Hydrophilic interaction liquid chromatography-mass spectrometry as a new tool for the characterization of intact semi-synthetic glycoproteins

Sara Tengattini[1,2,#], Elena Domínguez-Vega[2,#], Caterina Temporini[1*], Teodora Bavaro[1], Francesca Rinaldi[1], Luciano Piubelli[3], Loredano Pollegioni[3], Gabriella Massolini[1], Govert W. Somsen[2]

[1]*Department of Drug Sciences, University of Pavia, via Taramelli 12, I-27100 Pavia, Italy*

[2]*Division of BioAnalytical Chemistry, Department of Chemistry and Pharmaceutical Sciences, Vrije Universiteit Amsterdam, de Boelelaan 1085, 1081 HV Amsterdam, The Netherlands*

[3] *Department of Biotechnology and Life Sciences, University of Insubria, via J.H. Dunant 3, I-21100 Varese (Italy) and The Protein Factory Interuniversity Centre, Politecnico of Milano and University of Insubria, via Mancinelli 7, I-20131, Milano (Italy)*

[#]Contributed equally

*Corresponding author caterina.temporini@unipv.it

Abstract

Improved methods for detailed characterization of complex glycoproteins are required in the growing sector of biopharmaceuticals. Hydrophilic interaction liquid chromatography (HILIC) coupled to high resolution (HR) time-of-flight mass spectrometric (TOF-MS) detection was examined for the characterization of intact *neo*-glycoproteins prepared by chemical conjugation of synthetic saccharides to the lysine residues of selected recombinant proteins. The separation performances of three different amide HILIC columns (TSKgel

Amide-80, XBridge BEH and AdvanceBio Glycan Mapping) were tested. Water-acetonitrile gradients and volatile eluent additives have been explored. Addition of 0.05% (v/v) trifluoroacetic acid to the mobile phase appeared to be essential for achieving optimum resolution of intact glycoforms and minimal ion suppression effects. Gradient elution conditions were optimized for each protein on every column. HILIC stationary phases were evaluated for the analysis of highly heterogeneous semi-synthetic derivatives of the same protein (ribonuclease A), and in the enhanced characterization of TB10.4 and Ag85B glycoconjugates, selected antigens from *Mycobacterium tuberculosis* (MTB). HILIC-MS results indicated that the HILIC selectivity is predominantly governed by size of the conjugated glycans and number of glycans attached, providing efficient glycoform separation. Moreover, HILIC separation prior to HRMS detection allowed assignment of several product impurities. Additional top-down MS/MS experiments confirmed conjugation at the N-terminus of TB10.4 next to its lysine residue. Overall, the obtained results demonstrate that amide-stationary-phase based HILIC coupled to MS is highly useful for the characterization of intact *neo*-glycoproteins allowing assessment of the number, identity and relative abundance of glycoforms present in the semi-synthetic products.

Keywords

Hydrophilic interaction liquid chromatography; *neo*-glycoproteins; intact protein analysis; mass spectrometry.

1. Introduction

Biopharmaceutical products, drugs of biological origin, represent an important class of therapeutics. In particular, glycoproteins are among the most commercially successful biological drugs due to their numerous applications and indications ranging from cancer to inflammation to infectious diseases (glycovaccines). It is well known that glycans attached to a protein can influence the therapeutic efficacy of protein drugs and the protein physicochemical and pharmacological properties. Thus, there is a need for the production of well-defined glycoproteins and elegant synthetic strategies have been explored for the preparation of homogeneous glycoproteins with particular emphasis on glycovaccines.

An interesting approach, often used for the production of glycoconjugate vaccines, entails the extraction or the synthesis of saccharide (antigens) and their covalent linkage to a carrier protein produced by recombinant DNA technology. The conjugation is commonly achieved by chemical activation of the glycan portion allowing targeting of nucleophilic functional groups of specific amino acid residues (lysines, aspartic/glutamic acids or cysteines) in the protein [1, 2].

The characterization of *neo*-glycoprotein products, obtained as described above, is quite challenging. In fact, in addition to the inevitable heterogeneity of recombinant proteins due to the biotechnological manufacturing, the heterogeneity related to the carbohydrate component (structure, number, and position) has to be considered. Similarly to native glycoproteins, also for semisynthetic glycosylated proteins, the inherent structural complexity (i.e. sequence mutation, oxidation, deamidation, N- and C-terminal alteration, glycoform variety due to multiple conjugation sites) necessitates the development of analytical strategies for their characterization [3].

Liquid chromatography (LC) coupled to electrospray ionization mass spectrometry (ESI-MS) is widely used for glycoprotein characterization by bottom-up approaches, in which the glycoproteins are first digested into peptides, providing structural information and allowing localization of glycosylation sites [4, 5]. Nevertheless, quality control of the intact glycoproteins is of great added value in biopharmaceutical analysis, as it provides specific information on the presence of proteoforms as well as on the exact combinations of multiple modifications, which cannot be revealed by merely bottom-up approaches. Moreover, intact protein analysis is relatively fast, requires minimum sample treatment and prevents undesired modifications induced by enzymatic treatments [6, 7].

Hydrophilic interaction liquid chromatography (HILIC) has demonstrated to be a highly useful technique for analysis of amino acids, released glycans and glycopeptides [8]. In contrast to reversed-phase (RP) LC, HILIC has the ability to retain and resolve (highly) polar compounds, based on a complex retention mechanism, involving hydrophilic partitioning and polar interactions. As a consequence, glycans and glycopeptides show stronger HILIC retention with an increasing number and size of sugar units [9, 10].

Only few studies have been published on the use of HILIC methods for the characterization of intact proteins [8, 11, 12] and for the determination of their glycoform composition [11-14], indicating HILIC to be complementary and orthogonal to RPLC. So far, the application of HILIC for the separation of native glycoforms of intact proteins has been limited to the model glycoprotein ribonuclease B and monoclonal antibodies [11, 12]. Recently, we have reported a preliminary work where HILIC-UV was used for the monitoring of the glycosylation reaction of a model protein containing only one glycosylation site (ribonuclease A) with different glycans [15]. The developed analytical method was not compatible with MS detection, however, it was suitable for the analysis of intact *neo*-glycoproteins, allowing discrimination of non-glycosylated proteins from

glycoproteins and separation of different glycoforms, including isomers, as demonstrated by off-line MS analysis [15].

On-line MS detection after HILIC separation of glycoforms, would offer many advantages, such as the assignment of glycoform structures and other potential protein modifications and degradations. Moreover, following a top-down approach, HILIC-MS/MS could provide information on N-terminal sequence, including glycosylation occurrence at the α-amino group. Nevertheless, on-line HILIC-MS of intact glycoproteins has been reported only for ribonuclease B and for IdeS-digested and reduced monoclonal antibodies [11,12,14].

In order to expand the applicability of HILIC with UV and HRMS detection to the characterization of intact *neo*-glycoproteins, we studied the chromatographic performances of three different amide HILIC columns (TSKgel Amide-80, XBridge BEH and AdvanceBio Glycan Mapping). *Neo*-glycoproteins, obtained by chemical conjugation of synthetic mono-, di- and tri-mannose to lysine residues of ribonuclease A and of the two recombinant tuberculosis antigens TB10.4 and Ag85B [16], were considered as representative real-life proteins in this work. TB10.4 and Ag85B have been evaluated by our research group for the development of new effective vaccines against tuberculosis [3, 17]. Their semi-synthetic glycoconjugates comprise a large number of glycoforms with multiple conjugation sites, each carrying short carbohydrate chains. This intrinsic heterogeneity combined with the relatively small hydrophilic contribution of every individual carbohydrate chain, makes chromatographic resolution and characterization of these *neo*-glycoproteins an interesting case for the application of the three amide HILIC columns. MS-compatible mobile phase composition, gradients and column temperature were studied in order to achieve good separation of the semi-synthethic glycoprotein products. Moreover, the HILIC-MS results obtained aid in understanding structure-retention relationships of glycoproteins. The

potential of the developed HILIC-MS(/MS) methods for the detection of glycoconjugate product impurities and for the confirmation of the glycan position in conjugated TB10.4 was also examined.

2. Experimental

2.1 Reagents

RNase A and ribonuclease B (RNase B) from bovine pancreas were purchased from Sigma-Aldrich (St. Louis, MO, USA) and were used without further purification. Formic acid (FA), ammonium hydroxide (30%), sodium tetraborate, trifluoroacetic acid (TFA), phosphate-buffered saline, benzamidine chloride were from Sigma-Aldrich (St. Louis, MO, USA). Water was obtained from a Direct-QTM Millipore system Millipore (Millipore, Billerica, MA, USA). Acetonitrile (ACN) MS grade was purchased from Sigma-Aldrich (St. Louis, MO, USA).

In this study three IME-thioglycosides were used for protein glycosylation: Mannose-IME (Man-IME), Mannose(1-6)Mannose-IME (Man(1-6)Man-IME) and Mannose(1-6)Mannose(1-6)Mannose-IME (Man(1-6)Man(1-6)Man-IME). IME-thioglycosides were *ad hoc* synthetized according to the procedure previously reported [18] (see Supplementary Material). The TB10.4 and Ag85B immunogenic proteins were obtained as recombinant forms in *Escherichia coli* as reported in Piubelli et al. [16] and finally collected in 20 mM 3-(N-morpholino)propanesulfonic acid, 0.4 M NaCl, pH 7.0 at different concentrations.

2.2 Protein glycosylation

2.2.1 RNase A glycosylation procedure

According to the glycosylation protocol previously reported [15], the reaction was carried out in sodium tetraborate buffer, 100 mM, pH 9.5. RNase A was dissolved in the buffer to reach a

final concentration of 1.7 mg mL^{-1} and then the solution was mixed with IME-glycoside to a glycoside/protein molar ratio of 200/1. The reaction mixture was vortexed for 1 min and incubated for 24 h at 25 °C under continuous stirring. Based on thioglycosides used, the prepared glycoconjugates were named RNase-Man and RNase-Man(1-6)Man.

2.2.2 TB10.4 and Ag85B glycosylation procedure

The conditions used for the antigenic protein glycosylation required a protein concentration of 5.5 mg mL^{-1} and a glycoside/protein molar ratio of 200/1 in the same reaction buffer as above. Benzamidine chloride (1 mM) was added to prevent protein digestion by traces of enterokinase derived from protein production. Five *neo*-glycoproteins were synthesized, namely TB10.4-Man, TB10.4-Man(1-6)Man, Ag85B-Man, Ag85B-Man(1-6)Man and Ag85B-Man(1-6)Man(1-6)Man.

2.3 Sample preparation

Standard stock solutions of RNase B were prepared in pure water at a concentration of 2 mg mL^{-1} and then diluted with ACN/water to obtain working solutions at 0.5 mg mL^{-1} in ACN/water (50:50, v/v). *Neo*-glycoproteins, prepared from RNase A and antigenic proteins, were purified in order to remove reagents and salts. The reaction mixture was submitted to seven 20 min steps of ultrafiltration at 13000 g and 4 °C using centrifuge 5804-R (Eppendorf s.r.l., Milan, Italy) and Millipore's Amicon® Ultra filters with a nominal molecular weight limit of 3 or 10 kDa and load capacity of 500 µL. Proteins and *neo*-glycoproteins were finally collected and stored in PBS at concentration of 2 mg mL^{-1}. Prior HILIC-UV or HILIC-MS analysis samples were diluted to obtained a final concentration of 0.5 mg mL^{-1} in ACN/water/PBS (50:25:25, v/v/v).

2.4 Chromatographic equipment and conditions

Chromatographic separations of intact (*neo*-)glycoproteins were performed on an Agilent HPLC series 1200 system, equipped with a mobile-phase online degasser, quaternary pump, autosampler, column thermostated compartment, and diode array detector. For data acquisition and analysis, ChemStation software version Rev. B.04.01 was used. The HILIC columns studied were (1) TSKgel Amide-80 (150 x 2.0 mm, 3 μm, 80 Å) from Tosoh Biosciences Bioscience (Montgomeryville, PA, USA), (2) XBridge BEH Amide (150 x 3.0 mm, 2.5 μm, 130 Å) from Waters (Milford, MA, USA), and (3) AdvanceBio Glycan Mapping (150 x 2.1 mm, 2.7 μm) from Agilent Technologies (Palo Alto, CA, USA). The injection volume was fixed at 2 μL, the column temperature at 50 °C and the flow rate was set at 0.3 mL min^{-1} for the TSKgel Amide-80 column, and at 0.5 mL min^{-1} for the other two HILIC columns.

The final mobile phases were composed of ACN (A) and water (B) both containing 0.05% (v/v) TFA. Gradient elution conditions were optimized for each protein on every column. For RNase B and RNase glycoconjugates: with TSKgel Amide-80 from 32% to 42% B in 20 min followed by isocratic elution at 42% B for 10 min; with XBridge BEH Amide and AdvanceBio Glycan Mapping from 28% to 38% B in 20 min followed by isocratic elution at 38% B for 10 min. For TB10.4 glycoconjugates: with TSKgel Amide-80 from 21% to 31% B in 20 min followed by isocratic elution at 31% B for 10 min; with XBridge BEH Amide and AdvanceBio Glycan Mapping from 17% to 27% B in 20 min followed by isocratic elution at 27% B for 10 min. For Ag85B glycoconjugates: with TSKgel Amide-80 from 26% to 36% B in 20 min followed by isocratic elution at 36% B for 10 min; with XBridge BEH Amide and AdvanceBio Glycan Mapping from 22% to 32% B in 20 min followed by isocratic elution at 32% B for 10 min. UV absorbance was monitored at 214 nm.

2.5 MS detection

MS detection was performed using a maXis HD ultra-high resolution quadrupole time-of-flight (QTOF) mass spectrometer (Bruker Daltonics, Bremen, Germany) with an ESI source. The mass spectrometer was operated in positive-ion mode with an electrospray voltage of 4.5 kV. The nebulizer and drying gas conditions were 1.0 bar and 8.0 L/min nitrogen at 200 °C, respectively. Quadrupole ion and collision cell energy were 3.0 and 15.0 eV, respectively. Transfer and pre plus storage times were set at 190.0 and 20.0 μs, respectively. In source collision induced dissociation (ISCID) was set at 120 eV in order to dissociate protein-TFA adducts formed during ESI. The monitored mass range was 250−4000 m/z. Extracted ion chromatograms were obtained with an extraction window of ±0.5 m/z and using the smooth option of the software (Gaussian at 1 point). For fragmentation experiments, the most abundant ion of each protein was fragmented by collision ion dissociation (CID) using multiple reaction monitoring mode. The collision energy was separately optimized for each species and the final applied values were in the range 90-110 eV.

3. Results and discussion

In this work, HILIC-UV-MS was evaluated for the characterization of *neo*-glycoproteins. Three different commercial HILIC columns were investigated for their capacity to reduce heterogeneity complexity by resolving intact glycoforms as well as specific proteoforms to facilitate their assignment. Initial chromatographic optimization was carried out by HILIC-UV using MS-compatible conditions and RNase B as model glycoprotein. Then, HILIC-MS methods were applied to the analysis of semi-synthetic glycoproteins prepared starting from RNase A and from two recombinant tuberculosis antigen proteins, TB10.4 and Ag85B.

3.1 Development of a HILIC method for intact glycoform resolution

Previous studies of our research group have shown that a carbamoyl-silica-based stationary phase (TSKgel Amide-80) has good potential for the glycoform profiling of intact glycoproteins and *neo*-glycoproteins [15]. Encouraged by these interesting results, in the present study we also considered two other amide HILIC columns (XBridge BEH Amide and AdvanceBio Glycan Mapping) for glycoprotein analysis. The performance of the three HILIC columns was studied under MS-compatible conditions using the glycoprotein RNase B (MW 15 kDa) as test compound. RNase B is N-glycosylated at Asn 34 giving rise to five glycoforms differing in the number of mannose residues (5-9). Mobile phases consisting of ACN and water with different concentrations of FA, ammonium formate and TFA were tested. The injection volume was fixed at 2 μL in order to avoid the deterioration of peak shape due to the relatively high percentage of water (50% v/v) in the diluents necessary for full protein solubilization.

The eluent composition and gradient conditions were adjusted to obtain an appropriate elution window for RNase B. However, with the addition of 0.1% (v/v) FA or 10 mM ammonium formate (adjusted at pH 3.0) no separation among the five glycoforms was observed with each of the tested HILIC columns. Improved glycoform separation was obtained by increasing the percentage FA to 0.5% (v/v) or ammonium formate concentration to 100 mM (when allowed by column specifications), and by rising the column temperature from 25 to 50 °C. However, the obtained glycoform resolution was still unsatisfactory. Temperatures above 50 °C were tested for the XBridge BEH column (as permitted according to column specifications), but not significant increase in separation quality was observed.

Using a column temperature of 50 °C, the effect of TFA as alternative eluent additive [11, 14] on the three HILIC columns was studied. The addition of 0.025% (v/v) of TFA to the mobile phase resulted in an increased resolution of the five glycoforms of RNase B. The

concentration of TFA was next evaluated in the range 0.025%-0.1%. Higher percentages were not considered, as TFA is known to suppress ESI [19-21]. For all three HILIC columns, the overall peak profile of the RNase B glycoforms was hardly affected by increasing the concentration of TFA in the mobile phase. For instance, for the XBridge BEH column using a gradient of 30% to 40% water in 10 min followed by isocratic elution at 40% water for 10 min, the average distance between the apices of the consecutive glycoform peaks was 0.43 ± 0.06 min, 0.49 ± 0.06 min and 0.49 ± 0.05 min at 0.025%, 0.05% and 0.1% TFA, respectively. However, more narrow peaks were observed with rising TFA concentration, leading to an overall better glycoform resolution. Indeed, with the same column, the average resolution for consecutive peaks was 0.95, 1.19 and 1.25 using, respectively, 0.025%, 0.05% and 0.1% TFA (Fig. S1 in Supplementary data). From these experiments it is clear that TFA must be used to achieve satisfactory separation of the glycoforms. Thus, for further experiments we selected a TFA concentration of 0.05% to achieve good glycoform resolution while preventing excessive ion suppression in MS detection.

Figure 1 shows the HILIC-UV chromatograms for RNase B obtained under optimized conditions and Table 1 lists the corresponding Rs values for the separated glycoforms. A significantly lower resolution of the RNase B glycoforms was observed for the TSKgel Amide-80 column (average Rs 1.0 ± 0.1). Considering the similar peak spacing provided by the three stationary phases, the reduced separation power of this column is a result of the protein peak broadening, most probably due to the small pore size (80 Å) of the stationary phase. The XBridge BEH Amide and AdvanceBio Glycan Mapping materials provided baseline glycoform resolution with average Rs values of 1.5 ± 0.2 . The AdvanceBio Glycan Mapping column exhibited the shortest retention times and a somewhat better resolution for the low abundant glycoforms.

3.2 HILIC-UV-MS of *neo*-glycoproteins

3.2.1 RNAse glycoconjugates

In order to investigate the suitability of the developed HILIC methods for the characterization of semi-synthetic glycoproteins, glycoconjugates of RNase A were prepared and analysed. These conjugates were obtained by coupling a mono- or disaccharide activated with IME to the ε-amino group of lysine residues of RNase A (see preparation procedure in Section 2.2.1). RNase A (124 amino acids, MW 13,681 Da) contains ten lysines and, consequently, its glycosylation leads to the formation of a mixture of semi-synthetic glycoforms with different saccharide loading and positioning. These heterogeneous glycoconjugate products allow the assessment of the ability of the HILIC columns to discriminate among the different glycoforms. The *neo*-glycoproteins were dissolved in ACN-water (50:50, v/v) and analysed by HILIC-UV-MS using the three considered HILIC columns. MS detection was achieved using an ESI sprayer and a high resolution time-of-flight mass spectrometer operated in the positive-ion mode.

Figure 2 reports the UV chromatograms obtained for RNase A conjugated with Man-IME (upper panel) and Man(1-6)Man-IME (lower panel), showing clusters of partially resolved glycoform peaks. The median retention times and the width of the clusters are significantly larger for the disaccharide conjugate products, indicating that the HILIC retention of the intact glycoproteins increases with increasing glycan size. The HILIC-UV chromatograms obtained with the TSKgel Amide-80 column show less resolved glycoform profiles than the chromatograms acquired with the other two HILIC columns. Figure 3 shows the extracted-ion chromatograms (EICs) obtained for the conjugated RNase A products using the AdvanceBio Glycan Mapping column and depicting the intensity for the most abundant ion (charge state 6+) of each glycoform. The acquired mass data revealed that seven glycoforms carrying 4-10 saccharide units, were detected for the Man conjugates, whereas eight glycoforms carrying 2-

9 saccharide units were detected for the Man(1-6)Man conjugates. The results nicely show that the HILIC retention of the intact protein glycoforms increases proportionally with the number of conjugated saccharides, yielding a resolution which is based on the glycan load. Moreover, the increase of the HILIC retention *per* glycan is clearly larger for Man(1-6)Man (disaccharide) than for Man (monosaccharide). The observed glycoform peaks (Figure 3) are relatively broad, occasionally showing fine structure, particularly for the Man(1-6)Man conjugates. These wide, split peaks most probably are caused by the partial HILIC separation of positional isomers of *neo*-glycoproteins with the same number of conjugated saccharides, but different binding location on the protein backbone.

3.2.2 TB10.4 glycoconjugates

The tuberculosis antigen TB10.4 (103 amino acids, MW 11,076 Da) comprises one lysine residue in its amino acid sequence. The protein was glycosylated with Man-IME and Man(1-6)Man-IME using a previously optimized procedure [3], and the reaction products were purified and analysed. The HILIC gradient conditions were slightly adjusted in order to achieve appreciable retention and good glycoform separation. The starting water percentage was decreased from 32% to 21% for the TSKgel Amide-80 column, and from 28% to 17% for the XBridge BEH Amide and AdvanceBio Glycan Mapping columns, while maintaining the same gradient slope (0.5% min^{-1}) and time (20 min) as for the RNase B and A samples. Figure 4 shows the HILIC-UV chromatograms for TB10.4 conjugated with Man-IME using the various columns. The resolution for the main peaks observed for the TB10.4 samples after conjugation with Man-IME and Man(1-6)Man-IME, are reported in Table 2. Each HILIC column separated the TB10.4 variants (a, b, and c in Figs. 4A-C). However, due to narrower peaks obtained with the XBridge BEH Amide and AdvanceBio Glycan Map columns, obtained resolution values for these columns were significantly higher than for the TSKgel

Amide-80. The better separation also revealed the presence of several minor sample components (Figures 4B and 4C; see below for assignment). Comparison of Rs values obtained for the Man-IME and Man(1-6)Man-IME conjugates (Table 2) shows that doubling the size of the conjugated saccharide significantly enhances the HILIC resolution (factor 2.0). Clearly, the relative contribution of the size of the glycan moiety on the HILIC retention of the intact protein species is very large.

HILIC-MS allowed assignment of the peaks a, b and c to the non-glycosylated, monoglycosylated, and diglycosylated TB10.4 species, respectively. The presence of the latter can be explained by simultaneous conjugation of the amino groups of the lysine residue and the N-terminal alanine of TB10.4 [3, 22]. Glycosylation of the N-terminus of TB10.4 conjugated with Man-IME was confirmed by employing the MS/MS capabilities of the QTOF mass spectrometer after HILIC separation. The most abundant ion (charge state, 6+) obtained for TB10.4 (peak a) and mannosylated TB10.4 (peaks b and c) were fragmented by CID and the product-ion spectra were recorded (Figure 5). The b_6 fragment was used as diagnostic ion for the assignment of N-terminal glycosylation, as it was observed with good intensity for all fragmented species. The spectrum of TB10.4 (Figure 5A) shows an ion with m/z of 589.26 which corresponds to the expected value for the b_6 ion of TB10.4 that has a non-glycosylated N-terminus. The product-ion spectrum of TB10.4-$(Man)_2$ (Figure 5C) shows a clear mannosylated b_6 ion (m/z 824.31) indicating glycosylation at the N-terminus. Interestingly, the product-ion spectrum of TB10.4 (Figure 5B), shows b_6 ions (m/z 589.26 and 824.31) corresponding, respectively, to the non-glycosylated and glycosylated N-terminus of TB10.4. This indicates that the main mannosylated TB10.4 product actually is comprised of two isomeric glycoforms, which are not separated under the applied HILIC conditions, and which are conjugated either at the lysine or at the N-terminal alanine. Notably, this kind of structural information cannot be obtained by LC-MS using a bottom-up procedure.

Using HILIC-MS employing the AdvanceBio Glycan Mapping column, several minor components were observed next to the three TB10.4 species (Figure 6), which were formed during the glycosylation process. Deconvolution of the mass spectra obtained in the apices of the peaks revealed the presence of several degradation products formed during the glycosylation process. A large part of these products could be assigned to truncated protein forms that lost a number of amino acids from the C- or N-terminus side (Table 3). In addition, some oxidation products were observed showing larger retention than the native proteins. Oxidation is a common protein degradation occurring under the glycosylation conditions [22]. Notably, the HILIC column is capable to fully resolve the relatively small difference in protein structure induced by a single oxidation. The two peaks observed for oxidised TB10.4-Man most probably represent positional isomers.

3.2.3 Ag85B glycoconjugates

The antigenic protein Ag85B (292 amino acids, MW 31,346 Da) presents eight lysine residues as potential glycosylation sites. *Neo*-glycoproteins were produced by conjugation with Man-IME, Man(1-6)Man-IME and Man(1-6)Man(1-6)Man-IME. For larger proteins, solubility in the ACN-rich eluent and irreversible adsorption to HILIC materials might become an issue [23]. However, using gradient elution with a starting water percentage of 26% for the TSKgel Amide-80 column, and 22% for the XBridge BEH Amide and AdvanceBio Glycan Mapping columns, symmetric peaks and repeatable retention times and peak areas were obtained in the HILIC-UV traces of non-glycosylated Ag85B. For all columns the RSD values (n = 5) for retention time and peak area were below 0.4% and 2.5%, respectively, indicating proper performance.

Figure 7 shows the HILIC separations of Ag85B glycoforms obtained after Man-IME conjugation using the three columns. The resolutions obtained for the Ag85B *neo*-

glycoproteins differed considerably (Table 4). Using the AdvanceBio Glycan Mapping column (Figure 7C) an almost baseline separation of six glycoforms was achieved. The same number of peaks was detected with the XBridge BEH Amide column (Figure 7B), but the glycoforms partially co-eluted. The profile of overlapping peaks obtained with the TSKgel Amide-80 column (Figure 7A) did not allow calculation of peak resolutions. The relatively small pore size of the TSKgel Amide-80 particles clearly compromises the separation performance for proteins larger than 15k kDa.

In order to achieve glycoform assignment, Ag85B conjugated with the mono-, di- and tri-mannose saccharides were analysed by HILIC-MS employing the XBridge BEH Amide and AdvanceBio Glycan Mapping columns (Figure 8). MS detection enabled the unambiguous determination of the number of saccharide units conjugated. For Ag85B reacted with Man-IME, protein species bearing 3 to 8 glycans could be detected (Figures 8B and 8F), indicating all Ag85B molecules had been conjugated. Using Man(1-6)Man-IME (Figures 8C and 8G) Ag58B quantitatively reacted with the sugar, but the average degree of substitution (1 to 7 glycans attached) was lower. The sample coming from Ag85B conjugation with Man(1-6)Man(1-6)Man-IME (Figures 8D and 8H) contained a significant amount of unmodified antigen and only mono-, di-, and tri-glycosylated protein species. Evidently, the degree of substitution decreases with the size of the attached glycan, which is in accordance with the relative chemical reactivity of the activated sugars. From the overall results provided in Figure 8, it is very clear that the number and size of the glycans attached to the protein, strongly add to the HILIC retention and resolution of the intact species.

With the AdvanceBio Glycan Mapping column highest glycoform resolutions were obtained for the Ag58B *neo*-glycoproteins (Figures 8E-H). Interestingly, this good performance allowed detection of an extra low abundant component in the Ag58B sample conjugated with Man(1-6)Man(1-6)Man-IME (Figure 9A) which eluted just after the native Ag85B protein.

The deconvoluted spectrum recorded at the apex of the extra peak (Figure 9C) reveals two masses (31,361.8 and 31,580.4 Da) which are higher than the molecular mass of Ag85B (31,345 Da). The first mass might be due to mono-oxidised Ag85B. The second mass indicates the presence of Ag85B conjugated with Man-IME. The Man(1-6)Man(1-6)Man-IME reagent is not totally pure and may contain traces of Man-IME, leading to formation of minor amounts of Ag85B-Man conjugates. Reliable MS detection of these minor sample components among excess of other protein species is only possible due to the pre-separation provided by the HILIC selectivity.

4. Conclusion

HILIC-UV-MS(/MS) methods for the characterization of intact *neo*-glycoproteins were developed. Under optimized separation conditions, resolution and assignment of proteoforms and individual glycoforms of semi-synthetic intact glycoproteins were achieved. HILIC retention and selectivity appeared to be dominated by the size and number of the saccharides conjugated to the carrier protein. Three tested amide HILIC stationary phases showed similar selectivity for the glycosylated proteins. However, the XBridge BEH Amide and AdvanceBio Glycan Mapping columns showed better separation performance than the TSKgel Amide-80 column, leading to better glycoform resolution. For the larger glycoconjugated antigen Ag85B (MW, 31.3 kDa), optimum resolution was achieved with the AdvanceBio Glycan Mapping material, most likely due to the superficially porous nature of its particles. The HILIC-MS selectivity also permitted to detect and tentatively assign minor protein impurities and degradation products. Furthermore, MS/MS detection involving CID of glycoconjugates

provided information on glycan position, confirming N-terminal glycosylation of the TB10.4 antigen.

Overall, the HILIC-UV-MS results obtained for the different glycosylated proteins suggest that with proper selection of the HILIC stationary phase and optimization of MS-compatible chromatographic conditions, it is possible to achieve very good glycoconjugate resolution. Moreover, intact glycoform identities of semi-synthethic glycoprotein products can be assessed without the need for sample pre-treatment.

References

[1] R. Adamo, A. Nilo, B. Castagner, O. Boutureira, F. Berti, G.J.L. Bernardes, Synthetically defined glycoprotein vaccines: current status and future directions, Chem. Sci. 4 (2013) 2995-3008.

[2] M. Dalziel, M. Crispin, C.N. Scanlan, N. Zitzmann, R.A. Dwek, Emerging principles for the therapeutic exploitation of glycosylation, Science, 343 (2014) 1235681-1235688.

[3] C. Temporini, T. Bavaro, S. Tengattini, I. Serra, G. Marrubini, E. Calleri, F. Fasanella, L. Piubelli, F. Marinelli, L. Pollegioni, G. Speranza, G. Massolini, M. Terreni, Liquid chromatography–mass spectrometry structural characterization of *neo*-glycoproteins aiding the rational design and synthesis of a novel glycovaccine for protection against tuberculosis, J. Chromatogr. A, 1367 (2014) 57-67.

[4] K. Mariño, J. Bones, J.J. Kattla, P.M. Rudd. A systematic approach to protein glycosylation analysis: a path through the maze, Nat. Chem. Bio. 6 (2010) 713-723.

[5] M. Thaysen-Andersen, N.H. Packer, Advances in LC-MS/MS-based glycoproteomics: getting closer to system-wide site-specific mapping of the N- and O-glycoproteome, Biochim. Biophys. Acta, 1844 (2014) 1437-1452.

[6] K. Sandra, I. Vandenheede, P. Sandra, Modern chromatographic and mass spectrometric techniques for protein biopharmaceutical characterization, J. Chromatogr. A, 1335 (2014) 81-103.

[7] A. Staub, D. Guillarme, J. Schappler, J-L. Veuthey, Serge Rudaz, Intact protein analysis in the biopharmaceutical field, J. Pharm. Biomed. Anal. 55 (2011) 810–822.

[8] A. Periat, I.S. Krull, D. Guillarme, Applications of hydrophilic interaction chromatography to amino acids, peptides, and proteins, J. Sep. Sci. 38 (2015) 357-367.

[9] J. Bones, S. Mittermayr, N. O'Donoghue, A. Guttman, P.M. Rudd, Ultra Performance Liquid Chromatographic Profiling of Serum N-Glycans for Fast and Efficient Identification of Cancer Associated Alterations in Glycosylation, Anal. Chem. 82 (2010) 10208-10215.

[10] Y. Takegawa, K. Deguchi, T. Keira, H. Ito, H. Nakagawa, S.I. Nishimura, Separation of isomeric 2-aminopyridine derivatized N-glycans and N-glycopeptides of human serum immunoglobulin G by using a zwitterionic type of hydrophilic-interaction chromatography, J. Chromatogr. A, 1113 (2006) 177-181.

[11] 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.

[12] 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.

[13] T. Tetaz, S. Detzner, A. Friedlein, B. Molitor, J.L. Mary, Hydrophilic interaction chromatography of intact, soluble proteins, J. Chromatogr. A, 1218 (2011) 5892-5896.

[14] Z. Zhang, Z. Wu, M.J. Wirth, Polyacrylamide brush layer for hydrophilic interaction liquid chromatography of intact glycoproteins, J. Chromatogr. A, 1301 (2013) 156-161.

[15] A. Pedrali, S. Tengattini, G. Marrubini, T. Bavaro, P. Hemström, G. Massolini, M. Terreni, C. Temporini, Characterization of Intact Neo-Glycoproteins by Hydrophilic Interaction Liquid Chromatography, Molecules, 19 (2014) 9070-9088.

[16] L. Piubelli, M. Campa, C. Temporini, E. Binda, F. Mangione, M. Amicosante, M. Terreni, F. Marinelli, L. Pollegioni, Optimizing Escherichia coli as a protein expression platform to produce Mycobacterium tuberculosis immunogenic proteins, Microb. Cell Fact. 12 (2013) 115.

[17] G. Kallenius, A. Pawlowski, B. Hamasur, S.B. Svenson, Mycobacterial glycoconjugates as vaccine candidates against tuberculosis, Trends Microbiol. 16 (2008) 456-462.

[18] T. Bavaro, M. Filice, C. Temporini, S. Tengattini, I. Serra, C.F. Morelli, G. Massolini, M. Terreni, Chemoenzymatic synthesis of *neo*glycoproteins driven by the assessment of protein surface reactivity, RSC Adv. 4 (2014) 56455-56465.

[19] W.Z. Shou, W. Naidong, Simple means to alleviate sensitivity loss by trifluoroacetic acid (TFA) mobile phases in the hydrophilic interaction chromatography–electrospray tandem mass spectrometric (HILIC–ESI/MS/MS) bioanalysis of basic compounds, J. Chromatogr. B, 825 (2005) 186-192.

[20] R. Kostiainen, T.J. Kauppila, Effect of eluent on the ionization process in liquid chromatography–mass spectrometry, J. Chromatogr. A, 1216 (2009) 685-699.

[21] Y. Chen, A.R. Mehok, C.T. Mant, R.S. Hodges, Optimum concentration of trifluoroacetic acid for reversed-phase liquid chromatography of peptides revisited, J. Chromatogr. A, 1043 (2004) 9-18.

[22] S. Tengattini, E. Domínguez-Vega, L. Piubelli, C. Temporini, M. Terreni, G.W. Somsen, Monitoring antigenic protein integrity during glycoconjugate vaccine synthesis using capillary electrophoresis-mass spectrometry, Anal. Bioanal. Chem 408 (2016) 6123-6132.

Tables

Table 1. Retention times (RT), and resolutions (Rs) obtained for successively eluting glycoforms during HILIC-UV of RNase B on three columns. SD, standard deviation.

HILIC column	mean RT (min)	Rs (2,1)	Rs (3,2)	Rs (4,3)	Rs (5,4)	average Rs ± SD
TSK-gel Amide 80	16.84	1.1	1.0	1.0	0.9	1.0±0.1
XBridge BEH Amide	17.45	1.9	1.4	1.4	1.3	1.5±0.2
AdvanceBio Glycan Mapping	15.02	1.7	1.6	1.3	1.4	1.5±0.2

Table 2. Resolutions (Rs) obtained during HILIC-UV of TB10.4 conjugated with Man-IME and Man(1-6)Man-IME on three columns. Components, (a) non-glycosylated TB10.4, (b) monoglycosylated TB10.4, and (c) diglycosylated TB10.4.

	TB10.4-Man			**TB10.4-Man(1-6)Man**		
HILIC column	**RT a; b; c (min)**	**Rs (a,b)**	**Rs (b,c)**	**RT a; b; c (min)**	**Rs (a,b)**	**Rs (b,c)**
TSK-gel Amide 80	12.51; 14.33; 16.09	3.0	3.0	12.49; 16.43; 19.85	5.8	6.13
XBridge BEH Amide	16.04; 18.11; 19.82	8.1	6.2	16.01; 20.55; 24.23	17.3	12.7
AdvanceBio Glycan Mapping	12.02; 13.97; 15.65	7.4	7.0	11.99; 15.99; 19.30	15.6	13.7

Table 3. Species observed during HILIC-MS analysis of TB10.4 conjugated with Man-IME.

Peak	RT (min)	Experimental mass (Da)	Assignment	Theoretical mass (Da)
1	9.5	9933.4	A1-A92	9933.1
		9862.3	A1-M91	9862.0
2	9.7	10088.6	A1-R93	10088.8
3	10.2	9730.5	A1-M90	9730.8
4	10.7	9528.4	A1-M88	9528.5
		9599.5	A1-A89	9599.6
		10505.8	A1-E97	10505.6
		10646.9	A1-A99	10647.8
		11076.1	**TB10.4**	11076.3
5	11.2	7651.4	G36-G103	7651.1
		7979.6	Q33-G103	7979.5
		10097.4	A1-A91-Man	10097.1
		10168.5	A1-A92-Man	10168.2
6	11.6	6469.9	S48-G103-Man	6469.9
		6711.0	L46-G103-Man	6711.2
		9426.3	M18-G103-Man	9426.2
		8916.1	G23-G103-Man	8916.4
7	11.9	9965.5	A1-M90-Man	9965.8
8	12.4	11311.2	**TB10.4-Man**	11311.3
9	12.9	11011.0	A1-K100-Man	11011.0
10	13.1	10494.7	S9-G103-Man	10494.4
11	13.3	11327.2	Oxidised TB10.4-Man	11327.3
12	13.6	11327.2	Oxidised TB10.4-Man	11327.3
		9061.8	Q10-M90-Man	9061.8
13	14.0	11546.2	**TB10.4-Man$_2$**	11546.5
14	14.7	11562.3	Oxidised TB10.4-Man$_2$	11562.5

Table 4. Retention times (RT) and resolutions (Rs) obtained for successively eluting glycoforms during HILIC-UV on three columns of A85B conjugated with Man-IME. ND, not determinable.

HILIC column	mean RT (min)	Rs (2,1)	Rs (3,2)	Rs (4,3)	Rs (5,4)	Rs (6,5)	average α ± SD Rs ± SD
TSK-gel Amide 80	17.43	ND	ND	ND	ND	ND	ND
Waters XBridge BEH Amide	15.93	0.9	1.0	1.0	1.0	0.9	1.0 ± 0.1
Agilent AdvanceBio Glycan Map	15.01	1.1	1.1	1.0	1.2	1.2	1.1 ± 0.1

Figure Legends

Figure 1. HILIC-UV chromatograms obtained for RNase B (0.5 mg mL^{-1} in ACN-water (50:50, v/v)). Column, (A) TSKgel Amide-80, (B) Waters XBridge BEH Amide, and (C) Agilent AdvanceBio Glycan Mapping. Mobile phase comprised of ACN (solvent A) and water (solvent B) both containing 0.05% (v/v) TFA; injection volume, 2 µL; column temperature, 50 °C. Flow rate, (A) 0.3 mL min^{-1} and (B,C) 0.5 ml/min. Gradient, (A) from 32% to 42% B in 20 min followed by isocratic elution at 42% B for 10 min., and (B,C) from 28% to 38% B in 20 min followed by isocratic elution at 38% B for 10 min.

Figure 2. HILIC-UV chromatograms obtained for RNase A (0.5 mg mL^{-1}) conjugated with Man-IME (upper panel) and Man(1-6)Man-IME (lower panel) obtained using the (left) TSKgel Amide-80, (middle) XBridge BEH Amide, and (right) AdvanceBio Glycan Mapping columns. Further conditions, see Figure 1.

Figure 3. Extracted-ion chromatograms obtained during HILIC-UV-MS of RNase A conjugated with (A) Man and (B) Man(1-6)Man using the AdvanceBio Glycan Mapping column. The ion intensity of the 6+ charge state of each glycoform is depicted. Number of conjugated saccharides, 2 (dark red), 3 (yellow), 4 (pink), 5 (dark green), 6 (red), 7 (blue), 8 (orange), 9 (light green) and 10 (black). Further conditions, see Figure 1 and Experimental Section.

Figure 4. HILIC-UV chromatograms obtained for TB10.4 (0.5 mg mL^{-1}) conjugated with Man-IME. Column, (A) TSKgel Amide-80, (B) XBridge BEH Amide, and (C) AdvanceBio Glycan Mapping. Sample solvent, ACN-water (50:50, v/v); gradient, (A) from 21% to 31% B in 20 min followed by isocratic elution at 31% B for 10 min., and (B,C) from 17% to 27% B in 20 min followed by isocratic elution at 27% B for 10 min. Peak assignment, (a) TB10.4, (b) TB10.4-Man, (c) TB10.4-$(Man)_2$. Further conditions, see Experimental Section.

Figure 5. MS/MS spectra (400-800 *m/z*) of (A) TB10.4, (B) TB10.4-Man and (C) TB10.4-$(Man)_2$ obtained during HILIC-MS/MS of TB10.4 (0.5 mg mL^{-1}) conjugated with Man-IME. Parent ion, 6+ charge state of each protein species; column, AdvanceBio Glycan Mapping. Further conditions, see Figure 4 and Experimental Section.

Figure 6. HILIC-MS of TB10.4 (0.5 mg mL^{-1}) conjugated with Man-IME. Column, AdvanceBio Glycan Mapping; gradient, from 17% to 27% B in 20 min followed by isocratic elution at 27% B for 10 min. Further conditions, see Experimental Section.

Figure 7. HILIC-UV chromatograms obtained for Ag85B (0.5 mg mL^{-1}) conjugated with Man-IME. Column, (A) TSKgel Amide-80, (B) XBridge BEH Amide, and (C) AdvanceBio Glycan Mapping. Sample solvent, ACN-water (50:50, v/v); gradient, (A) from 26% to 36% B in 20 min followed by isocratic elution at 36% B for 10 min, and (B,C) from 22% to 32% B in 20 min followed by isocratic elution at 32% B for 10 min. Further conditions, see Figure 1.

Figure 8. HILIC-MS chromatograms obtained for Ag85B (A,E) conjugated with Man-IME (B,F), Man(1-6)Man-IME (C,G) and Man(1-6)Man(1-6)Man-IME (D,H). Column, (A-D) XBridge BEH Amide, and (E-H) AdvanceBio Glycan Mapping. Numbers indicate number of conjugated glycans. Further conditions, see Figure 7 and Experimental Section.

Figure 9. (A) Zoom of HILIC-MS base peak chromatogram of Ag85B conjugated with Man(1-6)Man(1-6)Man-IME. (B) Deconvoluted mass spectrum obtained at the apex of the minor peak eluting at 12.7 min. Conditions, see Figure 8.

Figures

Figure 1

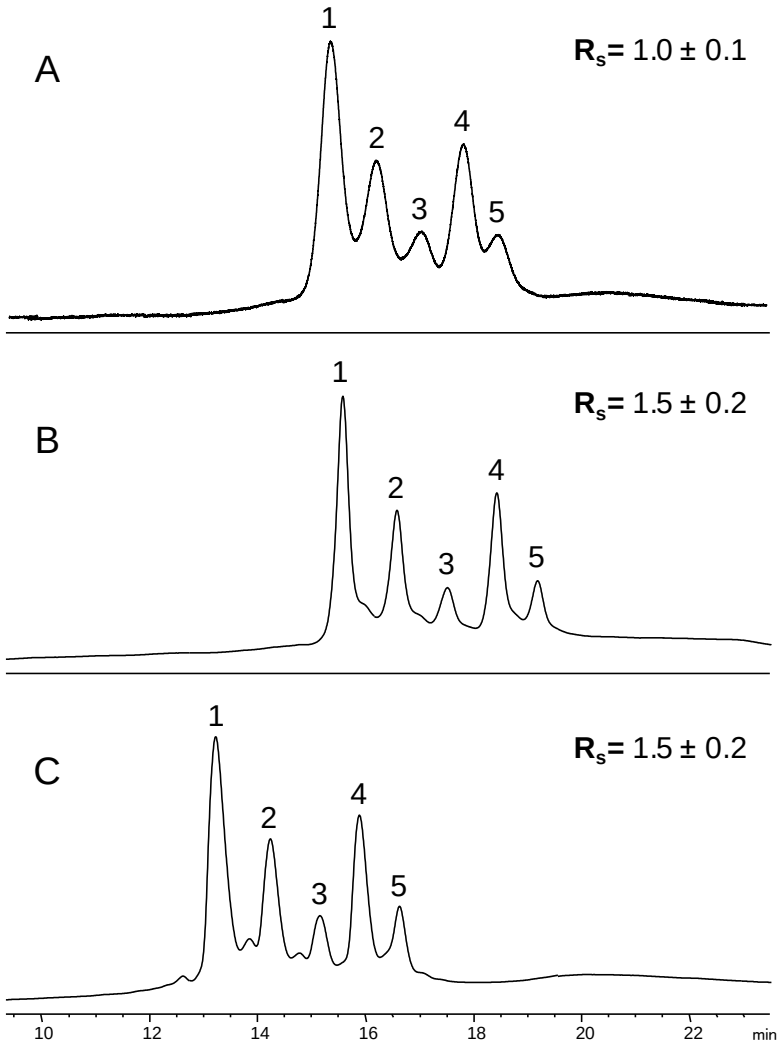


Figure 2

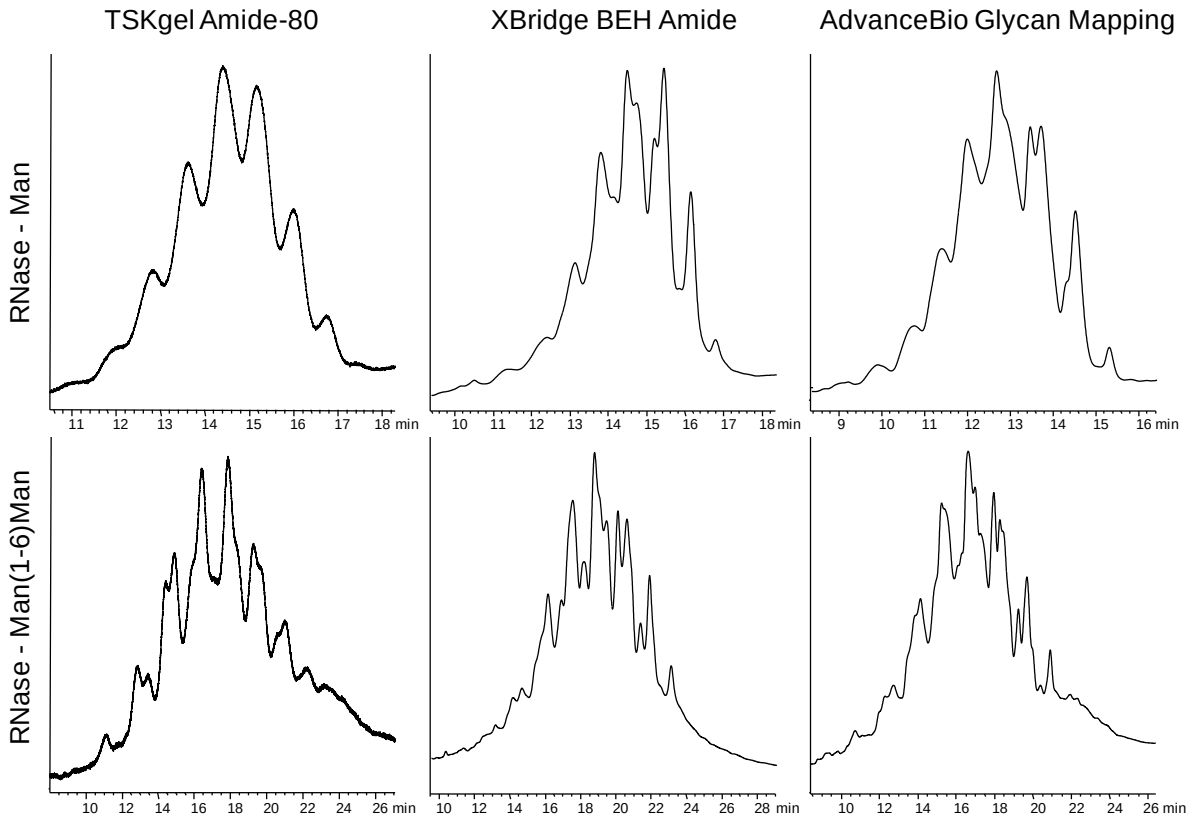

TSKgel Amide-80
XBridge BEH Amide
AdvanceBio Glycan Mapping
RNase - Man
RNase - Man(1-6)Man
11 12 13 14 15 16 17 18 min
10 11 12 13 14 15 16 17 18 min
9 10 11 12 13 14 15 16 min
10 12 14 16 18 20 22 24 26 min
10 12 14 16 18 20 22 24 26 28 min
10 12 14 16 18 20 22 24 26 min

Figure 3

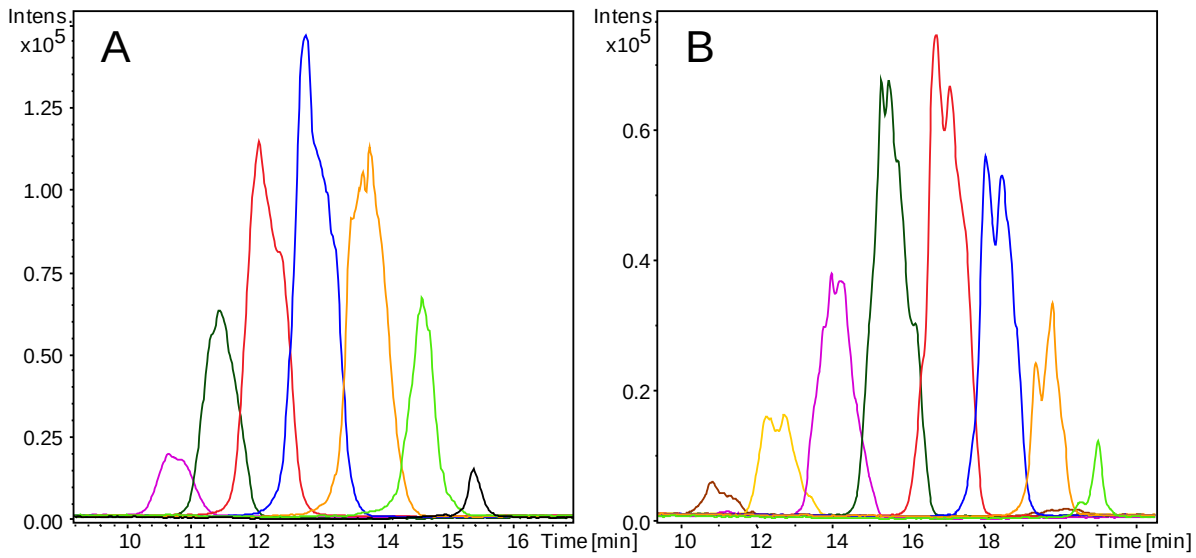

A
Intens
x10^5
1.25
1.00
0.75
0.50
0.25
0.00
10
11
12
13
14
15
16
Time [min]
B
Intens
x10^5
0.6
0.4
0.2
0.0
10
12
14
16
18
20
Time [min]

Figure 4

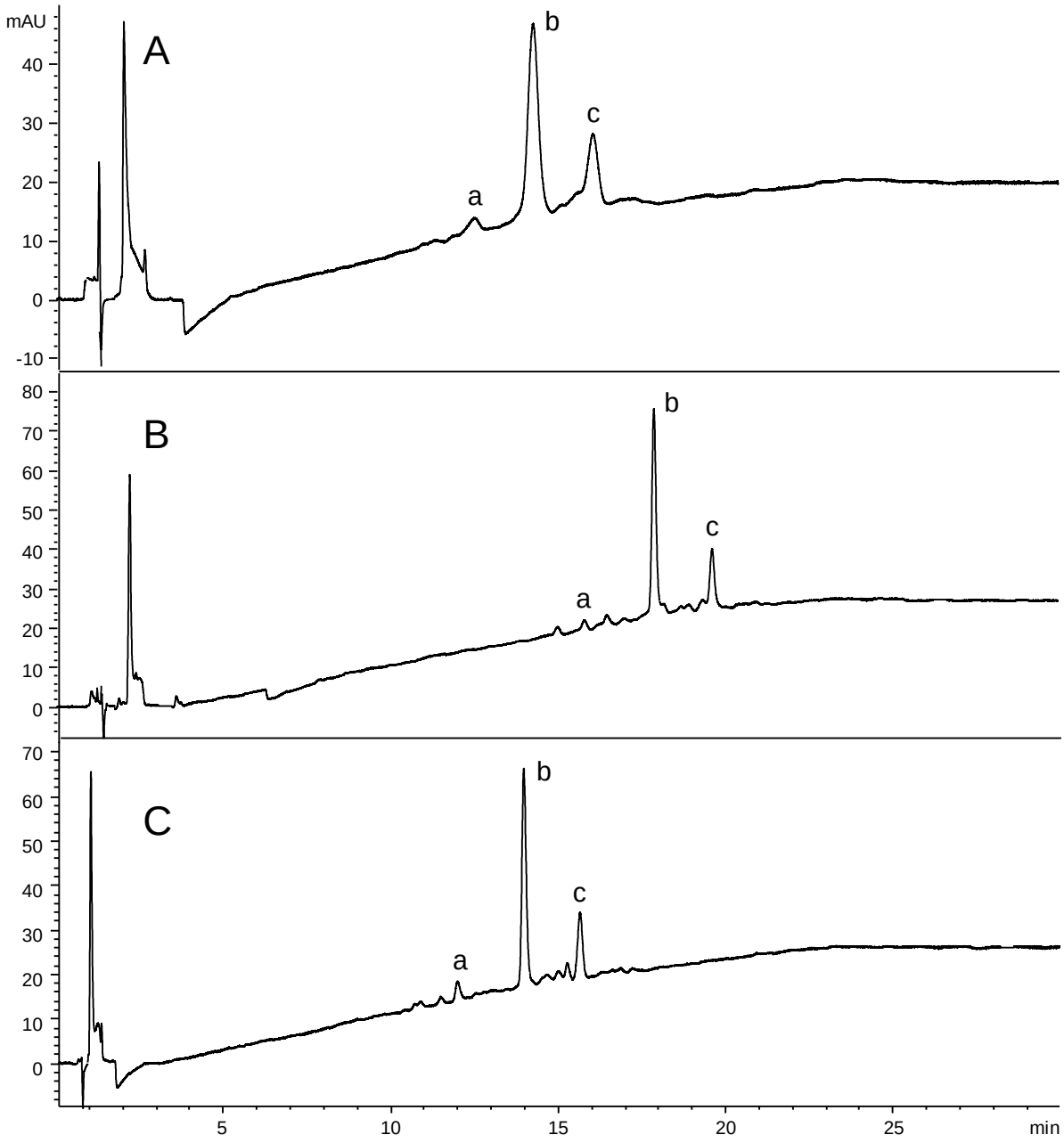


Figure 5

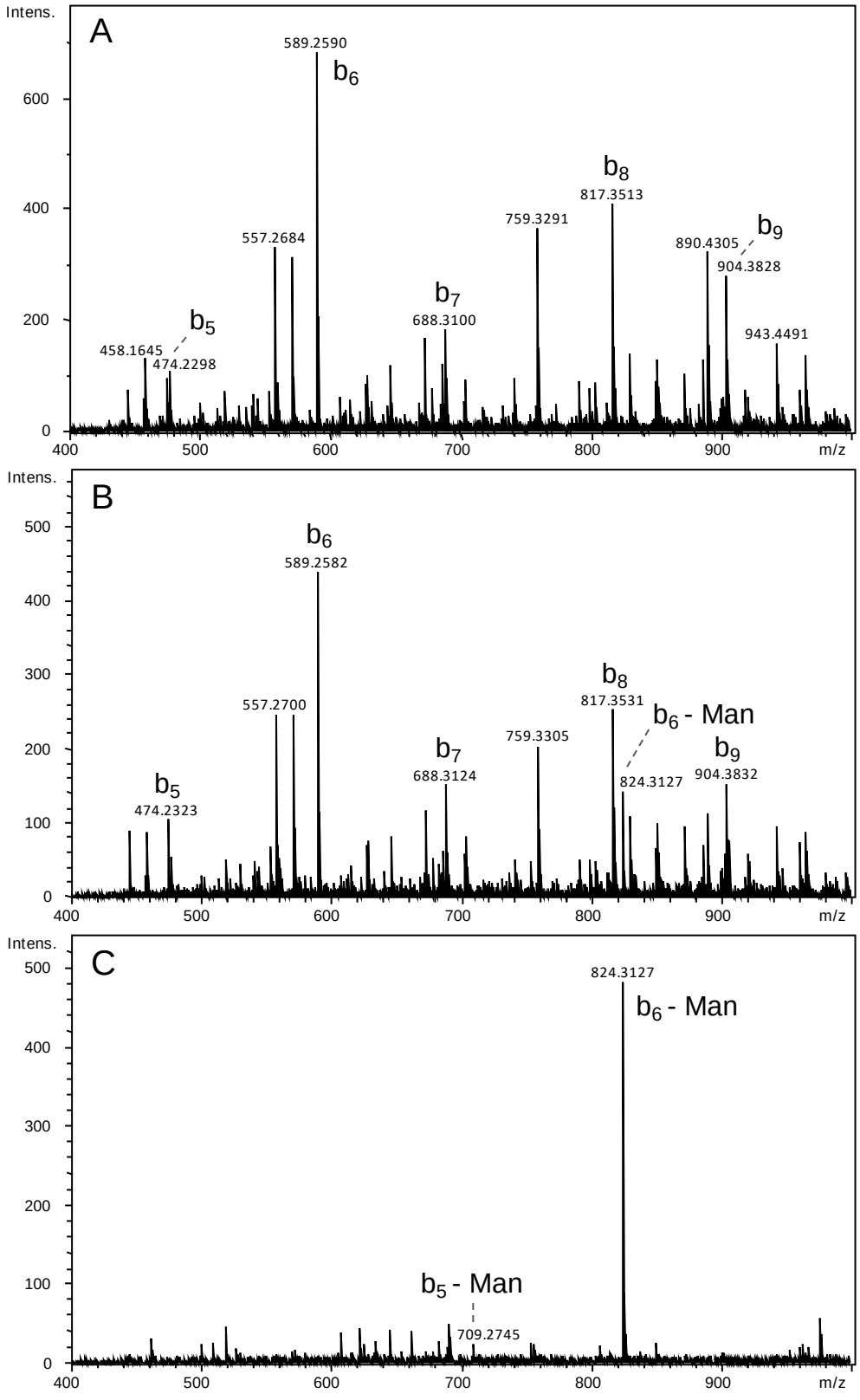

Intens.
A
589.2590
b6
b8
817.3513
759.3291
557.2684
b9
890.4305
904.3828
b7
688.3100
b5
458.1645
474.2298
943.4491
m/z
Intens.
B
b6
589.2582
557.2700
b8
817.3531
b6 - Man
759.3305
b7
688.3124
b9
904.3832
824.3127
b5
474.2323
m/z
Intens.
C
824.3127
b6 - Man
b5 - Man
709.2745
m/z

Figure 6

Figure 7

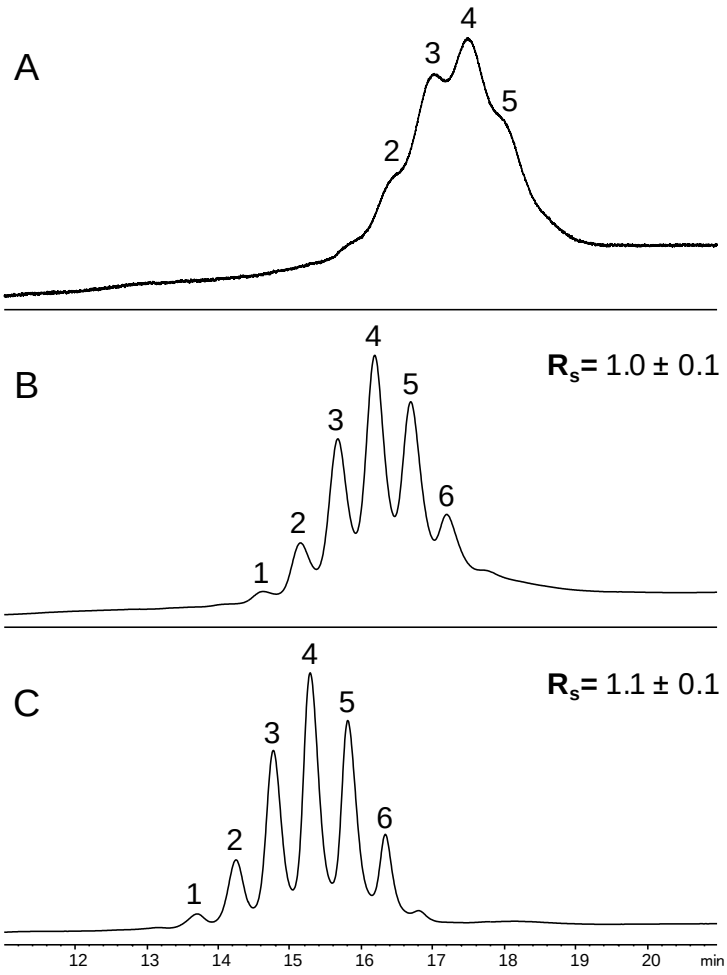

A
1
2
3
4
5
B
R_s= 1.0 ± 0.1
1
2
3
4
5
6
C
R_s= 1.1 ± 0.1
1
2
3
4
5
6
12
13
14
15
16
17
18
19
20
min

Figure 8

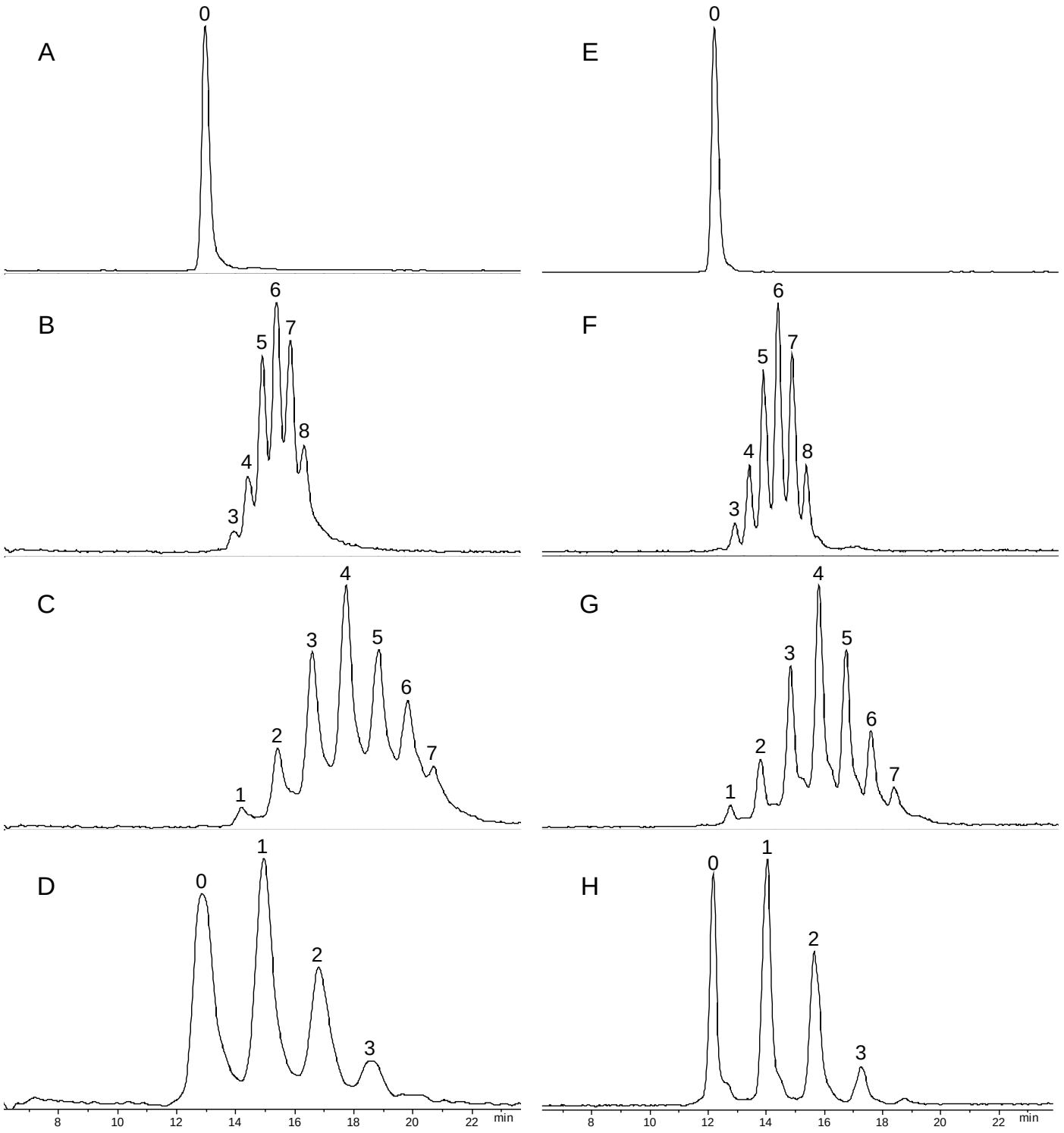
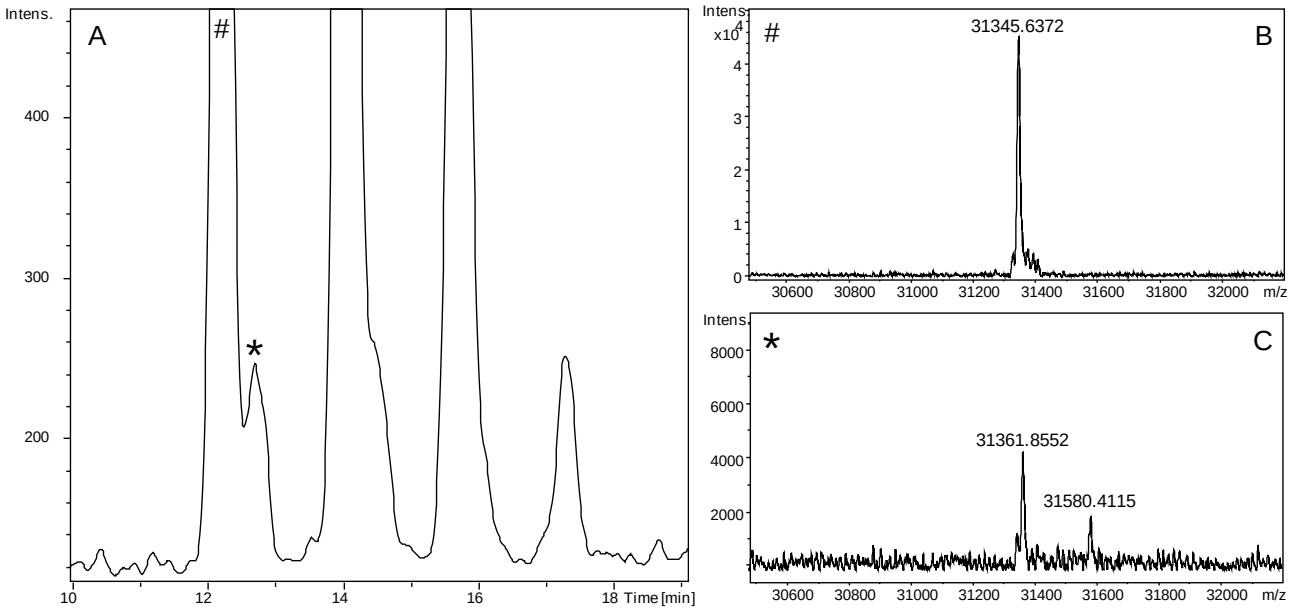

A
0
B
3
4
5
6
7
8
C
1
2
3
4
5
6
7
D
0
1
2
3
E
0
F
3
4
5
6
7
8
G
1
2
3
4
5
6
7
H
0
1
2
3
8
10
12
14
16
18
20
22
min

Figure 9


A
#
*
Intens.
400
300
200
10
12
14
16
18
Time [min]
B
#
31345.6372
x10 4
4
3
2
1
0
30600
30800
31000
31200
31400
31600
31800
32000
m/z
C
*
31361.8552
31580.4115
8000
6000
4000
2000