abundances of each N-glycan class: high mannose (HM); undecorated complex/hybrid (C/H); fucosylated complex/hybrid (C/H-Fuc); sialylated complex/hybrid (C/H-Sia); and fucosylated/sialylated complex/hybrid (C/H-Fuc&Sia). Nearly 300 individual N-glycan species (including isomers) were detected; however, not all are visible here due to the wide dynamic range of abundances. For the glycan cartoons, green circles denote mannose; yellow circles denote galactose; blue squares denote *N*-acetylglucosamine; red triangles denote fucose; light blue diamonds denote *N*-glycolylneuraminic acid.

The chromatographic traces in Figure 1 represent all of the mouse serum N-glycans identified by PGC nano-LC/MS. Sialylated glycans (including fucosylated/sialylated ones) make up the vast majority (83%). Significant isomer separation was achieved, with nearly 300 N-glycan compounds (including isomers) identified from just over 100 distinct N-glycan compositions. Mouse serum glycans spanned 5 orders of magnitude in abundance (by ion count), underscoring the necessity of the wide dynamic range provided by TOF MS detectors.

Glycan retention times followed a number of consistent trends. As seen in Figure 1, smaller and/or neutral glycans generally eluted earlier, while larger and/or acidic glycans generally eluted later. For example, high mannose glycans eluted first, followed by neutral (unsialylated) complex/hybrid glycans, and then acidic (sialylated) complex/hybrid glycans. Fucosylated glycans typically eluted later than their unfucosylated counterparts. These observations are consistent with previous studies on PGC separation of human serum N-glycans.^{21,22}

Compositions were initially assigned according to accurate mass, as described in the Experimental Section section. Primarily, mouse serum N-glycans are known to contain hexose (Hex), *N*-acetylhexosamine (HexNAc), fucose (Fuc), and *N*-glycolylneuraminic acid (NeuGc) as their basic building blocks. *N*-Acetylneuraminic acid (NeuAc) is also a biosynthetic possibility,²⁸ as are modifications such as dehydration, *O*-acetylation, or lactylation, though these are present only at low levels. Due to the vast number of compositional possibilities, extensive MS/MS experiments were performed in order to confirm accurate mass assignments and remove any compositional ambiguities.

MS/MS Screening for NeuAc. In mice (and, indeed, nearly all mammals, with the notable exception of humans), NeuAc is the biosynthetic precursor to NeuGc, the sialic acid that decorates the vast majority of mouse serum N-glycans.²⁸ Though NeuAc-containing N-glycans have rarely been reported in mouse serum except in cases of disease,²⁹ the role of NeuAc as a NeuGc precursor warranted its serious consideration as a compositional possibility even in healthy mouse serum N-glycans.

For glycan mass profiling, the possible presence of NeuAc, though rare, presents a significant issue. Monosaccharide combinations Fuc₁NeuGc₁ (C₆H₁₀O₄ + C₁₁H₁₇NO₉) and Hex₁NeuAc₁ (C₆H₁₀O₅ + C₁₁H₁₇NO₈) have identical molecular formulas of C₁₇H₂₇NO₁₃ and cannot be differentiated by mass alone. As a result, MS-only compositional assignments containing both Fuc and NeuGc (146.06 Da + 307.09 Da = 453.15 Da) are ambiguous, as they may also be alternately interpreted as containing Hex and NeuAc (162.05 Da + 291.10 Da = 453.15 Da). For example, the common mouse serum N-glycan Hex₃HexNAc₃Fuc₁NeuGc₁ (2296.82 Da) has exactly the same mass as Hex₆HexNAc₃NeuAc₁, despite vastly different biosynthetic origins. This compositional ambiguity can only be resolved with MS/MS, in which NeuGc (307.09 Da) and NeuAc (291.10 Da) fragments are prevalent and easily distinguished.

To establish the presence or absence of NeuAc-containing N-glycans, all mouse serum N-glycans were interrogated by tandem MS/MS. Spectra were quickly screened for the presence of typical NeuAc-associated fragments, including *m/z* 292.10 (NeuAc), *m/z* 274.09 (NeuAc-H₂O), and *m/z* 657.23 (Hex₁HexNAc₁NeuAc₁).

A representative MS/MS spectrum of NeuAc-containing glycan Hex₃HexNAc₄Fuc₁NeuAc₁ (*m/z* 1039.88) may be seen in Figure S1, Supporting Information. Consistent with the results of the screening process, typical NeuAc-associated fragments were found to be present in high abundance, while typical (or even atypical) NeuGc-associated fragments were completely absent. The mass 2077.75 Da, therefore, was associated *solely* with NeuAc-containing glycan Hex₃HexNAc₄Fuc₁NeuAc₁, without any contribution from the isomeric, NeuGc-containing glycan Hex₄HexNAc₄Fuc₂NeuGc₁ (also 2077.75 Da).

In all, only two N-glycans were found to contain NeuAc-associated fragments: biantennary fucosylated *di*-sialylated N-glycan Hex₃HexNAc₄Fuc₁NeuAc₂ (2368.84 Da) and biantennary fucosylated *mono*-sialylated N-glycan Hex₃HexNAc₄Fuc₁NeuAc₁ (2077.75 Da). No evidence was found, either by MS accurate mass assignment or by MS/MS fragment screening, of any N-glycans displaying a combination of NeuAc and NeuGc sialylation. Together, N-glycans containing NeuAc comprised

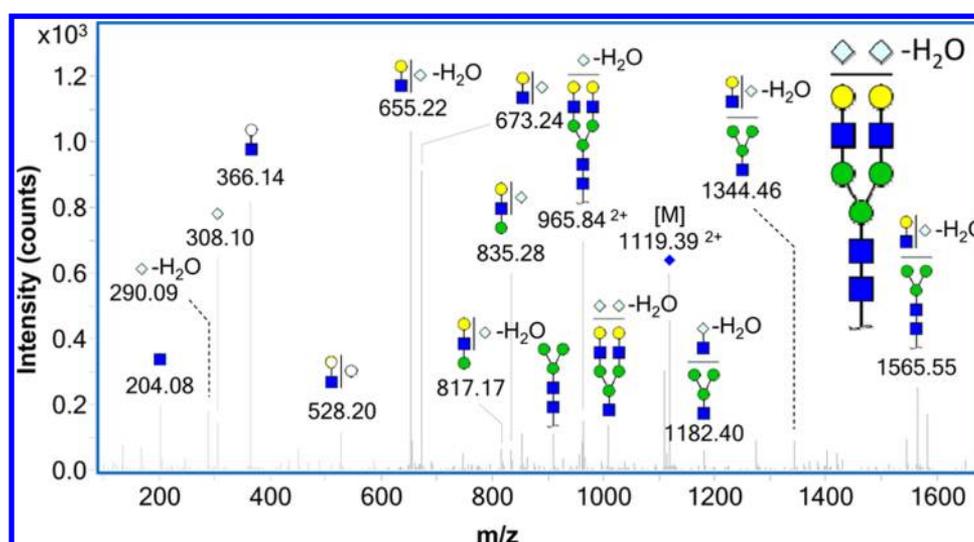


Figure 2. Positive mode CID MS/MS spectrum of mouse serum N-glycan Hex₃HexNAc₄NeuGc₂-H₂O (doubly protonated m/z 1119.39) confirming the presence of dehydration. Glycan cartoons follow earlier shape/color conventions with the addition of white circles, which denote hexose of an indeterminate type.

only 0.05% of the total glycan signal (by ion count), indicating the general rarity of NeuAc-containing N-glycans in mouse serum.

MS/MS Screening for Sialic Acid Dehydration. For sialic acids (NeuAc and NeuGc), dehydration ($-H_2O$) can occur during normal glycan biosynthesis as well as during ionization, immediately prior to MS analysis.³⁰ Thus, when dehydrated sialoforms are detected, great care must be taken to determine whether the origin of these modifications is instrumental or sample related. Isomer-sensitive PGC separation provides an easy route toward verification. Figure S2, Supporting Information, shows the overlaid chromatograms of *dehydrated* N-glycan Hex₃HexNAc₄NeuGc₂-H₂O (m/z 1119.39; solid pink line) and *nondehydrated* N-glycan Hex₃HexNAc₄NeuGc₂ (m/z 1128.39; dotted black line), both of which were detected in mouse serum. While both compounds have multiple isomers, these isomers elute at different times. In contrast, had dehydration occurred during ionization, peak shapes and retention times would be identical between the chromatograms (albeit at different intensities). The differing chromatographic profiles therefore suggest that the detected dehydration is indeed not of instrumental origin but rather represents real compounds found in the sample. These results are corroborated by previous analyses of human serum glycans under similar conditions,^{21,22} during which no sialic acid dehydration was detected whatsoever.

When profiling complex glycan mixtures, the presence of dehydrated sialic acids is easily detected by a combination of MS and MS/MS, simply by looking for neutral mass losses of 18.01 Da from sialylated precursor masses and fragments. Screening of MS/MS spectra for high abundances at m/z 655.22 (Hex₁HexNAc₁NeuGc₁-H₂O) quickly reveals glycans in which at least one sialic acid on the precursor is dehydrated. A subsequent match of the precursor mass to a dehydrated N-glycan mass is used to verify the assignment. This two-step verification process enables sensitive and specific detection of sialic acid dehydration.

For example, Figure 2 shows a representative MS/MS spectrum of biantennary disialylated dehydrated N-glycan Hex₃HexNAc₄NeuGc₂-H₂O (m/z 1119.39). The fragment at

m/z 655.22 (Hex₁HexNAc₁NeuGc₁-H₂O) is the most abundant peak in the spectrum, indicating the presence of a dehydrated NeuGc in the precursor. Fragments at m/z 1344.46 (Hex₄HexNAc₂NeuGc₁-H₂O), m/z 1182.40 (Hex₃HexNAc₂NeuGc₁-H₂O), m/z 965.84 (Hex₃HexNAc₄NeuGc₁-H₂O), and m/z 817.17 (Hex₂HexNAc₁NeuGc₁-H₂O) provide further support for the presence of a dehydrated NeuGc. Finally, the precursor mass (2236.76 Da) is exactly 18.01 Da from the mass of Hex₃HexNAc₄NeuGc₂ (2254.77 Da), the most abundant N-glycan in mouse serum, indicating that the precursor is indeed a dehydrated N-glycan.

It should be noted that, although the fragment at m/z 290.09 (NeuGc-H₂O) is also associated with dehydrated NeuGc, it is commonly present in MS/MS spectra of nondehydrated glycans, presumably as a result of MS/MS fragmentation. In contrast, the fragment at m/z 655.22 (Hex₁HexNAc₁NeuGc₁-H₂O) is found solely in MS/MS spectra of dehydrated glycans and thus serves as an extremely reliable marker of sialic acid dehydration.

Though the process of N-glycan dehydration has never been structurally characterized, the dehydration of ganglioside sialyloligosaccharide moieties has been explored in detail by ¹H NMR.³¹ On the basis of these findings, the dehydrated NeuGc₁Hex₁ structure proposed in Figure S4a, Supporting Information, is the product of a condensation reaction between the NeuGc carbonyl group and the spatially nearest hydroxyl group on an adjoining hexose. However, given the number of hydroxyl groups (and thus opportunities for dehydration) present on a typical N-glycan, further structural studies are clearly necessary to ascertain the precise nature and mechanism of this phenomenon.

In all, only two N-glycans were found to contain dehydrated NeuGc: biantennary *di*-sialylated N-glycan Hex₃HexNAc₄NeuGc₂-H₂O (2236.76 Da) and biantennary *tri*-sialylated N-glycan Hex₅HexNAc₄NeuGc₃-H₂O (2543.85 Da). The similarity of these two glycans (differing only by a single NeuGc) supports the accuracy of the assignment and suggests their biological association.

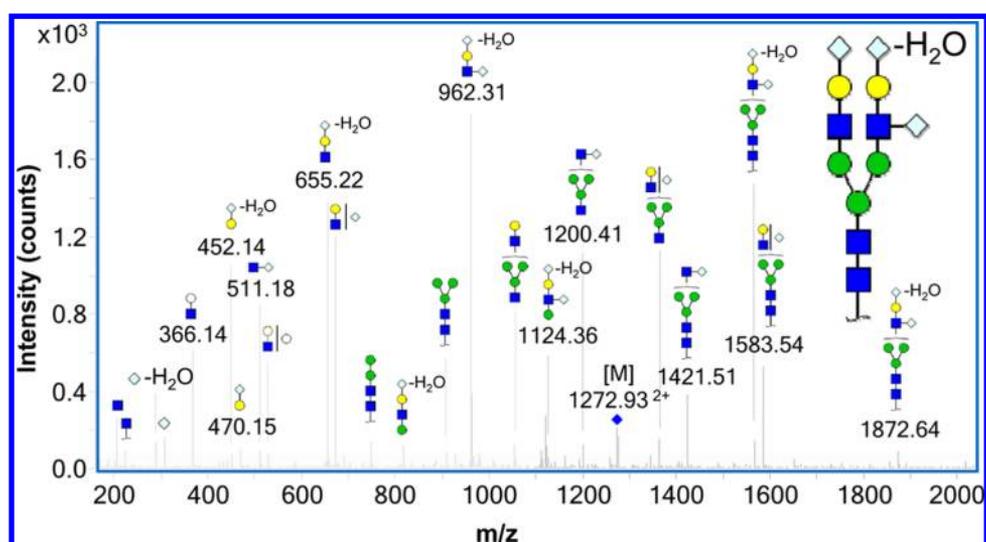


Figure 3. Positive mode CID MS/MS spectrum of mouse serum N-glycan Hex₃HexNAc₄NeuGc₃-H₂O (doubly protonated m/z 1272.93) confirming the presence of dehydration and antennal HexNAc sialylation.

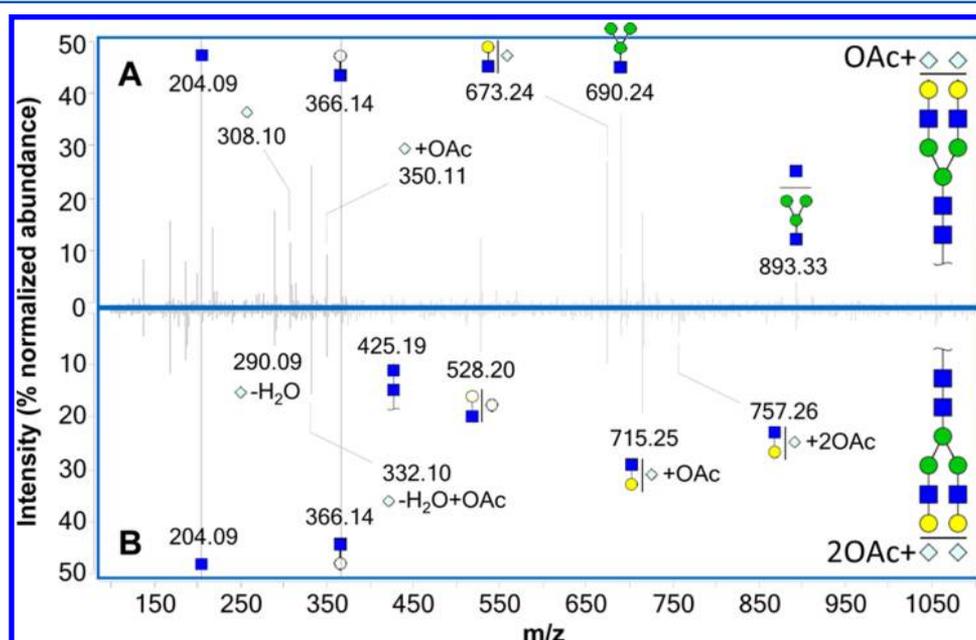


Figure 4. Positive mode CID MS/MS spectra of mouse serum N-glycans (A) Hex₃HexNAc₄NeuGc₂+OAc (doubly protonated m/z 1149.40), on the top, and (B) Hex₃HexNAc₄NeuGc₂+2 OAc (doubly protonated m/z 1170.40), on the bottom, confirming the presence of O-acetylation.

MS/MS Elucidation of Antennal HexNAc Sialylation.

Numerous studies have reported mouse serum N-glycans with more sialic acids than antennae,^{10,29} indicating that some of the antennae must be decorated with two (or more) sialic acids. However, these N-glycans are only rarely structurally explored; rather, putative structures are typically assigned on the basis of the assumption of common mammalian glycan motifs,³² with the first sialic acid attached to an antennal Hex and the second sialic acid attached to an antennal HexNAc. For this study, initially putative assignments of antennal HexNAc sialylation were subsequently confirmed and structurally elucidated by MS/MS.

Figure 3 shows a representative MS/MS spectrum of biantennary trisialylated dehydrated N-glycan Hex₃HexNAc₄NeuGc₃-H₂O (m/z 1272.93). Fragments at m/z 511.18 (HexNAc₁NeuGc₁), m/z 1200.41

(Hex₃HexNAc₂NeuGc₁), and m/z 1421.51 (Hex₃HexNAc₃NeuGc₁) confirm the attachment of one NeuGc to an antennal HexNAc. Additionally, fragments at m/z 452.14 (Hex₁NeuGc₁-H₂O) and m/z 470.15 (Hex₁NeuGc₁), combined with the absence of a fragment at m/z 493.17 (HexNAc₁NeuGc₁-H₂O), suggest that the remaining two sialic acids (consisting of one NeuGc and one NeuGc-H₂O) are both attached to an antennal Hex. Finally, fragments at m/z 962.31 (Hex₁HexNAc₁NeuGc₂-H₂O), m/z 1124.36 (Hex₂HexNAc₁NeuGc₂-H₂O), and m/z 1872.64 (Hex₄HexNAc₃NeuGc₂-H₂O) indicate that the HexNAc-attached NeuGc is on the same antenna as the Hex-attached NeuGc-H₂O. Together, these observations form a detailed structural image of biantennary trisialylated dehydrated N-glycan Hex₃HexNAc₄NeuGc₃-H₂O.

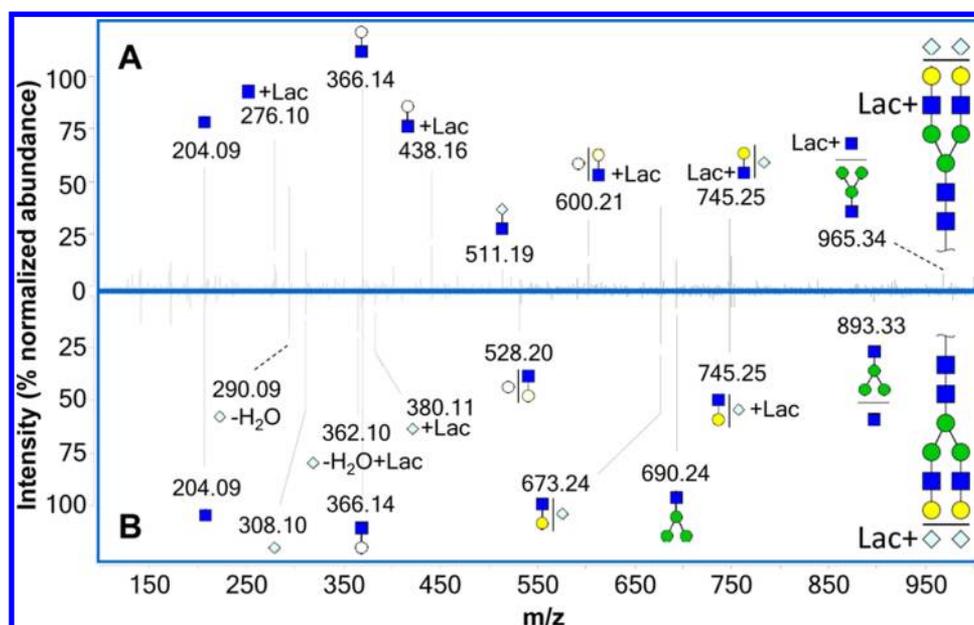


Figure 5. Positive mode CID MS/MS spectra of two isomers of mouse serum N-glycan Hex₅HexNAc₄NeuGc₂+Lac (doubly protonated m/z 1164.40) eluting at (A) 10.6 min, on the top, and (B) 11.7 min, on the bottom, confirming the presence of lactylation.

To further examine antennal HexNAc sialylation, MS/MS spectra were acquired of N-glycan Hex₃HexNAc₃NeuGc₁ (m/z 711.26), the smallest possible sialylated N-glycan in mouse serum, consisting of just the N-glycan core and a sialylated HexNAc (Figure S4, Supporting Information). Fragments at m/z 673.24 (Hex₁HexNAc₁NeuGc₁), m/z 835.30 (Hex₂HexNAc₁NeuGc₁), and m/z 997.33 (Hex₃HexNAc₁NeuGc₁) confirm the presence of NeuGc and suggest its attachment to the antennal HexNAc. Since sialylation of the N-glycan core is not biologically plausible, the antennal HexNAc is the only possible point of attachment for the NeuGc, effectively verifying this glycan motif in mice.

The presence of antennal HexNAc sialylation significantly increases the potential heterogeneity of mouse serum N-glycans. The number of possible N-glycan compositions rises simply because each N-glycan can be decorated with higher degrees of sialylation. Isomeric possibilities must also be considered, both as a result of differential placement (connected to either the antennal Hex or HexNAc) and with respect to the saccharide linkage.

O-Acetyl Modification of Sialylated N-Glycans. Sialic acids are known to exhibit a vast amount of micro-heterogeneity; however, few studies have been able to examine modified sialic acids in the context of the glycans to which they are attached. To further explore this area, sialylated N-glycans exhibiting modifications such as O-acetylation and O-lactylation were interrogated by MS/MS.

One of the known modifications of mouse serum N-glycans is O-acetylation (OAc), a modification of sialic acid in which an acetyl group (42.01 Da) is added to the hydroxyl groups at C4, C7, C8, and/or C9 of the sialic acid residue by O-acetyl transferases and 9-O-acetyl esterases.³³ In Figure 4, MS/MS spectra of (A) *mono*-O-acetylated biantennary disialylated N-glycan Hex₃HexNAc₄NeuGc₂+OAc (m/z 1149.40) and (B) *di*-O-acetylated Hex₃HexNAc₄NeuGc₂+2OAc (m/z 1170.40) are shown in parallel for easy comparison. Fragments at m/z 332.10 (NeuGc-H₂O+OAc), m/z 350.11 (NeuGc+OAc), and m/z 715.25 (Hex₁HexNAc₁NeuGc₁+OAc), common to both

N-glycans, confirm the presence of O-acetylation and its attachment to NeuGc. The fragment at m/z 757.26 (Hex₁HexNAc₁NeuGc₁+2OAc), present only in Figure 4B, further suggests that both O-acetyl modifications in this N-glycan are in fact attached to a single NeuGc (rather than one OAc modification on each NeuGc).

N-Glycans displaying di-O-acetylation of a single sialic acid have been identified previously²³ but never before in mouse serum. On the basis of known glycobiology, possible structures for singly and doubly O-acetylated NeuGc are proposed in Figure S3b,c, respectively, Supporting Information. Of course, since O-acetylation can occur at C4, C7, C8, and/or C9 of a sialic acid,³³ there is an abundance of other O-acetylated NeuGc isomers to consider as well.

O-Acetylation is important not just biologically but also medically. O-Acetylated sialic acids mediate binding of several viruses, including common cold viruses such as influenza C as well as certain coronaviruses.³⁴ Additionally, O-acetylation increases the half-life and potency of certain biopharmaceuticals by interfering with *in vivo* enzymatic desialylation of the N-glycans present on the drug.³⁵ Thus, detailed characterization of O-acetylation is of great concern to both producers and regulators of biopharmaceutical drugs and vaccines.

Novel O-Lactyl Modification of Sialylated N-Glycans. Close examination of the mouse serum N-glycan mass profile also revealed the presence of multiple glycan isomers with m/z 1164.40, corresponding to biantennary disialylated N-glycan Hex₅HexNAc₄NeuGc₂ with a mass shift of +72.02 Da, i.e., a lactyl group. Lactylation is not a well-studied phenomenon, perhaps due to the unavailability (until now) of suitable analytical tools and techniques. However, lactylated (and otherwise modified) sialic acid monosaccharides have previously been observed by GC/MS following acid hydrolysis and trimethylsilyl (TMS) derivatization.^{36,37} On the basis of these studies, a 9-O-lactylated NeuGc is proposed in Figure S3d, Supporting Information.

Despite their occasional presence on human and equine sialic acid monosaccharides^{36,37} as well as bacterial glycolipids,³⁸

lactylated N-glycans have never before been reported from any species. Isomer-specific MS/MS was therefore employed to investigate this phenomenon in detail. Figure 5 shows the MS/MS spectra of two isomers of lactylated biantennary disialylated N-glycan Hex₃HexNAc₄NeuGc₂+Lac (*m/z* 1164.40) eluting at (A) 10.6 min and (B) 11.7 min. Interestingly, though both isomers show multiple Lac-associated glycan fragments, there are few common Lac-associated fragments between the two isomers. For the isomer eluting at 10.6 min (Figure 5A), fragments at *m/z* 276.10 (HexNAc+Lac), *m/z* 438.16 (Hex₁HexNAc₁+Lac), *m/z* 600.21 (Hex₂HexNAc₁+Lac), *m/z* 745.25 (Hex₁HexNAc₁NeuGc₁+Lac), and *m/z* 965.34 (Hex₃HexNAc₂+Lac) confirm the presence of lactylation and suggest its attachment to an antennal HexNAc. Additionally, the fragment at *m/z* 511.19 (HexNAc₁NeuGc₁) suggests that at least one NeuGc is attached to an antennal HexNAc rather than an antennal Hex. For the isomer eluting at 11.7 min (Figure 5B), fragments at *m/z* 362.10 (NeuGc-H₂O+Lac), *m/z* 380.11 (NeuGc+Lac), and *m/z* 745.25 (Hex₁HexNAc₁NeuGc₁+Lac) likewise confirm the presence of lactylation but suggest its attachment to a NeuGc instead.

Glycan rearrangement or migration reactions have occasionally been reported under CID conditions and thus cannot be wholly discounted when CID is used for structural elucidation. However, previous reports of glycan rearrangement have largely involved derivatized rather than native glycans. Molecular modeling suggests that, in fact, glycan rearrangement is mechanistically associated with (and enhanced by) aromatic tags (such as 2-AB) attached at the glycan reducing end.^{39,40} The use of native glycans in this study, combined with the sheer number of isomer-specific fragments, makes significant rearrangement seem unlikely here. However, as this is the first study to detect lactylation on an intact N-glycan, chromatographically separate isomers of lactylated N-glycans, or structurally characterize these isomers by MS/MS, further experiments will be necessary to confirm and expand upon these results.

In Silico Theoretical Library of Mouse Serum N-Glycans. Theoretical libraries of human serum N-glycans have been applied extensively and with great success to aid glycomic analyses of human serum. Essentially, these libraries consist of a finite list of biologically possible N-glycans with which glycan mass spectral data can be automatically annotated.²⁴ For research applications requiring high-throughput processing of hundreds or thousands of samples, such as biomarker studies, automated N-glycan annotation has proven invaluable.^{21,22}

To similarly aid glycomic analyses of mouse serum, an in silico theoretical library of mouse serum N-glycans was created. As previously described for the human serum N-glycan library,²⁴ established knowledge of mouse serum N-glycan biosynthesis was used to determine the theoretical largest plausible N-glycan for each of the three main N-glycan types: high mannose, complex, and hybrid type N-glycans. Known glycobiochemistry was augmented with structural information gleaned from MS/MS interrogation of mouse serum N-glycans, described earlier, specifically, that (a) only minimal levels of NeuAc were present, with almost all sialylation being of the NeuGc variety and (b) antennal HexNAc sialylation was indeed a structural possibility. The resulting “maximum” N-glycans, shown in Figure S5, Supporting Information, were sequentially degraded, one monosaccharide at a time, all the way down to the N-glycan core. Each step along this retrosynthetic degradation pathway became a separate entry in the theoretical

library, for a grand total of 710 mouse serum N-glycan compositions.

Minor modification of the library was necessary in order to accommodate the atypical glycosylation found in the present study. MS/MS-verified library entries for NeuAc-containing N-glycans Hex₅HexNAc₄Fuc₁NeuAc₁ (2077.75 Da) and Hex₅HexNAc₄Fuc₁NeuAc₂ (2368.84 Da) replaced the putative entries for their isomeric NeuGc-containing counterparts Hex₄HexNAc₄Fuc₂NeuGc₁ (2077.75 Da) and Hex₃HexNAc₄Fuc₃NeuGc₂ (2368.84 Da), respectively. Similarly, entries for dehydrated, mono-O-acetylated, di-O-acetylated, and lactylated Hex₅HexNAc₄NeuGc₂ as well as dehydrated Hex₅HexNAc₄NeuGc₃ were added to the library. With these modifications, the size of the theoretical library increased to 715 mouse serum N-glycan compositions. Automated analysis of the raw LC/MS data using the theoretical library resulted in the detection of 133 glycan compositions, in line with the initial manual analysis, with 89 of the compositions detected at less than 5% abundance.

Since exploration of the mouse serum N-glycome is still in its very early stages, further research may expand the current version of the mouse serum N-glycan library, provided here as Table S1, Supporting Information. Though the number of potential N-glycans stemming from basic biosynthetic pathways is expected to remain slightly over 700, different mouse strains may yield low levels of uncharacterized glycans exhibiting NeuAc sialylation, dehydration, O-acetylation, lactylation, and/or other modifications. As shown above, MS/MS screening for diagnostic fragments can rapidly detect these modified N-glycans, which would then be added to later versions of the library.

CONCLUSIONS

We have comprehensively profiled the mouse serum N-glycome using isomer-sensitive PGC nano-LC/MS and -LC/MS/MS. In addition to structural elucidation of rare glycosylation such as antennal HexNAc sialylation or decoration with NeuAc (rather than NeuGc), isomer-specific MS/MS was used to identify and characterize a number of novel N-glycans exhibiting modifications such as dehydration, O-acetylation, and lactylation. This study represents the first time that lactylated N-glycans have been observed, let alone structurally characterized.

A theoretical library of mouse serum N-glycans was created, combining known glycobiochemistry with N-glycan structural information provided by MS/MS interrogation. Atypical N-glycosylation products were also screened for by MS/MS and added to the library. With the incorporation of future experimental results, this library has the capacity to gradually evolve from a (well-grounded but still theoretical) construct of glycobiological rules into an empirical database of verified N-glycan compositions.

Comprehensive annotation and characterization of the mouse serum N-glycome has far-reaching applications. Drug developers commonly use mice to rapidly develop biotherapeutic monoclonal antibodies (mAbs); however, proper glycosylation of mAbs is an ongoing concern that can be significantly addressed by a library of possible mouse N-glycans. Scientists working on biomarker studies using the mouse model can also apply this library to rapidly and confidently identify diagnostic N-glycan markers. Ultimately, this strategy may even be expanded and adapted to develop libraries for the N-

glycomes of other important organisms or subsets thereof, such as rats, CHO cells, etc.

■ ASSOCIATED CONTENT

■ Supporting Information

Additional information as noted in text. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: hjan@cnu.ac.kr. Tel: 82-42-821-8547. Fax: 82-42-821-8551.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We would like to thank Jane Qian Chen and Alexander D. Borowsky (Center for Comparative Medicine, UC Davis) for the mouse serum samples used in this study and Rudolf Grimm (Agilent Technologies) for instrumentation and technical support. We are also grateful for the support provided by the 2012 University-Institute Cooperation Program via the National Research Foundation of Korea as well as the Converging Research Center Program (2012K001505 for H.J.A.) via the Ministry of Education, Science, and Technology.

■ REFERENCES

- (1) Kim, J.; Coffey, D. M.; Creighton, C. J.; Yu, Z.; Hawkins, S. M.; Matzuk, M. M. *Proc. Natl. Acad. Sci.* **2012**, *109*, 3921–3926.
- (2) Atkinson, M. A.; Leiter, E. H. *Nat. Med.* **1999**, *5*, 601–604.
- (3) Ormrod, D. J.; Holmes, C. C.; Miller, T. E. *Atherosclerosis* **1998**, *138*, 329–334.
- (4) Blake, J. A.; Bult, C. J.; Kadin, J. A.; Richardson, J. E.; Eppig, J. T.; Mouse Genome Database Group. *Nucleic Acids Res.* **2011**, *39*, D842–D848.
- (5) Krüger, M.; Moser, M.; Ussar, S.; Thievensen, I.; Luber, C. A.; Forner, F.; Schmidt, S.; Zanivan, S.; Fässler, R.; Mann, M. *Cell* **2008**, *134*, 353–364.
- (6) Patterson, A. D.; Bonzo, J. A.; Li, F.; Krausz, K. W.; Eichler, G. S.; Aslam, S.; Tigno, X.; Weinstein, J. N.; Hansen, B. C.; Idle, J. R.; Gonzalez, F. J. *J. Biol. Chem.* **2011**, *286*, 19511–19522.
- (7) Adamczyk, B.; Tharmalingam, T.; Rudd, P. M. *Biochim. Biophys. Acta (BBA), Gen. Subj.* **2012**, *1820*, 1347–1353.
- (8) de Leoz, M. L. A.; Young, L. J. T.; An, H. J.; Kronewitter, S. R.; Kim, J.; Miyamoto, S.; Borowsky, A. D.; Chew, H. K.; Lebrilla, C. B. *Mol. Cell. Proteomics* **2011**, *10*, DOI: 10.1074/mcp.M110.002717.
- (9) Itoh, N.; Sakaue, S.; Nakagawa, H.; Kuroguchi, M.; Ohira, H.; Deguchi, K.; Nishimura, S.-I.; Nishimura, M. *Am. J. Physiol. - Endocrinol. Metab.* **2007**, *293*, E1069–E1077.
- (10) Chaze, T.; Slomianny, M.-C.; Milliat, F.; Tarlet, G.; Lefebvre-Darroman, T.; Gourmelon, P.; Bey, E.; Benderitter, M.; Michalski, J.-C.; Guipaud, O. *Mol. Cell. Proteomics* **2012**, DOI: 10.1074/mcp.M111.014639.
- (11) Hua, S.; Lebrilla, C.; An, H. J. *Bioanalysis* **2011**, *3*, 2573–2585.
- (12) Wuhrer, M.; Koeleman, C. A. M.; Deelder, A. M.; Hokke, C. H. *Anal. Chem.* **2003**, *76*, 833–838.
- (13) Pabst, M.; Bondili, J. S.; Stadlmann, J.; Mach, L.; Altmann, F. *Anal. Chem.* **2007**, *79*, 5051–5057.
- (14) Campbell, M. P.; Royle, L.; Radcliffe, C. M.; Dwek, R. A.; Rudd, P. M. *Bioinformatics* **2008**, *24*, 1214–1216.
- (15) Aldredge, D.; An, H. J.; Tang, N.; Waddell, K.; Lebrilla, C. B. *J. Proteome Res.* **2012**, *11*, 1958–1968.
- (16) Wu, S.; Salcedo, J.; Tang, N.; Waddell, K.; Grimm, R.; German, J. B.; Lebrilla, C. B. *Anal. Chem.* **2012**, *84*, 7456–7462.

- (17) An, H. J.; Gip, P.; Kim, J.; Wu, S.; Park, K. W.; McVaugh, C. T.; Schaffer, D. V.; Bertozzi, C. R.; Lebrilla, C. B. *Mol. Cell. Proteomics* **2012**, *11*, DOI: 10.1074/mcp.M111.010660.
- (18) Hua, S.; An, H. J. *BMB Rep.* **2012**, *45*, 323–330.
- (19) Totten, S. M.; Zivkovic, A. M.; Wu, S.; Ngyuen, U.; Freeman, S. L.; Ruhaak, L. R.; Darboe, M. K.; German, J. B.; Prentice, A. M.; Lebrilla, C. B. *J. Proteome Res.* **2012**, *11*, 6124–6133.
- (20) Tao, N.; Wu, S.; Kim, J.; An, H. J.; Hinde, K.; Power, M. L.; Gagneux, P.; German, J. B.; Lebrilla, C. B. *J. Proteome Res.* **2011**, *10*, 1548–1557.
- (21) Hua, S.; An, H. J.; Ozcan, S.; Ro, G. S.; Soares, S.; DeVere-White, R.; Lebrilla, C. B. *Analyst* **2011**, *136*, 3663–3671.
- (22) Hua, S.; Williams, C. C.; Dimapasoc, L. M.; Ro, G. S.; Ozcan, S.; Miyamoto, S.; Lebrilla, C. B.; An, H. J.; Leiserowitz, G. S. *J. Chromatogr., A* **2013**, DOI: <http://dx.doi.org/10.1016/j.chroma.2012.12.079>.
- (23) Oh, M. J.; Hua, S.; Kim, B. J.; Jeong, H. N.; Jeong, S. H.; Grimm, R.; Yoo, J. S.; An, H. J. *Bioanalysis* **2013**, *5*, 545–559.
- (24) Kronewitter, S. R.; An, H. J.; de Leoz, M. L.; Lebrilla, C. B.; Miyamoto, S.; Leiserowitz, G. S. *Proteomics* **2009**, *9*, 2986–2994.
- (25) Kronewitter, S. R.; de Leoz, M. L. A.; Peacock, K. S.; McBride, K. R.; An, H. J.; Miyamoto, S.; Leiserowitz, G. S.; Lebrilla, C. B. *J. Proteome Res.* **2010**, *9*, 4952–4959.
- (26) Hua, S.; Nwosu, C.; Strum, J.; Seipert, R.; An, H.; Zivkovic, A.; German, J.; Lebrilla, C. *Anal. Bioanal. Chem.* **2012**, *403*, 1291–1302.
- (27) Kronewitter, S. R.; De Leoz, M. L. A.; Strum, J. S.; An, H. J.; Dimapasoc, L. M.; Guerrero, A.; Miyamoto, S.; Lebrilla, C. B.; Leiserowitz, G. S. *Proteomics* **2012**, *12*, 2523–2538.
- (28) Kawano, T.; Koyama, S.; Takematsu, H.; Kozutsumi, Y.; Kawasaki, H.; Kawashima, S.; Kawasaki, T.; Suzuki, A. *J. Biol. Chem.* **1995**, *270*, 16458–16463.
- (29) Lin, S.-Y.; Chen, Y.-Y.; Fan, Y.-Y.; Lin, C.-W.; Chen, S.-T.; Wang, A. H. J.; Khoo, K.-H. *J. Proteome Res.* **2008**, *7*, 3293–3303.
- (30) Hsu, N.-Y.; Yang, W.-B.; Wong, C.-H.; Lee, Y.-C.; Lee, R. T.; Wang, Y.-S.; Chen, C.-H. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 2137–2146.
- (31) Ando, S.; Yu, R. K.; Scarsdale, J. N.; Kusunoki, S.; Prestegard, J. H. *J. Biol. Chem.* **1989**, *264*, 3478–3483.
- (32) Green, E. D.; Adelt, G.; Baenziger, J. U.; Wilson, S.; Van Halbeek, H. *J. Biol. Chem.* **1988**, *263*, 18253–18268.
- (33) Klein, A. R. P. *Biochimie* **1998**, *80*, 49–57.
- (34) Vlasak, R.; Luytjes, W.; Spaan, W.; Palese, P. *Proc. Natl. Acad. Sci.* **1988**, *85*, 4526–4529.
- (35) Llop, E.; Gutiérrez-Gallego, R.; Segura, J.; Mallorquí, J.; Pascual, J. A. *Anal. Biochem.* **2008**, *383*, 243–254.
- (36) Reuter, G.; Pfeil, R.; Kamerling, J. P.; Vliegthart, J. F. G.; Schauer, R. *Biochim. Biophys. Acta (BBA), Gen. Subj.* **1980**, *630*, 306–310.
- (37) Corfield, A. P.; Wagner, S. A.; Safe, A.; Mountford, R. A.; Clamp, J. R.; Kamerling, J. P.; Vliegthart, J. F. G.; Schauer, R. *Clin. Sci. (Lond)* **1993**, *84*, 573–579.
- (38) Mahapatra, S.; Yagi, T.; Belisle, J. T.; Espinosa, B. J.; Hill, P. J.; McNeil, M. R.; Brennan, P. J.; Crick, D. C. *J. Bacteriol.* **2005**, *187*, 2747–2757.
- (39) Harvey, D. J.; Mattu, T. S.; Wormald, M. R.; Royle, L.; Dwek, R. A.; Rudd, P. M. *Anal. Chem.* **2002**, *74*, 734–740.
- (40) Franz, A.; Lebrilla, C. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 325–337.