research articles Journal of proteome research

# Human Serum Processing and Analysis Methods for Rapid and Reproducible N-Glycan Mass Profiling

**Scott R. Kronewitter,[†] Maria Lorna A. de Leoz,[†] Kyle S. Peacock,[†] Kelly R. McBride,[†] Hyun Joo An,[†] Suzanne Miyamoto,[‡] Gary S. Leiserowitz,[‡] and Carlito B. Lebrilla*[,†]**

*Department of Chemistry, University of California Davis, One Shields Avenue, Davis, California 95616, and Division of Hematology/Oncology, UC Davis Cancer Center, 4501 X Street, Sacramento, California 95817*

**Received March 5, 2010**

Glycans constitute a new class of compounds for biomarker discovery. Glycosylation is a common post-translational modification and is often associated with transformation to malignancy. To analyze glycans, they are released from proteins, enriched, and measured with mass spectrometry. For biomarker discovery, repeatability at every step of the process is important. Locating and minimizing the process variability is key to establishing a robust platform stable enough for biomarker discovery. Understanding the variability of the measurement devices helps understand the variability associated with the chemical processing. This report explores the potential use of methods expediting the enzymatic release of glycans such as a microwave reactor and automation of the solid-phase extraction with a robotic liquid handler. The study employs matrix-assisted laser desorption/ionization-Fourier transform ion cyclotron resonance mass spectrometry but would be suitable with any mass spectrometry method. Methods for system-wide data analysis are examined because proper metrics for evaluating the performance of glycan sample preparation procedures are not well established.

**Keywords:** glycomics • interpretation of mass spectra • Fourier transform ion cyclotron resonance mass spectrometry • human serum

## Introduction

Glycans are complex linear and branched arrangements of monosaccharides that are essential to the vitality of organisms. It is projected that over 50% of all proteins are glycosylated based on the Swiss-Prot database of known protein sequences and putative binding sites.[1] N-linked glycans make up the most abundant group and are found only on an asparagine that is adjoined by a variable amino acid followed by serine, threonine, or less commonly cysteine amino acid (Asn-Xaa-Ser/Thr/Cys).[2–4] N-linked glycans have a large diversity in structures bound to a consensus sequence on a glycoprotein.

The biological importance of glycosylation in cellular recognition and communication has made them a prime target for discovering biomarkers for diseases.[5–10] N-glycans are typically isolated for analysis by cleaving off the glycan from glycoproteins with the endoglycosidase enzyme PNGase F (peptide: N-Glycosidase F). PNGase F is often the enzyme of choice because of its broad specificity toward almost all N-glycans. However, if the N-glycan core is modified with an α1−3 fucose, it will be resistant to PNGase F.[11] After release, mass spectrometry is often performed to obtain glycan compositions. Mass spectrometry is currently the best method for profiling glycans due to its sensitivity and speed of analysis. Furthermore, profiling glycan compositions is one key step toward understanding structure−function relationships of specific glycan structures.

Biomarker discovery relies on the reproducible release and quantitative analysis of glycans from tissue. The most direct method for discovery involves three major steps: (1) the release of the glycan, (2) enrichment, and (3) mass spectrometry. The reliability of the markers relies on each of these steps to be highly reproducible.

Global enzymatic release of N-glycans as performed with PNGase F requires treatment of serum in a physiological buffer and reacted isothermally at 37 °C. The shortest amount of time for digestion of pure glycoproteins is 2 h,[12,13] whereas digestion of complex mixtures, like serum, can take from 12 to 22 h.[14–16] However, three approaches have been implemented to alter the common PNGase F enzymatic release method: the creation of an immobilized PNGase F enzyme reactor,[17] high pressure-cycling of reaction conditions,[18] and the implementation of a microwave reactor.[15,19]

Palm et al.[17] built a monolithic PNGase F enzyme micro reactor by immobilizing the enzyme covalently onto a porous polymer-based monolith. They immobilized the PNGase F enzyme in a cross-linked acrylamide gel. Standard glycoproteins (chicken ovalbumin, bovine asialofetuin, bovine ribonuclease B, and bovine $\alpha_1$-acid glycoprotein) were shown to effectively deglycosylate in 3.5 min incubation times while maintaining constant activity for eight weeks.

* To whom correspondence should be addressed. Carlito B. Lebrilla, e-mail: cblebrilla@ucdavis.edu. Telephone: 1-530-752-0504. Fax: 1-530-752-8995. Postal Address: Department of Chemistry, University of California Davis, One Shields Avenue, Davis, California 95616.

[†] University of California Davis.

[‡] UC Davis Cancer Center.



10.1021/pr100202a 

Published on Web 08/10/2010

**Table 1.** Sample Breakdown[a]

| | samples | source | release method | solid phase extraction |
|---|---|---|---|---|
| Sample Set A | 4 | Sigma | Water Bath | Vacuum Manifold |
| Sample Set B | 4 | Sigma | Microwave | Gilson GX-274 |
| Sample Set C | 1 | Sigma | Microwave | Gilson GX-274 |
| Sample Set D | 48 | GOG | Microwave | Gilson GX-274 |

[a] Set A,B,C came from a single human serum sample purchased from Sigma Aldrich. Set D consists of 48 different healthy control individuals purchased from the Gynecologic Oncology Group.

Szabo et al.[18] used pressure cycling technology to enhance the release of N-glycans by PNGase F digestion. Pressure cycling between low (atmospheric) and high pressures (up to 30 kpsi) was shown to enhance the glycosidase activity requiring minutes for glycan release while requiring relatively low amounts of enzyme.

Sandoval et al.[19] employed microwave irradiation to enhance the enzymatic activity. Glycoproteins Herceptin (trastuzumab), Avastin (bevacizumab) (Genentech Inc.) and bovine ribonuclease B were deglycosylated with reaction times of 10 min. Bereman et al.[15] experimented with a liquid cooled microwave reactor that allowed for temperature control. They found that a fixed temperature of 25 °C for 20 min (~250 W) had the best performance based on the sum of the plasma glycan abundances.

This report examines a high-throughput sample preparation pipeline for global glycan mass profiling. It includes several metrics for method analysis including the percent coverage of glycan libraries, average glycan abundances, and coefficients of variation of glycan abundances. By isolating technical variables and experimental variables from sample variability, the variation of the glycan processing steps and spectra collection were studied. Microwave assisted enzymatic N-glycan release and automated solid phase extraction (performed by a liquid handler) were implemented in this workflow to speed up and stabilize the processing. A matrix-assisted laser desorption/ionization-Fourier transform ion cyclotron resonance (MALDI-FTICR) mass spectrometer was used to rapidly scan the glycan sample and capture high mass accuracy high-resolution mass spectra. The expedited automated sample preparation and mass spectra collection methods scale well for large clinical sample sets allowing for fast turnaround.

## Experimental Section

**Samples.** The human serum samples were purchased from Sigma Aldrich (Sigma) and the Gynecological Oncology Group (GOG) tissue bank. The number of technical replicates in the case of Sigma serum, and number of individual serum samples from the GOG are shown in Table 1. The pertinent sample processing parameters are also included. Human serum samples used for reproducibility studies were purchased from Sigma and all came from the same person and lot (Sets A,B,C) while samples for the broad human profile study (Set D) came from the GOG tissue bank. The GOG samples were organized and divided into aliquots at the University of California Davis Medical Center Clinical Laboratories. The serum processing at the GOG tissue bank was performed with an Internal Review Board (IRB) approved protocol. All samples were transported frozen and stored in a −75 °C freezer prior to processing.

Evaluation of the reproducibility of MALDI sample spot preparation within a spot (intraspot), and between spots (interspot) was performed on a single sample (Set A). Using a single sample decoupled the sample preparation variability from the variability due to mass spectra collection. Process reproducibility is evaluated with coefficients of variation (CV) and sample correlation coefficients. Standardization of the mass spectra collection process was also optimized on a set of four technical replicate serum samples.

A human serum sample set of four technical replicates from one serum sample were used to compare two sample processing methods and isolate the biological diversity from the experiments. The standard method (Set B) involves typical isothermal PNGase F N-glycan enzymatic release and manual vacuum manifold solid-phase extraction. The new method (Set C) replaces the glycan release with a more streamlined microwave-assisted PNGase F N-glycan release and exchanges the vacuum manifold with an automated Gilson GX-274 ASPEC liquid handler.

This method was further applied to a large set of healthy human serum samples ($n = 48$) and evaluated (Set D). Glycan profiles were employed for comparing spectra across large sample sets. Metrics used here for method evaluation include how often a glycan can be detected in a sample set and its average abundances.

**Protein Denaturation.** Serum proteins were denatured under mild conditions to cleave disulfide bonds while minimizing glycan degradation. Serum samples were mixed in equal parts with an aqueous solution of 200 mM ammonium bicarbonate and 10 mM dithiothreitol. Detergents and buffers used during this step can often interfere with the mass spectrometry analysis downstream.[20] The ammonium bicarbonate/dithiothreitol solution was chosen because it is easily removed during the solid-phase extraction step. Low salt levels produce samples that are easily ionized in the mass spectrometer. The thermal denaturation includes four cycles of alternating between boiling water (100 °C) and room temperature (25 °C) water for 15 s each. The whole process takes 2 min per set of 24 samples, which is favorable for high throughput applications.

**Rapid Enzymatic Glycan Release.** The CEM microwave reactor (CEM Corporation, North Carolina) has two operational modes: standard and solid-phase synthesis (SPS) mode. Both modes require the user to specify a temperature and microwave power set point. The temperature is monitored by a fiber optic probe and both temperature and microwave power levels are presented on the reactor display. Heat generated from the microwave radiation was removed by a stream of air around a water-filled sample cup. The water in the cup surrounds the sample tubes allowing for a mediated constant temperature across all samples. For the standard mode temperature control, the reactor uses process control algorithms to vary the microwave power to hold a selected temperature. Unfortunately, the algorithms are designed for high temperature reactions (i.e., organic reactions). Low temperature conditions produce a sinusoidal response in microwave power and sample temperature. The variability in power supplied to the sample from this method makes studying the microwave enhancement effect difficult. Standard mode temperature controlled microwave power and sample temperature vs. time were plotted and presented in Figure 1a. Alternately, the SPS mode has a temperature set point and uses a step function of power inputs to maintain the temperature. By relaxing the temperature set point in the SPS mode, the step function can be avoided so that well-controlled, constant power experiments can be conducted. Each constant power setting has a different equilibrium temperature that is balanced by the cooling capacity

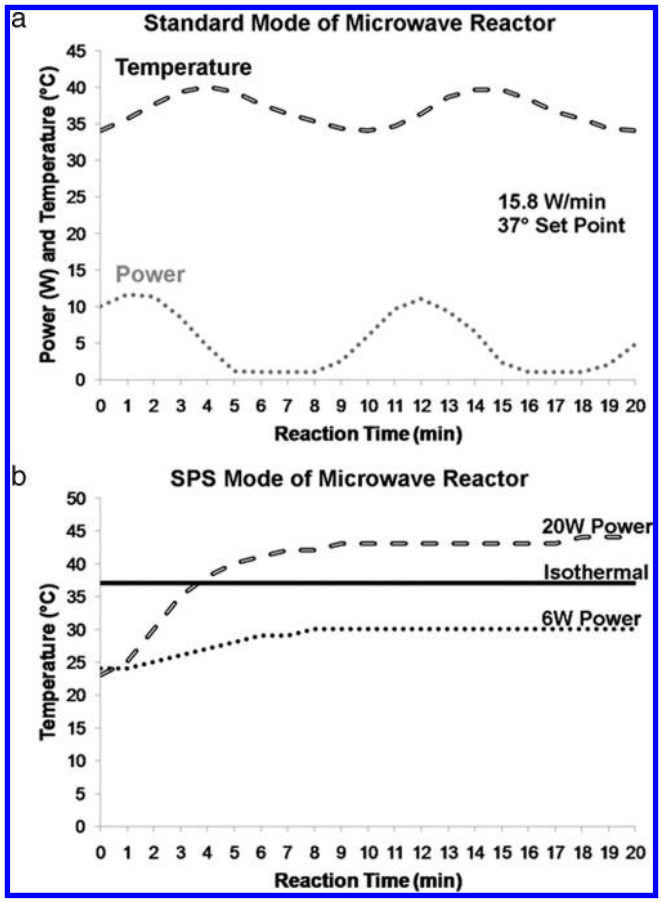


**Figure 1.** (a) Sinusoidal response of microwave power (watt) and sample temperature (degree Celsius) under standard mode reactor operation and a 37 °C set point. (b) PNGase F enzymatic digestion temperature profiles. The solid line represents the standard water bath incubation at 37 °C. The 6W and 20W air-cooled microwave reactor temperature profiles are presented. The temperature curves reach their equilibrium temperature eight minutes into the run.

of the reactor. Figure 1b depicts two stable equilibrium temperatures at 30 °C (6W) and 44 °C (20W). The optimal water bath temperatures for PNGase F release are plotted as a horizontal line. For a 10 min run, a 6W setting will input 3.5 KJ/run of microwave energy, while the 20W setting will input 12 KJ/run. However, much of this energy will be absorbed into the reactor walls and cooling water so that a 3-fold energy increase will keep the temperatures at a reasonable range (<45 °C).

Constant power experiments are preferred because they are better controlled, which should result in better reproducibility. Microwave-enhanced N-glycan release of 24 serum samples can be performed in 20 min (2 sets of 12 for 10 min). The optimal power setting was 20W for 10 min, which was preferred over the lower 6W setting because the former yielded better glycan library coverage and glycan abundances (data not shown).

**Ethanol Precipitation.** Ethanol precipitation is highly effective for removing proteins from the glycans. In 80% ethanol, the proteins precipitate and are centrifuged into a hard pellet. Chilled ethanol was used to precipitate residual proteins prior to solid-phase extraction. After the ethanol was added, the solution was further frozen in a −75 °C freezer for one hour to promote maximum precipitation. The samples were then centrifuged for thirty minutes at 13 200 rpm (16 110× *g*) in an Eppendorf 5415 D micro centrifuge and the supernatant containing the glycans was extracted and dried.

**Automated Graphitic Carbon Solid-Phase Extraction.** Cartridges filled with a stationary phase consisting of 150 mg of graphitic carbon are commonly used to desalt and fractionate glycans (Carbograph, Alltech Associates, Inc., Deerfield, IL). One challenge of graphitized carbon solid-phase extraction is the formation of bubbles on the high surface area carbon during the extraction process. Bubble formation causes flow channeling which changes the flow speeds through the cartridges and restricts the available area of the stationary phase. Altering the interaction area and flow rates will change the binding and elution patterns of the glycans. Various solid-phase extraction devices have been developed for small and large applications with varying capabilities for bubble removal.

For small sample sets, independent syringe plungers use positive pressure to force solutions through the cartridge. Positive pressure is preferred because it is capable of removing air bubbles that form on the carbon bed during the process. However, this method is not amenable to scale up because users can typically control only a few syringe plungers at a time. Reproducibility across large sample sets is also a problem because the force applied to the plunger varies greatly between operators.

When larger sample sets are analyzed, higher throughput vacuum manifolds are typically used. Vacuum manifolds pull a vacuum beneath the cartridges and use a pressure differential to pull solutions through the cartridges. Although vacuum manifolds scale well and typically hold up to 24 samples per manifold, they are intrinsically flawed because of their use of a negative pressure gradient. Bubbles that develop in the carbon beds cannot always be removed with the vacuum manifolds.

Automated rapid throughput solid-phase extraction was performed on a Gilson GX-274 ASPEC liquid handler. This liquid hander can be equipped with up to three sample cartridge racks allowing for 36 samples to be processed per run. The GX-274 is equipped with four probes backed by four discrete syringe pumps. This allows for constant positive pressure flows and precise fluid transfer volumes. The method programmed is as follows. The cartridge was washed with 6 mL nanopure water prior to use and conditioned with 6 mL 0.1% trifluoroacetic acid (TFA) in 80% acetonitrile (ACN)/$H_2O$ (v/v). Another 6 mL nanopure was used as a final rinse of the cartridges. Glycan solutions were applied to the GCC cartridge and subsequently washed with 12 mL of nanopure (E-Pure, Barnstead) water at a flow rate of 1 mL/minute for desalting. Glycans were eluted with 4.5 mL 10% ACN/$H_2O$ (v/v), 20% ACN/$H_2O$ (v/v), and 40% ACN/$H_2O$ (v/v) with 0.05% TFA. Advantages of the automation were immediately gained in reproducibility because of precisely controlled flow rates and delivered solution volumes. Furthermore, control over flow rate and volume variables allowed us to study the effects of perturbing the variables. The GX-274 can process 24 samples in approximately four hours with little user intervention. The processing time is only slightly faster than manual operations because of the direct mapping of the purification procedure to the robot. Much of the processing time is consumed by preparing and conditioning the graphitized carbon cartridges prior to use. After processing, the samples were collected and dried in a centrifugal evaporator apparatus. Each sample was reconstituted in nanopure water prior to MS analysis.

**MALDI FTICR Mass Spectrometry.** Mass spectra were recorded on an external source MALDI-FTICR instrument (HiResMALDI, IonSpec Corporation, Irvine, CA) equipped with a 7.0 T superconducting magnet and a pulsed Nd:YAG laser 355 nm. A solution of 2,5-dihydroxybenzoic acid (DHB)/NaCl was used as the matrix for all oligosaccharide analyses (5 mg/mL in 50% ACN, 25% 0.01 M NaCl, and 25% $H_2O$). For the negative mode analysis, the 0.01 M NaCl solution was replaced with nanopure water.

Benefits of MALDI mass spectrometry used for this purpose include better salt tolerances and the production of singly charged species. Sodium chloride was used as a doping agent to favor the sodiated form of the glycans and enhance ionization. The production of singly charged sodiated ions simplified the data analysis by eliminating the charge deconvolution and summation of several ions with different cation adducts. The FTICR detector further simplified the analysis by providing high mass accuracy assignments, fully baseline resolved monoisotopic ions, and low baselines for noise threshold cutoffs. MALDI spotting was performed on disposable MALDI plates (Hudson Surface technology, Inc., Newark, NJ) to prevent contamination from previous samples. Each MALDI spot was thoroughly mixed and dried rapidly under vacuum to reduce "sweet spot" formation during crystallization. This instrument is equipped with a hexapole for ion accumulation allowing for multiple lower energy laser shots to be accumulated prior to loading the ICR cell. Fifteen 355 nm Nd:YAG laser pulses were accumulated and then scanned. Three to five spectra were collected from different parts of the MALDI spot and averaged together during data analysis.

**Data Analysis Metrics.** Sample correlation coefficients are a measure of the overall distribution of the glycans in the spectra in comparison to an average distribution of the set. When the same ions present and of similar abundance, the correlation coefficients are close to unity. An ideal replicate spectrum will have a coefficient of 1.0 and dissimilar spectra will have a value of zero.

The coefficients of variations (CV, standard deviation/mean) were calculated for the abundance of each glycan and averaged across several areas within the MALDI sample spot. CV is preferred to standard deviation for reporting variability because it is scaled by the peak abundance. However, calculating the CV in this manner skews the average value to higher values because low abundant species with relatively larger CV values are more prevalent than highly abundant species.

## Results and Discussion

**Reproducibility of the MALDI MS.** A set of replicate samples (Set A) derived from a single serum sample processed similarly with the microwave reactor and the Gilson robotic liquid handler was examined. Five spectra were collected from each MALDI spot and their glycan profiles elucidated. Each spectrum contains ions accumulated from 15 laser shots from a single laser location. All spectra were normalized by a base peak, which was fixed for the three ACN fractions to ensure all ions are normalized to the same peak. The base peaks were: $m/z =$ 1485.534 for 10% ACN fraction, $m/z = 1647.586$ for 20% ACN fraction, and 1930.680 for the 40% ACN fraction. Replicate spectra from each sample area were averaged, and missing data were excluded from the average calculations.

The MALDI ionization is often criticized for having "sweet spots" that yield greater ionization of analyte than other locations within a MALDI spot as a consequence of heterogeneous crystallization. Solution heterogeneity caused by insufficient mixing of the matrix with the analyte and surface imperfections on the MALDI plate can cause these effects. Proper analyte dilutions and extensive mixing of the analyte with the matrix minimizes heterogeneity within the MALDI spot area. In addition, another cause of sample heterogeneity is the formation of large DHB crystals. Slow drying time, typical of air-dried samples, produce larger crystals. Smaller crystals were achieved by rapid drying through exposure to vacuum.

**Table 2.** Sample Correlation Coefficients of the Glycan Ion Abundance Distributions between 5 Spectra from Different Locations within the Same MALDI Spot[a]

| | sample 1 | sample 2 | sample 3 | sample 4 |
|---|---|---|---|---|
| Spectra 1 | 0.9982 | 0.9968 | 0.9991 | 0.9993 |
| Spectra 2 | 0.9985 | 0.9964 | 0.9984 | 0.9997 |
| Spectra 3 | 0.9997 | 0.9975 | 0.9991 | 0.9973 |
| Spectra 4 | 0.9994 | 0.9981 | 0.9979 | 0.9974 |
| Spectra 5 | 0.9981 | 0.9991 | 0.9995 | 0.9996 |
| CV | 0.16 | 0.17 | 0.12 | 0.12 |

[a] Four samples were examined and all samples are highly correlated within a spot. This similarity suggests spot heterogeneity has a minimal affect the relative ion distributions. The coefficient of variation (CV) within a spot is also shown. Even though the distributions are highly correlated, the glycan abundances vary 12−17%.

Each MALDI sample spot has a sufficiently large diameter to allow for several locations to be discretely targeted by the laser. Five locations were selected to minimize over sampling. The crystallized samples were sufficiently thick to allow for multiple laser shots per location during ion accumulation. For each spectra, 15 shots were accumulated in a hexapole prior to transfer to the ICR cell for analysis. A set of five spectra from one location is shown in Figure 4a to depict a typical response. The even responses of glycans abundances across all spectra demonstrate the homogeneity of the entire MALDI spot and reproducibility of the ionization process. An expanded section of low abundance glycan ions from the five spectra is shown in Figure 4b. Even at low abundances, the glycan response remains reproducible across the spot area.

Reproducibility of the MALDI process was evaluated for variability within a MALDI spot at different locations and between separate MALDI spots. Table 2 summarizes the results of the correlation coefficients for the five spectra and the different laser locations in the same MALDI spot area. The average correlation coefficient for spectra from the same spot is 0.998. There is high correlation between the spectra, but further averaging the spectra for each spot continues to decrease the overall coefficient of variation. By averaging five spectra instead of using one, the overall coefficient of variation is reduced from 21 to 10%. This decrease is in line with a decrease in standard deviation by the square root of n, where n is the number of spectra averaged ($n = 5$ spectra per MALDI spot). Decreasing the variation by averaging multiple spectra from a single MALDI spot is presented in Figure 5. If this square root of $n$ decrease in standard deviation can be maintained, more spectra should be collected (instrument time allowing).

Variation between different MALDI spots were studied by comparing spectra collected from separate MALDI spots on the probe. The reproducibility of the MALDI spots were examined by comparing the spectra from eight different MALDI spots from a single serum sample (10% ACN fraction). The average coefficient of variation across the MALDI spots was 10%, and the distribution for each ion is presented in Figure 6. Ion $m/z$

1850.66 has no deviation because it was chosen as the base peak for spectral normalization. Each ion was detected in all of the MALDI spots, and 66% of the ions have a CV less than 10%.

**Effect of FT-ICR Analyzer Pressure on MS Measurements.** A MALDI FT-ICR MS operating protocol was developed to maintain instrumental parameters between measurements. A key condition for the analysis was the pressure in the analyzer cell during the analysis. Collision gas was used to decrease the translational energy of the ions and produce optimal trapping in the ICR cell. Obtaining ultra high vacuum was necessary to decrease the damping of the transient in the ICR cell by reducing the number of the ion-neutral collisions in the cell. Rapid spectra collection increases the base pressure unless additional time for the pressure to return to baseline is allowed between samples. Monitoring the base pressure between $1.5-2.0 \times 10^{-10}$ Torr yielded the proper repetition rate for spectra collection.

To examine the effect of pressure related transient damping on reproducibility, a set of four technical replicate samples (10% ACN fraction) processed in the same manner as Set C was used. The samples were analyzed with, and without, the strict control over the base pressure and pump down rate of the instrument. We found that by consistently maintaining the base pressure prior to spectra collection, the average CV across the glycans decreased from 16 to 11%.

**Glycan Annotation.** Glycans were annotated with the in-house Glycolyzer analysis software. Briefly, the free induction decay signals from the ICR cell were fast Fourier transformed and internally calibrated using glycan signals that are present in all spectra. The noise was removed from the spectra by removing data that was below the lower limit of detection (LLOD). The LLOD was set to six standard deviations above the mean noise level. Resulting monoisotopic peaks were annotated with a retrosynthetic theoretical glycan library and exact mass assignments (<10 ppm mass error).[21] Averages of peak abundances were calculated by averaging all values detected above the noise level. Missing data was not included in the average values.

Each ACN fraction is measured separately with the high resolution MALDI FT-ICR. Theoretical retrosynthetic N-glycan networks were used as a metric for measuring how many glycans were detected in individual human glycan profile.[21] Each ACN fraction was separately analyzed and annotated resulting in 33 glycans being detected from the 10% fraction, 20 from the 20% fraction, and 29 from the 40% fraction. The summed total of all glycans detected is 89. However, several masses were detected across multiple ACN fractions. Tandem MS was performed on compositions that were found in at least two fractions. The results show that these compositions corresponded to different structures. The total number of distinct compositions overall fractions for the two methods was 64, resulting in a 20% coverage of the theoretical library.

**Comparison of Standard Release and Enrichment Method with a the Microwave Release and Robotic Enrichment Method.** Three criteria were used for comparing the performance of the methods used to process sample Set B and Set C: percent coverage of the theoretical retrosynthetic N-glycan network library, average glycan abundances, and the mean coefficient of variation (CV) of the glycan abundances.

Set B displayed 55 glycans (15% coverage) and Set C had 59 glycans (18% coverage). A grid diagram is presented in Figure 2 where each square represents a glycan and the shading

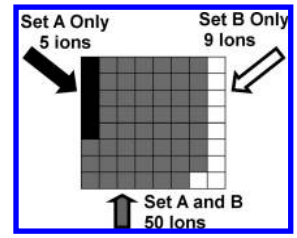


**Figure 2.** This grid diagram represents the number of glycans detected that are unique and in common between sample Set B and sample Set C. 50 glycans were detected in both sets while 5 were only detected in Set B and 9 were only detected in Set C.

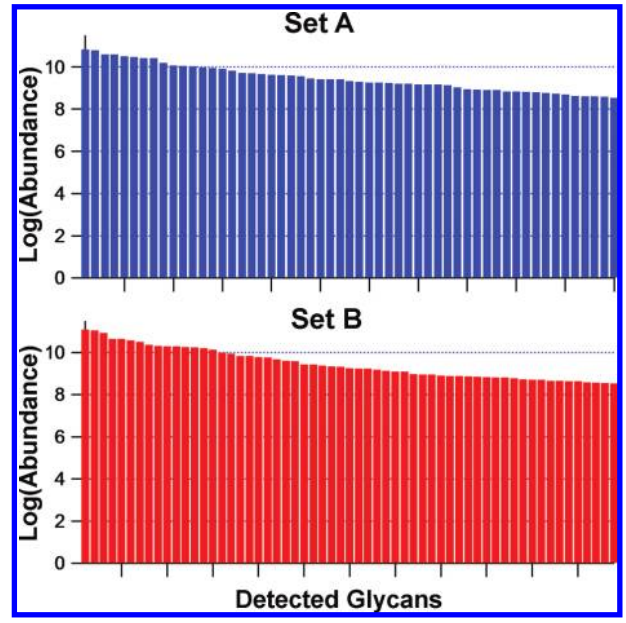


**Figure 3.** Specific glycan log-abundances from Set B and Set C are compared. Each bar represents a different glycan. Both plots are sorted by decreasing abundance.

represents its classification into one of three groups: only Set B, only Set C or both Set B and Set C. This figure shows the high percentage of common ions (78%) in Set B and Set C.

Measuring average glycan ion abundances is another metric for comparing data sets. Larger average ion abundances often result in more extensive glycan profiles because of the large dynamic range of glycan abundances in serum. These extended profiles may occur because low-abundance glycans can be detected above the lower limit of detection threshold in intense spectra. The average abundance ratio of Set B to Set C was 1.0 to 1.4 suggesting an overall increase in abundance for the new method. However, on an absolute scale a few dominant peaks will have a large contribution to the average and skew the results. To minimize this effect, logarithms of the abundances (base 10) are used for comparisons. The log average abundances for each glycan are presented in Figure 3. The *x*-axis represents individual glycans and the log abundances are plotted on the *y*-axis. The plots are sorted by decreasing abundances. The abundances of glycans from each method were comparable showing a similar response in the quantities of glycans ionized and detected.

The CV of each ion abundances was calculated in the following manner. The standard deviation of each normalized glycan abundance was calculated and subsequently divided by the respective mean ion abundance. Missing values for glycan

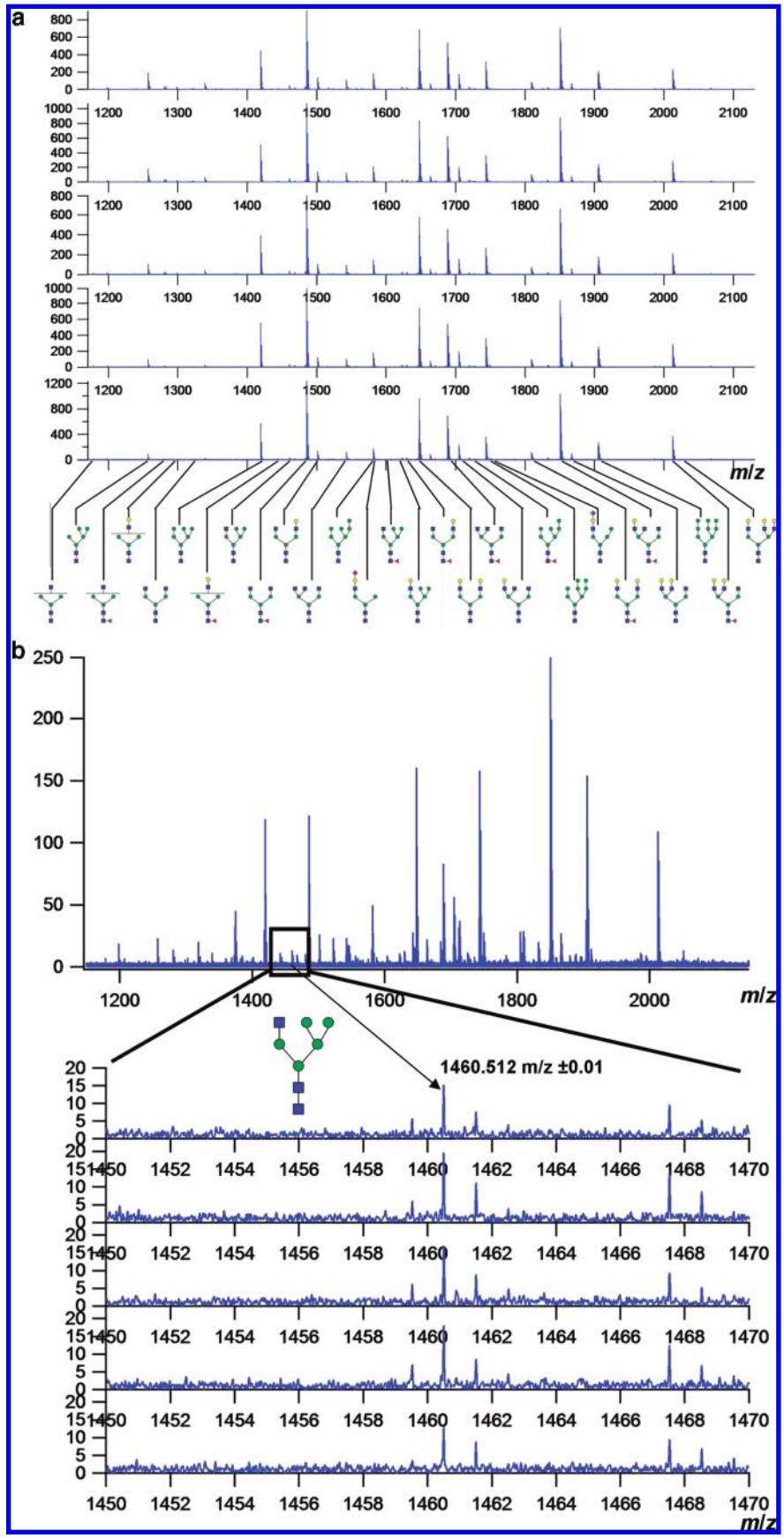


**Figure 4.** (a) Replicate mass spectra from a single MALDI spot from the 10% ACN elution fraction. The relative distribution of the glycans is highly conserved between the technical replicate spectra. Annotated putative structures are depicted based common serum glycobiology. (b) Expanded view of low abundance glycans detected in technical replicate spectra from the same MALDI spot from the 10% ACN fraction. All ions detected were above the lower limit of detection set at 6 S/N.

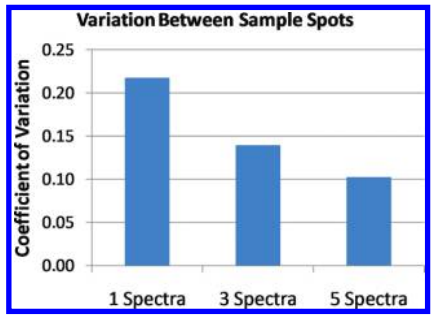


**Figure 5.** Coefficient of variation vs. number of spectra collected within a spot. The variation decrease by roughly the square root of *n* where *n* is the number of replicate spectra averaged.

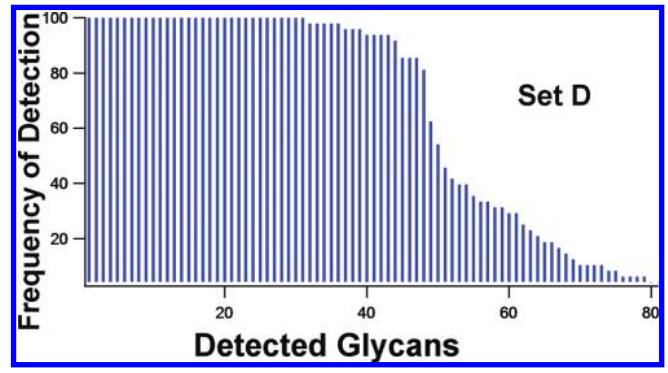


**Figure 7.** Frequency of detection of each glycan from Set D. Frequency of detection is the fraction of samples of the sample set that contained the glycan above the noise level multiplied by one hundred.

ions that were not detected above the lower limit of detection (6 S/N) were excluded from the calculation. The overall mean CV was calculated by taking the mean of all of the individual CV values calculated. This includes all CV values from each ACN fraction. Set B and Set C had a mean CV of 23% and 25%, respectively. The similarity of the mean CV calculated suggests proper automation of the manual process. The overall CV includes variation from the chemical (enzyme and enrichment) processing and mass spectrum collection.

**Application of the Rapid Processing Method to 48 Samples.** Evaluating large data sets (Set D) require additional metrics to handle the increased sample size dimension. Application of the theoretical retrosynthetic N-glycan library for spectral annotation allows for comparison across samples because each glycan ion is either detected above the LLOD noise threshold or not. This allows for the use of the frequency of glycan detection as a metric (frequency of detection). Glycans can generally be classified into two groups, glycans detected in all, or nearly all, of the samples and glycans that are less frequently detected. Frequently detected glycans are common to a population and constitute a common glycan profile. Common profile ions can be used specifically for calibration and method development as their response is predictable. The less frequent ions can typically be the result of ions that are near the LLOD noise threshold and are detected intermittently across a sample set. However, if a less frequent ion is detected well above the LLOD noise threshold, such as 10−20 sigma above the noise, it can be attributed to biological diversity.

The average frequency of detection of each glycan detected is plotted in Figure 7. The gradual decrease in detection frequencies can either be attributed to sample process variability or biological diversity of the sample set. The minimum frequency cutoff was established as the detection of a monoisotopic glycan ion above the noise threshold in at least two samples. Set D contained 31 ions detected in 100% of the spectra and the average frequency of detection is 67%. Across all samples, 81 N-glycan compositions were identified from the sample set and half were detected in at least 95% of the samples.

## Conclusions

A rapid, reproducible serum glycan purification pipeline was developed and its performance was evaluated. The faster

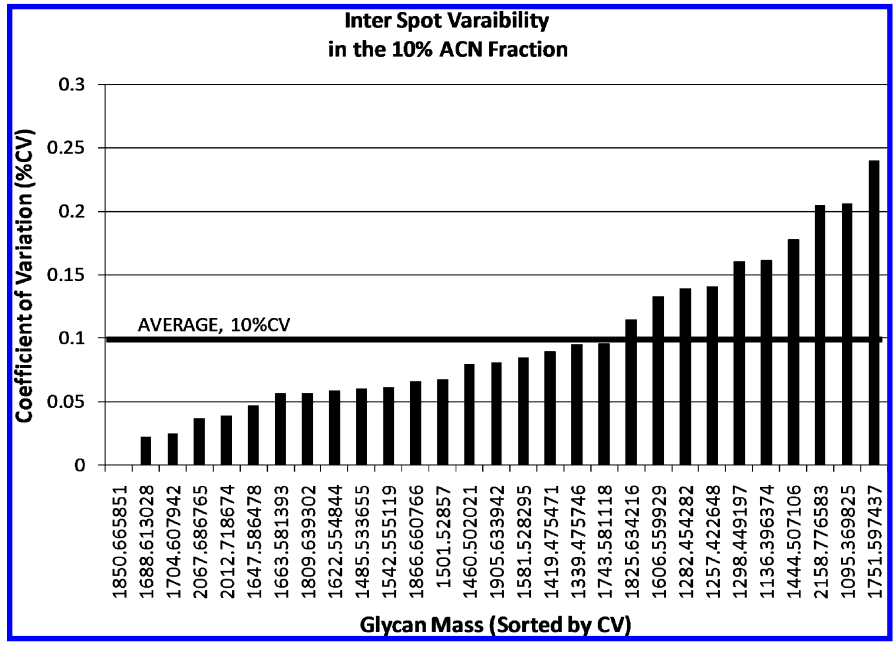


**Figure 6.** Coefficient of variation of each ion in the 10% ACN fraction vs. glycan mass. The plot is sorted by increasing variation. The coefficient of variation is calculated across a set of 8 MALDI spots from a single sample processed from Set C.

pipeline allows for 96 serum samples to be processed and ready for mass spectrometry analysis in one week. Faster turnaround is achieved since this timeline is inherently faster by one day when compared to the standard method. This new method is reproducible and applicable for large studies of healthy control and patient populations.

Substituting the vacuum manifolds with automated solid-phase extraction imparted the benefits of both positive pressure displacement and large sample processing capacity. The automation of the solid-phase extraction also creates free time for the researcher to perform concurrent tasks. By expediting, automating and standardizing the procedure, the coefficients of variation were decreased while maintaining glycan average abundances and profiles. The operator variability was minimized for the new method allowing for new operators to learn the procedure and help process larger sample sets.

**Acknowledgment.** We are grateful for funds provided by the National Institutes of Health (GM049077, HD061923), the Ovarian Cancer Research Fund, and the Ann Garat Fund. We also acknowledge the Gynecologic Oncology Group for providing serum samples used in this study.

## References

(1) Apweiler, R.; Hermjakob, H.; Sharon, N. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim. Biophys. Acta* **1999**, *1473* (1), 4–8.

(2) Vance, B. A.; Wu, W. Y.; Ribaudo, R. K.; Segal, D. M.; Kearse, K. P. Multiple dimeric forms of human CD69 result from differential addition of N-glycans to typical (Asn-X-Ser/Thr) and atypical (Asn-X-Cys) glycosylation motifs. *J. Biol. Chem.* **1997**, *272* (37), 23117–23122.

(3) Miletich, J. P.; Broze, G. J. Beta-Protein-C Is Not Glycosylated at Asparagine-329 - the Rate of Translation May Influence the Frequency of Usage at Asparagine-X-Cysteine Sites. *J. Biol. Chem.* **1990**, *265* (19), 11397–11404.

(4) Titani, K.; Kumar, S.; Takio, K.; Ericsson, L. H.; Wade, R. D.; Ashida, K.; Walsh, K. A.; Chopek, M. W.; Sadler, J. E.; Fujikawa, K. Amino-Acid-Sequences of Human Vonwillebrand-Factor. *Biochemistry* **1986**, *25* (11), 3171–3184.

(5) Lebrilla, C. B.; An, H. J. The prospects of glycan biomarkers for the diagnosis of diseases. *Mol. Biosyst.* **2009**, *5* (1), 17–20.

(6) Packer, N. H.; von der Lieth, C. W.; Aoki-Kinoshita, K. F.; Lebrilla, C. B.; Paulson, J. C.; Raman, R.; Rudd, P.; Sasisekharan, R.; Taniguchi, N.; York, W. S. Frontiers in glycomics: bioinformatics and biomarkers in disease. An NIH white paper prepared from discussions by the focus groups at a workshop on the NIH campus, Bethesda MD (September 11−13, 2006). *Proteomics* **2008**, *8* (1), 8–20.

(7) Turnbull, J. E.; Field, R. A. Emerging glycomics technologies. *Nat. Chem. Biol.* **2007**, *3* (2), 74–77.

(8) Barkauskas, D. A.; An, H. J.; Kronewitter, S. R.; de Leoz, M. L.; Chew, H. K.; White, R. W. D.; Leiserowitz, G. S.; Miyamoto, S.; Lebrilla, C. B.; Rocke, D. M. Detecting glycan cancer biomarkers in serum samples using MALDI FT-ICR mass spectrometry data. *Bioinformatics* **2009**, *25* (2), 251–257.

(9) An, H. J.; Miyamoto, S.; Lancaster, K. S.; Kirmiz, C.; Li, B.; Lam, K. S.; Leiserowitz, G. S.; Lebrilla, C. B. Profiling of glycans in serum for the discovery of potential biomarkers for ovarian cancer. *J. Proteome Res.* **2006**, *5* (7), 1626–1635.

(10) de Leoz, M. L. A.; An, H. J.; Kronewitter, S.; Kim, J.; Beecroft, S.; Vinall, R.; Miyamoto, S.; White, R. D.; Lam, K. S.; Lebrilla, C. Glycomic approach for potential biomarkers on prostate cancer: Profiling of N-linked glycans in human sera and pRNS cell lines. *Dis. Markers* **2008**, *25* (4−5), 243–258.

(11) Tretter, V.; Altmann, F.; Marz, L. Peptide-N4-(N-acetyl-beta-glucosaminyl)asparagine amidase F cannot release glycans with fucose attached alpha 1−3 to the asparagine-linked N-acetylglucosamine residue. *Eur. J. Biochem.* **1991**, *199* (3), 647–652.

(12) Ge, Y.; Gibbs, B. F.; Masse, R. Complete Chemical and Enzymatic Treatment of Phosphorylated and Glycosylated Proteins on ProteinChip Arrays. *Anal. Chem.* **2005**, *77* (11), 3644–3650.

(13) Papac, D.; Briggs, J.; Chin, E.; Jones, A. A high-throughput microscale method to release N-linked oligosaccharides from glycoproteins for matrix-assisted laser desorption/ionization time-of-flight mass spectrometric analysis. *Glycobiology* **1998**, *8* (5), 445–454.

(14) Lancaster, K. S.; An, H. J.; Li, B. S.; Lebrilla, C. B. Interrogation of N-linked oligosaccharides using infrared multiphoton dissociation in FT-ICR mass spectrometry. *Anal. Chem.* **2006**, *78* (14), 4990–4997.

(15) Bereman, M. S.; Young, D. D.; Deiters, A.; Muddiman, D. C. Development of a robust and high throughput method for profiling N-linked glycans derived from plasma glycoproteins by NanoLC-FTICR mass spectrometry. *J. Proteome Res.* **2009**, *8* (7), 3764–3770.

(16) Kyselova, Z.; Mechref, Y.; Kang, P.; Goetz, J. A.; Dobrolecki, L. E.; Sledge, G. W.; Schnaper, L.; Hickey, R. J.; Malkas, L. H.; Novotny, M. V. Breast cancer diagnosis and prognosis through quantitative measurements of serum glycan profiles. *Clin. Chem.* **2008**, *54* (7), 1166–1175.

(17) Palm, A. K.; Novotny, M. V. A monolithic PNGase F enzyme microreactor enabling glycan mass mapping of glycoproteins by mass spectrometry. *Rapid Commun. Mass Spectrom.* **2005**, *19* (12), 1730–1738.

(18) Szabo, Z.; Guttman, A.; Karger, B. L. Rapid release of N-linked glycans from glycoproteins by pressure-cycling technology. *Anal. Chem.* **2010**, *82* (6), 2588–2593.

(19) Sandoval, W. N.; Arellano, F.; Arnott, D.; Raab, H.; Vandlen, R.; Lill, J. R. Rapid removal of N-linked oligosaccharides using microwave assisted enzyme catalyzed deglycosylation. *Int. J. Mass Spectrom.* **2007**, *259* (1−3), 117–123.

(20) Bereman, M. S.; Williams, T. I.; Muddiman, D. C. Development of a nanoLC LTQ Orbitrap Mass Spectrometric Method for Profiling Glycans Derived from Plasma from Healthy, Benign Tumor Control, and Epithelial Ovarian Cancer Patients. *Anal. Chem.* **2009**, *81* (3), 1130–1136.

(21) Kronewitter, S. R.; Joo An, H.; Lorna de Leoz, M.; Lebrilla, C. B.; Miyamoto, S.; Leiserowitz, G. S. The development of retrosynthetic glycan libraries to profile and classify the human serum N-linked glycome. *PROTEOMICS* **2009**, *9* (11), 2986–2994.

PR100202A