# Profiling of a high mannose-type N-glycosylated lipase using hydrophilic interaction chromatography-mass spectrometry

Gargano, A.F.G.; Schouten, O.; van Schaick, G.; Roca, L.S.; van den Berg-Verleg, J.H.; Haselberg, R.; Akeroyd, M.; Abello, N.; Somsen, G.W.

**DOI**
10.1016/j.aca.2020.02.042

**Publication date**
2020

**Document Version**
Final published version

**Published in**
Analytica Chimica Acta

Analytica Chimica Acta 1109 (2020) 69–77

Contents lists available at ScienceDirect

Analytica Chimica Acta

journal homepage: www.elsevier.com/locate/aca

# Profiling of a high mannose-type N-glycosylated lipase using hydrophilic interaction chromatography-mass spectrometry

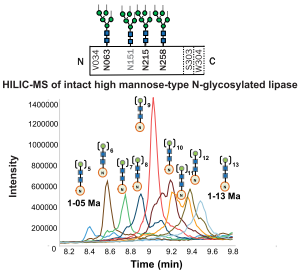
A.F.G. Gargano [a, b, c, *, 1], O. Schouten [d, 1], G. van Schaick [b, e], L.S. Roca [a, c], J.H. van den Berg-Verleg [d], R. Haselberg [a, b], M. Akeroyd [d], N. Abello [d], G.W. Somsen [a, b]

[a] Centre for Analytical Sciences Amsterdam, Amsterdam, the Netherlands
[b] Division of BioAnalytical Chemistry, Amsterdam Institute for Molecular and Life Sciences, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands
[c] van 't Hoff Institute for Molecular Science, University of Amsterdam, Amsterdam, the Netherlands
[d] DSM Biotechnology Center, Analysis Department, Delft, the Netherlands
[e] Center for Proteomics and Metabolomics, Leiden University Medical Center, Leiden, the Netherlands

## HIGHLIGHTS

- Comparison of SDS-PAGE, RPLC-MS and HILIC-MS of intact lipase.
- HILIC-MS allows for high-resolution separation of lipase glycoforms.
- In-depth glycoform characterization combining bottom-up and intact-protein methods.

## GRAPHICAL ABSTRACT




## ARTICLE INFO

*Article history:*
Received 12 November 2019
Received in revised form
17 February 2020
Accepted 23 February 2020
Available online 28 February 2020

*Keywords:*
High mannose-type N-Glycosylation
Intact protein analysis
HILIC
High-resolution mass spectrometry

## ABSTRACT

Many industrial enzymes exhibit macro- and micro-heterogeneity due to co-occurring post-translational modifications. The resulting proteoforms may have different activity and stability and, therefore, the characterization of their distributions is of interest in the development and monitoring of enzyme products. Protein glycosylation may play a critical role as it can influence the expression, physical and biochemical properties of an enzyme.

We report the use of hydrophilic interaction liquid chromatography-mass spectrometry (HILIC-MS) to profile intact glycoform distributions of high mannose-type N-glycosylated proteins, using an industrially produced fungal lipase for the food industry as an example. We compared these results with conventional reversed phase LC-MS (RPLC-MS) and sodium dodecyl sulfate–polyacrylamide gel-electrophoresis (SDS-PAGE). HILIC appeared superior in resolving lipase heterogeneity, facilitating mass assignment of N-glycoforms and sequence variants. In order to understand the glycoform selectivity provided by HILIC, fractions from the four main HILIC elution bands for lipase were taken and subjected to SDS-PAGE and bottom-up proteomic analysis. These analyses enabled the identification of the most abundant glycosylation sites present in each fraction and corroborated the capacity of HILIC to separate protein glycoforms based on the number of glycosylation sites occupied.

* Corresponding author. Center for Analytical Sciences Amsterdam, Universiteit van Amsterdam, Science Park 904, 1098 XH, Amsterdam, the Netherlands.
*E-mail address:* a.gargano@uva.nl (A.F.G. Gargano).

[1] A.F.G. Gargano and O. Schouten are both first authors of the manuscript as they contributed equally to the experimental work, data analysis and preparation of figures. The full list of the author's contribution is reported in the CRediT author statement.

https://doi.org/10.1016/j.aca.2020.02.042

Compared to RPLC-MS, HILIC-MS reducted the sample complexity delivered to the mass spectrometer, facilitating the assignment of the masses of glycoforms and sequence variants as well as increasing the number of glycoforms detected (69 more proteoforms, 177% increase). The HILIC-MS method required relatively short analysis time (<30 min), in which over 100 glycoforms were distinguished.

We suggest that HILIC(-MS) can be a valuable tool in characterizing bioengineering processes aimed at steering protein glycoform expression as well as to check the consistency of product batches.



## 1. Introduction

Glycosylation is one of the most common post-translational modifications (PTMs) of proteins. This PTM is generated by the covalent attachment of oligosaccharides (glycans) to the amino acid backbone of a protein, in particular to serine/threonine (O-glycosylation) or asparagine (N-glycosylation) residues. Depending on the type of organism, glycans are processed differently, giving rise to different N-glycosylation. In yeast and fungi, the high-mannose type (varying in the number of neutral mannose units) is dominant, whereas in vertebrates, complex types (including a variety of sugar units like sialic acid) are most prominent [1].

The evidence of the importance of N-glycosylation in regulating glycoprotein expression, structure and function is overwhelming [2]. N-glycosylation has shown to affect several enzyme properties, such as functional activity, conformation, stability and solubility [3]. For example, the partial or complete removal of the N-glycans in lipases resulted in reduced thermal stability and altered activity [2,4,5]. Glycosylation at N308 of human gastric lipase (one of its four N-glycosylation sites) appeared essential for full enzyme activity. Deglycosylated lipase was 50% less active than the wild-type enzyme. Interestingly the lack of glycosylation on the remaining glycosylation sites did not affect enzyme activity [6].

Therefore, profiling the glycosylation of enzymes such as lipases used in e.g. industrial food applications is highly relevant during the development and optimization of the enzyme production. Glycoproteins may be characterized at multiple levels using separation methods coupled to mass spectrometry (MS). Typically the analytical approaches used target either the glycans released from a protein [7], the glycopeptides [8] or the intact glycoprotein [9]. For the determination of the glycoforms expressed, the analysis of intact glycoproteins is the one conserving the highest amount of information but also the more difficult from an analytical chemistry perspective [10].

Several liquid-based hyphenated methods for glycoform profiling of intact proteins have been described, including capillary electrophoresis (CE) and liquid chromatography (LC) [9,11,12]. Coupling separation methods with high-resolution mass spectrometry (MS) enables to resolve the protein heterogeneity according to chemical differences (e.g. charge, hydrophobicity etc.) and measures the accurate masses of the proteoforms.

To date, RPLC, HILIC, IEC, SEC, HIC and CE methods have been applied to the online MS characterization of intact proteins (e.g. Refs. [13–18]). RPLC-based methods typically use a solvent combination that ensures compatibility with electrospray ionization MS and are therefore the most commonly used. However, these methods separate proteins mainly according to the amino acid composition and are typically not sufficiently selective to resolve glycoproteoforms. Recently, HILIC-MS has emerged as an analytical tool for protein analysis with complementary selectivity to RPLC-MS and capable of high-resolution separation of glycoforms [10,11,14,19,20]. HILIC methods for the characterization of glycoproteins are often using neutral stationary phases (e.g. using amide phases). Acetonitrile-water gradients are used with mobile phase additives allowing ion-pair formation with basic protein residues at low pH. As a result, ionic interactions of the protein with the stationary phase are minimized, leaving hydrophilic partitioning and hydrogen bonding as driving forces of retention. Under these conditions, the (neutral) sugars of protein glycans contribute substantially to retention, providing glycoform resolution according to the overall glycan size and composition [10,11,19]. Therefore, HILIC glycoform separations are complementary with respect to capillary electrophoresis (CE) where glycoforms are resolved on the basis of charge differences (e.g. differences in the number of sialic acid) [10].

In our investigation, we used a heavily N-glycosylated lipase as an example of profiling glycoforms with HILIC-MS. The high mannose-type glycosylation on this enzyme is particularly challenging to profile since it has high heterogeneity only based on neutral sugar. We adopted several analytical strategies to characterize this lipase, including intact protein approaches (SDS-PAGE, RPLC-MS and HILIC-MS) and a study of the N-glycosylation by protein digestion. HILIC demonstrated excellent resolving power for glycoforms of the enzyme (varying in the number of neutral sugar units) enabling an in-depth characterization of the proteoforms expressed.

## 2. Experimental

### 2.1. Chemicals

The lipase enzyme studied was provided by DSM Biotechnology Center as a formulation in wheat flour. The enzymes used in the processing described below were PNGase F, Trypsin Gold and EndoHf (Promega). M12 was used as molecular weight ladder (Invitrogen). All solvents and reagents were bought through Sigma Aldrich and of the highest quality available. Deionized water (18.2 MΩ) was obtained from a Milli-Q purification system.

### 2.2. Sample preparation

*Lipase sample*: 13.50 mg was weighed to which 1.35 mL Milli-Q water (deionized water at 18.2 MΩ cm) was added. The solution was shaken for 10 min at 1200 rpm in a thermomixer and centrifuged at room temperature (RT) for 10 min at 20817 rcf. The supernatant was removed and diluted 3x with Milli-Q water. To 65 μl of the diluted sample 65 μl 20% trichloroacetic acid (TCA) solution was added. This solution was put in the fridge at 4 °C for 1 h. After precipitation of the proteins, the sample was centrifuged for 20 min at 20817 rcf and 4 °C. After centrifugation, the supernatant was removed and the pellet was washed once with 500 μl acetone (−20 °C) and stored at −20 °C until use. Prior to use, the samples were centrifuged for 10 min at 20817 rcf and 4 °C. The supernatant was removed. After drying for 5 min at room temperature, the pellet was resolved in Milli-Q water to the concentration required.

*Deglycosylated lipase:* 10 mg of lipase powder samples were

dissolved in 1 mL Milli-Q water and mixed by shaking for 15 min at 25 °C. The samples were then spun down at 20,817 rcf and 4 °C in an Eppendorf centrifuge. A volume of 100 μl of supernatant was mixed with 100 μl 20% TCA. This solution was put on ice for 1 h to let proteins precipitate. Afterward, the samples were centrifuged for 15 min at 20,817 rcf and 4 °C. The supernatant was discarded, and the pellet was washed with 500 μl acetone (−20 °C) and centrifuged for 10 min at 20,817 rcf and 4 °C. The supernatant was discarded, and the pellet was left to dry under ambient air for 5 min, before being dissolved in 50 μl 50 mM sodium hydroxide, then supplemented with 350 μl 100 mM ammonium bicarbonate. A volume of 200 μl was placed in a new tube and deglycosylation was carried out by adding 10 μl PNGase F (500 units $mL^{-1}$, diluted in water) and incubating all samples overnight in an Eppendorf Thermomixer at 800 rpm and 37 °C.

Prior to HILIC-MS analysis, all lipase samples were diluted 4-fold with mobile phase B (see composition below). Before RPLC-MS, samples were diluted 5-fold with mobile phase A.

*Tryptic digestion of HILIC fractions of lipase:* Details of the HILIC fractionation method are reported in Section S2 of the supporting information. Freeze-dried fractions were dissolved in 20 μl 100 mM ammonium bicarbonate. 0.5 μl 200 mM tris(2-carboxyethyl)phosphine (TCEP) was added, samples were incubated for 30 min at room temperature. 0.5 μl 93.75 mM iodoacetamide (IAA) was added and samples were incubated for 30 min at room temperature in the dark. 0.5 μl 100 mM dithiothreitol (DTT) was added, samples were incubated for 30 min at room temperature. 1 μl 0.25 mg $mL^{-1}$ trypsin was added and samples were incubated for 4 h at 37 °C. Afterward, another 1 μl 0.25 mg $mL^{-1}$ trypsin was added and the samples were incubated overnight at 37 °C.

10 μl sample was reserved for analysis of the native peptides (non-deglycosylated). To a 10-μl aliquot 1 μl 1u $mL^{-1}$ PNGaseF was added and samples were deglycosylated by incubation overnight at 37 °C. All samples were acidified to 0.1% formic acid prior to RPLC-MS analysis.

### 2.3. Gel electrophoresis

65 μl solutions of either intact or deglycosylated lipase were added to 25 μl NuPAGE™ (Invitrogen) LDS sample buffer (4x) and 10 μl NuPAGE™ Sample Reducing Agent (10x) (Invitrogen). All samples were heated for 10 min at 70 °C.

Volumes of 15 μl of the HILIC fraction samples, 10 μl of the lipase control sample and 5 μl marker M12 (Invitrogen) were applied on NuPage 4–12% BisTris 10-well gel (Invitrogen). The gel was run at 200 V for 55 min, using a Xcell Surelock, with 600 mL 20x diluted MOPS SDS running buffer (Invitrogen) in the outer chamber. In the inner chamber, 200 mL of the running buffer was added of 0.5 mL of antioxidant (Invitrogen).

After running, the gels were fixed for 1 h with 50 mL 50% methanol/7% acetic acid fixation solution, rinsed twice with Milli-Q water and stained overnight with 50 mL Sypro Ruby staining solution (BioRad).

Images were made using the Typhoon FLA9500 (GE Healthcare) (610 BP 30, Blue (473 nm), PMT 650 V 100 μm) after washing the gel with demineralized water.

### 2.4. Liquid chromatography

LC separations of intact proteins were performed on an Acquity I-class instrument (Waters), equipped with a binary pump, autosampler, column thermostated compartment, and multiple wavelength absorbance detector.

RPLC analysis was performed using a Waters Acquity UPLC Protein BEH300C4 column (1.7 μm, 300 Å, 2.1 × 50 mm). Separations were done at 75 °C using gradient elution with mobile phase A being 0.1% formic acid (FA) in LC-MS grade water and mobile phase B being 0.1% FA in acetonitrile (ACN). The separations were performed at a flow rate of 0.2 mL $min^{-1}$ using an injection volume of 5 μl.

The gradient programming was as follows: 2% B for 2 min, 25% B at 2.2 min, 80% B at 8 min and 95% B at 8.2 min, 10% B at 9 min, 95% B at 10 min and 10% B at 11 min followed by 10% B for 5 min.

The HILIC column used was a Waters Acquity UPLC glycoprotein BEH Amide column (1.7 μm, 300 Å, 2.1 × 150 mm) equipped with a Waters Acquity UPLC glycoprotein BEH Amide guard column (1.7 μm 300 Å pore size 2.1 × 5 mm). Separations were done at 75 °C using a gradient elution with mobile phase A being ACN/water 98/2 v/v + 1% propionic acid and 0.025% trifluoroacetic acid (TFA) and mobile phase B being ACN/water 2/98 v/v + 1% propionic acid and 0.025% TFA. The injection volume was 5 μl.

For the HILIC separation, simultaneous flow rate and gradient programming elution was used. The sample was injected at 0.2 mL $min^{-1}$ using 10% B, and after 1.5 min the flow rate was increased to 0.4 mL $min^{-1}$ keeping 10% B, followed by a linear gradient from 10 to 30% B from min 2 to 3. At min 3 the flow rate was reduced to 0.2 mL $min^{-1}$. From min 3 to min 13 a linear gradient was applied from 30 to 45% B, then to 90% B at 14 min, keeping 90% till 15 min. The %B was then set to 10% B at min 15.5 and the flow rate was increased to 0.4 between min 18 and 18.2. The column was reequilibrated under these conditions for 7 min.

For the peptide level analysis, LC-MS/MS was performed using a Waters Acquity UPLC Peptide CSH C18 column (1.7 μm 130 Å pore size, 1 mm × 100 mm) at 50 °C. The flow rate was 0.15 mL $min^{-1}$ starting at 2% mobile phase B (0.1% FA in acetonitrile, mobile phase A being 0.1% formic acid). A linear gradient was made in 10 min from 2% B to 40% mobile phase B and in 2 min–70% mobile phase B. The column was flushed 1 min at 70% mobile phase B and re-equilibrated for 2 min at 2% mobile phase B.

### 2.5. Mass spectrometry

Two mass spectrometers were used in this study: a Synapt G2-S (Waters, Wilmslow, UK) and a Q-Exactive Plus (Thermo Fisher, Bremen, DE) were used to record RPLC-MS and HILIC-MS data and perform peak assignments and determine the mass error of our assignments. Intact protein data from the Q-Exactive plus are obtained with injection of about 4 times more sample on column with respect to the Waters Synapt G2-S. The Q-Exactive plus was used for bottom-up protein analysis.

Acquisition parameters for LC-MS intact protein measurements using the Synapt instrument: *m/z* range, 500–3500; scan time, 1 s; positive electrospray ionization (ESI); time-of-flight MS Resolution mode; capillary voltage, 2.5 kV; sampling cone, 40 V; source offset, 80 V; collision energy, 4 eV; source temperature 120 °C; desolvation temperature, 500 °C; desolvation gas flow, 500 l $h^{-1}$; nebulizer gas, flow 6 bar.

Acquisition parameters for the Q-Exactive Plus instrument for intact protein measurement: *m/z* range, 1100–2700 *m/z*; resolution, 17500; 7 microscans; max IT, 200 ms, AGC, 3e6, in source CID, 15 eV; intact protein mode activated (trapping gas 0.2); spray voltage, 3.50 kV; sheath gas flow, 60; auxiliary gas flow rate, 10; capillary temperature, 275 °C; S-lens RF, 70; aux gas temperature, 50 °C.

Peptide analysis was performed on the Q Exactive Plus instrument measuring 400–2000 *m/z* at 70000 resolution, automatic gain control (AGC) target was set to 3e6, maximum injection time to 75 ms, MS/MS was performed as Top 3 experiment at a resolution of 17500 with the AGC target of 1e5. 1+ and >8+ charge states were excluded, dynamic exclusion was of 5 s.

### 2.6. Data analysis

Extracted-ion chromatograms were obtained from the raw Synapt G2-S data, using the three most intense charge states observed (+15 to +20) and a 1-amu window. Time point/Intensity sum data were exported and plotted using Microsoft Excel and TIBCO Spotfire. Deconvoluted spectra were obtained using the maximum entropy deconvolution algorithm and MassLynx software.

For the Q-Exactive Plus data, deconvoluted masses were obtained using ThermoFisher Biopharma Finder software. The deconvolution was applied on the protein elution window (RPLC-MS, 4.5–8 min; HILIC-MS, 7–14 min), using the sliding windows option of the software and the following settings: target average spectrum width, 0.4 min; 25% offset, merge tolerance, 30 ppm; max retention time gap, 1 min; ReSpect (Isotopically Unresolved) as deconvolution algorithm; output mass range, 27000–37000 (Isotopic Profile as display mode), deconvolution tolerance, 30 ppm; peak model.

The mass spectrometry data (Q-Exactive plus) have been deposited to the PRIDE Archive (http://www.ebi.ac.uk/pride/archive/) via the PRIDE partner repository with the data set identifier PXD016215.

## 3. Results and discussion

This study aimed to profile glycoforms of a heavily N-glycosylated lipase protein. A commercially available lipase produced in *Aspergillus niger,* having a theoretical molecular weight in its mature form of about 28 kDa, was used as a proof of principle. This enzyme is a glycoprotein with four possible N-glycosylation sites (Fig. 1a) and four disulfide bridges (for amino acid sequence, see section S1 of the supporting information). It is highly heterogeneous due to variation in size and position of the glycans, which are dominated by high mannose types (i.e. mostly neutral glycans) [1]. Moreover, the enzyme exhibits potential truncation at the C-terminus of its sequence.

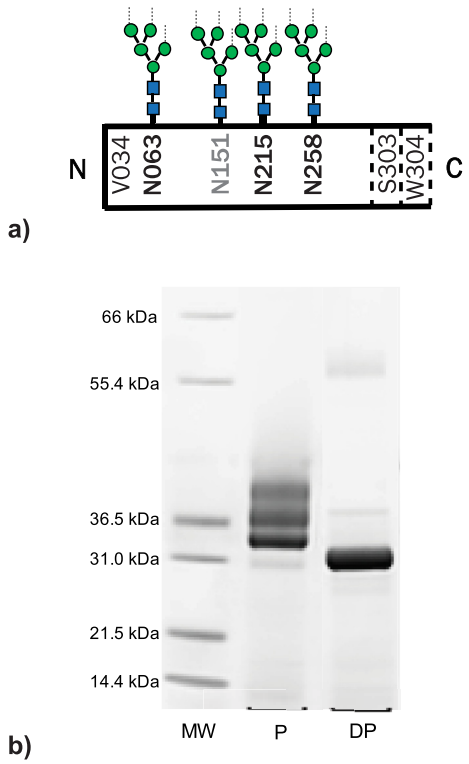



**Fig. 1.** a) Schematic representation of the lipase showing the sites of potential N-glycosylation and C-terminal truncation variants. b) SDS-PAGE on a 4–12% acrylamide gel; lanes represent the molecular weight marker (MW), lipase (P) and deglycosylated lipase (DP).

Currently, SDS PAGE and RPLC-MS are among the most common analytical methods to characterize intact biotechnological enzymes. However, these methods may have limited resolving power, particularly when protein heterogeneity originates from glycosylation. Here results for lipase obtained by SDS-PAGE and RPLC-MS are presented and compared with results obtained with HILIC-MS.

### 3.1. SDS-PAGE of lipase

SDS-PAGE offers a simple and direct method to obtain information on the size distribution of a protein sample. The analysis of lipase before and after deglycosylation (using PNGase F) is shown in Fig. 1b. The intact lipase (P) showed three main bands at approximately 33, 37 and 41 kDa, which most probably correspond to the protein with respectively one, two and three glycosylation sites occupied. A faint band for P observed at 30 kDa is likely due to non-glycosylated lipase, as confirmed by the single band observed for the deglycosylated protein (DP). The electrophoresis profile indicates that a large part of the protein sample is glycosylated. However, no further molecular information is obtained and the resolution provided by SDS-PAGE is too low to distinguish individual glycoforms.

### 3.2. RPLC-MS of lipase

RPLC employing stationary phases of relatively low hydrophobicity with wide pore sizes in combination with acidic water-ACN mobile phase gradients has shown useful for intact protein analysis, in particular in combination with ESI-MS. RPLC is MS compatible and can provide efficient sample desalting avoiding ionization suppression and source contamination. The lipase sample was analyzed by RPLC-MS using a wide-pore C4 column with a water-ACN gradient eluent comprising 0.1% FA (Fig. 2a). The relatively high lipophilicity of the enzyme required a high column oven temperature (75 °C) to reduce protein adsorption, increase recovery and minimize carryover. A steep gradient was used to maintain the width of the protein band within approximately 1 min.

When analyzing the profile of elution of the different glycoforms of this enzyme it is clear that the glycoforms are almost entirely coeluting. In the mass spectrum of the intact lipase predominantly proteoforms having one glycosylation site occupied (Fig. S4; the assigned masses are reported in Table S1) could be identified. These results seem to deviate from what was observed by SDS-PAGE, where the protein appeared to be distributed equally among glycoproteoforms having between one, two and three sites occupied (Fig. 1b). In RPLC-MS, coelution of the glycoforms most probably leads to severe ionization suppression of the less abundant proteoforms. The species having two or three glycosylation sites occupied statistically show a higher degree of heterogeneity giving rise to many (partially isobaric) proteoforms of relatively low abundance. Hence, these protein species suffer more from signal suppression during ESI and therefore are poorly or not detected in RPLC-MS.

Additionally, we tested longer gradients and other acid modifiers (data not shown) without significant improvement in the separation and identification of proteoforms. To reduce the complexity of the spectra, the enzyme was deglycosylated with PNGase F. This allowed to observe the truncation variants of the protein sequence. Results from the RPLC-MS analysis of lipase and

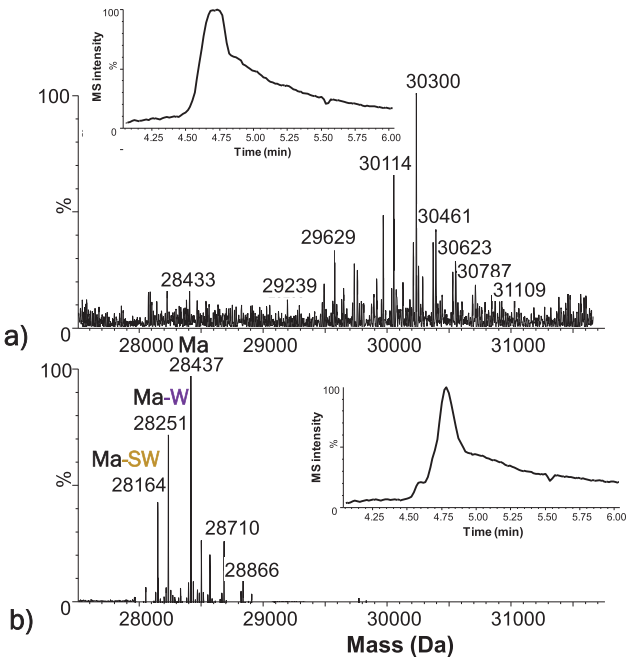



**Fig. 2.** Deconvoluted mass spectra obtained during RPLC-MS of a) lipase and b) deglycosylated lipase. The respective total-ion current chromatograms (TICs) are shown as inserts. In b) Ma indicates the mature form of the protein and Ma-W and Ma-SW two truncation variants.

deglycosylated lipase are summarized in Fig. 2. The most abundant protein sequence observed was the mature form of the protein (Ma), followed by a variant without a C-terminal tryptophan (Ma-W), and a form without a C-terminal serine and tryptophan (Ma-SW). Using this information we could assign 39 glycoproteoforms from the intact lipase measurement (the assigned masses are reported in Tables S1 and S2).

### 3.3. HILIC-MS of lipase

To develop a HILIC method for the characterization of the lipase sample, initially, a wide linear gradient was applied from about 90% to 40% ACN in 30 min. This yielded one main peak for the lipase sample. Focusing on the protein elution window, we confined the percentage change in ACN (ca. 70%–55%) and reduced the gradient steepness (13 min overall). This resulted in four major elution bands (Fig. 3a). The mobile phase contained 1% propionic acid and 0.025% of TFA in order to provide sufficient ion-pairing while largely avoiding signal suppression and adduct formation during ESI [21,22]. Details on HILIC method optimization for intact proteins can be found elsewhere [10,23,24].

The selective separation combined with high-resolution MS allowed the assignment of a multitude of lipase glycoforms, including their distribution in truncated forms. In total, 108 proteoforms were identified with mass accuracy within 60 ppm (data recorded at 17500 Rs and reported in Tables S1 and S2). Fig. 3 shows the annotated TIC (a) as well as the deconvoluted mass spectra of different elution zones (b). Similar to what was observed with SDS-PAGE, four successive protein bands were present, showing an increased degree of glycosylation (0–3 glycosylation site occupied). The similarity was stressed by the comparison of the distribution observed in the TIC with the one observed in the SDS-PAGE densitogram (Fig. 3c).

The lipase glycoforms are separated according to the number of mannose units present in the protein as underlined by the annotation provided in Fig. 3a. From the mass spectra obtained upon HILIC-MS analysis (Fig. 3b), the glycoform compositions could be determined, The glycoform retention clearly increases with size of glycan and number of glycosylation sites occupied. Moreover, the high-quality spectra allow distinction of C-terminal truncation variants per glycoform.

To visualize the obtained HILIC-MS separation in more detail, cumulative extracted-ion chromatograms from the signals of the three most intense charge states of each glycoform were constructed (Fig. 4a). Glycoforms differing one single mannose unit are clearly separated. Interestingly, for some of the glycoforms (e.g. 1N-8M and 1N-10 M), the cumulative extracted-ion traces showed multiple peaks. Most probably, isobaric glycoforms having the same overall composition of sugars units, but with different antenna lengths or at a different glycosylation site are starting to be resolved.

Cumulative extracted-ion chromatograms were also constructed for the truncation variants of single glycoforms (Fig. 4b and S6). This revealed that the truncation variants are slightly more retained than the intact Ma, with Ma-W showing the largest retention. Clearly, the HILIC method is more selective towards differences in neutral sugar content than towards variation in the amino acid composition. Still, as indicated above, because of the high glycoform resolution obtained, reliable annotation of the truncation variants becomes possible.

The elution conditions used in HILIC-MS reduced the sensitivity as compared to RPLC-MS, requiring about four-time more sample to be injected in order to obtain comparable signal intensities. However, about three times more proteoforms could be discerned when using HILIC (108 vs. 39). The deconvoluted mass spectra generated for the entire protein elution window with the two separation techniques are depicted in Fig. 5 and Table S1 provides the corresponding list of assigned masses. The two techniques give similar results for the distribution of glycoforms carrying one glycan only (Fig. 5b). However, with RPLC-MS the lower abundant proteoforms, such as the truncated variants of glycoforms with 2 and 3 glycosylation sites occupied, were hardly or not detected (Figs. 5a and 5c-d). This is most likely due to strong coelution of the present proteoforms which causes ionization suppression and reduction of dynamic range.

### 3.4. SDS-PAGE and bottom-up LC-MS of HILIC-fractionated lipase

In order to localize the sites of glycosylation of the different glycoforms and substantiate our observation on the similarity of the HILIC and SDS-PAGE profiles for lipase (Fig. 3c), we fractionated the lipase by HILIC and subsequentially performed SDS-PAGE on the separate fractions. As a compromise between sample loading and separation efficiency in HILIC, 3 μl of a 3 mg $mL^{-1}$ solution of lipase in 50:50 ACN-water containing 0.1% TFA was injected (for more details, see Section S2). Four successive fractions (F1-4) were collected, covering the elution windows indicated in the UV chromatogram (Fig. 6a). Subsequent SDS-PAGE of the fractions indeed revealed that the molecular weight of the eluting proteins increased with HILIC retention (Fig. 6b). The elution order followed the previous MS observations with non-glycosylated lipase (F1) eluting first, followed by lipase with one (F2), two (F3) and three (F4) glycosylation sites occupied.

To further study the specific glycosylation site occupancy of the lipase proteoforms, the proteins present in the HILIC fractions F1-4 of native and deglycosylated lipase were digested using trypsin. Deglycosylation of lipase was achieved using the amidase PNGase F, which cleaves the N-glycan bond resulting in the free glycan and deamidation of the corresponding asparagine residue,

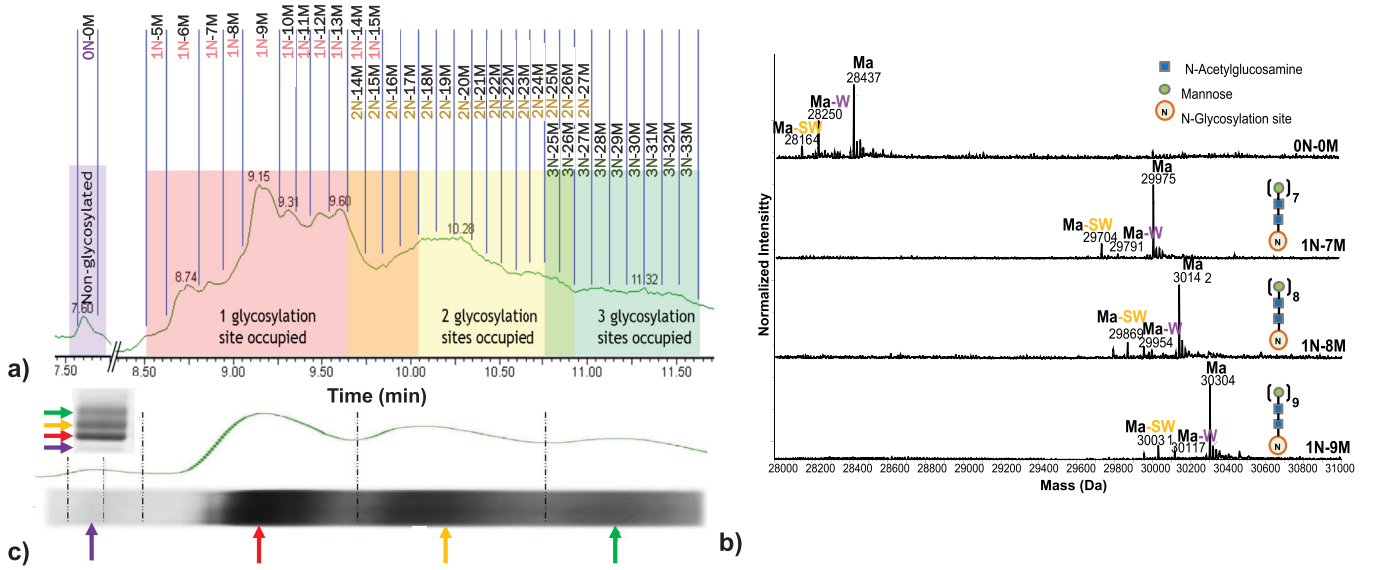



**Fig. 3.** a+b) HILIC-MS of lipase with a) TIC with annotated glycoform elution windows (xN-yM) indicating (x) the number of N-glycosylation sites occupied (0–3) and (y) the number of mannose units present next to a single N-acetylglucosamine present in each glycan; b) Deconvoluted mass spectra obtained for the corresponding elution windows (from above the average times were respectively 7.5–7.7 min; 8.7–8.8 min, 8.8–9.0 min and 9.0–9.3 min) which show signals for the mature form of the protein (intact sequence, Ma), mature form without a C-terminal tryptophan (Ma-W) and a mature form without a C-terminal serine and tryptophan (Ma-SW). c) SDS-PAGE of lipase (bottom trace and left top corner) with derived densitogram (top trace) aligned with the TIC trace obtained with HILIC-MS.

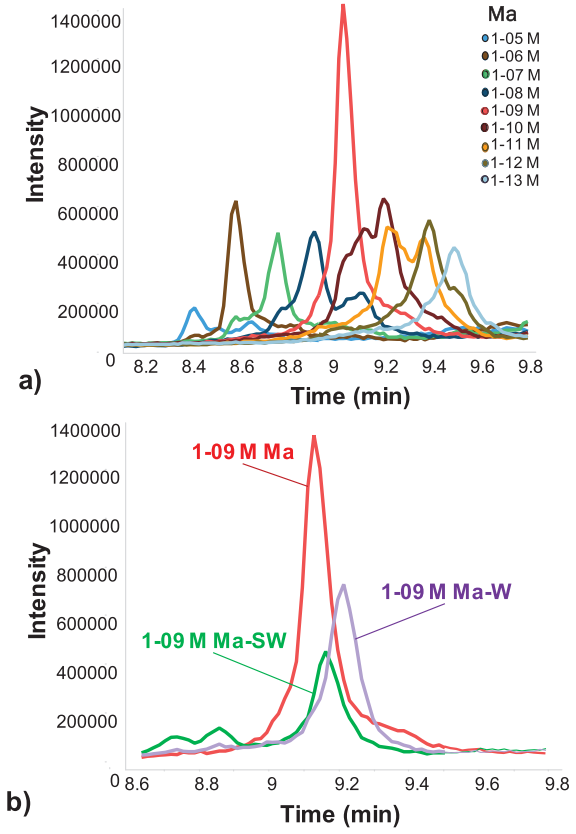



**Fig. 4.** HILIC-MS of lipase. a) EICs for non-truncated glycoforms having one glycan comprising 5–13 mannose units (1–05 M to 1–13 M). b) EICs of intact (1–09 M Ma) and truncated variants (1–09 Ma-W and 1–09 Ma-SW) of a glycoform having one glycan comprising 9 mannose units.

transforming it into an aspartic acid residue. The obtained eight protein digests were analyzed by RPLC-MS/MS, particularly aiming to characterize the specific four tryptic peptides which carry a lipase N-glycosylation site. When in the deglycosylated sample the respective peptide was detected in its deamidated form, the glycosylation site was considered occupied in the native lipase [25]. We preferred to use this simple approach for determining glycosylation site occupancy over the direct identification of the respective glycopeptides. The latter often show reduced ESI efficiency, especially when the glycans are large, potentially hindering their reliable detection.

The results of the bottom-up RPLC-MS approach are summarized in Table 1 and section S5 (Table S3, Figs. S5 and S6). In HILIC fraction F1 of both the native and deglycosylated lipase, no deamidated peptides were detected, confirming that non-glycosylated proteoforms of lipase elute first. Moreover, as for the native lipase no deamidated peptides were observed in F1, we concluded that the sample pretreatment does not cause deamidation, excluding false positives. The analysis of the subsequent fractions F2-4 revealed that the lipase proteoforms eluting in fraction F2 were predominantly singly glycosylated at N63 only. The proteoforms in fraction F3 were glycosylated at N63 and N215 (or, to a minor extent, N151), and the proteoforms in fraction F4 were glycosylated at N63, N215 and N258 (and to a minor extent at N151 instead of N215 or N258). The native form of the peptide comprising the N63 position was not detected in any of the fractions, however, its deamidated form was observed in fractions F2-4.

Finally, we tested whether the lipase could maintain its activity after being exposed to the HILIC conditions. This would allow probing of possible differences in activity among the glycoforms after HILIC fractionation. When reconstituting the lipase after its exposure to 70% ACN comprising 1% propionic acid and 0.025% TFA, the enzyme did not show a measurable activity (data not shown). These results confirm that the harsh HILIC separation conditions (high organic content, strong acid, and high column temperature) may, unfortunately, impair the functionality of proteins.

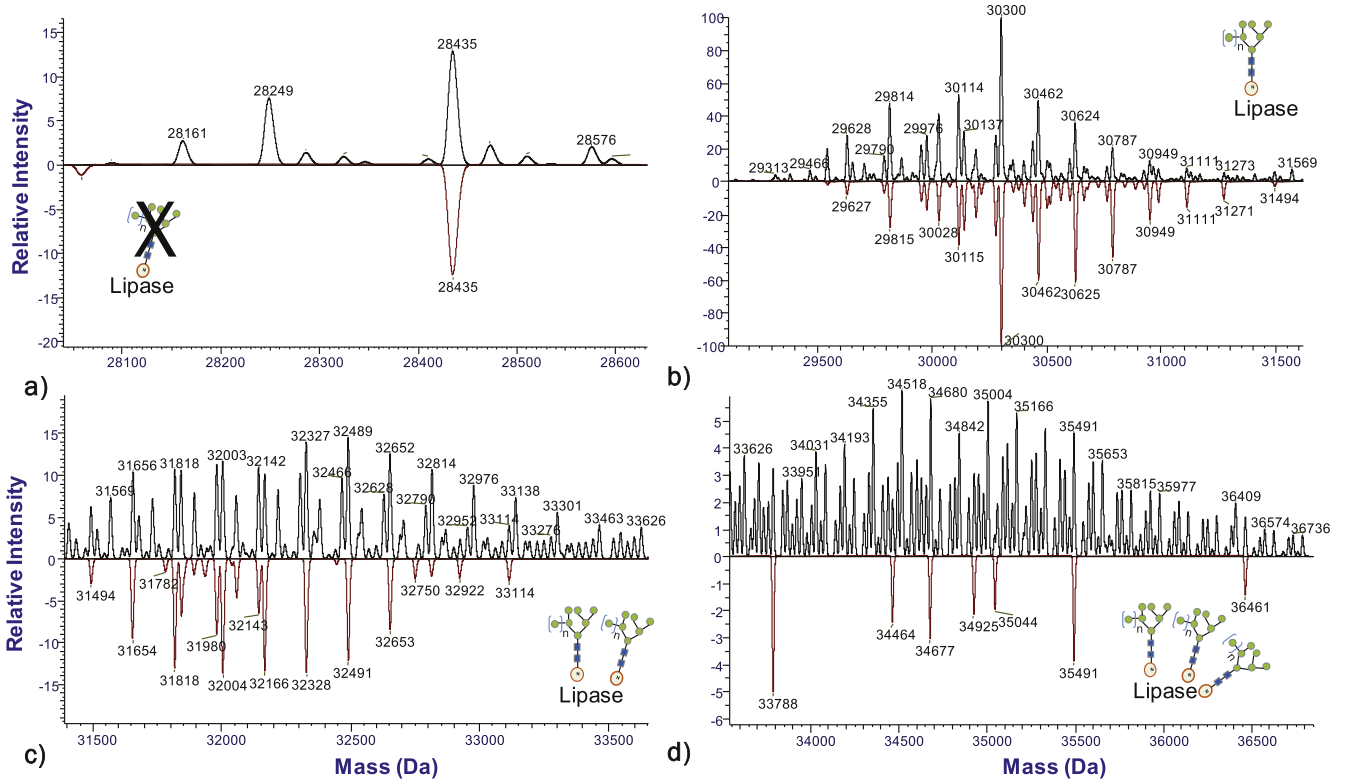



**Fig. 5.** Average lipase deconvoluted mass spectra constructed from RPLC-MS (bottom red) and HILIC-MS (top black) data covering the respective protein elution window. Figures a, b, c, d show the parts of the deconvoluted spectra corresponding to non-glycosylated, singly, doubly and triply glycosylated lipase, respectively. Details of the assignment of the singly glycosylated lipase forms is reported in Fig. S3; further comparison between HILIC-MS and RPLC-MS deconvoluted spectra can be found in Fig. S4. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

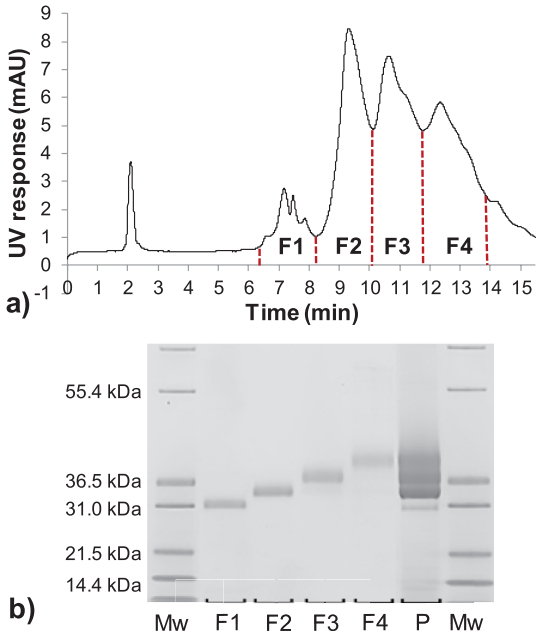



**Fig. 6.** a) HILIC-UV of lipase with indicated elution windows corresponding to fractions F1-4; b) SDS-PAGE of fractions F1-4 a molecular weight marker (Mw) and unfractionated lipase (P). To facilitate the detection of the band on the gel, F1 was collected from a HILIC-UV run of the deglycosylated protein. Detection wavelength, 280 nm; further conditions, see Experimental Section.

## 4. Conclusions

The detailed characterization of highly glycosylated proteins, such as industrial enzymes and biopharmaceutical products is essential, but challenging. Intact proteins analysis methods using RPLC-MS have limited resolving power for protein glycoforms, potentially leaving a significant number of them undetected.

This study shows that HILIC offers more detailed information on intact protein heterogeneity due to its unique selectivity towards glycoforms. When combined with MS, HILIC significantly reduces the complexity of protein sample delivered to the mass spectrometer at each time point, substantially increasing the number of glycoproteoforms mass analyzed. Here, we specifically show the potential of HILIC-MS to characterize a commercial lipase, also in comparison with RPLC-MS. Moreover, in order to better understand the elution of the lipase glycoforms in HILIC in terms of glycosylation site occupancy, we took HILIC fractions of lipase and subjected these to SDS-PAGE analysis as well as to tryptic digestion followed by LC-MS/MS to characterize the proteoforms at the peptide level.

Overall, HILIC-MS enabled assigning the full width of lipase glycoforms as well as pinpointing the presence of truncation variants which largely remained unrevealed in RPLC-MS. Close to three times more lipase (glyco)proteoforms could be observed with HILIC-MS as compared to RPLC-MS. The high resolving power and characterization capability of HILIC-MS clearly can be a valuable tool in monitoring the quality of glycosylated industrial enzymes during development and production, particularly for proteins with neutral high-mannose type of N-glycosylation.

**Table 1**
Native (non-glycosylated) peptides (A or a) and deamidated peptides (B or b) observed during RPLC-MS/MS of tryptic digests of HILIC fractions of native and deglycosylated (D) lipase (Fig. 6). Theoretical and observed masses are reported in Table S1. X indicates that the peptide was not detected, A and B indicate that the peptide was relatively abundant (abundance above 20%), a and b indicate that the peptide was present in relatively low abundance (below 20%).

| Peptide (AA) | Sequence | F1 | F1D | F2 | F2D | F3 | F3D | F4 | F4D |
|---|---|---|---|---|---|---|---|---|---|
| 1 (N63) | FYIQHGAAAYCNSETAAGA**NVT**CTGNA CPEIEANGVTVVASFTGTK[a] | X | X | X | **B** | X | **B** | X | **B** |
| 2 (N151) | AWEEIAD**NLT**AAVAK | A | A | A | A | A | **Ab** | A | **Ab** |
| 3 (N215) | VGNAELAEFIS**NQT**GGEFR | A | A | A | A | A | **aB** | A | **aB** |
| 4 (N258) | I**NYT**INDIK | A | A | A | A | A | A | A | **aB** |

[a] Cysteines have been treated with iodoacetamide to form carbamidomethyl-cysteine.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## CRediT authorship contribution statement

**A.F.G. Gargano:** Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization, Supervision, Funding acquisition, Project administration. **O. Schouten:** Resources, Visualization, Formal analysis, Investigation. **G. van Schaick:** Investigation, Visualization. **L.S. Roca:** Investigation. **J.H. van den Berg-Verleg:** Investigation, Visualization. **R. Haselberg:** Conceptualization, Writing - review & editing, Resources. **M. Akeroyd:** Conceptualization, Writing - review & editing, Funding acquisition, Project administration. **N. Abello:** Conceptualization, Visualization, Writing - review & editing, Funding acquisition, Project administration. **G.W. Somsen:** Conceptualization, Writing - original draft, Writing - review & editing, Funding acquisition, Project administration.

## Acknowledgments

AG acknowledges financial support by the Netherlands Organization for Scientific Research, NWO Veni grant IPA (722.015.009). LR acknowledges the STAMP project, which is funded under the Horizon 2020 – Excellent Science – European Research Council (ERC), Project 694151. The sole responsibility of this publication lies with the authors. The European Union is not responsible for any use that may be made of the information contained therein. The authors would like to thank Peter J. Schoenmakers, and Eva M. Johansson (University of Amsterdam) for their help and valuable discussions.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.aca.2020.02.042.

## References

[1] P. Gagneux, M. Aebi, A. Varki, Evolution of glycan diversity, in: Essentials Glycobiol, third ed., 2017, https://doi.org/10.1101/glycobiology.3e.020.
[2] D. Skropeta, The effect of individual N-glycans on enzyme activity, Bioorg. Med. Chem. 17 (2009) 2645–2653, https://doi.org/10.1016/j.bmc.2009.02.037.
[3] T. Jeoh, W. Michener, M.E. Himmel, S.R. Decker, W.S. Adney, Implications of cellobiohydrolase glycosylation for use in biomass conversion, Biotechnol, Biofuels 1 (2008) 10, https://doi.org/10.1186/1754-6834-1-10.
[4] M. Yang, X.W. Yu, H. Zheng, C. Sha, C. Zhao, M. Qian, Y. Xu, Role of N-linked glycosylation in the secretion and enzymatic properties of Rhizopus chinensis lipase expressed in Pichia pastoris, Microb. Cell Factories 14 (2015) 1–14, https://doi.org/10.1186/s12934-015-0225-5.
[5] J. Wolle, H. Jansen, L.C. Smith, L. Chanl, Functional role of N-linked glycosylation in human hepatic lipase: asparagine 56 is important for both enzyme activity and secretion 34 (1993) 2169–2176.
[6] C. Wicker-Planquart, S. Canaan, M. Rivière, L. Dupuis, Site-directed removal of N-glycosylation sites in human gastric lipase, Eur. J. Biochem. 262 (1999) 644–651, https://doi.org/10.1046/j.1432-1327.1999.00427.x.
[7] H.J. An, T.R. Peavy, J.L. Hedrick, C.B. Lebrilla, Determination of N-glycosylation sites and site heterogeneity in glycoproteins, Anal. Chem. 75 (2003) 5628–5637, https://doi.org/10.1021/ac034414x.
[8] M. Wuhrer, C.A.M. Koeleman, C.H. Hokke, A.M. Deelder, Protein glycosylation analyzed by normal-phase nano-liquid chromatography-mass spectrometry of glycopeptides, Anal. Chem. 77 (2005) 886–894, https://doi.org/10.1021/ac048619x.
[9] E. Balaguer, C. Neusüss, Glycoprotein characterization combining intact protein and glycan analysis by capillary electrophoresis-electrospray ionization-mass spectrometry, Anal. Chem. 78 (2006) 5384–5393, https://doi.org/10.1021/ac060376g.
[10] E. Domínguez-Vega, S. Tengattini, C. Peintner, J. van Angeren, C. Temporini, R. Haselberg, G. Massolini, G.W. Somsen, High-resolution glycoform profiling of intact therapeutic proteins by hydrophilic interaction chromatography-mass spectrometry, Talanta 184 (2018) 375–381, https://doi.org/10.1016/j.talanta.2018.03.015.
[11] G.W. Somsen, C. Temporini, G. Massolini, L. Pollegioni, E. Domínguez-Vega, F. Rinaldi, S. Tengattini, L. Piubelli, T. Bavaro, Hydrophilic interaction liquid chromatography-mass spectrometry as a new tool for the characterization of intact semi-synthetic glycoproteins, Anal. Chim. Acta 981 (2017) 94–105, https://doi.org/10.1016/j.aca.2017.05.020.
[12] R. Haselberg, G.J. De Jong, G.W. Somsen, Low-flow sheathless capillary electrophoresis-mass spectrometry for sensitive glycoform profiling of intact pharmaceutical proteins, Anal. Chem. 85 (2013) 2289–2296, https://doi.org/10.1021/ac303158f.
[13] Y. Jin, Z. Lin, Q. Xu, C. Fu, Z. Zhang, Q. Zhang, W.A. Pritts, Y. Ge, Comprehensive characterization of monoclonal antibody by Fourier transform ion cyclotron resonance mass spectrometry, mAbs 11 (2019) 106–115, https://doi.org/10.1080/19420862.2018.1525253.
[14] V. D'Atri, S. Fekete, A. Beck, M. Lauber, D. Guillarme, Hydrophilic interaction chromatography hyphenated with mass spectrometry: a powerful analytical tool for the comparison of originator and biosimilar therapeutic monoclonal antibodies at the middle-up level of analysis, Anal. Chem. 89 (2017) 2086–2092, https://doi.org/10.1021/acs.analchem.6b04726.
[15] A.F.G. Gargano, L.S. Roca, R.T. Fellers, M. Bocxe, E. Domínguez-Vega, G.W. Somsen, Capillary HILIC-MS: a new tool for sensitive top-down proteomics, Anal. Chem. 90 (2018) 6601–6609, https://doi.org/10.1021/acs.analchem.8b00382.
[16] F. Füssl, K. Cook, K. Scheffler, A. Farrell, S. Mittermayr, J. Bones, Charge variant analysis of monoclonal antibodies using direct coupled pH gradient cation exchange chromatography to high-resolution native mass spectrometry, Anal. Chem. 90 (2018) 4669–4676, https://doi.org/10.1021/acs.analchem.7b05241.
[17] A. Ehkirch, O. Hernandez-Alba, O. Colas, A. Beck, D. Guillarme, S. Cianférani, Hyphenation of size exclusion chromatography to native ion mobility mass spectrometry for the analytical characterization of therapeutic antibodies and related products, J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 1086 (2018) 176–183, https://doi.org/10.1016/j.jchromb.2018.04.010.
[18] B. Chen, Y. Peng, S.G. Valeja, L. Xiu, A.J. Alpert, Y. Ge, Online hydrophobic interaction chromatography-mass spectrometry for top-down proteomics, Anal. Chem. 88 (2016) 1885–1891, https://doi.org/10.1021/acs.analchem.5b04285.
[19] A. Periat, S. Fekete, A. Cusumano, J.-L. Veuthey, A. Beck, M. Lauber, D. Guillarme, Potential of hydrophilic interaction chromatography for the analytical characterization of protein biopharmaceuticals, J. Chromatogr., A 1448 (2016) 81–92, https://doi.org/10.1016/j.chroma.2016.04.056.
[20] V. D'Atri, L. Nováková, S. Fekete, D. Stoll, M. Lauber, A. Beck, D. Guillarme, Orthogonal middle-up approaches for characterization of the glycan heterogeneity of etanercept by hydrophilic interaction chromatography coupled to high-resolution mass spectrometry, Anal. Chem. 91 (2019) 873–880, https://doi.org/10.1021/acs.analchem.8b03584.
[21] F.E. Kuhlmann, A. Apffel, S.M. Fischer, G. Goldberg, P.C. Goodley, Signal enhancement for gradient reversed-phase high-performance liquid chromatography-electrospray ionization mass spectrometry analysis with

trifluoroacetic and other strong acid modifiers by postcolumn addition of propionic acid and isopropanol, J. Am. Soc. Mass Spectrom. 6 (1995) 1221–1225.
[22] A.F.G. Gargano, L.S. Roca, R.T. Fellers, M. Bocxe, E. Domínguez-Vega, G.W. Somsen, Capillary HILIC-MS: a new tool for sensitive top-down proteomics, Anal. Chem. 90 (2018) 6601–6609, https://doi.org/10.1021/acs.analchem.8b00382.
[23] G. van Schaick, B.W.J. Pirok, R. Haselberg, G.W. Somsen, A.F.G. Gargano, Computer-aided gradient optimization of hydrophilic interaction liquid chromatographic separations of intact proteins and protein glycoforms, J. Chromatogr., A (2019) 67–76, https://doi.org/10.1016/j.chroma.2019.03.038.
[24] B. Bobály, V. D'Atri, A. Beck, D. Guillarme, S. Fekete, Analysis of recombinant monoclonal antibodies in hydrophilic interaction chromatography: a generic method development approach, J. Pharmaceut. Biomed. Anal. 145 (2017) 24–32, https://doi.org/10.1016/j.jpba.2017.06.016.
[25] V. Shetty, Z. Nickens, P. Shah, G. Sinnathamby, O.J. Semmes, R. Philip, Investigation of sialylation aberration in N-linked glycopeptides by lectin and tandem labeling (LTL) quantitative proteomics, Anal. Chem. 82 (2010) 9201–9210, https://doi.org/10.1021/ac101486d.