# Polysialylated N-Glycans Identified in Human Serum Through Combined Developments in Sample Preparation, Separations, and Electrospray Ionization-Mass Spectrometry

Scott R. Kronewitter, Ioan Marginean, Jonathan T. Cox, Rui Zhao, Clay D. Hagler, Anil K. Shukla, Timothy S. Carlson, Joshua N. Adkins, David G. Camp II, Ronald J. Moore, Karin D. Rodland, and Richard D. Smith*

Biological Sciences Division, Pacific Northwest National Laboratory, P.O. Box 999, Richland, Washington 99352, United States

Ⓢ *Supporting Information*

ABSTRACT: The N-glycan diversity of human serum glycoproteins, i.e., the human blood serum N-glycome, is both complex and constrained by the range of glycan structures potentially synthesizable by human glycosylation enzymes. The known glycome, however, has been further limited by methods of sample preparation, available analytical platforms, e.g., based upon electrospray ionization-mass spectrometry (ESI-MS), and software tools for data analysis. In this report several improvements have been implemented in sample preparation and analysis to extend ESI-MS glycan characterization and to include polysialylated N-glycans. Sample preparation improvements included acidified, microwave-accelerated, PNGase F N-glycan release to promote lactonization,

and sodium borohydride reduction, that were both optimized to improve quantitative yields and conserve the number of glycoforms detected. Two-stage desalting (during solid phase extraction and on the analytical column) increased sensitivity by reducing analyte signal division between multiple reducing-end-forms or cation adducts. Online separations were improved by using extended length graphitized carbon columns and adding TFA as an acid modifier to a formic acid/reversed phase gradient, providing additional resolving power and significantly improved desorption of both large and heavily sialylated glycans. To improve MS sensitivity and provide gentler ionization conditions at the source-MS interface, subambient pressure ionization with nanoelectrospray (SPIN) was utilized. When these improved methods are combined together with the Glycomics Quintavariate Informed Quantification (GlyQ-IQ) recently described (Kronewitter et al. *Anal. Chem.* **2014**, *86*, 6268−6276), we are able to significantly extend glycan detection sensitivity and provide expanded glycan coverage. We demonstrated the application of these advances in the context of the human serum glycome, and for which our initial observations included the detection of a new class of heavily sialylated N-glycans, including polysialylated N-glycans.
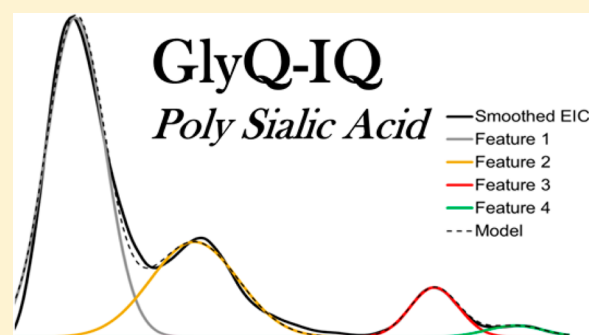
Glycosylation is an important, widespread protein modification in human biology and broadly present in most domains of life. In humans it is estimated that greater than 50% of proteins can be glycosylated,[2] and if O-glycosylation (e.g., O-GlcNAc and O-fucose) modifications were also considered, the number is indeed much higher. Glycans play highly specific, specialized roles at the molecular level, including cell-cell signaling, virus-receptor binding, immune responses, and protein folding and solubility. The glycome of human blood serum has been increasingly of interest for disease biomarker discovery because serum contains the host's response to systemic disease (e.g., immunoglobulin glycoproteins) and can be collected in a minimally invasive manner at relatively low cost. Indeed, the composition and structures of glycans released from human serum/plasma have provided candidate biomarkers for a wide range of cancer types.[3−14] Novel glycoforms have also been observed by profiling glycoproteins (e.g.,

epidermal growth factor receptor) excreted from cancer cell lines.[15]

Polysialic acid, a unique residue found on the antennae of glycans, is characterized by polymeric chains of sialic acid consisting of α2,8-linked or α2,9-linked (or alternating linkages[16]) sialic acid monomers with a degree of polymerization (DP) of 2 or more. Extended long chains of negatively charged glycopolymers exhibit unusual biological properties and have been shown to affect many biological systems,[17] including cell-cell adhesion,[18,19] multipotent stem cells,[20,21] brain plasticity,[22−24] voltage sensitive sodium channels,[25] cell-T lymphocyte interactions,[26] dendritic cell migration,[27] natural killer (NK) cell cytotoxicity,[28] circadian rhythm regulation,[29−31] and cancer tumor cells.[32,33]

The reported size of serum glycan profiles varies greatly (18-113 glycan compositions) as reviewed previously.[34] Recent papers utilizing liquid chromatography–mass spectrometry (LC–MS) have reported up to 73 glycan compositions for permethylated glycans[35] and up to 66 compositions[4] (300 isomers) for native, reduced glycans.[36,37] Earlier studies from our laboratory reported 126 compositions for native, reduced glycans based upon analyses using a subambient pressure nanoESI (SPIN) source[34] and we recently reported extension to 142 glycans and 645 isomer peaks.[1] Although polysialylated N-glycans have been reported on the human neural cell adhesion molecule (N-CAM), to our knowledge, there have been to date no reports describing polysialylated N-glycans in human serum or their direct detection using mass spectrometry (MS).

In this study improvements in sample preparation and informatics were shown to enable the observation of a subclass of "heavily" sialylated complex type glycans defined here as glycan compositions containing more sialic acids than expected antennae. N-glycan antennae are the branches of monosaccharides attached to the trimannose-chitobiose N-glycan core. In the case of complex type glycans, each antenna consists of a minimum of one N-acetylhexosamine and one hexose (typically a galactose). This criterion has previously been used to support the discrimination of polysialic acid because more sialic acids residues than antennae results in at least one antennae with more than one sialic acid attached.[38] Several of the polysialylated glycans were identified in their lactone form, including the formation of a lactone alongside the glycosidic bond connecting two polymerized sialic acid residues.[16] The additional lactone bond stabilizes the polysialic acid chains, increasing MS detectability. Use of the GlyQ-IQ software in combination with high performance computation facilitated both sample processing and instrumentation analysis for performing high-sensitivity MS glycan measurements. N-glycans released from serum proteins by Peptide: N-Glycosidase F (PNGase F) were profiled with graphitized carbon liquid chromatography (LC) coupled to a high-resolution, high-mass accuracy mass spectrometer retrofitted with a SPIN source.[39] The SPIN source increases both signal intensities and charge states and aids in the detection of labile sialylated glycans by providing gentler ionization conditions at the MS interface.[34,40] In addition to detecting standard high mannose, complex, and hybrid type glycans, families of glycans with polysialic acid extensions were also detected leading to expanded glycan profiles and libraries.

■ METHODS

Materials. A pooled reference human blood serum sample (male, blood type AB, not heat inactivated) was used for all glycomics analyses (Sigma-Aldrich, St. Louis, MO). Serial extraction cartridges were used for glycan purification: C8 (Discovery 500 mg, Sigma-Aldrich, St. Louis, MO) and graphitized carbon cartridges (Carbograph 150 mg, Alltech Associates, Inc., Deerfield, IL). The Hypercarb porous graphitized carbon (Hypercarb PGC) particles were used for the HPLC stationary phase (Thermo Fisher Scientific, Waltham, MA).

All other chemicals, sample preparation, and analysis were consistent with that previously reported.[1,34,41] The following reagents were used in sample preparation: PNGase F purified from *Flavobacterium meningosepticum* (New England BioLabs, Ipswich, MA) Nanopure or Milli-Q quality water (18 MΩ cm

or better), sodium phosphate (NaH_2PO_4), sodium borohydride ($NaBH_4$), dithiothreitol (DTT), ethanol (EtOH), trifluoroacetic acid (TFA), acetonitrile (AcN), hydrochloric acid (HCl), formic acid. All reagents were obtained from Sigma-Aldrich (St. Louis, MO), unless otherwise specified.

Sample Preparation. The overall pipeline for glycan profiling shown in Figure 1 spans glycan sample preparation
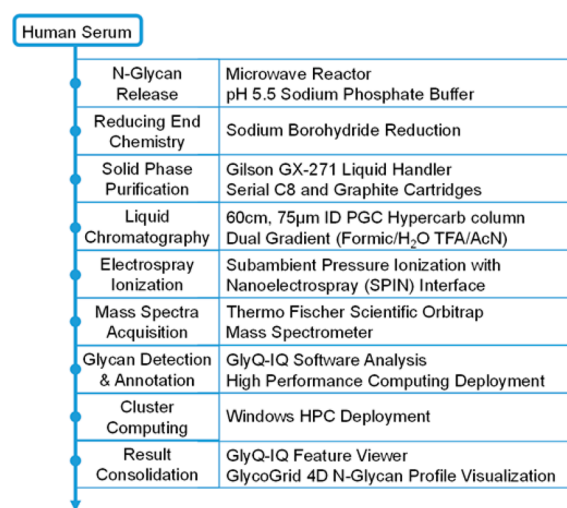

Figure 1. Experimental sample preparation and data analysis pipeline.

through LC–MS data acquisition and subsequent data informatics. Previous sample preparation methods were discussed in detail and reported by Kronewitter et al.[1,34,42] Key developments and features of this glycan release and purification protocol include an acidified (pH 5.5) PNGase F glycan release using a microwave reactor (CEM Discover, Matthews, NC), automated serial dual cartridge (C8 and graphitized carbon), glycan enrichment, and purification on a Gilson GX-274 ASPEC liquid handler (Gilson, Middleton, WI), and long gradient graphite HPLC (60 cm long, 75 μm i.d.).

Aliquots of serum (100 μL) were mixed with 50 μL of buffer (100 mM sodium phosphate with 20 mM DTT) and acidified to pH 5.5 using 1 M HCl. The samples were denatured by heating them in an Eppendorf 1.5 μL Thermomixer set at 95 °C for 2 min and let to cool before PNGase F was added. The samples were deglycosylated with 2 μL of PNGase F in a microwave reactor set to solid-phase synthesis (SPS) mode to allow for constant power experiments.[42] The catalyzed microwave reactor digestion and lactonization took place at constant power (20 W) for 40 min (4 × 10 min intervals), and the 50 mL cooling water in the reactor was exchanged with chilled water (4 °C) after each interval to keep the maximum temperature observed during the reaction less than 50 °C. EtOH (80% v/v) was used to precipitate deglycosylated proteins via a 60 min exposure in a −80 °C freezer followed by a 30 min centrifugation. The supernatant containing the glycans was isolated and dried before reduction. The samples were reduced with 1 M $NaBH_4$ at 60 °C for 60 min. The dual cartridge (C8, Graphite) solid phase extraction (SPE) and desalting were performed with customized algorithms optimized for decreased sample carryover and precise liquid transfer across SPE cartridges. The samples were first passed through the C8, and the flow-through was trapped on the graphite for glycan enrichment and desalting. Glycans were eluted from the graphite stationary phase with three aqueous
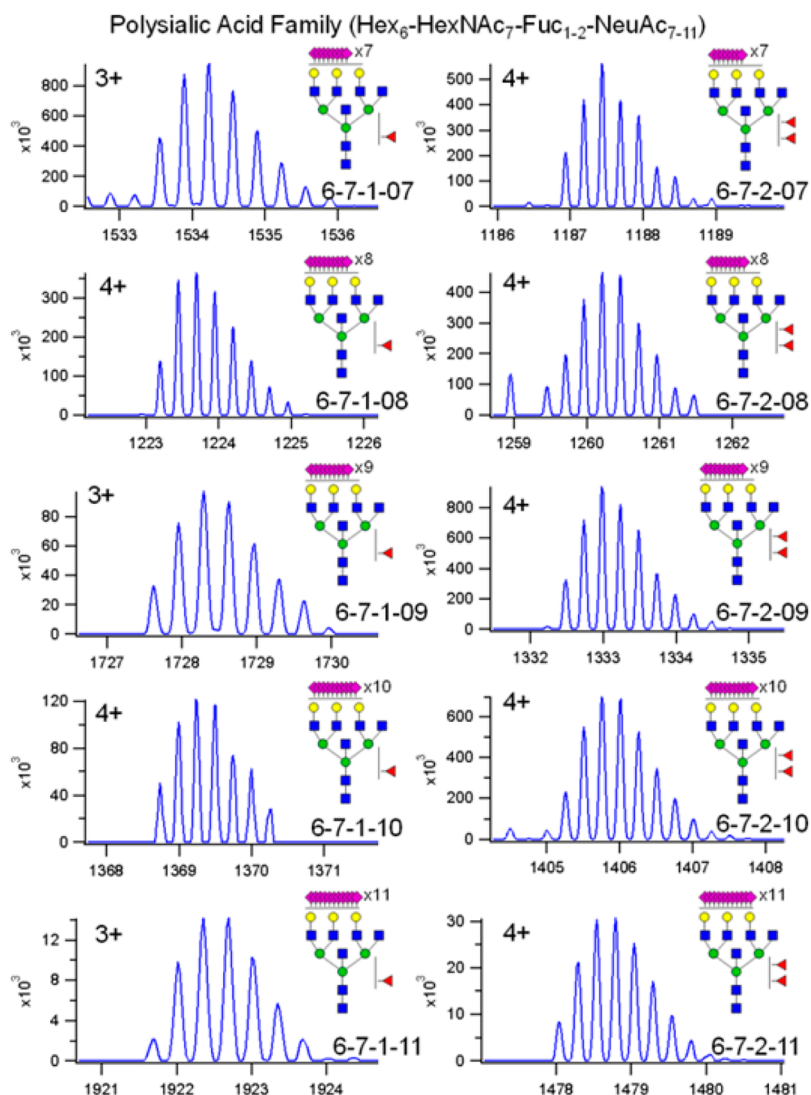
Figure 2. Example mass spectra of polysialylated glycans containing 7−11 sialic acids and 1−2 fucose residues. Putative glycan structures are shown and the compositions are consistent with number of sialic acid groups exceeding the possible number of N-glycan antennae. The glycan structure depictions are constant with the CFG nomenclature where the squares = N-acetylhexosamine, circles = hexose, triangles = fucose, and diamonds = sialic acid.

solutions comprised of different AcN and TFA concentrations. The fractionation included the following additions: 2 mL of 10% AcN, 2.25 mL of 20% AcN, and 1.75 mL of 40% AcN with 0.1% TFA, 0.5 mL of nanopure H_2O, and 2 mL of 40% AcN with 0.1% TFA. The eluted fractions were combined and dried.

HPLC and Mass Spectrometry. The in-house built LC system used two Agilent 1200 nanoflow pumps (Agilent Technologies, Santa Clara, CA), Valco valves (Valco Instruments Co., Houston, TX), and a PAL autosampler (Leap Technologies, Carrboro, NC) and was automated using custom software (LCMSNet) that allowed for parallel event coordination and therefore, approximately, 100% MS duty cycle was achieved through the use of two analytical columns. Two-column operation permitted columns to be "washed" and reconditioned off-line without any loss of duty cycle. Graphite capillary columns were used with the following specifications: 60 cm long × 360 μm o.d. × 75 μm i.d. (Polymicro Technologies Inc., Phoenix, AZ) with a 1 cm sol−gel frit for media retention.[43] The fused silica columns were slurry packed in-house with 3 μm diameter Hypercarb PGC particles. Each

column was electrically isolated from the switching valves with a 15 cm, 30 μm i.d. empty fused silica capillary tube.[44] Adding empty isolator columns did not adversely affect the chromatography or result in significant peak broadening.

A dual acid modifier reversed phase gradient was used for liquid chromatographic separation. The mobile phase consisted of mobile phase A, 0.1% formic acid in H_2O, and mobile phase B, 0.2% trifluoroacetic acid in AcN. The samples were loaded at 1% B and the 90 min analytical gradient used was (time in minutes:% mobile phase B): 0:1%, 1:4%, 2:6%, 91:30%, 95:95%, 96:95%, 100:1.%. In between each run a wash and condition gradient was applied: 0:1%, 1:5%, 6:40%, 10:80%, 12:90%, 32:100%, 41:100%, 45:1%, 69:1%.

An Exactive Orbitrap MS (Thermo Scientific, San Jose, CA) was used for glycan profiling and a LTQ-Orbitrap Velos Pro MS (Thermo Scientific) was used for tandem MS measurements. Both instruments were operated in ultrahigh resolution mode (100 K resolution at m/z 400) with an m/z range of 200−3000 and three microscans were averaged per spectrum. The AGC was set for high dynamic range (3×10^6) with a
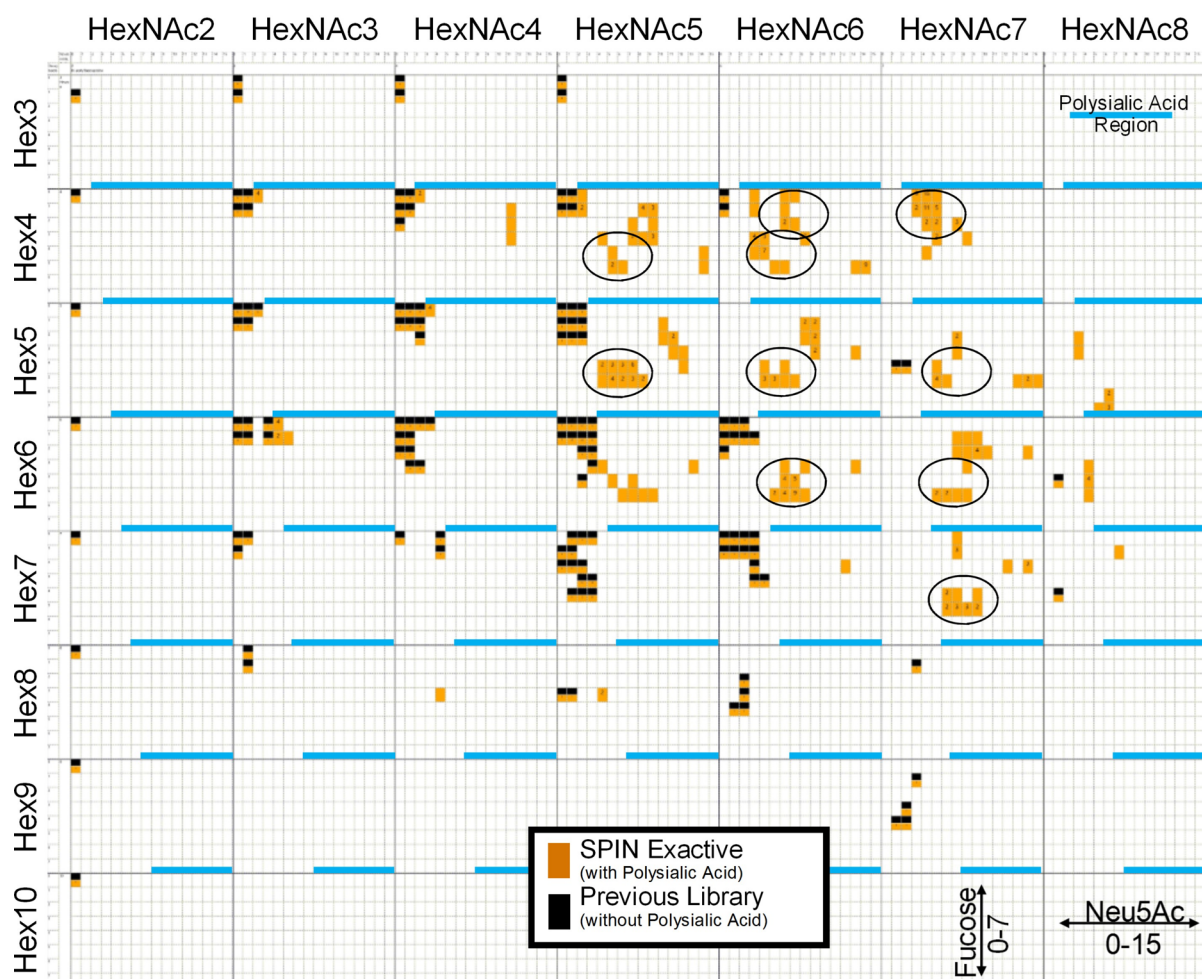
Figure 3. GlycoGrid Comparison between the current glycan library that allowed for polysialic acid and lactonization (orange) and the former library that did not (black). The regions circled indicated several polysialylated glycans that differed by either a sialic acid or a fucose residue. Blue horizontal lines indicate regions where polysialylated glycans were plotted.

maximum ion injection time of 100 ms. The front end ion optics on both instruments were modified to include a SPIN-MS interface.[39] In the SPIN-MS interface the ESI emitter is brought into the first vacuum region of the MS directly adjacent to an electrodynamic ion funnel that has been described in detail previously.[39,45] For clarity, the systems will be referred to as the SPIN Exactive and SPIN Velos systems and the data used to complement each other rather than for juxtaposition.

Lactonized Polysialic Acid Chains. Evidence of lactone bond formation during the acidified glycan release helped detect polysialylation by stabilizing some or all of the sialic acid glycosidic bonds so the initial glycan remain intact throughout the experiment. Lactonization occurs through a condensation reaction between the carboxyl group and hydroxyl group on adjacent sialic acid monomers yielding the loss of water (−18.010 Da) per lactone bond formed and decreased monoisotopic mass observed in the mass spectra.[16] The combination of the lactone bond stabilization and gentler ionization conditions present in the SPIN source[40] aided the MS analysis. The lactone group can also be used to differentiate polysialylation from monosialylated antennae because the presence of lactones indicated a chain length of at least two.

Data Analysis. High-resolution LC−MS data sets were analyzed with the GlyQ-IQ annotation software with high performance computing as described elsewhere.[1] Briefly, the

GlyQ-IQ software uses five measurement metrics (exact mass, isotope fit scoring, LC peak modeling and fitting, glycan family relationships, and in-source fragmentation) to identify and annotate glycans in LC−MS data sets. The GlyQ-IQ Viewer and GlycoGrid 4D visualization were used for manual annotation and glycan profile visualization.[1,34] Manual inspection of the isotopic envelopes was used to ensure that each feature reported was also supported by high-resolution spectral evidence (i.e., a confident and accurate MS molecular weight assignments and isotopic envelope profiles). High-resolution tandem mass spectrometry (MS/MS) data collected on the SPIN Velos system were manually annotated.

■ RESULTS AND DISCUSSION

Results. *Applying the Suite of Refined Methods.* We were able to evaluate and improve the experimental coverage of the serum glycome by a combination of (1) acidified PNGase F digestion, (2) increasing sample signal by sample processing improvements and gentle ionization conditions provided by a SPIN ionization source, and (3) using high-sensitivity, high-specificity software.[1] The present results support the observations of polysialylated N-glycans from serum glycoproteins with strong corroboration between the detection of several glycan families grouped together by differing amounts of sialic acid, MS/MS fragmentation, and in-source fragmentation.
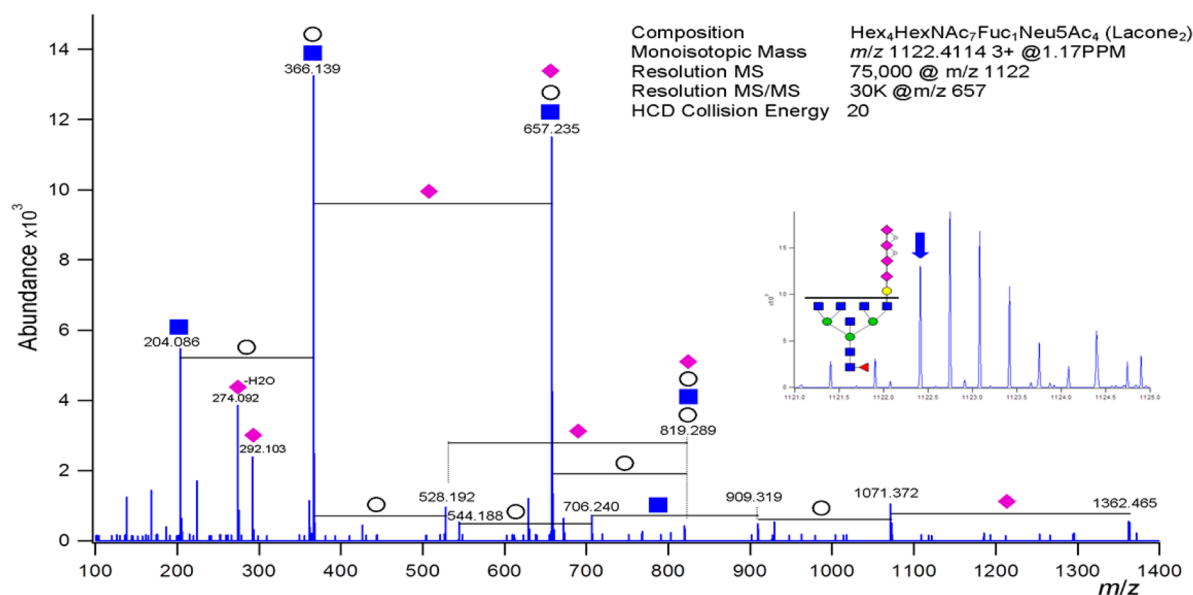
Figure 4. High-resolution HCD spectrum of fragments from N-glycan containing polysialic acid. Inset: high-resolution MS1 spectrum indicating monoisotopic mass and a putative cartoon of the glycan structure. CFG nomenclature was used to denote mass differences with the following modification: Open circle symbols represent generic hexose mass differences.
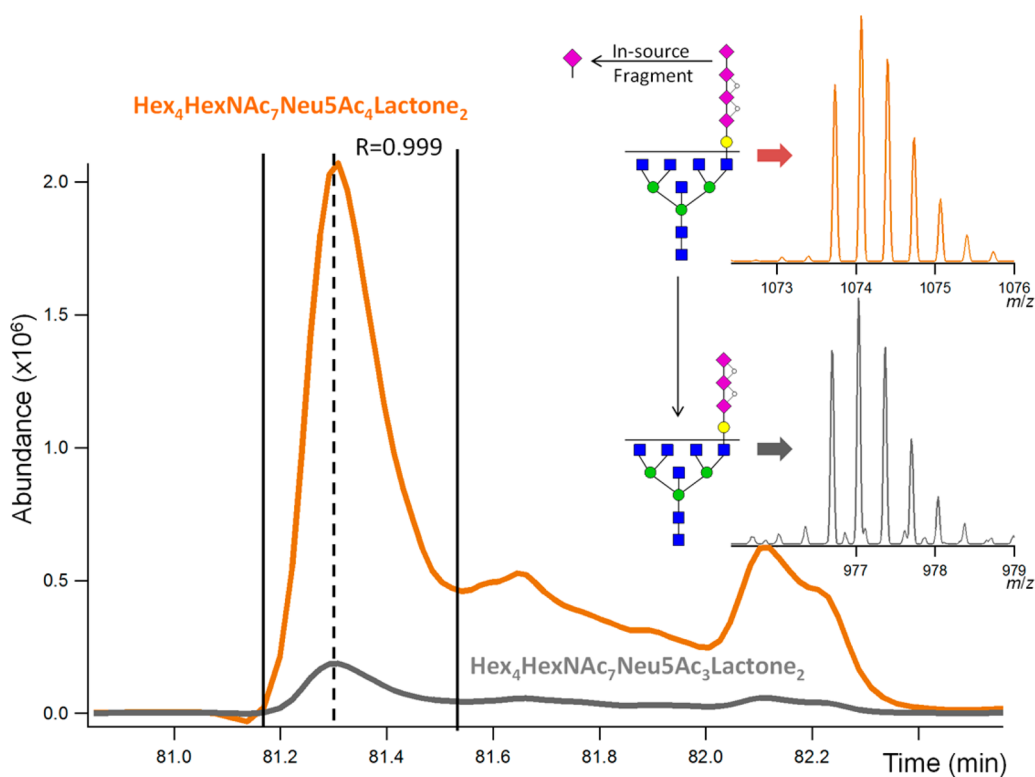

Figure 5. In-source fragmentation detected linking $Hex_4HexNAc_7Neu5Ac_3$ (Lactone$_2$) and $Hex_4HexNAc_7Neu5Ac_4$ (Lactone$_2$) by one Neu5Ac residue. The first chromatographic peak shown on the left between the vertical lines had a correlation coefficient of 0.999 after the peaks were modeled and correlated. The inlayed mass spectrum corresponded to the high-resolution isotope profiles of the parent (orange) and fragment (gray) glycans. Cartoon structures (CFG nomenclature) of the parent and fragment illustrate the in-source fragmentation loss of a sialic acid.

Polysialylated Glycan Families. Interglycan evidence of polysialylated N-glycans included the high-resolution, high-mass accuracy spectra and clusters of glycans with large degrees of sialic acid (more sialic acid than branches) that differed in mass by one sialic acid monomer. An exemplary glycan family supporting polysialic acid identification contained 7−11 sialic acids, 1−2 fucose monosaccharides, and no lactonization.

Respective mass spectra for the sialylated glycan family are presented in Figure 2 along with putative structures using the CFG nomenclature.[46] In addition to the sialic acid and fucose differences presented in Figure 2, there are also related comparable clusters found with one more hexose or N-acetylhexosamine as shown circled in Figure 3. Almost all of the polysialylated glycans detected here contained at least one

fucose and were consistent with the well-studied N-glycans detected on the neural cell adhesion molecule (N-CAM) in the brain suggesting similar glycosyltransferase and glycosidase machinery was used.[17,38,47−49] Several more hybrid and bisecting HexNAc structures were observed that were also supported by prior reported structural analysis.[38]

High-Resolution, Higher-Energy Collisional Dissociation (HCD) Glycan Fragmentation. To explore the use of HCD spectra for improved characterization of glycan fragmentation patterns and elucidate structural information, Figure 4 displays the spectrum acquired for a polysialylated glycan with composition $Hex_4HexNAc_7Fuc_1Neu5Ac_4$ (Lacone$_2$). This structure was chosen as an example because of the limited number of hexose monosaccharides present, the large number of sialic acid residues, the multiple lactone groups, and its HCD fragmentation pattern. Figure 4 shows an inlayed "zoomed-in" mass spectrum and a putative glycan structure.

When subtracting out the N-glycan pentasaccharide core, only one hexose (presumably galactose) remains and restricts the antennae extension count to one. Glycans with only one antenna extended beyond the N-acetylhexosamine residues yielded one attachment site for enzymatic addition of sialic acid and subsequently, a single, polysialylated chain. Thus, a chain of 4 sialic acids (with 2 lactone bonds) is favored over several individual sialic acid monosaccharides. Although single antennae can be predicted, additional information is needed to determine to which branch of the core the antenna is attached.

Diagnostic ions at *m/z* 292.103 and 274.092 (water loss) indicated sialic acid was present in the compound, but alone was not diagnostic for polysialic acid, since both ions were also observed in nonpolysialylated glycan fragmentation spectra. In addition the presence of sialic acid as part of ions at *m/z* 657.235, 819.289, and 1362.465 indicated that a sialic acid monosaccharide was attached to the terminal hexose, and it was postulated that it does not directly contain one of the two stabilizing lactone bonds which could be present between the terminal three sialic acids.

The reduced alditol terminus of the core N-acetylhexosamine (+2 hydrogen atoms) was not detected nor was any fucose differences detected on the antennae fragments. This combined absence suggested the fucose was bonded to the core and this group was lost in the isolation and fragmentation process.

In-Source Fragmentation Sialic Acid Verification. The nonfucosylated form of the glycan in Figure 5, $Hex_4HexNAc_7Neu5Ac_4$ (Lactone$_2$), shows an example of in-source fragmentation detection indicating the presence of Neu5Ac in the composition. The GlyQ-IQ software identified the matched pair by searching for features one sialic acid greater than or less than the target mass (±291.095 Da, adjusted for charge states).

The extracted ion chromatograms were extracted, bounded by the overlapping region, each peak was modeled, and the model peaks were correlated together.[1] The correlation information not only confirmed that the mass was glycan related, but it also confirmed one of the four common monosaccharides searched was present in the compound without using MS/MS. In addition to the main peak correlation at 81.3 min, other isomeric structures were also conceivable at elution times 81.65, 82.1, and 82.2 min.

Human Serum Glycan Profile. Overall, 290 glycan compositions were experimentally annotated and of those, 68 confirmed by all 5 variables of the GlyQ-IQ paradigm.[1] Each

reported composition has been manually verified using the GlyQ-IQ Viewer[1] by reviewing the glycan isotopic profile to ensure the correct monoisotopic mass was chosen. GlyQ-IQ was able to identify and remove 81 detected species (8% of the total) that corresponded to in-source fragmentation products. The 290 experientially detected glycan compositions and 994 chromatography separated isomers are presented in Supplementary Table 2 in the Supporting Information. The applied data processing methods resulted in the detection of an additional 134 identifications and 22 in-source fragmentation confirmations beyond the former retrosynthetic glycan library results where polysialic acid was not included.[1]

Figure 3 includes glycan profiles resulting from applying the latest polysialylated glycan library developed in this work (orange) and our previous results (black) with no polysialic acid considerations. The number of hexose residues in a composition can be used to determine the maximum possible number of antennae that can be sialylated because the terminal galactose is a required substrate for sialic acid glycosyltransferase activity. When the number of sialic acid residues is greater than the number of antennae in complex glycans (total hexose − (3) core hexose), polysialylation is likely to occur on one more antennae and the display regions are indicated with a blue horizontal bar. While 99% of nonheavily sialylated annotated features passed the GlyQ-IQ feature validation process, only 93% of the heavily sialylated glycans passed due to the 17-fold larger library size and larger monoisotopic masses.

Observed Lactonization. The percent lactonization can be calculated by dividing the number of lactone groups by the total number of sialic acids residues in the composition. Interestingly, we observed significant "heavily" sialylated glycans containing a low level of lactones (1−2 lactones), as indicated in Figure 6 showing 20% lactonized at this level. This was also
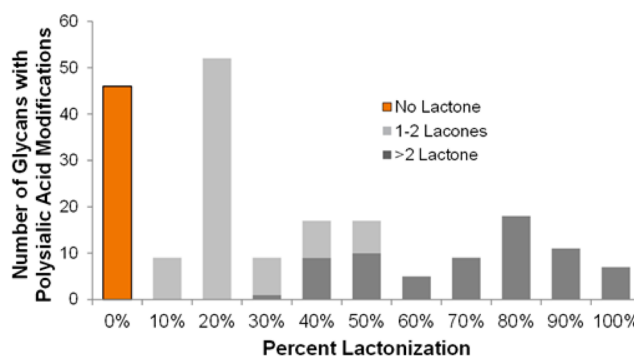

Figure 6. Percent lactonization (number of lactone bonds divided by the number of sialic acid residues) distribution of all "heavily" sialylated glycans. Orange denotes glycans with no lactonization detected. Light gray indicates glycans with 1−2 lactones. Dark gray indicates glycans with more than 2 lactones.

consistent with polysialylated glycans that were in the initial stage of lactonization.[50] In contrast, the glycans with a large number of lactones (more than 2) had lactonization peaking at ∼80% (Figure 6). Despite our use of gentle ionization, in-source fragmentation detection, lactonization, and sensitive software, it remains likely that even longer chains of polysialic acid existed in the sample studied.

Combining the Evidence. Multiple pieces of evidence were used to support the presence of polysialylated N-glycans in human serum. Although only 46% of the mass was covered by MS/MS analysis, it greatly reduced the number of combina-

torial possibilities that can explain the remainder of the parent mass. The unknown mass difference between the intact parent molecule and the minimum N-glycan composition stipulated by the fragmentation spectrum (N-glycan core + sialylated lactosamine branch, $Hex_4HexNAc_3NeuAc$ at mass 1568.571 Da) is 1795.641 Da.

The polysialylation hypothesis used to explain the mass difference included hits from a combinatorial expansion comprised of hexose (ranges 4−12), HexNAc (ranges 3−8), fucose (ranges 1−7), and Neu5Ac (ranges 1−15). The allowable number of sialic acid lactonization ranged from 0 to 14 ($N − 1$, where N is the number of sialic acid residues polymerized). The polysialylated glycan $Hex_4HexNAc_7FucNeuAc_4$ with 2 lactones was the only glycan hit using an excessively large 30 ppm window. The relaxed tolerance was used to indicate no other close hits were possible using this parameter set.

Figure 7 includes a summary of information used from the glycan family relationships, fragmentation ions, and accurate
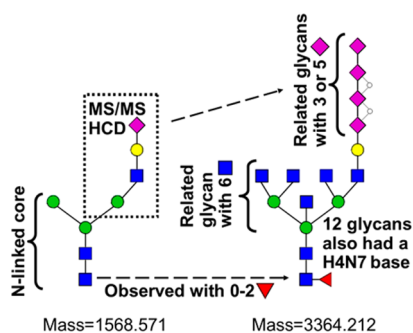

Figure 7. Extending MS/MS fragmentation data with glycan families and accurate mass.

mass to support the heavily sialylated glycan compositions assignment and how the limited number of hexose residues and lactonization were used to predict the polysialylated sialic acid extension on the antennae. The polysialylated glycan $Hex_4HexNAc_7FucNeuAc_4$ (2 Lactones) was detected in a glycan family with 7 primary family members with the same $Hex_4HexNAc_7$ base and differ by a single monosaccharide (and an additional 5 secondary members with disaccharide relationships) with the following fucose and Neu5Ac combinations: $Fuc_0(Neu5Ac_3, Neu5Ac_4, Neu5Ac_5)$, $Fuc_1(Neu5Ac_3, Neu5Ac_4, Neu5Ac_5)$, and $Fuc_2(Neu5Ac_4, Neu5Ac_5)$. Diagnostic ions detected in the fragmentation data and monosaccharide differences between fragment ions were used to confirm at least one sialylated LacNAc branch. Combining this information with accurate match only allowed for one composition to explain the parent mass. No other combinations of these monosaccharides could explain the parent mass allowing for a generous 30 ppm mass tolerance.

Discussion. *Refined Sample Preparation.* In this work, we have refined a reduced, nonderivatized N-glycan sample preparation pipeline to improve MS detection by coalescing glycan peaks into one reduced, desalinated glycoform with intrinsic signal intensity since the glycan signal was not divided between reducing-end-forms or cation adducts. The high performance cation removal is further explained in the section S1 in the Supporting Information and a desalinated spectrum is presented in Figure S1 in the Supporting Information. Multiple charge states were detected per glycan as a product of the

electrospray ionization, and the feature abundances were added together via the GlyQ-IQ software.[1] The nonderivatized glycomics approach implemented here can be performed in relatively mild conditions, with reduced sample handling, and subsequently will produce glycan profiles with minimal perturbations. Positive mode, proton adducted N-glycans analyzed by LC−ESI-MS has been shown to be a desirable form for quantification since the ion abundances correlate well with fluorescently labeled HPLC results.[51] The high sensitivity of this current method is achieved by combining quantitative chemical reactions, searching for lactone stabilized polysialic acid chains, improved chromatographic stationary phase desorption, chromatographic isomer peak separation, and gentle electrospray ionization conditions. The details involved in refining several aspects of the workflow are discussed below.

One Step, Multiple Effects. Lowering the pH of the PNGase F digestion to 5.5 from the optimal (pH 7.5−8.0)[52] had two beneficial effects, while maintaining the fast, efficient release with the microwave reactor. First, reduced glycans were obtained by first quantitatively converting the reducing end of the glycans to the water-stable aldehyde form during the acidified PNGase F glycan release step. The homogeneous form allows for following a then predictable conversion to desired reduced alditol product during the reducing-end aldehyde reduction step. Second, the lower pH in conjunction with the microwave-assisted reactions[53] (at relatively elevated temperatures ~50 °C)[54] favors lactone formation and leads to sialic acid stabilization. Although it is hard to distinguish native lactone formation from sample handling induced lactonization, the acidified release conditions may have helped stabilize the polysialylated glycan modifications by forming lactone bonds between sialic acid residues through acid catalyzed condensation reactions.

Effect 1: Acid Catalyzed Glycan Deamination Mechanism. The acidified digestion conditions developed for this research simplified the reducing-end mixture so glycans were in a single form prior to reduction. This approach overcame any reducing-end heterogeneity that divides the glycan signals and complicates the data analysis. As the glycans were initially released, their N-glycan core was terminated with a β-glycosylamine conjugated with ammonia.[55] Particular attention was spent on the accelerated PNGase F release[56] and the respective mixture of reducing-end glycoforms produced when the glycan-asparagine bond was cleaved. The amine readily hydrolyzed in aqueous conditions over the course of minutes to hours[57] to form an aldehyde. This resulted in producing a mixture of the amine and aldehyde forms in solution when accelerated reactions were used, such as using immobilized enzyme[57] or microwave reactors.[41,56]

Multiple glycoforms behaved differently in subsequent reaction and purification steps, and such mixtures should be avoided[57] to ensure the glycans are fully in the alditol form and display good chromatographic separations. The acidic conditions were sufficient to convert all released glycans to their aldehyde form prior to reduction, as determined indirectly by monitoring the completeness of the reducing-end reduction reaction as the aldehyde form needs to be present for conversion to the alditol form in a reducing environment. Any residual amino forms present contributed to the unconverted aldehyde quantity, since they will be gradually hydrolyzed into the aldehyde form via water hydrolysis during the subsequent sample preparation. The quantitative con-

version is depicted in Figure 8, which shows the alditol (A) and trace aldehyde (B) forms detected.
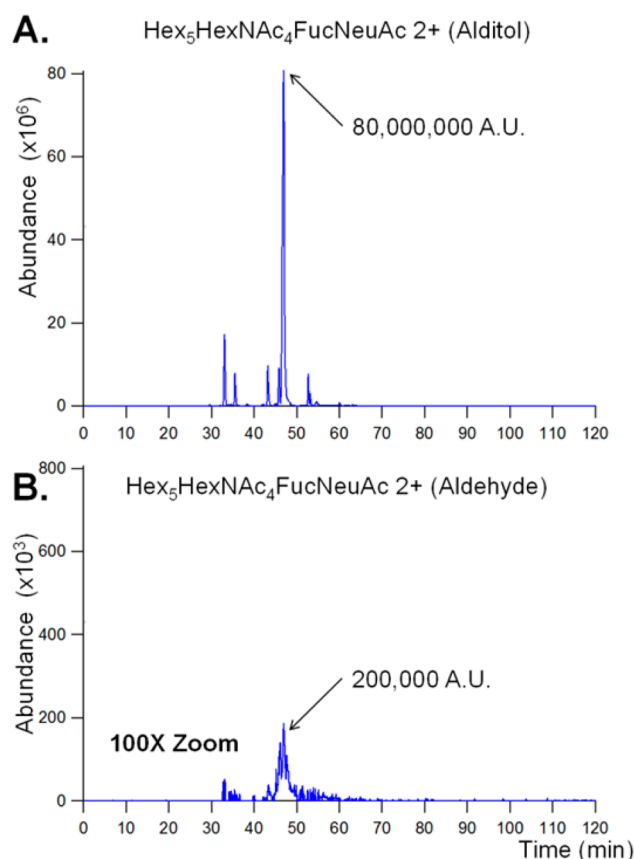

Figure 8. Extracted ion chromatograms (m/z 1040.89) showed the quantitative reduction to alditol form achieved by quantitative hydrolysis of β-glycosylamine groups first. 99.7% conversion to the alditol form was achieved.

Once all of the glycans were in one reducing end form prior to downstream reduction, the reduction was performed quantitatively. This provided an optimized signal since there was no glycoform partitioning. In addition reducing the pH of the reaction buffer prior to digestion decreased the overall processing time by eliminating the need for sequential acid treatments.

Effect 2: Microwave-Assisted Sialic Acid Lactonization. The acidic digestion served a secondary purpose of starting the acid lactonization stabilization of polysialic acid residues. The presence of a lactone was detected by the mass loss corresponding to one water molecule from the intact mass of a parent glycan containing two or more sialic acid residues. There are three stages of lactonization for polysialylated glycans which include initialization or precursory (1−2 lactone modifications), middle stage (multiple lactone species up to DP-2), and final stage (complete lactonization).[50]

Surprisingly, lactonized sialic acid has previously been reported to be stable at 4 °C in 1.0 M HCl solutions overnight, which is helpful in preserving the sialylation structure through sample preparation.[50] Several heavily sialylated N-glycans were detected with 1−2 lactone modifications that are consistent with the lactonization initiation process. Figure 9 depicts a distribution of detected lactone modifications in the heavily sialylated glycans that indicated a large percentage of
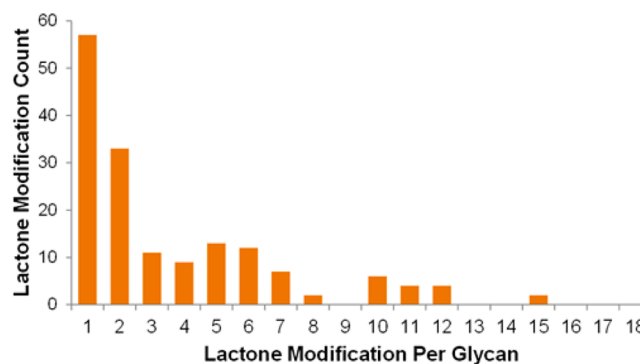

Figure 9. Number of lactones detected per "heavily" sialylated glycan.

polysialylated species present in the precursory lactonization stage (1−2 lactones detected).

However, lactonization is susceptible to environmental conditions (pH changes), and it is difficult to differentiate lactonization artifacts introduced during sample preparation from physiological lactonization in the serum.

Long Column Chromatography. HPLC performance on slurry packed, graphitized carbon nanoLC columns was improved by increasing the column length to improve chromatographic peak resolution and by adding TFA to the mobile phase gradient. Adding the TFA to the mobile phase during LC−MS data acquisition was not only more consistent with the elution of fractions used in the sample preparation solid phase extraction step but also produced narrower, taller peaks[58] and was amenable to desorbing acidic sialylated glycans from the graphite stationary phase.[59] Additional enhanced performance observed from the chromatography setup is expanded in section S2 in the Supporting Information.

Gentle Ionization Conditions and Polysialylated Glycans. The SPIN source facilitated sialic acid detection by providing enhanced signal intensities, higher charge states, and apparently, gentler ionization and ion transfer.[34] The gentle electrospray conditions provided by the SPIN ionization source helped to reduce the internal energy of the molecules and minimize in-source fragmentation. The additional glycan coverage of labile heavily sialylated glycan species provided by this method have been presented in Figure 10 where the nonsialylated (column 0) and common sialylated glycans detected were colored in gray and the heavily sialylated glycans with more sialic acids than predicted branches were in orange.
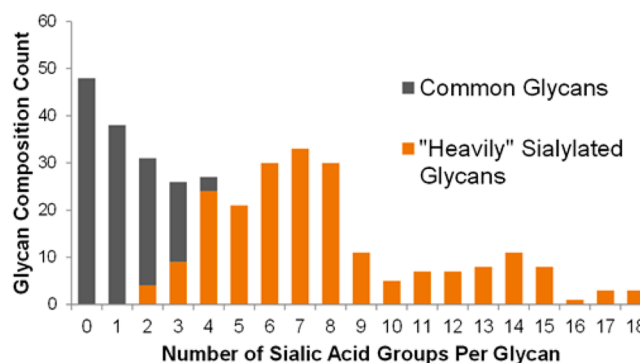

Figure 10. Number of sialic acid residues per glycan. Orange denotes "heavily" sialylated glycans. Gray includes typical monosialylation on the terminus of the galactose residue. Nonsialylated glycans, such as high mannose glycans, are included in the 0 column.

The gentle ionization conditions were particularly important for polysialylated glycans because they were low in abundance, relatively high mass, and fragment in-source under energetic ionization conditions.

One possible explanation for the augmented polysialic acid detection was that the SPIN source has a different (and faster/shorter) desolvation profile than, e.g., a conventional heated capillary ESI interface where charged droplets are desolvated more rapidly in a lower pressure region rather than, e.g., a heated capillary at near atmospheric pressure. Shortening the capillary has previously been shown to enhance detection of higher mobility ions (e.g., highly charged glycans).[60] This effect is further enhanced in the SPIN interface because the sample solution is electrosprayed directly into an electrodynamic ion funnel in a region at ~20 Torr, negating both the conventional period at atmospheric pressure as well as the period in transit through a higher pressure heated region as desolvation and ionization occur. The carboxylic acid groups present in sialic acid monosaccharides have relatively high hydrogen bonding stabilization energy with the carbonyl and hydroxyl groups, which can form hydrogen bonding interactions with the H and O atoms of the water molecules.[61] This stabilization effect causes the carboxylic acid groups to desolvate a water molecule later than other sites under desolvating conditions. Delaying this process was important for sialic acid ion stability because if desolvated too early, where the desolvation environments are often harshest, the fully desolvated molecule is prone to in-source fragmentation via proton transfer from the carboxylic acid to the glycosidic bond. We postulated that by delaying the desolvation to cooler regions of the source, such as inside the electrodynamic ion funnel, and by reaching this stage more rapidly, the stability of the sialic acid residues is increased as the final desolvation occurs in a gentler environment.

Expanded Theoretical Retrosynthetic Glycan Library. Utilizing the increased measurement sensitivity, we expanded a former retrosynthetic glycan library from its conception in 2009[62] and minor expansion in 2012,[34] to improve the coverage of the human serum glycome. The initial library was bounded in glycan size by tetra-antennary, penta-fucosylated, tetra-sialylated glycans with a bisecting N-acetylhexosamine (HexNAc). These bounds were set because much larger glycans were rarely detected, e.g., using label-free matrix-assisted laser desorption ionization (MALDI) conditions. However, applying gentle LC−ESI-MS platforms, larger, multiply charged, and more labile glycans can be detected, and we recently demonstrated that the SPIN source interface was able to detect additional glycans generally not observed using conventional ESI-MS platforms.[1,34]

The expanded retrosynthetic glycan library now includes fucosylated high mannose structures (Man3-12) and an overall degree of fucosylation up to 7 (increased from 5).[34] N-Acetyllactosamine (LacNAc) groups were expanded to include branching up to penta-antennary and containing poly-N-acetyllactosamine (polyLacNAc) groups, so that 7 LacNAc in total are considered (increased from 4).[34] This allowed the glycan sizes extended to those with 10 hexose and 8 HexNAc monosaccharides (increased from 7 hexose and 7 HexNAc).[34] Although glycans with 8 LacNAc groups (tetra antennary N-glycan with 4 polyLacNAc) have been reported at low abundance on recombinant BRP 3 erythropoietin applying a 2-D separation (anion exchange, HILIC),[63] the largest number of LacNAc groups that we were able to detect in this study in our serum sample was 5.

In the present work the allowable amount of sialic acid residues was expanded to include polysialylation and set to a limit of 20 degrees of polymerization (increased from 5 nonpolymerized sialic acid monosaccharides).[34] Although the actual maximum chain length is currently unknown, previous HPLC-based studies reported degrees of sialic acid polymerization of 15−35 for protein glycosylation and up to 50 on N-CAM.[35] However, the largest sialylated species detected here with an acceptable GlyQ-IQ profile contained 18 sialic acids while the majority of the heavily sialylated group contained 4−9 sialic acid residues.

Lactonization was addressed by including 0 to 100% lactonization of sialylated glycans in the library. The latest theoretical retrosynthetic glycan library has been included as Supplementary Table 1 in the Supporting Information and contains 3257 theoretical glycan compositions and 19 383 glycan masses (with lactonization modifications allowed). Formerly, only 436 glycans were considered in the library.[34] The scalability and high specificity of the GlyQ-IQ software and HPC implementation was amenable to the relatively large, sparse, theoretical retrosynthetic glycan library presented here.

■ CONCLUSIONS

We have developed an approach for improved glycoform characterization based upon several developments that included increases in detectable analyte (reduced signal division from quantitative reactions and methodical desalting), more sensitive ionization (high ion transmission and softer ionization using the SPIN source), enhanced liquid chromatography (60 cm graphite fused silica columns with TFA modified formic acid AcN/water gradient), and high sensitivity and specificity software (GlyQ-IQ characterization leveraging theoretical retrosynthetic glycan libraries and high performance computing).[1] Using this composite method for human serum glycan profiling, 290 glycan compositions were elucidated and 68 were confirmed by in-source fragmentation without the use of serum depletion or multidimensional chromatography. In addition we identified 994 isomers peaks (based on LC separation) and confirmed 143 as glycans with in-source fragmentation verification ions. The GlyQ-IQ software also enabled confirmation of several glycan compositions in addition to those with enough intensity for MS/MS fragmentation. These advances enabled the first report and direct measurement of families of polysialylated glycans in human serum using mass spectrometry. Measuring polysialylated glycan compositions provided a foundation for follow-up glycan structure and glycoprotein site localization studies. Serum polysialylated glycans or the attached glycoproteins may serve as a source for disease related biomarkers circulating in the blood, as the disruption of metabolism observed in diabetes, cancer,[18,19,32,33] and other disease may be reflected in altered glycosylation of circulating proteins.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Description of two stage desalting and acid gradient chromatography and tables of the glycan library and human serum glycans. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: rds@pnnl.gov.

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Kronewitter, S. R.; Slysz, G. W.; Marginean, I.; Hagler, C. D.; LaMarche, B. L.; Zhao, R.; Harris, M. Y.; Monroe, M. E.; Polyukh, C. A.; Crowell, K. L.; Fillmore, T. L.; Carlson, T. S.; Camp, D. G., 2nd; Moore, R. J.; Payne, S. H.; Anderson, G. A.; Smith, R. D. *Anal. Chem.* **2014**, *86*, 6268−6276.

(2) Apweiler, R.; Hermjakob, H.; Sharon, N. *Biochim. Biophys. Acta: Gen. Subj.* **1999**, *1473*, 4−8.

(3) Ruhaak, L. R.; Miyamoto, S.; Lebrilla, C. B. *Mol. Cell. Proteomics* **2013**, *12*, 846−855.

(4) Kronewitter, S. R.; De Leoz, M. L.; Strum, J. S.; An, H. J.; Dimapasoc, L. M.; Guerrero, A.; Miyamoto, S.; Lebrilla, C. B.; Leiserowitz, G. S. *Proteomics* **2012**, *12*, 2523−2538.

(5) 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**, *1279*, 58−67.

(6) Hua, S.; An, H. J.; Ozcan, S.; Ro, G. S.; Soares, S.; DeVere-White, R.; Lebrilla, C. B. *Analyst* **2011**, *136*, 3663−3671.

(7) Hu, Y.; Desantos-Garcia, J. L.; Mechref, Y. *Rapid Commun. Mass Spectrom.* **2013**, *27*, 865−877.

(8) Gaye, M. M.; Valentine, S. J.; Hu, Y.; Mirjankar, N.; Hammoud, Z. T.; Mechref, Y.; Lavine, B. K.; Clemmer, D. E. *J. Proteome Res.* **2012**, *11*, 6102−6110.

(9) Mechref, Y.; Hu, Y.; Garcia, A.; Hussein, A. *Electrophoresis* **2012**, *33*, 1755−1767.

(10) Mitra, I.; Zhuang, Z.; Zhang, Y.; Yu, C. Y.; Hammoud, Z. T.; Tang, H.; Mechref, Y.; Jacobson, S. C. *Anal. Chem.* **2012**, *84*, 3621−3627.

(11) Alley, W. R., Jr.; Vasseur, J. A.; Goetz, J. A.; Svoboda, M.; Mann, B. F.; Matei, D. E.; Menning, N.; Hussein, A.; Mechref, Y.; Novotny, M. V. *J. Proteome Res.* **2012**, *11*, 2282−2300.

(12) Kang, P.; Madera, M.; Alley, W. R., Jr.; Goldman, R.; Mechref, Y.; Novotny, M. V. *Int. J. Mass Spectrom.* **2011**, *305*, 185−198.

(13) Alley, W. R., Jr.; Madera, M.; Mechref, Y.; Novotny, M. V. *Anal. Chem.* **2010**, *82*, 5095−5106.

(14) Tousi, F.; Bones, J.; Iliopoulos, O.; Hancock, W. S.; Hincapie, M. *J. Chromatogr., A* **2012**, *1256*, 121−128.

(15) Wu, S. L.; Taylor, A. D.; Lu, Q.; Hanash, S. M.; Im, H.; Snyder, M.; Hancock, W. S. *Mol. Cell. Proteomics* **2013**, *12*, 1239−1249.

(16) Cheng, M. C.; Lin, C. H.; Lin, H. J.; Yu, Y. P.; Wu, S. H. *Glycobiology* **2004**, *14*, 147−155.

(17) Rutishauser, U. *Nat. Rev. Neurosci.* **2008**, *9*, 26−35.

(18) Cunningham, B. A.; Hoffman, S.; Rutishauser, U.; Hemperly, J. J.; Edelman, G. M. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 3116−3120.

(19) Sadoul, R.; Hirn, M.; Deagostini-Bazin, H.; Rougon, G.; Goridis, C. *Nature* **1983**, *304*, 347−349.

(20) Kim, D. S.; Lee, D. R.; Kim, H. S.; Yoo, J. E.; Jung, S. J.; Lim, B. Y.; Jang, J.; Kang, H. C.; You, S.; Hwang, D. Y.; Leem, J. W.; Nam, T. S.; Cho, S. R.; Kim, D. W. *PLoS One* **2012**, *7*, e39715.

(21) Raitano, S.; Verfaillie, C. M.; Petryk, A. *Front. Physiol.* **2013**, *4*, 49.

(22) Gascon, E.; Vutskits, L.; Kiss, J. Z. *Brain Res. Rev.* **2007**, *56*, 101−118.

(23) Gomez-Climent, M. A.; Guirado, R.; Castillo-Gomez, E.; Varea, E.; Gutierrez-Mecinas, M.; Gilabert-Juan, J.; Garcia-Mompo, C.; Vidueira, S.; Sanchez-Mataredona, D.; Hernandez, S.; Blasco-Ibanez, J. M.; Crespo, C.; Rutishauser, U.; Schachner, M.; Nacher, J. *Cereb. Cortex* **2011**, *21*, 1028−1041.

(24) Bonfanti, L. *Prog. Neurobiol.* **2006**, *80*, 129−164.

(25) James, W. M.; Agnew, W. S. *Biochem. Biophys. Res. Commun.* **1987**, *148*, 817−826.

(26) Curreli, S.; Arany, Z.; Gerardy-Schahn, R.; Mann, D.; Stamatos, N. M. *J. Biol. Chem.* **2007**, *282*, 30346−30356.

(27) Rey-Gallardo, A.; Escribano, C.; Delgado-Martin, C.; Rodriguez-Fernandez, J. L.; Gerardy-Schahn, R.; Rutishauser, U.; Corbi, A. L.; Vega, M. A. *Glycobiology* **2010**, *20*, 1139−1146.

(28) Moebius, J. M.; Widera, D.; Schmitz, J.; Kaltschmidt, C.; Piechaczek, C. *BMC Immunol.* **2007**, *8*, 13.

(29) Fedorkova, L.; Rutishauser, U.; Prosser, R.; Shen, H.; Glass, J. D. *Physiol. Behav.* **2002**, *77*, 361−369.

(30) Glass, J. D.; Watanabe, M.; Fedorkova, L.; Shen, H.; Ungers, G.; Rutishauser, U. *Neuroscience* **2003**, *117*, 203−211.

(31) Glass, J. D.; Shen, H.; Fedorkova, L.; Chen, L.; Tomasiewicz, H.; Watanabe, M. *Neurosci. Lett.* **2000**, *280*, 207−210.

(32) Jimbo, T.; Nakayama, J.; Akahane, K.; Fukuda, M. *Int. J. Cancer* **2001**, *94*, 192−199.

(33) Suzuki, M.; Suzuki, M.; Nakayama, J.; Suzuki, A.; Angata, K.; Chen, S.; Sakai, K.; Hagihara, K.; Yamaguchi, Y.; Fukuda, M. *Glycobiology* **2005**, *15*, 887−894.

(34) Marginean, I.; Kronewitter, S. R.; Moore, R. J.; Slysz, G. W.; Monroe, M. E.; Anderson, G.; Tang, K.; Smith, R. D. *Anal. Chem.* **2012**, *84*, 9208−9213.

(35) Inoue, S.; Lin, S. L.; Inoue, Y. *J. Biol. Chem.* **2000**, *275*, 29968−29979.

(36) Hu, Y.; Mechref, Y. *Electrophoresis* **2012**, *33*, 1768−1777.

(37) Aldredge, D.; An, H. J.; Tang, N.; Waddell, K.; Lebrilla, C. B. *J. Proteome Res.* **2012**, *11*, 1958−1968.

(38) von Der Ohe, M.; Wheeler, S. F.; Wuhrer, M.; Harvey, D. J.; Liedtke, S.; Muhlenhoff, M.; Gerardy-Schahn, R.; Geyer, H.; Dwek, R. A.; Geyer, R.; Wing, D. R.; Schachner, M. *Glycobiology* **2002**, *12*, 47−63.

(39) Page, J. S.; Tang, K.; Kelly, R. T.; Smith, R. D. *Anal. Chem.* **2008**, *80*, 1800−1805.

(40) Cox, J. T. 2014, unpublished work.

(41) Maniatis, S.; Zhou, H.; Reinhold, V. *Anal. Chem.* **2010**, *82*, 2421−2425.

(42) Kronewitter, S. R.; de Leoz, M. L.; Peacock, K. S.; McBride, K. R.; An, H. J.; Miyamoto, S.; Leiserowitz, G. S.; Lebrilla, C. B. *J. Proteome Res.* **2010**, *9*, 4952−4959.

(43) Maiolica, A.; Borsotti, D.; Rappsilber, J. *Proteomics* **2005**, *5*, 3847−3850.

(44) Tornkvist, A.; Nilsson, S.; Amirkhani, A.; Nyholm, L. M.; Nyholm, L. *J. Mass Spectrom.* **2004**, *39*, 216−222.

(45) Marginean, I.; Page, J. S.; Tolmachev, A. V.; Tang, K.; Smith, R. D. *Anal. Chem.* **2010**, *82*, 9344−9349.

(46) Varki, A.; Cummings, R. D.; Esko, J. D.; Freeze, H. H.; Stanley, P.; Marth, J. D.; Bertozzi, C. R.; Hart, G. W.; Etzler, M. E. *Proteomics* **2009**, *9*, 5398−5399.

(47) Liedtke, S.; Geyer, H.; Wuhrer, M.; Geyer, R.; Frank, G.; Gerardy-Schahn, R.; Zahringer, U.; Schachner, M. *Glycobiology* **2001**, *11*, 373−384.

(48) Kudo, M.; Kitajima, K.; Inoue, S.; Shiokawa, K.; Morris, H. R.; Dell, A.; Inoue, Y. *J. Biol. Chem.* **1996**, *271*, 32667−32677.

(49) Sato, C.; Kitajima, K. *J. Biol. Chem.* **2013**, *154*, 115−136.

(50) Zhang, Y.; Lee, Y. C. *J. Biol. Chem.* **1999**, *274*, 6183−6189.

(51) Pabst, M.; Grass, J.; Toegel, S.; Liebminger, E.; Strasser, R.; Altmann, F. *Glycobiology* **2012**, *22*, 389−399.

(52) Plummer, T. H., Jr.; Elder, J. H.; Alexander, S.; Phelan, A. W.; Tarentino, A. L. *J. Biol. Chem.* **1984**, *259*, 10700−10704.

(53) Roberts, B. A.; Strauss, C. R. *ChemInform* **2005**, *36*, 653−661.

(54) Otera, J.; Nishikido, J. In *Esterification: Methods, Reactions, and Applications*, 2nd ed.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2010; pp 3−157.

(55) Maley, F.; Trimble, R. B.; Tarentino, A. L.; Plummer, T. H., Jr. *Anal. Biochem.* **1989**, *180*, 195−204.

(56) 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.

(57) Bynum, M. A.; Yin, H.; Felts, K.; Lee, Y. M.; Monell, C. R.; Killeen, K. *Anal. Chem.* **2009**, *81*, 8818−8825.

(58) Chakraborty, A. B.; Berger, S. J. *J. Biomol. Technol.* **2005**, *16*, 327−335.

(59) Melmer, M.; Stangler, T.; Premstaller, A.; Lindner, W. *J. Chromatogr., A* **2011**, *1218*, 118−123.

(60) Page, J. S.; Marginean, I.; Baker, E. S.; Kelly, R. T.; Tang, K.; Smith, R. D. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 2265−2272.

(61) Gao, Q.; Leung, K. T. *J. Chem. Phys.* **2005**, *123*, 074325.

(62) Kronewitter, S. R.; An, H. J.; de Leoz, M. L.; Lebrilla, C. B.; Miyamoto, S.; Leiserowitz, G. S. *Proteomics* **2009**, *9*, 2986−2994.

(63) Bones, J.; McLoughlin, N.; Hilliard, M.; Wynne, K.; Karger, B. L.; Rudd, P. M. *Anal. Chem.* **2011**, *83*, 4154−4162.