Vol. 9, No. 3, 2018
ISSN 2233-4203/ e-ISSN 2093-8950
INVITED ARTICLE www.msletters.org | Mass Spectrometry Letters

Mass Spectrometry in the Determination of Glycosylation Site and *N*-Glycan Structures of Human Placental Alkaline Phosphatase

Kemal Solakyildirim[1,*], Lingyun Li[2], and Robert J. Linhardt[2,3,4]

[1]*Department of Chemistry, Faculty of Arts and Science, Erzincan University, Erzincan, Turkey*
[2]*Center for Biotechnology and Interdiscıplinary Studies, Rensselaer Polytechnic Institute, Troy, NY, USA*
[3]*Department of Chemistry and Chemical Biology, Rensselaer Polytechnic Institute, Troy, NY, USA*
[4]*Department of Chemical and Biological Engineering and Biology, Rensselaer Polytechnic Institute, Troy, NY, USA*

Received April 4, 2018; Revised May 29, 2018; Accepted May 30, 2018
First published on the web September 30, 2018; DOI: 10.5478/MSL.2018.9.3.67

Abstract : Alkaline phosphatase (AP) is a membrane-bound glycoprotein that is widely distributed in the plasma membrane of cells of various organs and also found in many organisms from bacteria to humans. The complete amino acid sequence and three-dimensional structure of human placental alkaline phosphatase have been reported. Based on the literature data, AP consists of two presumptive glycosylation sites, at Asn-144 and Asn-271. However, it only contains a single occupied *N*-linked glycosylation site and no occupied *O*-linked glycosylation sites. Hydrophilic interaction chromatography (HILIC) has been primarily employed for the characterization of the glycan structures derived from glycoproteins. *N*-glycan structures from human placental alkaline phosphatase (PLAP) were investigated using HILIC-Orbitrap MS, and subsequent data processing and glycan assignment software. 16 structures including 10 sialylated *N*-glycans were identified from PLAP.

Keywords : Alkaline phosphatase, glycosylation, *N*-glycans, Mass spectrometry, HILIC

Introduction

Alkaline phosphatase (AP) is a membrane-bound glycoprotein that is widely distributed in the plasma membrane of cells of various organs and also found in many organisms from bacteria to humans.[1] Most APs has catalytic sites that contain three metal ions (two Zn and one Mg), which are necessary for their enzyme activity. The complete amino acid sequence and three-dimensional structure of human placental alkaline phosphatase have been reported.[2,3] Based on the literature data, AP has two known glycosylation sites, at Asn-144 and Asn-271.[1-3] However, it only contains a single occupied *N*-linked glycosylation site without any occupied *O*-linked glycosylation sites.[4,5]

Although AP can be found in nearly every tissue in the body, one of the most clinically relevant forms comes from human placenta.[6] Alkaline phosphatase is a very useful serum marker, with wide applications from diagnosing hepatoma to predicting mortality in dialysis patients.[7-9]

Glycosylation is one of the most important and most complex post-translational modifications (PTMs) of proteins and plays a significant role in many biological processes, including protein stability, cellular adhesion, signaling and most importantly diseases [10]. It represents the most widespread PTM found in natural and biopharmaceutical proteins.[11] For instance, more than half of the human proteins are glycosylated and their function depends on particular glycoforms. *N*-linked glycosylation is one of the major types of glycosylation. Their monosaccharide compositions, branching, and glycosidic linkages complicate *N*-linked glycans. They can be divided into three main groups based on monosaccharide composition and branching: 1) high mannose, 2) complex, and 3) hybrid types.[12] One glycosylation site on glycoprotein can be associated with these different types of glycan structures, which makes their structural analysis a very challenging task.

Mass spectrometry (MS) has been widely employed for the characterization of the glycan structures derived from glycoproteins.[13-15] However, in most cases, MS analysis alone is not sufficient to elucidate very complex structures due to extensive *N*-glycan heterogeneity. Therefore, many research groups apply a combination of different type of

*Reprint requests to Kemal Solakyildirim
E-mail: ksolakyildirim@erzincan.edu.tr

separation techniques such as capillary electrophoresis (CE),[16] liquid chromatography (LC),[17] hydrophilic interaction liquid chromatography (HILIC)[18,19] with either electrospray ionization (ESI) or matrix-assisted laser desorption ionization (MALDI) MS techniques. HILIC is widely used for the analysis of glycans from glycoproteins.[20]

PLAP is one of three tissue-specific human APs extensively studied because of its ectopic expression in tumors. Also, it is known that many glycoproteins produced in tumor cells have unusual sugar chains despite that their polypeptide portions are the same as those produced in normal cells. Unlocking the structural code of glycosylation offers another dimension toward understanding protein glycosylation and biological roles of carbohydrates. Therefore, it was aim to determine glycosylation site and characterization of *N*-glycan structures of PLAP in order to reveal any unusual sugar moiety exists.

In this study, we investigated *N*-glycan structures from human placental alkaline phosphatase (PLAP). First, we used in-gel tryptic digestion of PLAP to determine the glycosylation site within the peptide sequence of the glycoprotein.[21,22] Second, *N*-linked glycans were released and purified from PLAP glycoprotein by in-gel PNGase F digestion.[23] Third, purified *N*-glycans were analyzed by HILIC-ESI-FT-MS to determine the structures and finally, the PLAP *N*-glycans were identified by data processing and glycan assignment software based on their mass and established biological rules.

Experimental

Materials

Recombinant secreted human alkaline phosphatase (PLAP, expressed in human embryonic kidney 293 cell line) was purchased from U-Protein Express B.V. (Science Park Utrecht, The Netherlands). PNGase F was purchased from Sigma-Aldrich (St. Louis, MO, USA). All reagents used were either analytical grade or better.

Methods

In-gel de-glycosylation and extraction of *N*-glycans

PLAP (5 μg) from human was subjected to a linear gradient SDS-PAGE (4-20% (w/v) gel) under reducing conditions and stained with Coomassie blue. Gel bands of interest from the SDS-PAGE gel (Bio-Rad Laboratories, Hercules, CA, USA) were excised and cut into small cubes (ca. 1 × 1 mm). The Coomassie blue stained pieces of gel were then de-stained with 100 mM ammonium hydrogen carbonate in 50% aqueous acetonitrile (1:1, vol/vol) with gently mixing overnight at 4°C. The gel pieces were washed 3-times with pure water followed by 3-times washed with pure acetonitrile for 10-20 min. The gel pieces were then dried at room temperature for 30 min in a SpeedVac concentrator (Thermo Fisher Scientific, San Jose, CA, USA). Dried gel pieces were covered with reduction buffer (10 mM DTT in 100 mM NH_4NHCO_3) and incubated at 56°C for 30 min. After removal of reduction buffer, the gel pieces were covered with alkylation buffer (55 mM iodoacetamide in 100 mM NH_4HCO_3) and incubated in the dark at 20°C for 30 min. Alkylation buffer was removed and the samples were washed with water and then with pure acetonitrile. After reduction and alkylation, the gel pieces were dried and 3U PNGase F (*N*-glycosidase F, Sigma-Aldrich St. Louis, MO) enzyme solution was added to the dried gel pieces, which were later covered with 100 mM NH_4HCO_3 solution and incubated at 37°C for 10 h. After that, the incubation buffer was removed to extract *N*-glycans from the gel pieces followed by three extractions with 100-μl of pure water and three extractions with 100-μl acetonitrile. After combining extracts and the removed incubation buffer, a SpeedVac concentrator was used at room temperature to dry them all. Released *N*-glycans were recovered in the supernatant following protein precipitation with 70% ice-cold ethanol at -20°C for 60 min. and centrifugation at 15,000 for 10 min. The supernatant containing the *N*-glycans was transferred to a fresh tube and dried by SpeedVac. Released *N*-linked glycan samples were re-dissolved in 10 μl of water prior to further use.

HILIC-FT-MS analysis of released N-glycans

HILIC HPLC experiments were performed using the following operating procedure: the 50 mm × 2.0 mm (i.d.), 3 μm Luna column was obtained from Waters (Milford, MA, USA). Purified *N*-glycans were analyzed using mobile phase A, 5 mM ammonium acetate in 100% water, and mobile phase B, 5 mM ammonium acetate in 90% acetonitrile. The gradient was: 0 min, 100%B, 10 min, 50% B, 12 min, 20% B, 15 min, 20% B, 20 min, 100% B at a flow rate of 20 μl/min and a column temperature of 25°C. Samples (5 μl) were diluted in 20 μl of mobile phase B. Injection amount was 8 μl. The nano-HILIC chromatography system (Agilent Technologies, Santa Clara, CA, USA) was connected to an LTQ-Orbitrap-FT mass spectrometer (Thermo Fisher Scientific, San Jose, CA, USA) using a standard ESI source equipped with a stainless-steel capillary. Negative mode mass spectrometry was operated in with an ESI voltage at -4.5 kV. A capillary voltage of -40 V and heated capillary temperature of 275°C was used. A sheath gas flow and auxiliary gas flow were 30 and 6 arbitrary units. A nominal resolution setting of 100,000 was used.

Results and Discussion

RP-LC-MS analysis of in-gel trypsin digested PLAP

Glycosylation is one of the most important PTMs, which is very crucial for the cellular processes of proteins, including the cellular localization and the regulation of protein function. Analysis of glycosylation is a very challenging task because of the variability in the attached glycans. Generally, the main strategy of determining glycosylation is a combination of several steps, including

removing and purifying glycans and analyzing their structures by sequential utilization of multiple analytical techniques. However, the mass shift in the peptide molecular weight, the abundance of the glycans in the protein and the stability of the glycans during MS ionization can make the analysis of glycosylation very challenging.

The complete amino acid sequence and crystal structure of PLAP have been studied.[3,24] Asn-144 and Asn-271 are the two potential glycosylation sites. The peptide sequence, positions and possible modifications with theoretical and determined masses of trypsin-digested PLAP can be seen in supplementary data 1. There are four disulfide bridges of which one of these is close to the first glycosylation site and might rigidify the loop carrying the carbohydrate chain.

In-gel tryptic digestion of PLAP was used to determine the glycosylation site within the peptide sequence of the glycoprotein. Because of its capability of separating glycopeptides, reversed-phase high performance liquid chromatography (HPLC) chromatography is a widely used technique for site-specific analysis of glycopeptides after trypsin digestion of a glycoprotein.[25] Figure 1A demonstrates the total ion chromatogram (TIC) of in-gel trypsin digested PLAP. As it is shown, almost all peptide sequences (supplementary data 1) were well separated.

A representative example of an extracted ion chromatogram (EIC) is shown for the peptide sequence [ALTETIMFDDAIER] with its corresponding mass spectrum containing a peak at *m/z* 812.898 (Figure 1B). All other peptide sequences with very high accuracy and selectivity were determined (data not shown).

The EIC and the corresponding mass spectra of peptide sequence between 268-272, which contains one of the putative glycosylation sites in PLAP (Figure 2A). If the glycosylation were to occur on the Asn-271 residue, we should not be able to observe the peak at *m/z* 369.190. However, we determined a very well separated and strong peak at *m/z* 369.190, which suggests that there is no/partial glycosylation at Asn-271. In contrast, when *m/z* 520.732 was extracted, there was no single peak observed and extracted mass spectra did not show the mass unit of the peptide between residues 140-147, which indicates the full glycosylation at Asn-144 (Figure 2B).

Trypsin digested sample was further treated with PNGase F for release of *N*-glycans from the glycopeptide to confirm glycosylation at Asn-144 of PLAP. The EIC spectra for the peptide sequence between 140-147 with *m/z* 521.776, which correspondence the peptide sequence [FNQC*N*TTR] after *N*-glycan release (Figure 3). The reason for observing *m/z* 521.7 instead of *m/z* 520.7 is that on PNGase F catalyzed cleavage of *N*-linked glycans from the glycopeptide Asn-144 (MW 132.1) is deaminated to afford Asp-144 (MW 133.1). These results definitively establish glycosylation at Asn-144 in PLAP.

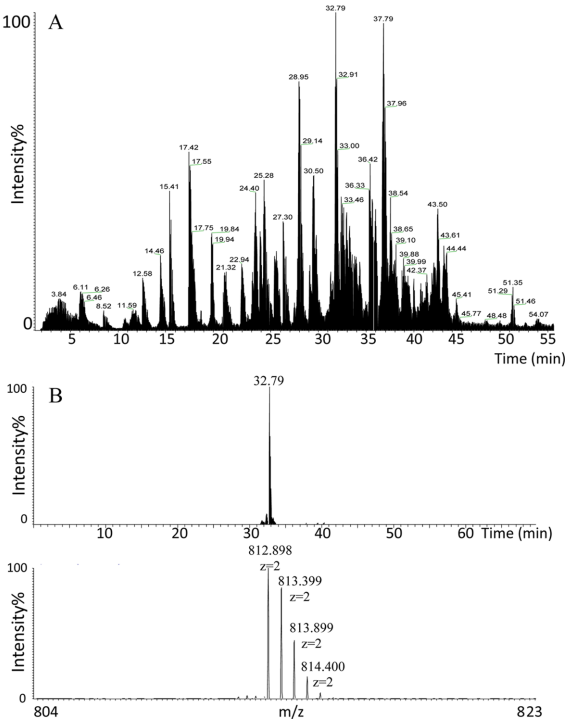

Figure 1. RP-LC-Orbitrap mass spectra of in-gel trypsin digested PLAP. A shows TIC of trypsin digested PLAP, B shows an example of EIC for the peptide sequence of ALTETIMFDDAIER (*m/z* 812.898).

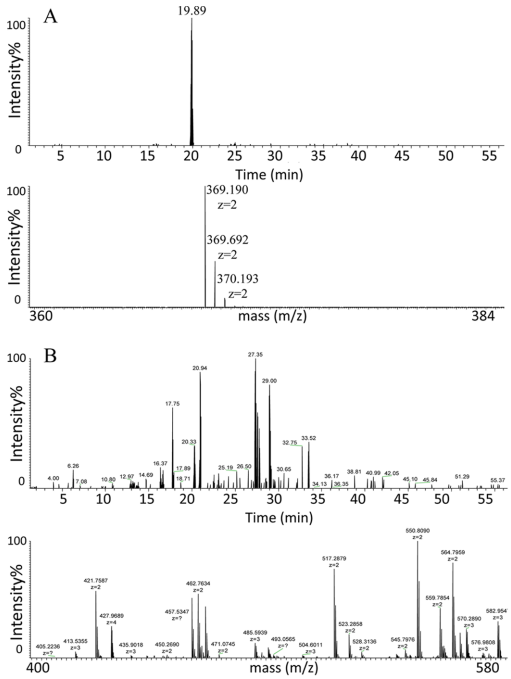

Figure 2. RP-LCOrbitrap mass spectra of in-gel trypsin digested PLAP. A shows EIC of *m/z* 369.190 (peptide sequence 268-272, possible glycosylation at 271), B shows EIC of *m/z* 520.733 (peptide sequence 140-147, possible glycosylation at 144).

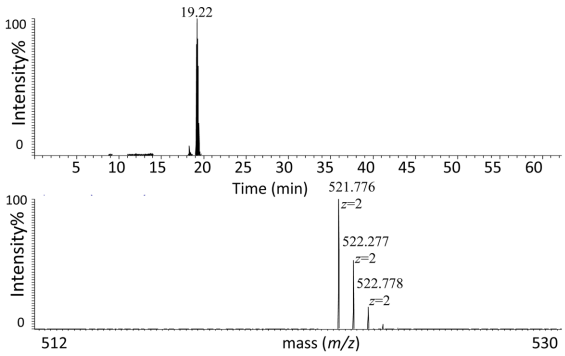

Figure 3. EIC of PNGase F digestion after in-gel trypsin digestion of PLAP. EIC of *m/z* 521.776 shows peptide sequence mass 140-147 after releasing glycans from glycopeptide.

HILIC-MS analysis of in-gel PNGase F digested PLAP

HILIC has been successfully applied to the analysis of carbohydrates. It is commonly used for the separation and analysis of very polar molecules such as glycans. Different types of stationary columns can also be employed in HILIC systems. HILIC also does not require expensive ion-pairing reagents unlike reversed phase-HPLC and it can be conveniently coupled to MS, especially in the ESI mode. Thus, a combination of HILIC and ESI-MS is a very efficient technique for the analysis of glycan profiles and structures, demonstrating very high resolution and mass accuracy with the capacity to identify structure.

After determining the *N*-glycosylation site of PLAP, we used HILIC-Orbitrap MS platform to determine the *N*-glycan structures. First, *N*-linked glycans were released and purified from PLAP glycoprotein by in-gel PNGase F digestion. The glycans released were then separated from the protein chains by ultrafiltration using 30 kDa MWCO spin columns. Samples were then analyzed by HILIC-Orbitrap MS. Orbitrap mass analyzer singles itself out from other mass analyzers providing very high-mass resolution and mass accuracy. A HILIC chromatogram of purified *N*-glycans from the glycoprotein is shown in Figure 4A. Some impurities were observed in first 4 min. The *N*-glycans eluted from the HILIC column starting from 5 min. Most of the *N*-glycans came out between 5 min and 7 min as it is indicated as a box in Figure 4A. A representative HILIC-MS spectrum of *N*-glycans is shown in Figure 4B-C. All glycans observed on MS spectrum were $[M-2H]^{2-}$ and we did not observe any adducts associated with metal ions (*i.e.*, $[M+Na-2H]^{-}$), thus, the spectra were simple and easy to interpret. The MS spectra were divided into two portions (Figure 4B and C) to show the *N*-glycan structures in clear fashion. The *N*-glycan structures in Figure 4B are shown in Figure 5A while the *N*-glycan structures in Figure 4 C are shown in Figure 5B.

PLAP *N*-glycans were identified by data processing and glycan assignment software based on their mass

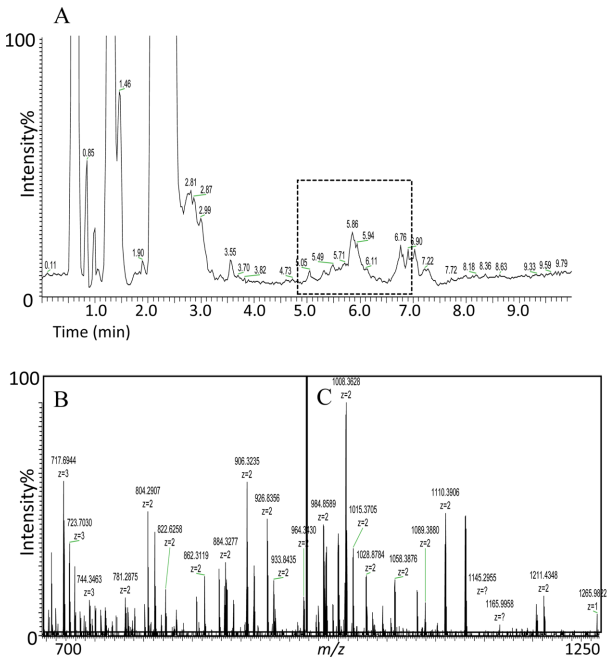

Figure 4. HILIC-Orbitrap mass spectra of in-gel PNGaseF digested PLAP. A shows HILIC chromatogram of *N*-glycan from PLAP, B and C shows Orbitrap MS spectra (averaged from

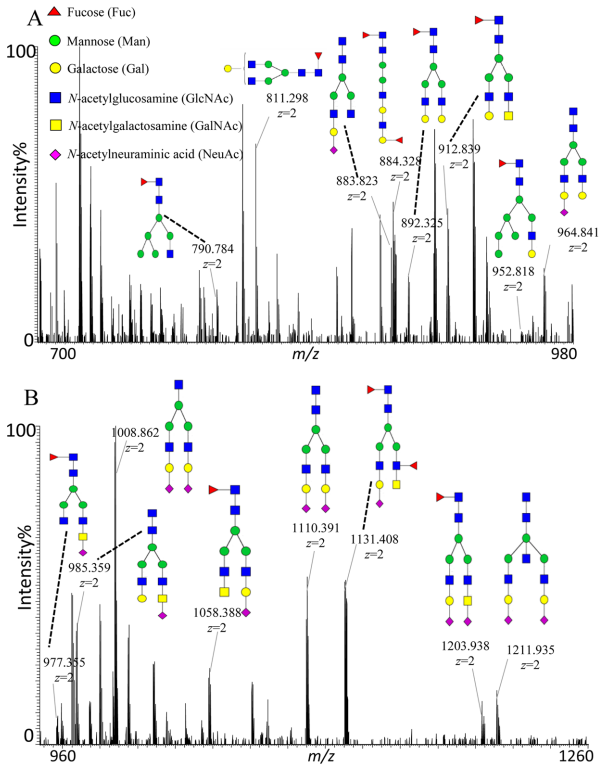

Figure 5. HILIC-Orbitrap mass spectra of in-gel PNGaseF digested PLAP. A shows HILIC MS spectrum *m/z* between 700-980, B shows HILIC-MS spectrum of *m/z* between 960-1260.

Table 1. Comprehensive list of PLAP *N*-glycans detailing glycan masses, monosaccharide compositions, and glycan types.

			glycan composition				
m/z	z	m	hexose	HexNAc	Fucose	NeuAc	glycan type
790.784	2	1583.585	5	3	1		Hybrid type
811.298	2	1624.611	4	4	1		Biantennary-complex
883.823	2	1769.662	4	4		1	Biantennary-complex
884.328	2	1770.671	4	4	2		Hybrid type
892.325	2	1786.665	5	4	1		Biantennary-complex
912.839	2	1827.694	4	5	1		Biantennary-complex
952.818	2	1907.652	7	3	1		Hybrid type
964.841	2	1931.698	5	4		1	Biantennary-complex
977.355	2	1956.726	6	2	1	1	Biantennary-complex
985.359	2	1972.733	4	5		1	Biantennary-complex
1008.862	2	2019.740	5	3		2	Biantennary-complex
1058.388	2	2118.791	4	5	1	1	Biantennary-complex
1110.391	2	2222.797	5	4		2	Biantennary-complex
1131.408	2	2264.832	4	5	2	1	Biantennary-complex
1203.938	2	2409.891	4	5	1	2	Biantennary-complex
1211.935	2	2425.886	5	5		2	Triantennary-complex

composition and established biological rules. A list of *N*-glycans with their *m/z* value, neutral mass, monosaccharide composition and type of glycan is provided in Table 1. As summarized in Table 1, 16 structures including 10 sialylated *N*-glycans were identified from PLAP. Sialic acid is widely distributed in tissues of human and sialylation plays a crucial role in glycoprotein structure and function. It can influence protein-protein interaction, as well as the physical and chemical properties of glycoproteins. Most of the *N*-glycan structures of alkaline phosphatase identified here are biantennary complex-type of structures in agreement with the results of the previous studies.[1,3]

Conclusions

A method for *N*-glycan analysis, consisting of determining *N*-glycosylation site, then releasing of glycans from the PLAP by PNGase F, analysis of *N*-glycan structures by HILIC-Orbitrap MS, and subsequent data processing and glycan assignment software based on their mass composition and established rules of biosynthesis was described in this paper. HILIC mobile phases contain volatile solvents and a high content of organic solvents such as acetonitrile and methanol with low viscosity allows higher flow rates that make HILIC ideal separation technique prior to ESI-MS. The LTQ-Orbitrap, which was used for this study, provides a very high accurate mass measurement that leads to high confidence in the reliable identification of glycan compositional assignment.

Supporting Information

Supplementary information is available at https://drive.google.com/file/d/1rMJQEyKMpB4Dli-OhpGpUMUbUcQAtCrJ/view?usp=sharing.

Acknowledgments

This work was supported by the US National Institute of Health (grant number HL096972).

References

1. Endo, T.; Ohbayashi, H.; Hayashi, Y.; Ikehara, Y.; Kochibe, N.; Kobata, A. *J. Biochem.* **2010**, 103, 182.
2. Millán, J. L. *Purinergic Signal.* **2006**, 2, 335.
3. Le Du, M. H.; Stigbrand, T.; Taussig, M. J.; Menez, A.; Stura, E. A. *J. Biol. Chem.* **2001**, 276, 9158.
4. Zhang, F.; Murhammer, D. W.; Linhardt, R. J. *Appl. Biochem. Biotechnol.* **2002**, 101, 197.
5. Nam, J. H.; Zhang, F.; Ermonval, M.; Linhardt, R. J.; Sharfstein, S. T. *Biotechnol. Bioeng.* **2008**, 100, 1178.
6. Chen, Y. H.; Chang, T. C.; Chang, G. G. *Protein Expr. Purif.* **2004**, 36, 90.
7. Suzuki, A.; Lymp, J.; Donlinger, J.; Mendes, F.; Angulo, P.; Lindor, K. *Clin. Gastroenterol. Hepatol.* **2007**, 5, 259.
8. Regidor, D. L.; Kovesdy, C. P.; Mehrotra, R.; Rambod, M.; Jing, J.; McAllister, C. J.; Wyck, D. V.; Kopple, J. D.; Kalantar-Zadeh, K. *J. Am. Soc. Nephrol.?***2008**, 19, 2193.
9. Dziedziejko, V.; Safranow, K.; Slowik-Zylka, D.; Machoy-Mokrzynska, A.; Millo, B.; Machoy, Z.; Chlubek, D. *Biochime* **2009**, 91, 445.
10. Oliveira-Ferrer, L.; Legler, K.; Milde-Langosch, K. *Semin. Cancer Biol.* **2017**, 44, 141.
11. Haltiwanger, R. S.; Feizi, T. *Curr. Opin. Struct. Biol.* **2011**, 21, 573.

12. Szabo, Z.; Guttman, A.; Karger, B. L. *Anal. Chem.* **2010**, 82, 2588.
13. North, S. J.; Hitchen, P. G.; Haslam, S. M.; Dell, A. *Curr. Opin. Struct. Biol.* **2009**, 19, 498.
14. Kim, Y. -G.; Gil, G.; Jang, K.; Lee, S.; Kim, H.; Kim, J.; Chung, J.; Park, C.; Harvey, D. J.; Kim, B. *J. Mass Spectrom.* **2009**, 44, 1087.
15. Gil, G. -C.; Iliff, B.; Cerny, R.; Velander, W. H.; Van Cott, K. E. *Anal. Chem.* **2010**, 82, 6613.
16. Mechref, Y.; Muzikar, J.; Novotny, M. V. *Electrophoresis* **2005**, 26, 2034.
17. Nwosu, C. C.; Aldredge, D. L.; Lee, H.; Lerno, L. A.; Zivkovic, A. M.; German, J. B.; Lebrilla, C. B. *J. Proteome Res.* **2012**, 11, 2912.
18. Melmer, M.; Stangler, T.; Schiefermeier, M.; Brunner, W.; Toll, H.; Rupprecther, A.; Linder, W.; Premstaller, A. *Anal. Bioanal. Chem.* **2010**, 398, 905.
19. Ruhaak, L. R.; Huhn, C.; Waterreus. W.; De Boer, A. R.; Neususs, C.; Hokke, C. H.; Deelder, A. M.; Wuhrer, M. *Anal. Chem.* **2008**, 80, 6119.
20. Buszewski, B.; Noga, S. *Anal. Bioanal. Chem.* **2012**, 402, 231.
21. Shevchenko, A.; Tomas, H.; Havlis, J.; Olsen, J. V.; Mann, M. *Nat. Protoc.* **2006**, 1, 2856.
22. Morelle, W.; Michalski, J. -C. *Nat. Protoc.* **2007**, 2, 1585.
23. Tarentino, A. L.; Gomez, C. M.; Plummer, T. H. *Biochemistry* **1985**, 24, 4665.
24. Harris, H. *Clin. Chim. Acta* **1990**, 186, 133.
25. Desaire, H. *Mol. Cell. Proteomics* **2013**, 12, 893.

