Published in final edited form as:

Anal Chem. 2015 November 3; 87(21): 10830–10838. doi:10.1021/acs.analchem.5b02063.

Multiplexed Targeted Mass Spectrometry-Based Assays for the Quantification of N-Linked Glycosite-Containing Peptides in Serum

Stefani N. Thomas^{*,†,‡}, Robert Harlan^{†,§,‡}, Jing Chen^{†,‡}, Paul Aiyetan[†], Yansheng Liu[‡], Lori J. Sokoll[†], Ruedi Aebersold^{‡,§}, Daniel W. Chan[†], and Hui Zhang[†]

[†]Department of Pathology, Clinical Chemistry Division, Johns Hopkins University School of Medicine, 1550 Orleans Street, CRBII Room 3M03, Baltimore, Maryland 21231, United States

[‡]Department of Biology, Institute of Molecular Systems Biology, ETH Zurich, 8093 Zurich, Switzerland

[§]Faculty of Science, University of Zurich, 8057 Zurich, Switzerland

Abstract

Protein glycosylation is one of the most common protein modifications, and the quantitative analysis of glycoproteins has the potential to reveal biological functions and their association with disease. However, the high throughput accurate quantification of glycoproteins is technically challenging due to the scarcity of robust assays to detect and quantify glycoproteins. Here we describe the development of multiplexed targeted MS assays to quantify N-linked glycosite-containing peptides in serum using parallel reaction monitoring (PRM). Each assay was characterized by its performance metrics and criteria established by the National Cancer Institute's Clinical Proteomic Tumor Analysis Consortium (NCI CPTAC) to facilitate the widespread adoption of the assays in studies designed to confidently detect changes in the relative abundance of these analytes. An in-house developed software program, MRMPPlus, was used to compute assay performance parameters including specificity, precision, and repeatability. We show that 43 selected N-linked glycosite-containing peptides identified in prostate cancer tissue studies carried out in our group were detected in the sera of prostate cancer patients within the quantitative range of the developed PRM assays. A total of 41 of these formerly N-linked glycosite-containing peptides (corresponding to 37 proteins) were reproducibly quantified based on their relative peak area ratios in human serum during PRM assay development, with 4 proteins showing differential significance in serum from nonaggressive (NAG) vs aggressive (AG) prostate cancer patient serum ($n = 50$, NAG vs AG). The data demonstrate that the assays can be used for the high throughput and reproducible quantification of a panel of formerly N-linked glycosite-containing

*Corresponding Author: Phone: +1 410-955-7138. Fax: +1 443-287-6388. sthoma92@jhmi.edu.

[†]Present Address: R.H.: Center for Resources in Integrative Biology, Johns Hopkins University Schools of Medicine and Public Health, Baltimore, MD.

[‡]S.N.T., R.H., and J.C. contributed equally to the study.

Supporting Information: The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.analchem.5b02063.

Author Contributions: The manuscript was written through contributions of all authors. All authors have given approval of the final version of the manuscript.

Notes: The authors declare no competing financial interest.

peptides. The developed assays can also be used for the quantification of formerly N-linked glycosite-containing peptides in human serum irrespective of disease state.

Glycosylation is one of the most common protein modifications, and aberrant glycosylation has been implicated in carcinogenesis via various mechanisms including growth factor receptor regulation, growth factor modulation, cell–cell adhesion, immune system modulation, cell motility, and adhesion to endothelium.^{1–5} The quantification of specific post-translational modifications (PTMs) in individual proteins is technically challenging, in part due to the scarce availability of specific assays required to confidently detect the respective site of modification. The development of such assays, specifically those based on site-specific antibodies, is resource intensive. In contrast, multiple reaction monitoring (MRM), or selected reaction monitoring (SRM), mass spectrometry (MS) assays for the site-specific quantification of protein PTMs can be developed with relative ease. Such assays consist of the mass-to-charge ratio and relative intensity of specific fragment ions that indicate the sequence position of the modified amino acid residue and additional information such as the elution time and precursor ion mass of the respective analyte. From the first application of MRM to the quantification of peptides in biological tissues by Desiderio et al. in 1983,⁶ MRM-based assays have been developed in recent years for the quantification of PTMs such as glycosylation,⁷ phosphorylation,⁸ and ubiquitylation.⁹

MRM MS enables the targeting of specific analytes of interest, provides high specificity and sensitivity,^{10–13} and it is presently the most widely used MS-based targeted proteomic approach. MRM measurements are typically carried out in triple quadrupole (QQQ) mass spectrometers. The advantages of MRM compared to other quantitative analytical methods such as Western blotting, ELISA, and immunohistochemistry include multiplexed detection and the ability to use spiked-in, stable isotope-labeled standards to foster the absolute or precise relative quantification of endogenous analytes. MRM-based targeted protein assays do not require an antibody, and they can be used to detect either the unmodified or post-translationally modified forms of proteins.

Parallel reaction monitoring (PRM), first published in 2012,¹⁴ is a targeted proteomics strategy where all product ions of the target peptides are simultaneously monitored at high resolution and high mass accuracy. In PRM, the third quadrupole of a QQQ mass spectrometer is substituted with a high-resolution and accurate mass analyzer to permit the

parallel detection of all target product ions in one high-resolution mass analysis. PRM analyses exhibit performance characteristics (dynamic range and lower limits of detection and quantification) that are similar to those of MRM.¹⁵ Some advantages of PRM compared to MRM include (1) PRM spectra are highly specific because all of the potential product ions of a peptide, instead of just 3–5 transitions as in MRM, are recorded to confirm peptide identity; (2) high-resolution mass analysis can separate coisolated background ions from the peptide ions of interest which increases selectivity; and (3) the *a priori* selection of target transitions is not required, therefore, requiring minimal upfront method development and facilitating automated data analysis.¹⁶ Thus, PRM enables high-quality quantitative measurements, comparable in performance to those conducted using MRM performed on a triple quadrupole mass spectrometer, while simplifying method development.

The widespread availability of validated targeted MS-based assays has been recognized as a critical prerequisite to quantify proteins and to generally increase the reproducibility of data between laboratories and studies. The National Cancer Institute (NCI) of the United States National Institutes of Health has promoted the standardization and analytical validation of targeted MS-based quantification of peptides through the Clinical Proteomic Tumor Analysis Consortium (CPTAC).¹⁷ Toward this end, the CPTAC Assay Portal (<https://assays.cancer.gov>) was developed to provide a public repository of well-characterized, MS-based, targeted proteomic assays.¹⁸ The successful large-scale development and robust analytical performance of targeted MS-based analyses has been demonstrated across multiple laboratories.^{19,20}

Following the framework for targeted MS-based assay “fit-for-purpose” validation established by CPTAC with input from the outside community,²¹ we describe, in this study, the characterization of “Tier 2” PRM assays targeting N-linked glycosite-containing peptides from serum proteins. Tier 2 targeted MS-based assays are precise, relative quantitative assays with established performance that are most suitable for target verification, wherein relative changes in the levels of large numbers of targeted analytes are precisely and consistently measured across samples for nonclinical purposes. We developed PRM assays targeting formerly N-linked glycosite-containing peptides from serum glycoproteins because changes in the abundance or the glycosylation of serum glycoproteins have been shown to correlate with the glycoproteins identified from the disease site in cancer and other diseases.^{22,23} Other studies have provided MS-based evidence of the benefit of incorporating enrichment strategies targeting glycosite-specific, cancer-related carbohydrate structures into cancer biomarker discovery pipelines.^{24,25}

Because serum can be collected from patients in a minimally invasive manner there is great potential clinical utility for biomarkers that are validated in serum,^{26–29} and the majority of established serum biomarkers are glycoproteins secreted or leaked from the diseased tissue.^{30,31} Serum, instead of plasma, was selected for the assays we developed because of its lack of coagulant factors and the absence of the high molecular weight abundant protein fibrinogen.

Prostate cancer carcinogenesis is associated with aberrant glycosylation, and the majority of prostate cancer biomarkers are glycoproteins. According to 2014 data from the American

Cancer Society, prostate cancer is the second-leading cause of cancer death in men, and it is the most frequently diagnosed cancer in men aside from skin cancer.³² A significant challenge in the development of specific treatments for prostate cancer is the inability of currently available diagnostic biomarkers and histological examination of tumor tissue biopsies to distinguish aggressive (AG) from nonaggressive (NAG) prostate cancer.³³ Consequently, the under-treatment of AG and the overtreatment of NAG prostate cancer occur frequently. The identification of proteins whose relative levels of abundance can differentiate AG from NAG prostate cancer is an important step in prostate cancer biomarker development.

To test the ability of our fully characterized PRM-based assays to measure the relative abundance of N-linked glycosite-containing peptides in different disease states, 43 N-linked glycosite-containing peptides that were previously identified from prostate cancer tissue studies conducted by our group^{34,35} were quantified in a cohort of 75 serum samples. The cohort consisted of 25 patients with NAG prostate cancer, 25 patients with AG prostate cancer, and 25 patients without cancer as determined by a negative prostate cancer biopsy. Formerly N-linked glycosite-containing peptides were enriched from serum using the solid phase extraction of N-linked glycopeptides (SPEG) method,³⁶ wherein glycoproteins are conjugated to a solid support using hydrazide chemistry followed by the specific release of formerly N-linked glycosylated peptides via PNGaseF. The results indicate that 4 of the selected target peptides quantified with the PRM assays developed in this study have significantly different levels of relative abundance between the NAG and AG prostate cancer patients. The assays described in the current study can be adopted for the relative abundance analysis of N-linked glycosite-containing peptides in any human-derived serum sample.

Experimental Section

Clinical Specimens

Serum samples were from the Johns Hopkins Clinical Chemistry laboratory specimen bank. The race of the patients was not recorded during sample collection. The serum samples for assay deployment using biological samples were obtained from three groups of a total of 75 patients: Men who underwent radical prostatectomy (RP) for prostate cancer and had characteristics suggestive of either aggressive disease or nonaggressive disease and noncancer/biopsy negative. The clinical characteristics of these patients are detailed in Table 1. Aggressive (AG): Serum from 25 male patients with an RP Gleason score ≥ 7 and positive seminal vesicles and/or lymph node involvement and/or prostate cancer biochemical or other progression (positive bone scan or death from prostate cancer). Six of the patients in the Aggressive prostate cancer group died of prostate cancer. Nonaggressive (NAG): Serum from 25 male patients with RP Gleason score ≤ 6 and organ-confined disease (negative for capsular penetration, seminal vesicles, and lymph nodes). All had negative surgical margins. These patients did not exhibit biochemical or clinical prostate cancer progression for a minimum of 10 years following surgery. Negative (Neg): Serum from 25 male patients whose prostate biopsies were negative. Specimens were obtained prior to surgery or biopsy, they were not treated with protease inhibitors, and they were stored at -80°C in the

specimen bank for an average of 12 years with a maximum of 2 freeze–thaw cycles. For the preliminary studies to test the serum detection of the assay targets that were selected from the discovery-phase proteomic studies of prostate cancer tissues, serum from 38 patients was utilized: 10 normal men (age 23 ± 1 years) and 28 men with prostate cancer (13 with PSA concentrations between 4 and 10 ng/mL and 15 with PSA concentrations between 50 and 100 ng/mL). These specimens were stored at -80°C in the specimen bank for <1 year and underwent 1 freeze–thaw cycle.

Isolation of N-Linked Glycopeptides from Human Serum

N-linked glycopeptides were isolated from human serum based on the SPEG procedure using hydrazide tips integrated into an automated liquid handler (Versette, Thermo Fisher Scientific) as previously described.³⁷ A volume of 40 μL of serum was diluted 1:1 with oxidation buffer (500 mM sodium acetate, 0.3 mM sodium chloride, pH 5), oxidized with 15 mM sodium periodate, and buffer exchanged into coupling buffer. The serum sample was processed further using the Versette liquid handling robotic system.

LC-PRM Analysis

The formerly N-linked glycosite-containing peptides that were selected for assay development were identified in our previous prostate cancer tissue proteomic studies.^{34,35} Peptides with oxidizable (e.g., methionine and tryptophan) and modifiable (e.g., glutamine) residues were not excluded. Assay development was conducted using crude stable isotope-labeled (SIS) N-linked glycosite-containing heavy-isotope-labeled peptide standards ($\sim 60\%$ chemical purity, $>99\%$ isotopic purity, Thermo Fisher Scientific PEPotec SRM peptide library) and endogenous N-linked glycosite-containing peptides enriched from commercially available human serum (Sigma-Aldrich). Approximate concentrations of the crude peptides were determined via UV–vis absorption measurements at a wavelength of 280 nm using a NanoDrop spectrophotometer. SIS peptides incorporated a fully atom-labeled ^{13}C and ^{15}N isotope at the C-terminal lysine (K) or arginine (R) position of each tryptic peptide, resulting in a mass shift of +8 or +10 Da, respectively. Deamidated Asn residues corresponding to N-glycosylation sites were synthesized as Asp residues. Peptides were provided in 0.1% TFA/50% ACN and stored at -80°C until use. A stock SIS mix was cleaned via strong cation exchange (SCX) to rid the peptides of contaminants such as polymers and salt. The peptide recovery following SCX cleanup was approximately 50%. Because of the use of crude as opposed to purified peptides, the peptide concentrations reported in this study are approximate values and are not to be interpreted as absolute values.

The mixture of endogenous and SIS peptides was analyzed by LC-PRM using a Dionex UltiMate 3000 RSLCnano LC system (Thermo Fisher Scientific) coupled to a Q-Exactive mass spectrometer (Thermo Fisher Scientific). The peptides were injected (6 μL) onto a C18 trap column (300 μm i.d. $\times$ 5 mm packed with Acclaim PepMap 100, 5 μm , 100 $\text{\AA}$ C18; Thermo Fisher Scientific) at a loading pump flow rate of 5 $\mu\text{L}/\text{min}$, followed by separation on a 75 μm i.d. $\times$ 25 cm EASY-Spray analytical column packed with 2 μm Acclaim PepMap RSLC C18 (Thermo Fisher Scientific). Mobile phase A was 2% ACN/0.1% formic acid in water, and mobile phase B was 90% ACN/0.1% formic acid. The column was heated to 42

°C. Separations were performed at 500 nL/min across a 59 min linear gradient from 5–40% B.

An EASY-Spray source (Thermo Fisher Scientific) with zero dead volume nanoViper fittings was used with the Q-Exactive. The spray voltage was 1.8 kV and the capillary temperature was 250 °C. The mass spectrometer was operated in a targeted-MS² acquisition mode with a maximum IT of 100 ms, 1 microscan, 70 000 resolution, 1×10^5 AGC target, 2.0 *m/z* isolation window, and 28% normalized collision energy. Intrarun mass calibration was conducted using lock masses of 445.12003 *m/z* and 371.10123 *m/z*. A scheduled multiplexed PRM method was created to monitor the most abundant charge states of each of the 43 heavy SIS peptides, their cognate light forms, and 11 indexed retention time (iRT) standards (Biognosys; Zurich, Switzerland) to monitor RT drift across the runs. The PRM detection windows were 240 s.

Assay Characterization: Reverse Response Curves

The assays were characterized based on several metrics including the LOD, lowest analyte concentration at which the signal is distinguishable from the noise,³⁸ LLOQ, lowest concentration of the analyte at which quantitative measurements can be made, ULOQ, highest concentration of analyte above which the signal is not linear, linearity, carry-over, partial validation of specificity, intraday assay CV, interday assay CV, and total assay CV. SIS peptides were spiked into and serially diluted with a biological matrix consisting of N-linked glycosite-containing peptides enriched from commercially available human serum using an automated format of the SPEG method.^{36,37} As previously demonstrated, the average specificity of the automated hydrazide tip-based isolation of N-linked glycosite-containing peptides is 88.6%.³⁷ Reverse response curves, varying amount of a heavy peptide spiked into matrix with a constant amount of light peptide to construct a calibration curve, were generated for each peptide to determine the linear range of its corresponding assay. The biological matrix (serum) was the source of background analytes and the light (endogenous) peptides. Because unpurified synthetic peptides (~60% purity) were used as the spiked-in heavy isotope-labeled standards, the precise amount of each peptide standard was unknown (approximate amounts were used for calculation purposes), and the reported values for LOD, LLOQ, and ULOQ are not absolute values; however, the values are accurate across a dynamic range of at least 3 orders of magnitude based on the linearity of the response curves. Peak area ratios (heavy/light) were used as the dependent variables to generate the response curves.

The seven-point response curves (0.0576, 0.288, 1.44, 7.2, 36, 180, and 900 pmol on column) for each assay covered 4 orders of magnitude in abundance range, and they were run in triplicate in order of increasing concentration with 3 blank runs prior to the first replicate run of the curve and 2 blank runs following each curve. Using this run order scheme, the maximum carry-over was 9.14% with an average carry-over of 0.079% (Supplemental Table 2). Carryover was calculated by dividing the peak area of the analyte peptide in the blank after the highest concentration point on the response curve by the peak area of the highest concentration point on the response curve; the blanks after each high concentration point were averaged across the triplicate runs of the curve. To assess the

linearity, we fitted a simple linear regression model to the data using 3 middle calibration points other than the middle calibration point. The observation was considered to be linear if the average of the middle calibration point concentration was within 5% of that predicted from the best fit line passing through the other points. The slope of the calibration curve is representative of the analytical sensitivity of the method for the analyte; the steeper the slope, the more sensitive the assay or the stronger the mass spectrometer's response to a change in the concentration of the analyte.

Assay Reproducibility

Assay performance reproducibility (represented by the technical CV) was determined by measuring replicates of spiked serum samples in the same manner. Reproducibility was measured across 5 days at 3 levels, Lo, Med, and Hi, to approximate 2× LLOQ, 50× LLOQ, and 100× LLOQ, respectively. The order in which the samples were run was randomized to more accurately reflect the variability in assay performance. To determine the intra-assay variability, the CV for the triplicate analyses of each concentration level on each of the 5 days was calculated. The interassay variability was calculated at each concentration by determining the CV of each injection (first, second, and third) across the 5 days. Finally, the total assay CV was calculated based on the square root of the sum of the squares of the average intra-assay CV and the average interassay CV. Transitions with peak area ratio CVs > 20% were determined to be problematic because such a large variation is unexpected in the linear-response region in targeted MS assays.

Data and Statistical Analysis

The raw PRM data were processed using Skyline,³⁹ a vendor-neutral tool for targeted MS assay development and data collection that facilitates peptide and transition selection, collision energy optimization, method export, peak detection, and peak integration. All Skyline-processed data for the assays that passed the precision criteria of CV ≤ 20% are available at https://panoramaweb.org/labkey/project/CPTAC%20Assay%20Portal/JHU_DChan_HZhang_ZZhang/Serum_QExactive_GlycopeptideEnrichedPRM/begin.view?. Assay details, assay parameters, response curves, repeatability data, detailed standard operating protocols, and additional assay-specific resources can be located on the CPTAC assay portal <https://assays.cancer.gov/> using the search term “Johns Hopkins University.”

In Skyline, peaks were automatically integrated followed by manual inspection. Initially, the top three ranked transitions for each precursor were selected. If a transition was determined to be an interfering ion based on its lack of coelution with the other transitions and an inconsistent relative intensity compared to the other transitions, it was replaced by the next highest ranking transition. The assay characterization data exported from Skyline were processed using MRMPPlus, an in-house developed computational program to compute QC metrics for targeted MS-based assays (Supplemental Methods). Example MRMPPlus input files for the calculation of assay characterization parameters based on response curves are included in Supplemental Table 4.

Exported Skyline data from the PRM analysis of the noncancer, nonaggressive, and aggressive prostate cancer serum samples were analyzed by a Mann–Whitney U-test to determine the statistical significance of the peak area ratios. *p*-values <0.05 were considered significant.

Results and Discussion

Assay Development and Analytical Performance

We developed targeted MS assays using a Q-Exactive mass spectrometer to take advantage of the unique features of PRM including the lack of a requirement for *a priori* target transition selection and the ability to detect all product ions of the target peptide in one high-resolution mass spectrum. A schematic overview of our assay development and characterization workflow is presented in Supplemental Figure 1. Two discovery-phase proteomic studies of prostate cancer tissues conducted by our group resulted in the identification of 377 unique N-linked glycosite-containing peptides.^{34,35} Preliminary MRM assays using a microflow LC system and a QQQ mass spectrometer were developed to detect these peptides in serum. A total of 56 N-linked glycosite-containing peptides were detected in serum from a patient cohort consisting of 10 normal men (age 23 ± 1 years) and 28 men with prostate cancer (13 with PSA concentrations between 4 and 10 ng/mL and 15 with PSA concentrations between 50 and 100 ng/mL) (data not shown). The initial criteria for the detection of these peptides using the preliminary uncharacterized MRM assays were coelution of the endogenous light peptide with the heavy peptide and identical rank-order of the transitions between the light and heavy peptide. Upon transferring these 56 assays to a Q-Exactive mass spectrometer using a nanoflow LC system to enable the use of PRM and to perform full assay development and characterization, there was a 23% attrition rate due to poor/low peptide ionization efficiency and a high limit of detection. From our initial group of 56 assay targets, 43 assays were fully developed and characterized using PRM. Supplemental Table 1 contains the list of the 43 assays, including their transitions and PRM scheduled windows. The assays were developed as Tier 2 assays according to the framework for targeted MS-based assay “fit-for-purpose” validation established by CPTAC with input from the outside community.²¹ Tier 2 assays have a moderate-to-high degree of analytical validation, include labeled internal standards for every analyte, and they have high specificity, moderate-to-high precision (20–35% CV), and high repeatability. They are, however, validated to a degree that would allow their use for clinical decision making.

Targeted MS data require a significant amount of manual review, which is a time-intensive process that is subject to human error. Tools have been developed to facilitate targeted MS data analysis including the algorithm for Automated Detection of Inaccurate and Imprecise Transitions (AuDIT),⁴⁰ MSstats,⁴¹ and mProphet.⁴² Therefore, we developed and applied here an in-house tool, MRMPplus, to compute QC metrics for targeted MS-based assays that are specifically developed in accordance with the guidelines established by the NCI CPTAC Assay Development Working Group (<https://assays.cancer.gov/guidance-document/>).¹⁸

Figure 1 shows the extracted ion chromatograms (XICs), RT stability, response curve, and repeatability of a representative assay, Clusterin (82EDALNETR₈₉), deamidated at Asn86 as a consequence of the SPEG method. The red and blue traces in Figure 1A demonstrate the

coelution of the light and heavy forms of the peptide, and the insets show the XICs of the transitions that appear in the same order of relative abundance for the light (upper inset) and heavy peptide (lower inset). The average retention times of both peptides for all 7 points of the response curve across the triplicate runs of the response curve were consistent with an average RT CV of 2.0% for the light peptide and 2.4% for the heavy peptide (Figure 1B). The lack of detection of the heavy peptide in one of the three replicate injections of the lowest concentration point on the response curve (57.6 fmol) resulted in a relatively high CV of the retention time at this concentration (4.6% vs 2.2% for the light peptide). The response curve for the assay was linear with an LLOQ of 57.6 fmol (Figure 1C). Demonstrating the high repeatability of the assay, the Total Assay CVs at the Low, Medium, and High concentration levels were 4.84%, 3.27%, and 4.08%, respectively (Figure 1D).

An abridged version of the 43 assay characterization parameters calculated by MRMPPlus is shown in Supplemental Table 2. Because of the low blank sample background noise level of the Orbitrap mass analyzer of the Q-Exactive mass spectrometer, the calculation of LOD resulted in some assays having LODs in the attomole range, which is below the lowest concentration point (57.6 fmol). A caveat to consider when evaluating the reported LOD, LLOQ, and upper limit of quantification (ULOQ) values for these assays is that they cannot be accurately determined given that the precise concentrations of each crude peptide product is unknown, but they are typically within a factor of 2–3 of the nominal value.⁴³ However, the main purpose of these Tier 2 assays is not to provide actual concentrations at the peptide level but rather to precisely and consistently measure relative changes in the levels of large numbers of targeted analytes across samples.²¹

MRMPPlus was unable to calculate values for all of the transitions for peptides AFNSTLPTMAQMEK and SYNVTSVLFR (Supplemental Table 2, highlighted in red) based on the response curve assay data; hence, the LOD and LLOQ values for these peptides are high compared to the values for the other 41 peptides. It should be noted that even though the response curve analytical performance of peptide AFNSTLPTMAQMEK was not robust, its repeatability measures of Intra-Assay CV, Inter-Assay CV, and Total Assay CV < 20% were acceptable. As an indication of the specificity of the assays, the transitions from each endogenous (light) peptide coeluted with the transitions from its cognate stable isotope-labeled peptide. Given that these were PRM-based assays, all of the potential fragment ions for each peptide were recorded. The 3 fragment ions that were selected to quantify each peptide had the best overall performance even though 6 of the 123 transitions from the 41 peptides with acceptable repeatability data had peak area ratios that deviated >30% from the mean. These 6 transitions had less than ideal specificity that was possibly caused by ion suppression or isobars of the target analyte.

The results from the reproducibility assessment of the 43 assays are presented in Supplemental Table 3. All assays listed in Supplemental Table 3 except Lysosome-associated membrane glycoprotein 2 LNSSTIK and Afamin FNETTEK passed the criteria of Intra-, Inter-, or Total Assay CV < 20% at all of the three concentration levels (the mass spectrometer response of these two peptides was comparatively low). Among the other 41 assays, only AAPAPQEATATFNSTADR had a Total Assay CV > 20% (Total assay CV = 20.33% at the medium concentration level).

Assay Deployment Using Biological Samples

To test the performance of our fully characterized targeted MS-based assays in detecting differences in the relative abundance of N-linked glycosite-containing peptides in biological samples, the PRM assays were deployed for the analysis of serum samples from prostate cancer patients with different disease states. These assays represent 37 glycoproteins.

All 41 assays that had Intra-, Inter-, and Total Assay CVs < 20% were deployed using serum from males without prostate cancer with elevated serum PSA levels but with negative prostate biopsies (Neg.; $n = 25$), serum from males with low PSA and with nonaggressive prostate cancer (NAG; $n = 25$), and serum from males with high PSA and with aggressive prostate cancer (AG; $n = 25$). The clinical characteristics of these patients are shown in Table 1.

The patient-derived serum samples were processed via SPEG in a manner identical to the commercial human serum used for the assay characterization experiments to enrich N-linked glycosite-containing peptides. Prior to PRM analysis, each sample was spiked with a mixture of the 41 heavy-isotope labeled peptide standards to enable peak area comparison. The inclusion of stable isotope-labeled peptides in each sample permitted the evaluation of instrument-related issues that could affect analyte detection.

Among the 41 assay targets, 37 were detected with no significant differences between the NAG group and the AG group. There were 4 N-linked glycosite-containing peptides with significantly higher levels ($p < 0.05$) in the serum from the NAG vs AG patient groups: AFNSTLPTMAQMEK (CD44 antigen; $p = 0.040$); EEQFNSTFR (Immunoglobulin γ -2 heavy chain; $p = 0.041$); GAFISNFSMTVDGK (Interalpha-trypsin inhibitor heavy chain H2, ITIH2; $p = 0.010$); and INNTHALVSLQLNLK (Cadherin-13; $p = 0.016$) (Figure 2). Representative extracted ion chromatograms of these peptides are included in Supplemental Figure 2. Although the fold-changes of the peak area ratios of these 4 peptides were all <1.5-fold, the fold-changes were significant (Supplemental Table 5). None of the 41 N-linked glycosite-containing peptides had relative abundance levels that were significantly different between the prostate cancer (AG, NAG) and noncancer (Neg.) groups.

Although these results are from a discovery phase study and a small cohort and therefore require further verification and validation using a larger number of independent specimens, the serum levels of these proteins have been previously implicated in prostate cancer biology. CD44 is a receptor for hyaluronic acid (HA) that mediates cell–cell and cell–matrix interactions through its affinity for HA, and CD44+ cells have been linked to the epithelial-mesenchymal transition in prostate cancer metastasis. Interestingly, immunoglobulin γ expression has been shown to be positively correlated with the Gleason score and histological grade in prostate tissues; however, the results from our data indicate that a glycosite-containing peptide of this protein (Asn176) was relatively more highly expressed in the serum from patients diagnosed with NAG prostate cancer (Gleason score 6) vs those diagnosed with AG prostate cancer (Gleason score 7). A variant of interalpha-trypsin inhibitor heavy chain, ITIH4, was found to be significantly elevated in the urine of prostate cancer patients compared to those with benign prostatic hyperplasia, and previous studies found that fragmentation patterns of human serum ITIH4 are associated with prostate cancer

and could hold important diagnostic information. Down-regulation of Cadherin-13 has been associated with prostate cancer, and Cadherin-13 re-expression in most cancer cell lines inhibits cell proliferation and invasiveness, increases susceptibility to apoptosis, and reduces tumor growth in *in vivo* models.

Conclusions

PRM is becoming widely used for quantitative MS-based assays because of its high-quality quantitative measurements and simplified method development. In this study, PRM assays were developed for the quantification of serum-derived N-linked glycosite-containing peptides based on relative peak area ratios. The assays were characterized according to the requirements for Tier 2 research use assays, and the corresponding response curve data, repeatability data, as well as standard operating protocols have been deposited and are publicly available on the CPTAC Assay Portal (<https://assays.cancer.gov>) to facilitate the dissemination and adoption of these assays. The availability of the assay performance data will enable other researchers to evaluate the assay performance prior to deploying the assays in their own laboratories.

In total, 41 assays were developed to measure N-linked glycosite-containing peptides that were identified from discovery-phase proteomic studies conducted by our group using tissue from prostate cancer patients.^{34,35} We then deployed the assays to determine the relative levels of N-linked glycosite-containing peptides in serum from patients with negative, nonaggressive, and aggressive prostate cancer biopsies. The relative levels of four N-linked glycosite-containing peptides were significantly different between the nonaggressive and aggressive prostate cancer groups, and the relative abundance of 37 formerly N-linked glycopeptides was unchanged between the groups.

There is an expectation for recent advances in MS technology and the associated analytical approaches to drive improvements in healthcare that are reliant upon the development of sensitive and specific biomarkers.⁴⁹ The assays we developed in this study can be used to evaluate the relative level of N-linked glycosite-containing peptides in human serum regardless of the disease state of the patients from which the serum is obtained. The data generated from the application of these Tier 2 targeted MS assays could be used to develop clinical bioanalytical or diagnostic laboratory tests that comply with Food and Drug Administration (FDA) and Clinical Laboratory Improvement Amendments (CLIA) guidance toward the development and characterization of Tier 1 assays using purified peptides that provide accurate, precise, clinically actionable information for medical practitioners or that inform decision-making in the development of drugs for human use.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported in part by the National Institutes of Health under grants and contracts from the National Cancer Institute Clinical Proteomics Tumor Analysis Consortium (CPTAC, Grant U24CA160036 to H.Z. and D.W.C.) and the Early Detection Research Network (EDRN, Grant U01CA152813 to H.Z. and Grant

U24CA115102 to D.W.C.); the National Heart, Lung and Blood Institute Programs of Excellence in Glycosciences (Grant P01HL107153 to H.Z.) and NHLBI Proteomic Center (Grant N01-HV-00240 to H.Z.); and from the Swiss National Science Foundation (Grant No. 3100A0-107679 to R.A.). The authors would like to thank the CPTAC Assay Development Working Group for guidance regarding assay design and data interpretation.

References

1. Boligan KF, Mesa C, Fernandez LE, von Gunten S. *Cell Mol Life Sci.* 2015; 72:1231. [PubMed: 25487607]
2. Kaprio T, Satomaa T, Heiskanen A, Hokke CH, Deelder AM, Mustonen H, Hagstrom J, Carpen O, Saarinen J, Haglund C. *Mol Cell Proteomics.* 2015; 14:277. [PubMed: 25452313]
3. Lee J, Kim C, Yang H, Park I, Oh N, Hua S, Jeong H, An H, Kim SC, Lee GM, Koh GY, Kim HM. *Mol Cancer Ther.* 2015; 14:470. [PubMed: 25534360]
4. Lee LY, Thaysen-Andersen M, Baker MS, Packer NH, Hancock WS, Fanayan SJ. *Proteome Res.* 2014; 13:4783.
5. Liang Y, Ma T, Thakur A, Yu H, Gao L, Shi P, Li X, Ren H, Jia L, Zhang S, Li Z, Chen M. *Glycobiology.* 2015; 25:331. [PubMed: 25347993]
6. Desiderio DM, Kai M. *Biomed Mass Spectrom.* 1983; 10:471. [PubMed: 6616023]
7. Pan S, Zhang H, Rush J, Eng J, Zhang N, Patterson D, Comb MJ, Aebersold R. *Mol Cell Proteomics.* 2005; 4:182–190. [PubMed: 15637048]
8. Domanski D, Murphy LC, Borchers CH. *Anal Chem.* 2010; 82:5610. [PubMed: 20524616]
9. Kirkpatrick DS, Hathaway NA, Hanna J, Elsasser S, Rush J, Finley D, King RW, Gygi SP. *Nat Cell Biol.* 2006; 8:700. [PubMed: 16799550]
10. Lange V, Picotti P, Domon B, Aebersold R. *Mol Syst Biol.* 2008; 4:222. [PubMed: 18854821]
11. Liebler DC, Zimmerman LJ. *Biochemistry.* 2013; 52:3797. [PubMed: 23517332]
12. Pan S, Aebersold R, Chen R, Rush J, Goodlett DR, McIntosh MW, Zhang J, Brentnall TA. *J Proteome Res.* 2009; 8:787. [PubMed: 19105742]
13. Picotti P, Aebersold R. *Nat Methods.* 2012; 9:555. [PubMed: 22669653]
14. Peterson AC, Russell JD, Bailey DJ, Westphall MS, Coon JJ. *Mol Cell Proteomics.* 2012; 11:1475. [PubMed: 22865924]
15. Gallien S, Duriez E, Crone C, Kellmann M, Moehring T, Domon B. *Mol Cell Proteomics.* 2012; 11:1709. [PubMed: 22962056]
16. Gallien S, Bourmaud A, Kim SY, Domon BJ. *Proteomics.* 2014; 100:147. [PubMed: 24200835]
17. Rodriguez H, Rivers R, Kinsinger C, Mesri M, Hiltke T, Rahbar A, Boja E. *Proteomics: Clin Appl.* 2010; 4:904. [PubMed: 21137031]
18. Whiteaker JR, Halusa GN, Hoofnagle AN, Sharma V, MacLean B, Yan P, Wrobel JA, Kennedy J, Mani DR, Zimmerman LJ, Meyer MR, Mesri M, Rodriguez H, Paulovich AG. *Nat Methods.* 2014; 11:703–704. [PubMed: 24972168]
19. Addona TA, Abbatiello SE, Schilling B, Skates SJ, Mani DR, Bunk DM, Spiegelman CH, Zimmerman LJ, Ham AJ, Keshishian H, Hall SC, Allen S, Blackman RK, Borchers CH, Buck C, Cardasis HL, Cusack MP, Dodder NG, Gibson BW, Held JM, Hiltke T, Jackson A, Johansen EB, Kinsinger CR, Li J, Mesri M, Neubert TA, Niles RK, Pulsipher TC, Ransohoff D, Rodriguez H, Rudnick PA, Smith D, Tabb DL, Tegeler TJ, Variyath AM, Vega-Montoto LJ, Wahlander A, Waldemarson S, Wang M, Whiteaker JR, Zhao L, Anderson NL, Fisher SJ, Liebler DC, Paulovich AG, Regnier FE, Tempst P, Carr SA. *Nat Biotechnol.* 2009; 27:633. [PubMed: 19561596]
20. Kennedy JJ, Abbatiello SE, Kim K, Yan P, Whiteaker JR, Lin C, Kim JS, Zhang Y, Wang X, Ivey RG, Zhao L, Min H, Lee Y, Yu MH, Yang EG, Lee C, Wang P, Rodriguez H, Kim Y, Carr SA, Paulovich AG. *Nat Methods.* 2014; 11:149–155. [PubMed: 24317253]
21. Carr SA, Abbatiello SE, Ackermann BL, Borchers C, Domon B, Deutsch EW, Grant RP, Hoofnagle AN, Huttenhain R, Koomen JM, Liebler DC, Liu T, Mac Lean B, Mani DR, Mansfield E, Neubert H, Paulovich AG, Reiter L, Vitek O, Aebersold R, Anderson L, Bethem R, Blonder J, Boja E, Botelho J, Boyne M, Bradshaw RA, Burlingame AL, Chan D, Keshishian H, Kuhn E, Kinsinger C, Lee JS, Lee SW, Moritz R, Oses-Prieto J, Rifai N, Ritchie J, Rodriguez H, Srinivas

- PR, Townsend RR, Van Eyk J, Whiteley G, Wiita A, Weintraub S. *Mol Cell Proteomics*. 2014; 13:907. [PubMed: 24443746]
22. Dube DH, Bertozzi CR. *Nat Rev Drug Discovery*. 2005; 4:477. [PubMed: 15931257]
23. Kim YS, Hwang SY, Oh S, Sohn H, Kang HY, Lee JH, Cho EW, Kim JY, Yoo JS, Kim NS, Kim CH, Miyoshi E, Taniguchi N, Ko JH. *Proteomics*. 2004; 4:3353. [PubMed: 15529413]
24. Drake PM, Schilling B, Niles RK, Prakobphol A, Li B, Jung K, Cho W, Braten M, Inerowicz HD, Williams K, Albertolle M, Held JM, Iacovides D, Sorensen DJ, Griffith OL, Johansen E, Zawadzka AM, Cusack MP, Allen S, Gormley M, Hall SC, Witkowska HE, Gray JW, Regnier F, Gibson BW, Fisher SJ. *J Proteome Res*. 2012; 11:2508. [PubMed: 22309216]
25. Drake PM, Schilling B, Niles RK, Braten M, Johansen E, Liu H, Lerch M, Sorensen DJ, Li B, Allen S, Hall SC, Witkowska HE, Regnier FE, Gibson BW, Fisher SJ. *Anal Biochem*. 2011; 408:71. [PubMed: 20705048]
26. Chambers AG, Percy AJ, Simon R, Borchers CH. *Expert Rev Proteomics*. 2014; 11:137. [PubMed: 24476379]
27. Garbis SD, Townsend PA. *Expert Rev Proteomics*. 2013; 10:337. [PubMed: 23992417]
28. Huijbers A, Velstra B, Dekker TJ, Mesker WE, van der Burgt YE, Mertens BJ, Deelder AM, Tollenaar RA. *Int J Mol Sci*. 2010; 11:4175. [PubMed: 21151433]
29. Pin E, Fredolini C, Petricoin EF 3rd. *Clin Biochem*. 2013; 46:524. [PubMed: 23266295]
30. Meany DL, Chan DW. *Clin Proteomics*. 2011; 8:7. [PubMed: 21906357]
31. Zhang H, Chan DW. *Cancer Epidemiol, Biomarkers Prev*. 2007; 16:1915. [PubMed: 17932335]
32. Siegel R, Ma J, Zou Z, Jemal A. *Ca-Cancer J Clin*. 2014; 64:9. [PubMed: 24399786]
33. Andriole GL, Crawford ED, Grubb RL 3rd, Buys SS, Chia D, Church TR, Fouad MN, Gelmann EP, Kvale PA, Reding DJ, Weissfeld JL, Yokochi LA, O'Brien B, Clapp JD, Rathmell JM, Riley TL, Hayes RB, Kramer BS, Izmirlan G, Miller AB, Pinsky PF, Prorok PC, Gohagan JK, Berg CD. *N Engl J Med*. 2009; 360:1310. [PubMed: 19297565]
34. Chen J, Xi J, Tian Y, Bova GS, Zhang H. *Proteomics*. 2013; 13:2268. [PubMed: 23716368]
35. Tian Y, Bova GS, Zhang H. *Anal Chem*. 2011; 83:7013. [PubMed: 21780747]
36. Zhang H, Li XJ, Martin DB, Aebersold R. *Nat Biotechnol*. 2003; 21:660. [PubMed: 12754519]
37. Chen J, Shah P, Zhang H. *Anal Chem*. 2013; 85:10670. [PubMed: 24079330]
38. Lavagnini I, Magno F. *Mass Spectrom Rev*. 2007; 26:1. [PubMed: 16788893]
39. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, Kern R, Tabb DL, Liebler DC, MacCoss MJ. *Bioinformatics*. 2010; 26:966. [PubMed: 20147306]
40. Abbatiello SE, Mani DR, Keshishian H, Carr SA. *Clin Chem*. 2010; 56:291. [PubMed: 20022980]
41. Choi M, Chang CY, Clough T, Broudy D, Killeen T, MacLean B, Vitek O. *Bioinformatics*. 2014; 30:2524. [PubMed: 24794931]
42. Reiter L, Rinner O, Picotti P, Huttenhain R, Beck M, Brusniak MY, Hengartner MO, Aebersold R. *Nat Methods*. 2011; 8:430. [PubMed: 21423193]
43. Verbeken M, Wynendaele E, Lefebvre RA, Goossens E, Spiegeleer BD. *Anal Biochem*. 2012; 421:547. [PubMed: 22033292]
44. Shang Z, Cai Q, Zhang M, Zhu S, Ma Y, Sun L, Jiang N, Tian J, Niu X, Chen J, Sun Y, Niu Y. *Oncotarget*. 2015; 6:1202–1216. [PubMed: 25483103]
45. Liu Y, Chen Z, Niu N, Chang Q, Deng R, Korteweg C, Gu J. *Prostate*. 2012; 72:690. [PubMed: 22430367]
46. Jayapalan JJ, Ng KL, Shuib AS, Razack AH, Hashim OH. *Electrophoresis*. 2013; 34:1663. [PubMed: 23417432]
47. Song J, Patel M, Rosenzweig CN, Chan-Li Y, Sokoll LJ, Fung ET, Choi-Miura NH, Goggins M, Chan DW, Zhang Z. *Clin Chem*. 2006; 52:1045. [PubMed: 16574760]
48. Andreeva AV, Kutuzov MA. *Genes, Chromosomes Cancer*. 2010; 49:775. [PubMed: 20607704]
49. Paik YK, Hancock WS. *Nat Biotechnol*. 2012; 30:1065. [PubMed: 23138303]

Clusterin (aa 82-89); Dea-Asn86

Figure 1.

Representative LC-PRM assay characterization data for clusterin glycopeptide (aa 82–89; deamidated Asn86). (A) XICs of endogenous light peptide (red) and stable isotope-labeled heavy peptide (blue) indicating coelution. The insets show the XICs of the y6, y5, and y4 transitions for the light (top) and heavy (bottom) peptides. (B) Retention times of the light and heavy peptides across triplicate runs of the response curve. (C) Response curve. The data from the summed transitions and the individual transitions are shown. The low region of the response curve (0.0576–1.44 pmol) is enlarged in the inset. (D) Repeatability data demonstrating reproducible heavy/light peak area ratios of the clusterin glycopeptide analyzed at three levels (Lo, Med, and Hi) in triplicate across 5 days. $n = 3$ for all experiments. Data shown in parts B–D are the means $\pm$ s.d.

Figure 2.

Statistical analysis of formerly N-linked glycosite-containing peptides with different levels of relative abundance in serum from patients with AG ($n = 25$) and NAG ($n = 25$). Box-and-whisker plots represent the minimum, 1st quartile, median, 3rd quartile, and maximum values. (A) Interalpha-trypsin inhibitor heavy chain H2 peptide (GAFISNFSMTVDGK). (B) CD44 antigen peptide (AFNSTLPTMAQMEK). (C) Immunoglobulin gamma 2 heavy chain peptide (EEQFNSTFR). (D) Cadherin-13 peptide (INNTHALVSLQLNLNK). p -values were determined from a Mann–Whitney U-test.

Table 1
Summary of Clinical Information for Patient Serum Specimens^a

group	<i>n</i>	age	PSA level
non-cancer	25	67.6 ± 6.6	N/A
biopsy negative (Neg.; RP Gleason score N/A)			
prostate cancer	25	58.6 ± 5.2	5.3 (3.4–7.2)
nonaggressive (NAG; RP Gleason score ≤ 6)			
prostate cancer	25	59.0 ± 7.7	11.5 (7.5–22.6)
aggressive (AG; RP Gleason score ≥ 7)			

^aN/A = not applicable. Age is mean ± s.d. PSA level (ng/mL) is median (inter-quartile range).