

Three-Dimensionally Functionalized Reverse Phase Glycoprotein Array for Cancer Biomarker Discovery and Validation

Li Pan,[†] Hillary Andaluz Aguilar,[‡] Linna Wang,[§] Anton Iliuk,^{§,||} and W. Andy Tao^{*,†,‡,§,||,⊥}

[†]Department of Medicinal Chemistry & Molecular Pharmacology, [‡]Department of Chemistry, and [§]Department of Biochemistry, Purdue University, West Lafayette, Indiana 47907, United States

^{||}Tymora Analytical Operations, West Lafayette, Indiana 47906, United States

[⊥]Center for Cancer Research, Purdue University, West Lafayette, Indiana 47907, United States

S Supporting Information

ABSTRACT: Glycoproteins have vast structural diversity that plays an important role in many biological processes and have great potential as disease biomarkers. Here, we report a novel functionalized reverse phase protein array (RPPA), termed polymer-based reverse phase glycoprotein array (polyGPA), to capture and profile glycoproteomes specifically, and validate glycoproteins. Nitrocellulose membrane functionalized with globular hydroxyaminodendrimers was used to covalently capture preoxidized glycans on glycoproteins from complex protein samples such as biofluids. The captured glycoproteins were subsequently detected using the same validated antibodies as in RPPA. We demonstrated the outstanding specificity, sensitivity, and quantitative capabilities of polyGPA by capturing and detecting purified as well as endogenous α -1-acid glycoprotein (AGP) in human plasma. We further applied quantitative N-glycoproteomics and the strategy to validate a panel of glycoproteins identified as potential biomarkers for bladder cancer by analyzing urine glycoproteins from bladder cancer patients or matched healthy individuals.

Protein glycosylation is the most structurally complex form of post-translational modifications (PTMs) and the resulting enormous structural diversity of glycoproteins plays an essential role in a wide variety of biological processes.^{1,2} Consequently, aberrant glycosylation has been found closely associated with key pathological steps of cancer development and progression.^{3–5} Glycoproteins have great potential as cancer biomarkers. In fact, more than half of the FDA-approved cancer biomarkers used in clinical practice are glycoproteins.⁴ However, only a handful of them are actually detecting protein glycosylation,^{6,7} due to the lack of specific and sensitive methods for quantifying specific glycoproteins. Despite the challenges, profiling of human glycoproteomes in search of potential cancer biomarkers has remained a primary interest of biomedical research.

Reverse phase protein array (RPPA) has emerged as a powerful high-throughput platform to profile protein expressions in a large number of biological samples simultaneously.^{8–11} With only a minimal amount of starting material at sub-nL volume, RPPA is an ideal platform for clinical samples such as needle biopsies, biological fluids, and microdissected tissues.^{8–10} RPPA requires only one validated detection

antibody per target compared to sandwich immunoassays, which need antibody pairs.^{12–16} Multiple studies such as the Human Protein Atlas project have validated over 1000 detection antibodies suitable for array analysis.^{17–19} A wide range of applications have described the use of RPPA for clinical sample analysis, including leukemia, breast cancer, prostate cancer studies, and many others.^{20–23}

Detection of functional proteins, in particular protein PTMs by RPPA, however, has been extremely limited.^{8,10,24} This primarily is due to low availability and quality of PTM-specific antibodies compared to those for total proteins, and an overwhelming number of PTM sites present with most not having any commercial antibodies. This problem is even more evident in glycoproteomics, where glycosylation-specific antibodies are virtually nonexistent.⁷ Moreover, the high complexity of thousands of proteins in an unenriched mixture makes it extremely challenging to detect specific glycoproteins. Other forward microarray formats, such as glycan or lectin arrays, have been previously used for profiling of glycosylation.^{25–27} However, glycan or lectin arrays typically are not applicable in RPPA format due to relatively weak interactions and that they cannot work under the denaturing conditions.^{28,29}

In the past decades, mass spectrometry (MS) has been the driving force in profiling glycoproteomes for biomarker research.^{30,31} Prior to MS analysis, it is often necessary to enrich either glycoproteins or glycopeptides through lectin affinity chromatography (LAC),^{32–34} borate chelating,^{35,36} or hydrazide chemistry-based capture.^{37,38} Because it requires a fair amount of sample, multiple steps for sample preparation, and the commitment of a high performance instrument, MS-based glycoproteomics is typically used for in-depth profiling of glycoproteins during the discovery stage.

Here, we introduce a novel platform, termed polymer-based reverse phase glycoprotein array (polyGPA), to capture and profile glycoproteomes. In the first step, the surface of the globular molecule, polyamidoamine generation 4 (PAMAM G4) dendrimer, was derivatized with multiple hydroxyamine groups (Supporting Information). Next, nitrocellulose membrane was coated with the synthesized hydroxyaminodendrimers to generate three-dimensionally functionalized RPPA, polyGPA, to enable the capture of glycoproteome through the

Received: September 30, 2016

Published: November 14, 2016

formation of covalent oxime bonds between hydroxyamine and aldehyde groups generated after oxidation of glycoproteins (Figure 1).^{37,39} The chemical conjugation step has few side

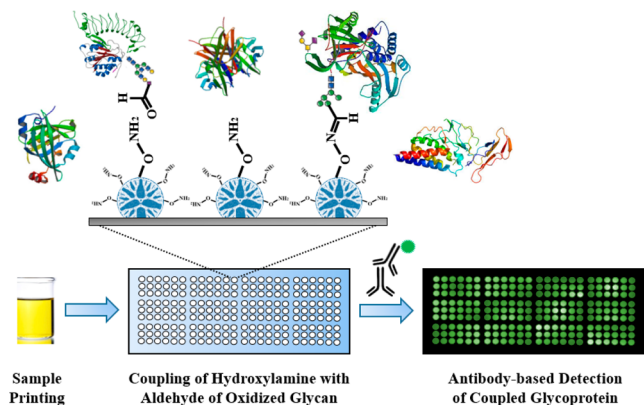


Figure 1. Design and experimental workflow for polyGPA.

reactions, occurs immediately during sample printing, and can tolerate denaturing conditions. The captured denatured glycoproteins can then be detected using the same validated antibodies with superior specificity as those used in RPPA. Any changes in polyGPA signal can then be attributed to changes in overall glycosylation of the target when normalized to protein level.

Similar to RPPA, polyGPA requires a minimal amount of starting material, making it ideal for analyzing clinical samples. Moreover, polyGPA overcomes the limitation of RPPA in detecting proteins with PTMs. Currently, RPPA cannot analyze glycoproteins due to the lack of glycosite-specific antibodies.⁷ By capturing glycoproteins particularly, polyGPA not only eliminates the need of glycosite-specific antibodies for detection but also significantly reduces the complexity of the crude sample to facilitate the detection of low abundance glycoproteins using validated protein antibodies.

To evaluate initially the specificity, sensitivity, and quantitative capabilities of polyGPA, we tested the method using the α -1-acid glycoprotein (AGP) standard. The sub-nL volume of the AGP standard in oxidized or unoxidized form in different concentrations ranging from 0.2 to 20 pg/nL was printed in quadruplicate onto the polyGPA membrane, and the AGP signal was detected using anti-AGP primary and fluorophore-linked secondary antibodies. As a control, the identical printing, incubation, and detection procedure was carried out with uncoated nitrocellulose membrane (i.e., regular RPPA). The comparative results at the same detection intensity are shown in Figure 2A. As can be seen from the image scans, the polyGPA-coated membrane displayed over 99% specificity in selecting the oxidized form of AGP (with aldehyde groups) over the unoxidized control AGP. The results also demonstrated outstanding quantitative capabilities and high reproducibility (Figure 2B). Surprisingly, polyGPA's sensitivity is much higher than RPPA (over 10-fold signal increase) for the same protein concentration. We attribute this phenomenon to improved orientation of the glycoprotein AGP during its glycan binding to the polyGPA membrane, exposing more epitopes on the protein and increasing overall signal. This feature can be particularly important during the analysis of low abundant clinical samples.

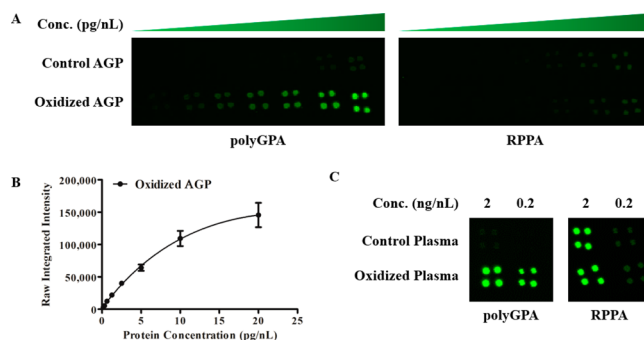


Figure 2. (A) Specific capture and detection of standard AGP by polyGPA. (B) Quantification of fluorescent signals from oxidized AGP in panel A by polyGPA. (C) Detection of endogenous AGP in human plasma by polyGPA and RPPA.

We further tested our hypothesis that polyGPA has high binding capacity due to its ability to reduce significantly the sample complexity by enriching glycoproteins on the membrane. Thus, higher concentrations of starting materials could be used for polyGPA than ever possible for regular RPPA. A typical starting sample for standard RPPA is 1 $\mu\text{g}/\mu\text{L}$ due to the limited binding capacity of unmodified nitrocellulose. Specifically targeting glycoproteins with polyGPA, we expect to be able to start with much higher protein concentration, thus ensuring improved detection of low-abundant glycoproteins. For initial testing of polyGPA with crude samples, we attempted to detect endogenous AGP in one of the most complex samples, human plasma. We diluted the plasma sample with the oxidation buffer to final 2 $\mu\text{g}/\mu\text{L}$ to accommodate the upper limit of binding capacity of RPPA (Supporting Information). The data demonstrated the detection of endogenous AGP from complex oxidized plasma sample with high signal/noise ratio (Figure 2C). As expected, the unoxidized plasma produced little signal, showcasing polyGPA's outstanding specificity. Consistent with the above experiment, the signal produced by polyGPA for the same protein concentration was significantly stronger than that from uncoated nitrocellulose membrane in RPPA.

The true value of polyGPA would be demonstrated during its applications in discovery or validation of potential disease biomarkers. Liquid biopsies—analysis of biofluids such as plasma, serum, and urine—have recently gained much attention as a likely useful source of diagnostic biomarkers. Compared to plasma and serum, the urinary proteome is much less complex. Nevertheless, urine can contain important biomarkers, especially for urological diseases such as bladder cancer, where secreted proteins from the diseased organs are directly released to the urinary tract.^{40,41} In this study, we applied quantitative N-glycoproteomics and polyGPA to profile and validate potential bladder cancer biomarkers using human urine samples of bladder cancer patients and healthy control individuals. Although urinary proteomics experiments have been previously carried out in search of biomarkers for bladder cancer,^{41–44} attempts on urinary glycoproteins have been limited.⁴⁵ Yang et al. identified 265 glycoproteins from 54 bladder cancer and 46 control urine samples using lectin affinity chromatography for enrichment, of which only α -1-antitrypsin (A1AT) was found significantly increased in bladder cancer (FDR < 1%) based on the spectral counting method.⁴⁵

Herein, urine samples from 16 bladder cancer patients and 8 age- and sex-matched healthy volunteers were obtained from

Indiana Biobank, pooled equally into the cancer or control group, respectively, and concentrated as described in SI. After digestion, N-linked glycopeptides were enriched using hydrazide chemistry and N-glycans were cleaved by PNGase F. After desalting, the released former N-glycopeptides from each group (enriched from 100 μg of peptides) were analyzed by LC-MS/MS. From a total of 8 LC-MS/MS runs (4 technical replicates per group), MaxQuant identified 1072 unique glycosylation sites with localization probability larger than 0.75 and the consensus motif N-X-S/T/C for N-linked glycosylation (where X is any amino acid except proline), representing 553 glycoproteins with 1% FDR. Overall, we were able to quantify 898 unique glycosites from 493 glycoproteins (Table S1), among which 41 unique glycosites from 25 glycoproteins were found significantly enriched in bladder cancer with a ratio of at least 3 and p -value < 0.01 (Figure 3A and Table S2).

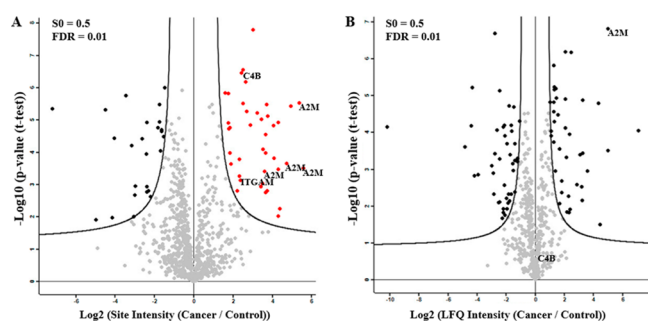


Figure 3. (A) Glycosylation sites or (B) proteins statistically enriched in bladder cancer by permutation-based FDR-corrected t test ($S_0 = 0.5$, FDR = 0.01).

To determine further whether the increase in glycosite abundance in bladder cancer occurred at the protein expression level or glycosylation occupancy level, 1 μg of the total peptides from the bladder cancer or control group was also analyzed by LC-MS/MS. From a total of 6 LC-MS/MS runs (3 technical replicates per group), MaxQuant identified 748 unique proteins with 1% FDR, out of which 518 proteins were quantified (Table S3). Among them, 38 proteins were found as significantly enriched in bladder cancer with at least a 2-fold difference at p -value < 0.05 (Figure 3B and Table S3). As expected, most glycosites found enriched in bladder cancer also showed increases in their protein expression levels, including α -2-macroglobulin (A2M), apolipoprotein B (APOB), complement C3 (C3), complement factor H (CFH), haptoglobin (HP), neutrophil collagenase (MMP8), serum paraoxonase (PON1), and others. However, we were also interested in identifying glycoproteins whose protein expression remained unchanged but N-glycosylation increased in cancer samples, such as complement C4-B (C4B). We also noticed a few glycosites enriched in bladder cancer from our glycoproteomic data that were not identified in our proteomic experiments, likely due to their low abundance preventing them to be identified without glycopeptide enrichment, such as integrin α -M (ITGAM).

As a proof-of-principle, we applied polyGPA to validate one candidate from each category, A2M, C4B, and ITGAM, with another cohort of urine samples from 8 bladder cancer patients and 5 healthy controls. Urine samples were concentrated, preoxidized and each individual sample was printed onto the polyGPA membranes and noncoated membranes as in regular

RPPA. Specific protein antibodies were then used to detect and quantify endogenous A2M, C4B, and ITGAM signals in individual samples. Consistent with the glycoproteomic and proteomic results, A2M signals were significantly higher in cancer patients than in controls in both polyGPA and RPPA experiments (Figure 4A). However, polyGPA showed much

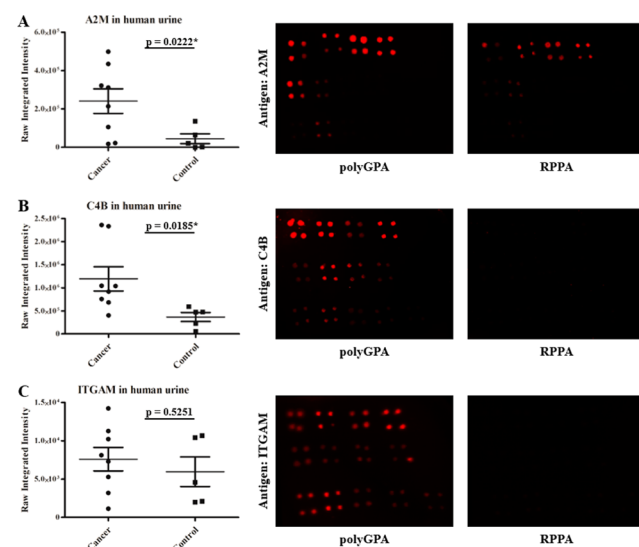


Figure 4. Quantification of endogenous (A) A2M, (B) C4B, and (C) ITGAM in human urine. For each membrane, top two rows were printed with 8 cancer patient samples and bottom row was printed with 5 healthy control samples, 4 prints per individual sample. For quantitation of signals in polyGPA, mean intensity of 4 prints per individual was used.

better sensitivity because enrichment of glycoproteins in polyGPA significantly reduced sample complexity and facilitated antibody binding. This enhanced sensitivity proved to be critical for the detection of proteins with much lower abundances than A2M, such as C4B and ITGAM, because their protein signals were barely detectable in RPPA (Figure 4B,C). More importantly, C4B displayed significantly higher signal in cancer samples in polyGPA. Combined with the glycoproteomic and proteomic results, the data suggest that monitoring of urinary C4B glycosylation could potentially serve as a marker for bladder cancer, although this remains to be further validated with a larger sample size.

In summary, we present a novel reverse phase glycoprotein array functionalized with globular hydroxyaminodendrimers to covalently capture glycoproteomes from complex biological samples including human plasma and urine, with subsequent detection using validated antibodies. The platform utilizes a similar procedure that is widely implemented in RPPA, thus maintaining extremely high sensitivity, reproducibility, and adaptability for clinical applications. We demonstrated polyGPA's high specificity toward glycoproteins, superior sensitivity, and quantitative capabilities. Combined with a quantitative N-glycoproteomic study on urine samples of bladder cancer patients and healthy individuals, polyGPA was successfully applied to validate glycoproteins A2M and C4B as potential markers for bladder cancer. Although polyGPA alone cannot differentiate changes at the glycosylation level or at the protein expression level, this strategy could become a powerful technique for sensitive screening of glycoproteins for biomarker discovery and validation studies, and is particularly appealing

due to severe lack of commercially available glycosite-specific antibodies.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.6b10239.

Materials, synthesis of the reagent, the polyGPA assay, sample preparations for LC-MS/MS analysis, MS data analysis (PDF)

Supplementary table (XLSX)

Supplementary table (XLSX)

Supplementary table (XLSX)

■ AUTHOR INFORMATION

Corresponding Author

*watao@purdue.edu

ORCID

W. Andy Tao: 0000-0002-5535-5517

Notes

The authors declare the following competing financial interest(s): Co-authors A.I. and W.A.T. are co-founders of Tymora Analytical.

■ ACKNOWLEDGMENTS

This project has been funded by NIH grants 5R41GM109626, 1R01GM111788, and 1R41CA210772, and by NSF grant 1506752. Additional support was provided by the Purdue University Center for Cancer Research (P30 CA023168).

■ REFERENCES

- (1) Marth, J. D.; Grewal, P. K. *Nat. Rev. Immunol.* **2008**, *8*, 874.
- (2) Moremen, K. W.; Tiemeyer, M.; Nairn, A. V. *Nat. Rev. Mol. Cell Biol.* **2012**, *13*, 448.
- (3) Stowell, S. R.; Ju, T.; Cummings, R. D. *Annu. Rev. Pathol.: Mech. Dis.* **2015**, *10*, 473.
- (4) Almeida, A.; Kolarich, D. *Biochim. Biophys. Acta, Gen. Subj.* **2016**, *1860*, 1583.
- (5) Munkley, J.; Elliott, D. J. *Oncotarget* **2016**, *7*, 35478.
- (6) Chen, K.; Gentry-Maharaj, A.; Burnell, M.; Steentoft, C.; Marcos-Silva, L.; Mandel, U.; Jacobs, I.; Dawnay, A.; Menon, U.; Blixt, O. *J. Proteome Res.* **2013**, *12*, 1408.
- (7) Blixt, O.; Westerlind, U. *Curr. Opin. Chem. Biol.* **2014**, *18*, 62.
- (8) Paweletz, C. P.; Charboneau, L.; Bichsel, V. E.; Simone, N. L.; Chen, T.; Gillespie, J. W.; Emmert-Buck, M. R.; Roth, M. J.; Petricoin, I. E.; Liotta, L. A. *Oncogene* **2001**, *20*, 1981.
- (9) Charboneau, L.; Tory, H.; Chen, T.; Winters, M.; Petricoin, E. F., 3rd; Liotta, L. A.; Paweletz, C. P. *Briefings Funct. Genomic Proteomics* **2002**, *1*, 305.
- (10) Nishizuka, S.; Charboneau, L.; Young, L.; Major, S.; Reinhold, W. C.; Waltham, M.; Kouros-Mehr, H.; Bussey, K. J.; Lee, J. K.; Espina, V.; Munson, P. J.; Petricoin, E., 3rd; Liotta, L. A.; Weinstein, J. N. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 14229.
- (11) Akbani, R.; Becker, K. F.; Carragher, N.; Goldstein, T.; de Koning, L.; Korf, U.; Liotta, L.; Mills, G. B.; Nishizuka, S. S.; Pawlak, M.; Petricoin, E. F., 3rd; Pollard, H. B.; Serrels, B.; Zhu, J. *Mol. Cell. Proteomics* **2014**, *13*, 1625.
- (12) MacBeath, G. *Nat. Genet.* **2002**, *32* (Suppl), S26.
- (13) Gonzalez, R. M.; Seurynck-Servoss, S. L.; Crowley, S. A.; Brown, M.; Omenn, G. S.; Hayes, D. F.; Zangar, R. C. *J. Proteome Res.* **2008**, *7*, 2406.
- (14) Pla-Roca, M.; Leulmi, R. F.; Tourekhanova, S.; Bergeron, S.; Laforte, V.; Moreau, E.; Gosline, S. J.; Bertos, N.; Hallett, M.; Park, M.; Juncker, D. *Mol. Cell. Proteomics* **2012**, *11*, M111.011460.
- (15) Holm, A.; Wu, W.; Lund-Johansen, F. *New Biotechnol.* **2012**, *29*, 578.
- (16) Cretich, M.; Damin, F.; Chiari, M. *Analyst* **2014**, *139*, S28.
- (17) Schwenk, J. M.; Igel, U.; Neiman, M.; Langen, H.; Becker, C.; Bjartell, A.; Ponten, F.; Wiklund, F.; Gronberg, H.; Nilsson, P.; Uhlen, M. *Mol. Cell. Proteomics* **2010**, *9*, 2497.
- (18) Uhlen, M.; Oksvold, P.; Fagerberg, L.; Lundberg, E.; Jonasson, K.; Forsberg, M.; Zwahlen, M.; Kampf, C.; Wester, K.; Hober, S.; Wernerus, H.; Bjorling, L.; Ponten, F. *Nat. Biotechnol.* **2010**, *28*, 1248.
- (19) Sevecka, M.; Wolf-Yadlin, A.; MacBeath, G. *Mol. Cell. Proteomics* **2011**, *10*, M110.005363.
- (20) Nanjundan, M.; Byers, L. A.; Carey, M. S.; Siwak, D. R.; Raso, M. G.; Diao, L.; Wang, J.; Coombes, K. R.; Roth, J. A.; Mills, G. B.; Wistuba, II; Minna, J. D.; Heymach, J. V. *J. Thorac. Oncol.* **2010**, *5*, 1894.
- (21) Cheng, S.; Serra, S.; Mercado, M.; Ezzat, S.; Asa, S. L. *Clin. Cancer Res.* **2011**, *17*, 2385.
- (22) Irwin, M. E.; Nelson, L. D.; Santiago-O'Farrill, J. M.; Knouse, P. D.; Miller, C. P.; Palla, S. L.; Siwak, D. R.; Mills, G. B.; Estrov, Z.; Li, S.; Kornblau, S. M.; Hughes, D. P.; Chandra, J. *PLoS One* **2013**, *8*, e70608.
- (23) Gollapudi, K.; Galet, C.; Grogan, T.; Zhang, H.; Said, J. W.; Huang, J.; Elashoff, D.; Freedland, S. J.; Rettig, M.; Aronson, W. J. *Am. J. Cancer Res.* **2013**, *3*, 523.
- (24) Ummanni, R.; Mannsperger, H. A.; Sonntag, J.; Oswald, M.; Sharma, A. K.; Konig, R.; Korf, U. *Biochim. Biophys. Acta, Proteins Proteomics* **2014**, *1844*, 950.
- (25) Hsu, K. L.; Pilobello, K. T.; Mahal, L. K. *Nat. Chem. Biol.* **2006**, *2*, 153.
- (26) Pilobello, K. T.; Slawek, D. E.; Mahal, L. K. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 11534.
- (27) Syed, P.; Gidwani, K.; Kekki, H.; Leivo, J.; Pettersson, K.; Lamminmaki, U. *Proteomics* **2016**, *16*, 1257.
- (28) Yang, Z.; Hancock, W. S.; Chew, T. R.; Bonilla, L. *Proteomics* **2005**, *5*, 3353.
- (29) Gupta, G.; Suroli, A.; Sampathkumar, S. G. *OMICS* **2010**, *14*, 419.
- (30) Frost, D. C.; Li, L. *Adv. Protein Chem. Struct. Biol.* **2014**, *95*, 71.
- (31) Zhang, Y.; Jiao, J.; Yang, P.; Lu, H. *Clin Proteomics* **2014**, *11*, 18.
- (32) Kaji, H.; Saito, H.; Yamauchi, Y.; Shinkawa, T.; Taoka, M.; Hirabayashi, J.; Kasai, K.; Takahashi, N.; Isobe, T. *Nat. Biotechnol.* **2003**, *21*, 667.
- (33) Wang, Y.; Wu, S. L.; Hancock, W. S. *Glycobiology* **2006**, *16*, 514.
- (34) Krishnamoorthy, L.; Mahal, L. K. *ACS Chem. Biol.* **2009**, *4*, 715.
- (35) Sparbier, K.; Wenzel, T.; Kostrzewa, M. *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.* **2006**, *840*, 29.
- (36) Xu, Y.; Wu, Z.; Zhang, L.; Lu, H.; Yang, P.; Webley, P. A.; Zhao, D. *Anal. Chem.* **2009**, *81*, S03.
- (37) Zhang, H.; Li, X. J.; Martin, D. B.; Aebersold, R. *Nat. Biotechnol.* **2003**, *21*, 660.
- (38) Tian, Y.; Zhou, Y.; Elliott, S.; Aebersold, R.; Zhang, H. *Nat. Protoc.* **2007**, *2*, 334.
- (39) Klement, E.; Lipinszki, Z.; Kupihar, Z.; Udvardy, A.; Medzihradsky, K. F. *J. Proteome Res.* **2010**, *9*, 2200.
- (40) Chen, Z.; Kim, J. *BMC Urol.* **2016**, *16*, 11.
- (41) Guo, J.; Ren, Y.; Hou, G.; Wen, B.; Xian, F.; Chen, Z.; Cui, P.; Xie, Y.; Zi, J.; Lin, L.; Wu, S.; Li, Z.; Wu, L.; Lou, X.; Liu, S. *J. Proteome Res.* **2016**, *15*, 2164.
- (42) Linden, M.; Lind, S. B.; Mayrhofer, C.; Segersten, U.; Wester, K.; Lyutvinskiy, Y.; Zubarev, R.; Malmstrom, P. U.; Pettersson, U. *Proteomics* **2012**, *12*, 135.
- (43) Chen, C. L.; Lin, T. S.; Tsai, C. H.; Wu, C. C.; Chung, T.; Chien, K. Y.; Wu, M.; Chang, Y. S.; Yu, J. S.; Chen, Y. T. *J. Proteomics* **2013**, *85*, 28.
- (44) Li, C.; Li, H.; Zhang, T.; Li, J.; Liu, L.; Chang, J. *Biochem. Biophys. Res. Commun.* **2014**, *446*, 1047.
- (45) Yang, N.; Feng, S.; Shedden, K.; Xie, X.; Liu, Y.; Rosser, C. J.; Lubman, D. M.; Goodison, S. *Clin. Cancer Res.* **2011**, *17*, 3349.