Glycoproteomics assays for prostate cancer biomarker discovery

Wang, W.
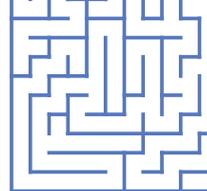
Chapter 2

# In-depth Glycoproteomic Assay of Urinary Prostatic Acid Phosphatase

Wei Wang[1], Carmen R. de Nier[1], Manfred Wuhrer[1], Guinevere S.M. Lageveen-Kammeijer[1,2]*

1 Leiden University Medical Center, Center for Proteomics and Metabolomics, Leiden, 2300 RC, the Netherlands

2 University of Groningen, Groningen Research Institute of Pharmacy, Groningen, 9713 AV, The Netherlands
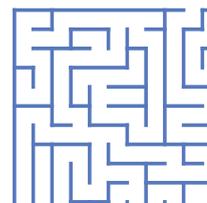
Abstract: Prostate-specific antigen (PSA) is a well-known clinical biomarker in prostate cancer (PCa) diagnosis, but a better test is still needed as the serum level-based PSA quantification exhibits limited specificity and comes with poor predictive value. Prior to PSA, prostatic acid phosphatase (PAP) was used but it was replaced by PSA as it improved the early detection of PCa. Revisiting PAP, and then in specific its glycosylation thereof, it appears to be a promising new biomarker candidate. Namely, previous studies have indicated that PAP glycoproteoform differ between PCa versus non-PCa individuals. However, an in-depth characterization of PAP glycosylation is still lacking. In this study, we established an in-depth glycoproteomic assay for urinary PAP by characterizing both micro- and macro-heterogeneity of the PAP glycoprofile. For this purpose PAP samples were analyzed by capillary electrophoresis coupled to mass spectrometry after affinity purification from urine and proteolytic digestion. The developed urinary PAP assay was applied on a pooled DRE (digital rectal examination) urine sample from nine individuals. Three glycosylation sites were characterized, namely N_{94}, N_{220} and N_{333}, via *N*-glycopeptide analysis. Taking sialic acid linkage-isomers into account, a total of 63, 27 and 4 *N*-glycan structures were identified, respectively. The presented PAP glycoproteomic assay will enable the determination of potential glycomic biomarkers for early detection and prognosis of PCa in cohort studies.

Keywords: Diagnostic marker, Prognostic marker; Glycosylation; CE-MS; Prostate Cancer; Prostatic Acid Phosphatase (PAP); Urinary Assay.

2.1 Introduction

Prostate cancer (PCa) is the second most common cancer in men. In 2020, approximately 1.4 million men were diagnosed with PCa and around 375,000 men died of it worldwide (Data source: GLOBOCAN 2020).[212] In the majority of the early stage PCa cases, no symptoms are observed.[41] Early detection is essential to reduce the mortality of PCa and to improve the quality of patients' life as well as treatability.[213] As a biomarker for the early detection of PCa, the serum concentration of prostate-specific antigen (PSA) is clinically used, when elevated concentrations are found (>3 ng/mL (the Netherlands)), follow-up examination is advised, such as digital rectal examination (DRE), ultrasonography and invasive prostate biopsy.[214] Unfortunately, the PSA test is of limited value and tends to result in overdiagnosis and overtreatment of patients,[215] as its exhibits low specificity and poor predictive value to discriminate low-risk PCa from high-risk PCa as well as its inability to differentiate between PCa and other (benign) prostate related diseases.[216-219] Therefore, there is an urgent need for a more advanced non-invasive biomarker for the early detection of PCa.

Prior to the discovery of PSA, prostatic acid phosphatase (PAP) was used as a marker for the detection of PCa but was found to exhibit nonspecific expression in multiple organs.[220] Similar to PSA, PAP is a glycoprotein mainly synthesized in prostate epithelial cells. Interestingly, its glycosylation -rather than its enzyme activity and concentration in serum- has been correlated to PCa in several studies pointing out the potential of PAP glycosylation as a diagnostic biomarker of PCa. PAP is present in two different forms; cellular PAP (cPAP) and secreted PAP (sPAP). The two forms of PAP differ in glycosylation and have different isoelectric point (pI) values which could be related to their sialylation levels.[221] The sPAP normally forms a dimer and shows phosphatase activity.[222] A series of Asn (N) to Gln (Q) point mutations of PAP proves that N-linked glycosylation is of critical importance for PAP stability ($N_{333}Q$ mutant) and its catalytic activity ($N_{94}Q$ single mutant, $N_{220}Q$ single mutant and $N_{94}Q$ and $N_{220}Q$ double mutant).[223] It was observed that mutation of the N-linked glycosylation site N_{333} that closest to the His_{12} in active site of PAP had a greater effect on protein activity and stability than mutation of sites N_{94} and N_{220} which are further away.

Various N-glycan structures on PAP derived from human seminal plasma were first elucidated by ^1H Nuclear Magnetic Resonance spectroscopy in 1987.[207] In this study, one oligomannosidic N-glycan and six complex-type N-glycans were reported. The complex-type N-glycans were fucosylated and partially sialylated, of which four were diantennary structures and two were triantennary structures. Later Jakob *et al.*, studied the crystal structure of seminal PAP and revealed that actually two N-glycosylation sites (N_{94} and N_{333}) are occupied with oligomannosidic N-glycans and only site N_{220} is occupied with complex-type N-glycans.[206] In 1997, Yoshida *et al.,* reported altered N-linked sugar chains of PAP during oncogenesis in the prostate tissue by means of lectin affinity chromatography. In this study,

low levels of oligomannosidic *N*-glycans and hybrid-type *N*-glycans with core fucose were observed in PCa comparing to benign prostatic hyperplasia (BPH) as well as high levels of hybrid-type *N*-glycans without core fucose.[204] With the rapid development of mass spectrometry (MS), a more in-depth analysis of glycosylation became available. White *et al.*, studied the glycomic profile of PAP derived from seminal plasma and reported 21 *N*-glycan structures (*N*-glycan release) by matrix-assisted laser desorption-ionization (MALDI) coupled to time-of-flight mass spectrometry (TOF-MS) in 2009.[203] To identify the site occupancy, glycopeptide analysis was performed, with similar findings as found in previous studies, except for *N*-glycosylation site N_{94}, which appeared to be occupied with complex type *N*-glycans rather than oligomannosidic ones reported by Jakob et. al.[206] Based on released *N*-glycan analysis of PAP, lower levels of fucosylated di- and tetraantennary as well as oligomannosidic *N*-glycans were found while an increase was found of smaller monoantennary *N*-glycans for the pool of PCa patients (n=92) as compared to normal control (n=65) and BPH (n=59) pool.[203] Another MS study in 2013, demonstrated similar glycosylation features for *N*-glycosylation sites N_{94} and N_{220} (occupied with complex type *N*-glycans) on PAP derived from DRE urine, but also an increase of bisected structures. Interestingly, in this study *N*-glycosylation site N_{333} was not only occupied with oligomannosidic *N*-glycans but also with complex-type *N*-glycans.[202] Additionally, by using sialic acid- and fucose-binding lectins, a significant decrease in captured PAP was observed in the DRE urine pool of aggressive PCa (n=10) compared to indolent PCa (n=10). The comparison of indolent PCa to non-PCa (n=10) pooled DRE urine samples revealed a significant decrease in captured PAP.[202] A more recent study by Sugár *et. al.,* in 2021, analyzed cancerous (n=49) and healthy (n=46) prostate tissues from biopsies and characterized the differences in their glycosylation using an untargeted bottom-up approach and revealed an increase in fucosylation and a decrease in sialylation for *N*-glycosylation site N_{94} of PAP.[224]

While these MS studies explore PAP glycosylation to a certain extent, limitations were encountered regarding the coverage of site N_{220} as well as the coverage of minor glycoforms and, for the glycoproteomic studies missed proteolytic cleavages for all three glycosylation sites resulted in complex data sets.[202, 203] Moreover, no distinction could be made between isomeric sialylated species, despite indications of profoundly different roles of those isomers in cancer in general: α2,6-linked isomers have been implicated in blocking galectin binding and enhancing tumor cell survival while α2,3-linked isomers are considered as hallmark of malignant types of cancers. [208-211] Therefore, to gain a further insight into PAP glycosylation and its biomarker potential in relation to PCa, an in-depth characterization of the PAP *N*-glycome is of critical importance.

In this study, we report on an in-depth urinary PAP glycoproteomic assay (uPGA). The assay was successfully applied on a pooled urine sample collected from nine males after DRE. The developed method is a stepping stone to define the biomarker potential of PAP glycosylation for patient stratification in relation to

PCa.

2.2 Materials and methods

2.2.1 Chemicals. See Supporting Information, S-1.1.

2.2.2 Samples. Urine samples were collected from healthy female volunteers at the Leiden University Medical Center. One female urine pool (FUP) was formed by pooling all collected urine samples together. Before pooling, the urine samples were centrifuged for 5 min at 2500 x g at 5 °C and the pH was measured. The obtained FUP was aliquoted and stored at −20 °C. Prior to use, the FUP was thawed to room temperature (RT), followed by a centrifugation step (5 min at 500 x g) to fasten the precipitation of urinary sediments. The DRE urine pool was prepared at Roche (Penzberg, Germany). DRE urine samples were collected according to the appropriate standard operating procedures approved by local ethic committees (sample information see **Supporting Information, Table S-1**). Samples were stored at <-70°C until analysis. Repeated freeze–thawing was avoided. First, all nine individual samples were diluted by a factor of 1,000 or 10,000 to fit in with the measurement range of the PAP concentration assay (developed by Roche). Then one mL from each sample was used to form the urine pool (**Supporting Information, Table S-1**).

2.2.3 Anti-PAP beads. Two biotinylated anti-PAP antibodies, anti-PAP I (1.39 mg/mL) and anti-PAP II (1.08 mg/mL), were provided by Roche. Antibodies are Roche inhouse generated mouse monoclonal antibodies against PAP antigen using hybridoma technology [225]. The two different antibodies were coupled to high-capacity streptavidin agarose resins (HCS beads) separately. The HCS beads were pre-washed four times with 1x phosphate-buffered saline (PBS). Then, 500 μL of drained beads was added to 5 mg of anti-PAP antibodies. Coupling of the antibody with beads was achieved by incubating overnight (ON) on a roller shaker (Roller 10 digital, IKA Laboratory Technology, Staufen, Germany) at 4 °C. The coupled beads were washed twice with 1x PBS to remove unbound antibodies. To remove unspecific bindings, the beads were incubated for 3 min with 100 mM formic acid (FA), followed by a wash with 1x PBS to adjust pH to 7. The supernatant was removed after centrifugation (2 min at 100 x g). A 50% bead suspension was made by resuspending the beads in 1x PBS with 0.02% NaN$_3$ (*v/v*). Prior to usage, the coupled anti-PAP beads were stored at 4 °C. The coupling efficiency was determined by comparing the same volume prior (approximately 9 μg of antibodies) and after coupling the antibodies by SDS-PAGE analysis (**Supporting Information, S-1.2**).

2.2.4 PAP capture. To develop the PAP capture protocol, 250 μL of 5x PBS spiked with 1.5 μg of PAP standard, was added to one mL of FUP. Male urine was considered unsuitable to use during the development of the PAP capturing procedure as it contains PAP. In contrast, FUP contains no PAP and the matrix is the closest to that of male urine samples, hence it was chosen to use for protocol development. By spiking a PAP standard to the FUP we were able to mimic male

urine samples. Different amounts and density of beads suspension (2 μL, 4 μL and 10 μL of 25% beads suspension and 2 μL and 5 μL of 50% beads suspension) were evaluated for an optimal affinity purification procedure as well as a series of FA concentration (5 mM, 10 mM, 25 mM, 50 mM, 100 mM and 150 mM) for the elution buffer. All experiments were performed in duplicates. In all cases, the yield was evaluated by SDS-PAGE.

The optimized protocol uses 2 μL of 50% anti-PAP I beads suspension which was added to a mixture containing one mL of FUP and 250 μL of 5x PBS with 1.5 μg of PAP standard or one mL of diluted, pooled DRE urine (~3.2 μg/mL) and 250 μL of 5x PBS. Incubation was performed ON at 4 °C by constantly rotating the tubes horizontally. Afterwards, the beads were washed with 1x PBS and transferred to 96-well polypropylene filter plate with a 10 μm PE frit (Orochem, Naperville, IL). By using a vacuum manifold (Merck Millipore, Darmstadt, Germany), the beads were sequentially washed once with 600 μL of 1x PBS and twice with 600 μL of 50 mM ammonium bicarbonate. For elution, 200 μL of 150 mM FA was added followed by 5 min on a plate shaker (1000 rpm, Heidolph Tiramax 100 platform shaker, Heidolph, Schwabach, Germany). The eluate was collected in a low protein binding flat bottom plate (polypropylene plate, Greiner Bio-One, 120 Kremsmünster, Austria) by centrifugation (2 min, 100 x g). Samples were then dried by vacuum concentration at 45 °C for 2 h.

2.2.5 Tryptic digestion. Various conditions and enzymes were explored for the most optimal digestion of PAP, different reduction and alkylation strategies were compared. Briefly, three different enzymes, with different properties were explored (TPCK treated trypsin from Sigma-Aldrich (Steinheim, Germany), trypsin gold, trypsin platinum and sequence grade modified (SGM) trypsin from Promega (Madison, WI)) as well as different enzyme:protein ratios (1:5, 1:10, 1:20 and 1:50). For all digestion experiments 200 ng PAP standard was used. In the case of dried samples, reconstitution was performed in 5 μL of 25 mM ammonium bicarbonate.

The final procedure used the following conditions: reduction was performed by adding 1 μL of 12 mM DL-dithiothreitol (DTT) to the samples, followed by a heating step at 60 °C for 30 min. Afterwards, the samples were alkylated by adding 1 μL of 48 mM iodoacetamide (IAA) and left in the dark for 30 min at RT. Next, 1 μL of 48 mM DTT was added and samples were incubated in light for 20 min at RT. Then samples were proteolytically digested with SGM trypsin by adding 1 μL of 200 ng/μL SGM trypsin ON. Digestion was quenched by adding 1 μL of 5% FA to the sample.

2.2.6 CE-MS/MS. The identification of PAP glycopeptides was performed on a bare-fused silica capillary (i.d. 30 μm, o.d. 91 cm) using a CESI 8000 system (SCIEX, Brea, CA) that was coupled to a UHR-QqTOF-MS (Impact; Bruker Daltonics, Bremen, Germany). Prior to injection, 4 μL of tryptic PAP glycopeptides (20 ng/μL) was mixed with 2 μL leading electrolyte (LE, 1200 mM ammonium acetate at pH 3.2). The capillary voltage was set at 20 kV and the capillary temperature at

15 °C. Injection was executed by applying 25 psi pressure for 24 s, corresponding to 13.5% of the capillary volume (87 nL). All experiments were performed in positive ionization mode and intensities were enhanced by the usage of DEN-gas (nanoBooster, Bruker Daltonics) with acetonitrile as a dopant at 0.2 bar.[226] The temperature and flow rate of the drying gas was set at 150 °C and 1.2 L/min, respectively. Fragmentation data was acquired on the three most abundant precursor ions between *m/z* 150 and 2500 with a 1 Hz spectral acquisition frequency and a minimum intensity of 3757. Depending on the *m/z* values, the precursor ions were isolated with a width of 8–15 Th. The collision energies were set as a linear curve in a *m/z* dependent manner, ranging from 20 eV at *m/z* 500 to 70 eV at *m/z* 2000 for all charge states (1–5), applying a basic stepping mode with collision energies of 100% (80% of the time) to 70% (20% of the time). A mass exclusion list was applied during fragmentation to avoid the influence of co-captured proteins (**Supporting Information, Table S-2**).

LC-MS/MS. See Supporting Information, S-1.3 and S-1.4.

2.2.7 Assay assessment. The limit of detection (LOD) of the established assay was determined by spiking the PAP standard (0 ng, 20 ng, 50 ng, 100 ng and 200 ng) into one mL FUP, followed by executing the complete protocol including PAP capturing, tryptic digestion, and CE-MS analysis. The inter- and intraday validation of the assay was assessed by spiking 200 ng PAP standard to one mL FUP over three separate days, with four replicates per day, and executing the complete protocol.

2.2.8 Data analysis. Both LC-MS(/MS) (**Supporting Information, S-1.3**) and CE-MS(/MS) data were analyzed with Compass DataAnalysis 5.0 SR1 (x64) (Build 203.2.3586, Bruker Daltonics). With DataAnalysis all MS and tandem MS data were manually screened for glycopeptides based on exact *m/z* values, migration/elution order and relative abundance of the analytes. Extracted ion
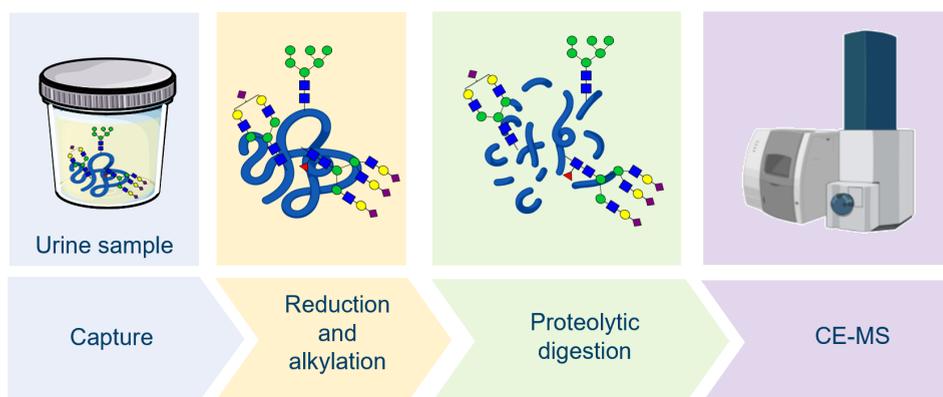

Figure 1. In-depth urinary PAP glycoproteomic workflow. PAP was enriched from DRE urine by immunoaffinity capturing overnight. After capturing, PAP was proteolytic digested and directly subjected to CE-MS measurement. Three glycosylation sites were identified and characterized in an in-depth manner.

electropherograms (EIEs)/extracted ion chromatographs (EICs) were acquired with the first three isotopes of the singly, doubly, triply and quadruply charged analytes ranging from 200 Da to 2500 Da using a width of ± m/z 0.02 Da. Tandem MS data acquired from LC-MS (Orbitrap) (**Supporting Information, S-1.4**) were analyzed with Xcalibur (Thermo, 2.2 SP1.48) for structural elucidation of glycopeptides. Besides, Byonic (v4.6.1, Protein Metrics Inc.) searching was performed using a homo sapience protein database acquired from Uniprot.

2.3 Results and discussion

A urinary PAP glycoproteomics assay (uPGA) was established (**Figure 1**). To begin with, PAP was captured from one mL of urine, followed by proteolytic digestion and direct measurement by CE-MS. The established uPGA resolves all three glycosylation sites on DRE urinary PAP and reveals multiple glycoforms per site including the distinction of sialic acid linkage-specific isomers on glycopeptide level. Signature glycosylation features are identified (e.g. sialylation, fucosylation, branching), characterized and their relative abundance being determined in an in-depth manner (rel. abundance < 0.1%). (**Figure 2, Supporting Information, Figure S-1, Table S-3** and **S-4**).

2.3.1 PAP capturing. For the optimization of the capture procedure, several parameters were tested. First of all, a successful coupling of the two anti-PAP antibodies was assessed by SDS-PAGE (**Supporting Information, Figure S-2**). Followed by determining the optimal elution buffer. For this purpose, different FA concentrations (5 mM, 10 mM, 25 mM, 50 mM, 100 mM and 150 mM) were investigated by SDS-PAGE (N=2; **Supporting Information, Figure S-3**). A higher concentration of FA resulted in a higher yield for the anti-PAP I beads with 150 mM FA being the most optimal elution buffer. As for anti-PAP II beads, with the exception of 5 mM FA, the yields with the different FA concentrations were comparable. Thus, 150 mM FA and 50 mM FA was chosen as elution buffer for anti-PAP I beads and anti-PAP II beads, respectively. Next, the bead volume was evaluated in combination with different quantities of anti-PAP beads (2 µL, 4 µL and 10 µL of 25% beads suspension and 2 µL and 5 µL of 50% beads suspension, N=2) (**Supporting Information, Figure S-4**). The highest capturing efficiency with the lowest variation was found with 2 µL of 50% beads suspension for both anti-PAP I and anti-PAP II. Capture efficiencies were determined to be 75% ± 1% and 73% ± 2% for anti-PAP I and anti-PAP II, respectively. To investigate whether any capturing bias existed for the two different antibodies in regard to glycosylation profile the glycomic profiles were compared before and after capturing by CE-MS (N=3; **Supporting Information, Figure S-5**). No significant differences were observed between the two anti-PAP beads. Based on yields, anti-PAP I was selected for further experiments.

2.3.2 Proteolytic cleavage. Initially, 1.5 µg PAP standard was used for the digestion with TPCK treated trypsin. Subsequent experiments, to optimize the digestion, started with a decreased amount of PAP standard (200 ng). A total
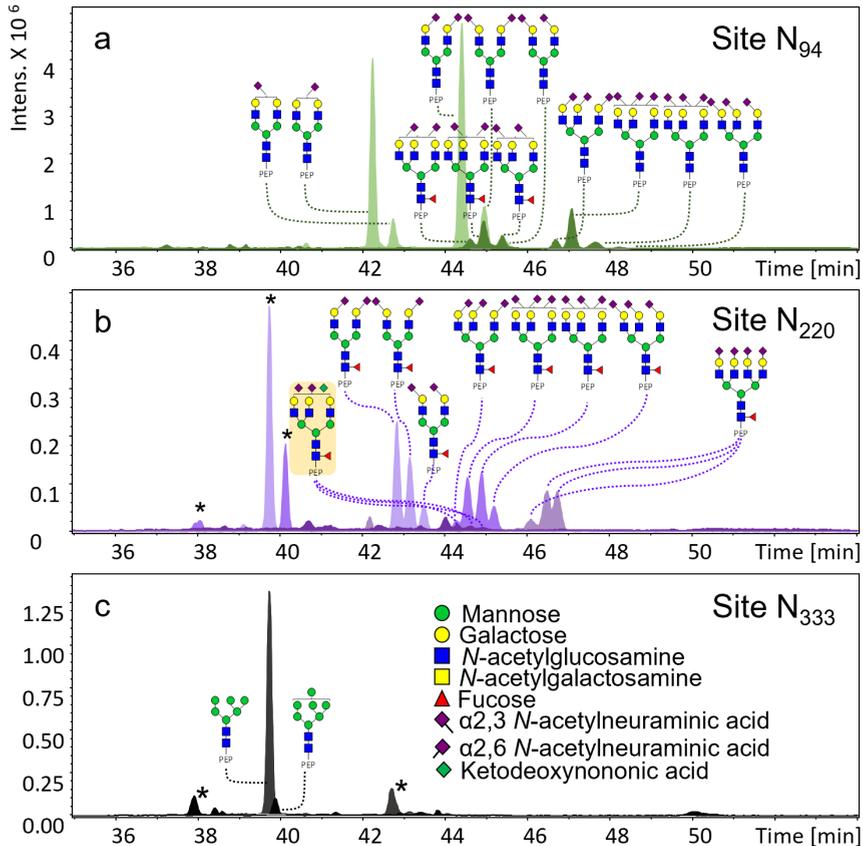
Figure 2. Extracted ion electrograms (EIEs) of the most abundant glycans on DRE urinary PAP on (a) site N_{94}, (b) site N_{220} and (c) site N_{333}. Glycan structure with yellow background color in figure (b) highlights the Kdn-containing glycan identified on PAP. Three isomers were observed of the Kdn-containing glycan. PEP: peptide backbone. Asterisk indicates non-glycopeptide analytes. Kdn: ketodeoxynononic acid. Peptide backbone of site N_{94} is $FLN_{94}ESYK$. Peptide backbone of site N_{220} is $VYDPLYCESVHN_{220}FTLPSWATEDTMTK$. Peptide backbone of site N_{333} is $N_{333}ETQHEPYPLMLPGCSPSCPLER$.

of four different peptide backbones were found for *N*-glycosylation site N_{94}, including no, one or two missed cleavages. Two peptide variants were found for sites N_{220} and N_{333} with no or one missed cleavage (**Supporting Information, Table S-5**). The observed missed cleavages on peptide backbones are probably due to multiple cleavage sites which are in close position to each other, such as, LGEYIRKRYRKFLN_{94}ESY̲KHE (cleavage sites are underlined). Another cause might be the proximity of the cleavage sites close to the *N*-glycosylation site, meaning the protease could be hampered by the large glycan moiety.[203] Similarly, the study by White *et al.,* analyzed PAP glycosylation after digestion with trypsin or chymotrypsin and both proteases resulted in missed cleavages.[203]

As the high variation in peptide backbones increases data complexity, different

types of trypsin and different enzyme:protein ratios were investigated to potentially improve the digestion efficiency and specificity. For this purpose, three different trypsin preparations were selected with higher specificity and stability compared to the conventional TPCK treated trypsin. Whereas, SGM trypsin has been developed to be highly active, with a high specificity and stability compared to TPCK trypsin, trypsin gold is designed to have a maximum digestion specificity in combination with a high resistance for autolytic cleavage. The final selected enzyme, trypsin platinum, is free of any detectable nonspecific proteolytic activity and has maximal autoproteolytic resistance, besides its high digestion efficiency. For all four types of trypsin, various ratios of the enzyme:protein were investigated (1:5, 1:10, 1:20 and 1:50). N-glycosylation site N_{220} was not detected using TPCK trypsin which was therefore excluded for further experiments. Among all different ratios, 1:5 presented the most number of PAP glycopeptides independently of the selected enzyme (**Supporting Information, Table S-6**). A more in-depth comparison was performed on this ratio (**Supporting Information, Figure S-6**). For N-glycosylation site N_{94}, four peptide backbones were observed for all enzymes (**Supporting Information, Table S-6**). The most abundant peptide backbone was found to be $FLN_{94}ESYKHEQVYIR$ (one missed cleavage) with a relative abundance of 63%, 71% and 60% for SGM trypsin, trypsin gold and trypsin platinum, respectively. The highest analyte area was observed with trypsin platinum, followed by trypsin gold and SGM trypsin, indicating trypsin platinum has highest digestion efficiency. While SGM trypsin showed the smallest variation (average relative standard deviation (RSD) 8.0%) indicating higher robustness and higher repeatability, followed by trypsin platinum showed average RSD 9% and the largest variation was found for trypsin gold with average RSD 21%. Relative abundancies obtained from all three types of trypsin were comparable with average RSD of 8%, 6% and 8% for SGM trypsin, trypsin gold and trypsin platinum, respectively. For N-glycosylation site N_{220}, SGM trypsin outperformed the two other enzymes with a subsequently higher area for most analytes and small variation between replicates. Regarding the analytes area, the average RSDs are 13%, 34% and 41% for SGM trypsin, trypsin gold and trypsin platinum, respectively. However, among the different trypsin types a visible difference in relative abundancies was found. Comparing with SGM trypsin, the glycomic profile acquired from trypsin platinum presented a higher abundancy for diantennary N-glycans, while lower abundancies were found for tri- and tetraantennary N-glycans. This indicated that trypsin platinum and trypsin gold were likely hampered in their access to N-glycosylation site N_{220}, especially when occupied with larger N-glycans, such as tri- and tetraantennary N-glycans. Regarding their relative area, average RSDs of 7%, 15% and 27 % for SGM trypsin, trypsin gold and trypsin platinum were found, respectively. For N-glycosylation site N_{333}, the highest number of N-glycans and highest area were found using trypsin platinum, although SGM trypsin showed smallest variation with average RSD of 3% in analytes area. Average RSDs of trypsin gold and trypsin

platinum are 36% and 20% respectively, which is roughly ten times higher than SGM trypsin implying the digestion efficiency of trypsin gold and platinum is of poor repeatability. For the two *N*-glycans detected by all types of trypsin, the relative abundances were comparable.

Considering the coverage for all three *N*-glycosylation sites, SGM trypsin was considered for further optimization, especially as SGM trypsin could potentially provide more information in case of low concentrated PAP samples. Namely, *N*-glycans attached to *N*-glycosylation site N_{220} have lower recovery than the other two glycosylation sites and, more importantly, *N*-glycans occupying this site possess important *N*-glycomic features, such as, fucosylation, sialylation and a higher level of branching which could be interesting for cancer biomarker discovery studies (**Supporting Information, Figure S-7**).[57, 227-229] Furthermore, PAP has three disulfide bonds. To facilitate digestion, the disulfide bonds are generally reduced with a reducing reagent. Subsequently, an alkylation step should be accomplished to prevent the broken disulfide bonds from re-forming. This process can be accomplished with different strategies. The most commonly used reducing reagent is DTT and the alkylation is usually performed by IAA. To optimize the digestion of PAP, the method of DTT as a thiol reducing agent together with IAA for alkylation was examined. Next to this, TCEP was also assessed for PAP reduction in combination with chloroacetamide (CAA) for alkylation. However, TCEP treated samples did not show better digestion and an extra sample clean-up was needed prior MS analysis. Upon the reduction and alkylation, RapiGest SF Surfactant was also evaluated for PAP digestion as it is commonly used to enhance enzymatic digestion of proteins. It is a mild denaturant that helps to unfold proteins and expose the proteolytic sites the enzymatic cleavage sites. However, in the case of PAP digestion, no significant improvements in digestion efficiency or specificity were observed (*data not shown*).

2.3.3 Assay assessment and data analysis. To explore the glycosylation of PAP, tryptic PAP was measured both with LC-MS and CE-MS to find the most suitable analytical platform. The data acquired from LC-MS showed no separation on sialic acid linkage-specific isomers for sites N_{94} and N_{220} and was achieved by CE-MS. A high level of sialylation had been observed on PAP *N*-glycans in our study (**Supporting Information, Figure S-7**) and differentiation of sialylation isomers was considered a critical requirement as this important glycosylation feature is a known hallmark of cancer and differently linked sialic acid may play different roles in different types of cancer.[230-234] Thus, CE-MS was chosen for its excellent glycoform separation power.

To test the limit of detection (LOD) of the developed assay, different amounts of PAP standard were spiked, individually, to one mL FUP (0, 20, 50, 100 and 200 ng) (**Supporting Information, Figure S-8**). In total, 59, 29, 14 and no *N*-glycopeptides with isomer distinction were observed with 200, 100, 50 and 20 ng spiked PAP, respectively. The result showed that the lower LOD of the established assay is
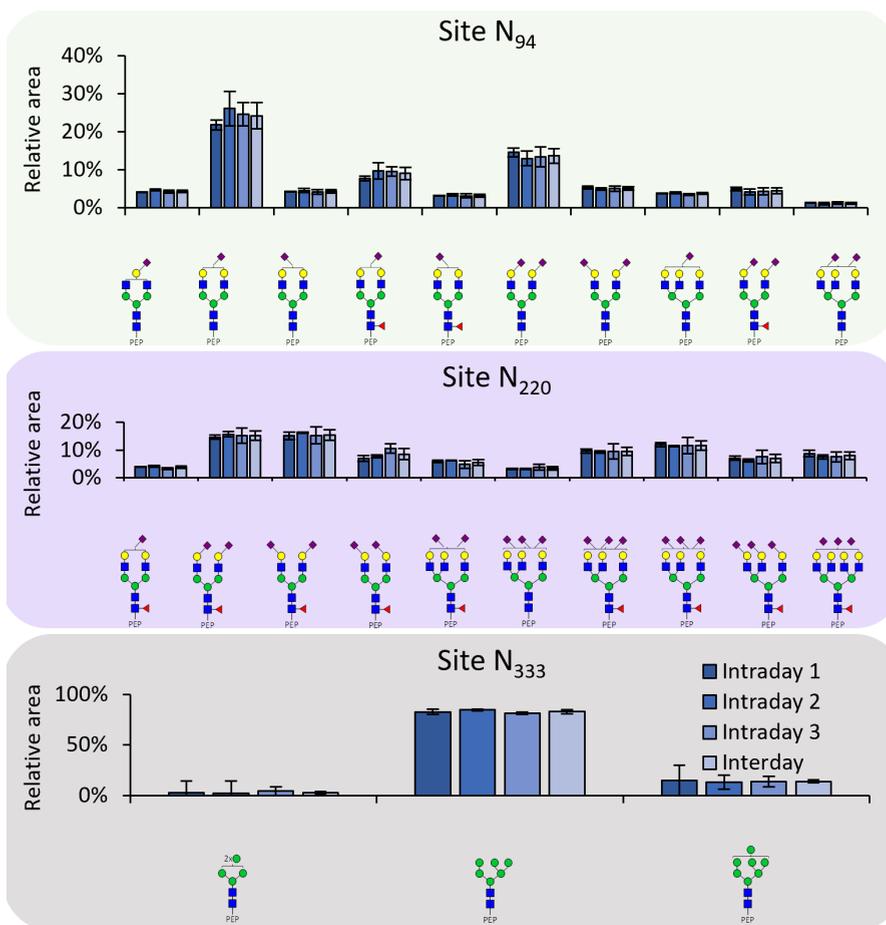
between 20 and 50 ng of PAP. To further evaluate the robustness and repeatability of the assay, an intra- and inter-day validation was performed (n = 4 and n = 12, respectively). Briefly, 200 ng of PAP standard was spiked to one mL FUP, followed by an immunoaffinity capture of PAP, tryptic digestion and measurement by CE-MS. Relative abundance of PAP glycopeptides were comparable for all three N-glycosylation sites (**Figure 3, Supporting Information, Figure S-9**). For site N_{94}, average relative standard deviations (RSD) of the top 10 analytes were 5%, 12% and 14% for day one, two and three, respectively. For site N_{220}, average RSD of the top 10 analytes were 8%, 6% and 25% on day one, two and three respectively. For site N_{333}, average RSD were 10%, 7% and 3% on day one, two and three respectively.

To demonstrate the applicability of the established assay and to investigate the N-glycome of PAP derived from urine, the assay was applied on a pooled DRE urine sample (**Figure 2, Supporting Information, Table S-3** and **Figure S-1**). The glycosylation features of urinary PAP was described above. In addition, the Byonic search (v4.6.1, Protein Metrics Inc.) of captured DRE urinary PAP by LC-MS/MS (Orbitrap) showed that PAP was the main protein captured, confirming the success of the urinary PAP assay. Other observed proteins were co-captured from urine, including hemoglobin subunit OS, semenogelin-1, albumin, semenogelin-2, etc. On the co-captured proteins, no glycans were observed based on Byonic search (*data not shown*).

2.3.4 PAP glycosylation. A PAP standard derived from seminal plasma was used for assay development. PAP glycoprofiles of all three sites were characterized by MS and compared to literature reports (**Supporting Information, Table S-3** and **Figure S-10**). Consistent with White *et al.,*[203] we found that N-glycosylation sites N_{94} and N_{220} were occupied with complex glycans while site N_{333} contained oligomannosidic glycans. A total of 17 and 13 unique N-glycan compositions were identified in seminal PAP for N-glycosylation sites N_{94} and N_{220}, respectively. These were all complex-type di-, tri- and tetraantennary N-glycans, with and without core fucosylation, with high levels of sialylation. As for site N_{333}, four oligomannosidic N-glycans were found ranging from Man 5 to Man 8. Interestingly, ketodeoxynononic acid (Kdn) on human PAP was observed similar to our first discovery of Kdn-containing N-glycans on PSA (**Figure 4** and **Supporting Information, Figure S-11**). [235] One Kdn-containing glycan was identified on site N_{220} and assigned to a glycan composition H6N5F1S2K1 on the basis of MS/MS data (observed *m/z* value is 1518.356 [4+], theoretical *m/z* value is 1518.364 [4+]). Three isomers were observed at relative abundance of 0.8%, 1.2% and 0.8%. Having found Kdn on PSA [235] and now on PAP, we expect that more human glycoproteins may carry this modification.

In this study, we characterized the glycosylation of DRE urinary PAP for the first time using the established uPGA (**Figure 2, Supporting Information, Figure S-1** and **S-7**). Similar to seminal plasma PAP, urinary PAP glycoprofile has a high micro-heterogeneity and macro-heterogeneity. Three N-glycosylation sites were

Figure 3. Intra-and interday validation of the developed prostatic acid phosphatase (PAP) Glycoproteomics Assay showing relative abundance of the most abundant 10 glycopeptides per glycosylation sites. Seminal plasma PAP standard was spiked into a female urine pool (FUP), an in-solution digestion was performed and was used to test the repeatability (intraday, N = 4) and intermediate precision (interday, N = 12) of the PAP Glycoproteomics Analysis by CE-MS. Note: for site N_{333}, the average relative standard deviation (AVG RSD) is showing the top three analytes.

AVG RSD (top 10 analytes)	Intraday 1	Intraday 2	Intraday 3	Interday
Site N_{94}	5.3%	12.1%	14.2%	12.1%
Site N_{220}	8.1%	5.5%	25.0%	16.5%
Site N_{333}	9.9%	6.7%	3.4%	15.5%

detected with complex *N*-glycans attached on glycosylation sites N_{94} and N_{220}, while site N_{333} carried oligomannosidic *N*-glycans. A total of 28 and 12 unique *N*-glycan compositions were identified in urinary PAP for *N*-glycosylation site N_{94} and $_{220}$, respectively, comprising only complex di-, tri- or tetraantennary *N*-glycans, with and without core fucosylation and with overall high levels of
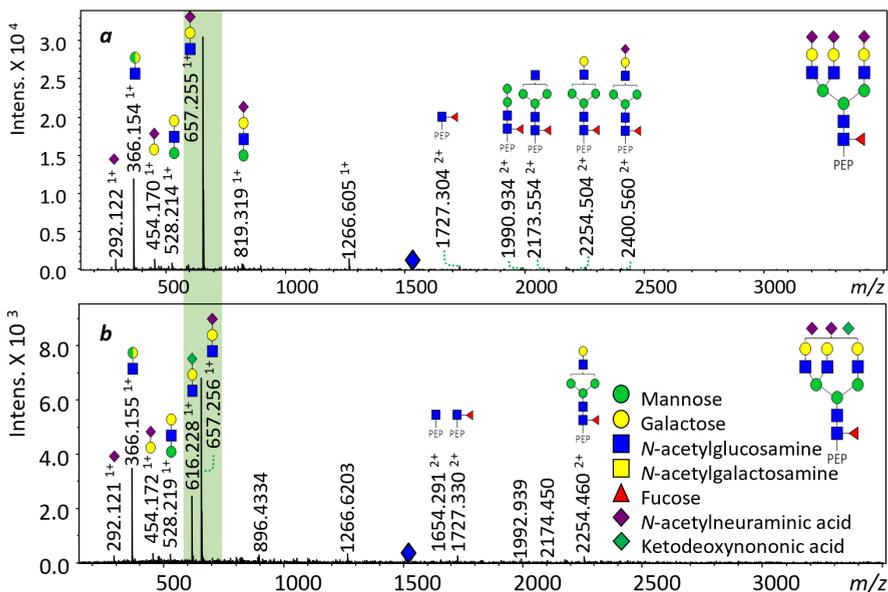
Figure 4. The identification of Kdn-containing *N*-glycans in seminal plasma PAP glycopeptides via tandem MS using CE-MS. (a, b) Tandem MS spectra of glycopeptides H6N5F1S3 and H6N5F1S2K1 on glycosylation site N_{220} with peptide backbone VYDPLYCESVHNFTLPSWATEDTMTK. H: hexose. N: *N*-acetylglycosamine. F: fucose. S: *N*-acetylneuraminic acid (Neu5Ac). K: ketodeoxynononic acid (Kdn). Pep: peptide backbone. Green background highlighted the significant ions H1N1K1 for Kdn identification. Blue diamonds indicate for parent ions (quadruply charged) of the glycopeptides. The assignment of glycan structures is based on tandem MS spectra.

sialylation. Based on analyte area, 99% of glycans on site N_{94} were sialylated, among them 34% were mono-sialylated, 52% were di-sialylated, and 13% were tri-sialylated. The fucosylation level on site N_{94} was 38%. An overall high amount of branching was observed on site N_{94} with 60% diantennary, 29% triantennary and 8% tetraantennary *N*-glycans. Site N_{220} showed glycosylation features similar to site N_{94} with high levels of sialylation, fucosylation and high branching. It fucosylation level is more than two times higher than site N_{94}. All *N*-glycans on site N_{220} were found to be sialylated with 20% being tetra-sialylated. Fucosylation (89%) and branching (36% diantennary, 36% triantennary, 28% tetraantennary *N*-glycans) were found to be higher than for site N_{94}. As for site N_{333}, four unique *N*-glycan compositions were found with only oligomannosidic type *N*-glycans were found, varying from Man 5 till Man 8. This is in discrepancy with a previous study where *N*-glycosylation site N_{333} was not only shown to be occupied with oligomannosidic type *N*-glycans but also with complex type *N*-glycans.[202] This discrepancy could be related to PCa, as study by Nyalwidhe *et al.,* investigated PAP from DRE urine pools of aggressive PCa. In our study, more information in regard to sialic acid linkages was acquired by CE-MS via isomer separation [177].

Sialic acid linkage-specific isomers were separated in the CE capillary based on their size and charge. In which, α 2,6-sialylated glycans/glycopeptides migrates earlier than its α 2,3-sialylated variant. This migration behavior might be related to the slight difference in their pK_a values (Δ 3.4·10^{-2}) [156]. This resulted in the identification of 63, 27 and 4 unique *N*-glycan structures on site N_{94}, N_{220} and N_{333}, respectively (**Supporting Information, Table S-4 and Figure S-12**). Our analyses provide considerably more depth and coverage than the report by both by White *et al.,* and Nyalwidhe *et al.,*. Namely, the latter study analyzed a PAP tryptic digest and observed only three, two and five *N*-glycan structures for sites N_{94}, N_{220} and N_{333}, respectively. White *et al.,* also analyzed a PAP tryptic digest and observed only three and two *N*-glycan structures for sites N_{94} and N_{333}, respectively. No information was acquired for site N_{220}. Their result of chymotrypsin digestion showed three, two and two glycoforms detected on sites N_{94}, N_{220} and N_{333}, respectively. [203] Next to glycopeptide analysis, the glycosylation of PAP was further examined by enzymatically releasing the *N*-glycans on three different seminal plasma pools (PCa, BPH and normal seminal). They reported 21 *N*-glycan structures, and similar to our results, this included two oligomannosidic *N*-glycans and 19 complex type of *N*-glycans. The identified complex glycans were di-, tri-and tetraantennary *N*-glycans which were highly sialylated (mono-, di-, tri- and tetrasialylation) and core-fucosylated.[203] Besides, bisecting structures were observed for *N*-glycosylation site N_{94} and site N_{220} by Nyalwidhe *et al.,* (glycopeptide level).[202] Tandem MS data was shown on glycan H6N6F1S3 attaching to peptide backbone FLNESYK. However, the significant ions of bisecting structure (H1N3 as well as H1N3F1 being attached to the peptide backbone) were not present. The same glycan was also detected in our study, however, our tandem MS data cannot confirm whether this glycan is a hybrid or complex type *N*-glycan. Interestingly, in our study the Kdn-containing glycan (H6N5F1S2K1) identified on site N_{220} in seminal plasma PAP was also detected in urinary PAP. Similar to seminal PAP, three isomers were observed all at low relative abundance (0.9%, 0.9% and 0.6%) (**Supporting Information, Figure S-11**).

Relatively minor glycosylation differences were observed between seminal plasma PAP (commercial PAP standard) and PAP derived from DRE urine (**Supporting Information, Figure S-1**). The most abundant *N*-glycans on site N_{94} of PAP were H5N4S2 for DRE urinary PAP followed by H5N4S1 and H5N4F1S2 , the monosialylated nonfucosylated variant was most abundant (H5N4S1) for seminal plasma followed by H5N4S2 and H5N4F1S1. In general it seems that the monosialylated structures were more abundant in the seminal pool and a higher sialylation for the urinary pool; trisialylation (**Supporting Information, Figure S-1 and S-7**). Whether the differences observed are due to different biofluids type (seminal plasma versus urine) or the health conditions of the urine sample donors remains unresolved. Paired analysis of biofluids from the same individuals would allow to address this.

During the study, minor questions were prompted and deserve consideration in upcoming investigations. For example, the general signal intensity of PAP dropped 2-3 fold after two weeks storage at -20 °C (*data not shown*), particularly affecting the coverage of site N_{220}, of which the glycopeptide signals already tend to be low after fresh sample preparation. It is unclear whether the signal drop is due to instability of PAP glycopeptides and degradation during storage or rather due to aspecific binding and adsorption effects. As a consequence, PAP digests should be measured immediately or stored at -80 °C for short time after processing.

Apart from the altered glycosylation, PAP is regaining attention as cPAP also plays important role in prostate carcinogenesis and progression,[221] as cPAP can inhibit the growth of androgen-independent prostate cells. As demonstrated by and increased prostate epithelial cell proliferation and subsequently developed invasive adenocarcinomas in PAP knockout mice.[236] A study using xenograft animal models further indicate the potential of cPAP as a tumor suppressor in PCa.[221] Gunia *et al.,* also showed a significant inverse correlation between the expression of cytoplasmic PAP and histopathologic staging of incidental PCa.[237] Furthermore, PAP is successfully applied as a therapeutic target for the treatment of patients with castration-resistant prostate cancer (CRPCa).[238, 239] The autologous cellular immunotherapy named Sipuleucel-T which uses PAP as a target was approved by the US Food and Drug Administration (FDA) for the treatment of patients with CRPCa.[238, 239] As described previously, cPAP and sPAP differ in glycosylation[221] and there is a lacking of study on characterizing the glycosylation of cPAP. Therefore, it is of interest to investigate the glycosylation features of cPAP and explore their biological roles.

2.4 Conclusions

Several studies have implicated the diagnostic biomarker potential of an altered PAP glycosylation in relation to PCa. However, these studies lacked the ability to perform an in-depth characterization, identifying low abundant species. In this study, we established an uPGA that allowed to distinguish α2,3 from α2,6 linked sialylation and its potential was demonstrated on pooled samples, revealing a LOD between 20-50 ng/mL and a high glycosylation complexity of sites N94 and N220. Overall, we provide an important stepping stone to further verify the previous discoveries and evaluate the clinical and diagnostic potential of PAP glycosylation features on an individual level in large sample cohorts.

Supporting Information

The Supporting Information is available via https://pubs.acs.org/doi/10.1021/acsmeasuresciau.3c00055.