Analytical Biochemistry 417 (2011) 80–88

Contents lists available at ScienceDirect

Analytical Biochemistry

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

Characterization of glycoprotein digests with hydrophilic interaction chromatography and mass spectrometry

Martin Gilar [a,*], Ying-Qing Yu [a], Joomi Ahn [a], Hongwei Xie [a], Huanhuan Han [b], Wantao Ying [b], Xiaohong Qian [b]

[a] Waters Corporation, 34 Maple Street, Milford, MA 01757, USA
[b] Beijing Proteome Research Center, Beijing Institute of Radiation Medicine, 33 Life Science Park Road, Changping District, Beijing 102206, China

ARTICLE INFO

Article history:
Received 4 March 2011
Received in revised form 13 May 2011
Accepted 17 May 2011
Available online 26 May 2011

Keywords:
Hydrophilic interaction chromatography
Mass spectrometry
Glycoprotein
Characterization
Glycopeptides
Glycans

ABSTRACT

A new hydrophilic interaction chromatography (HILIC) column packed with amide 1.7 μm sorbent was applied to the characterization of glycoprotein digests. Due to the impact of the hydrophilic carbohydrate moiety, glycopeptides were more strongly retained on the column and separated from the remaining nonglycosylated peptides present in the digest. The glycoforms of the same parent peptide were also chromatographically resolved and analyzed using ultraviolet and mass spectrometry detectors. The HILIC method was applied to glyco-profiling of a therapeutic monoclonal antibody and proteins with several N-linked and O-linked glycosylation sites. For characterization of complex proteins with multiple glycosylation sites we utilized 2D LC, where RP separation dimension was used for isolation of glycopeptides and HILIC for resolution of peptide glycoforms. The analysis of site-specific glycan microheterogeneity was illustrated for the CD44 fusion protein.

Many therapeutic proteins, including monoclonal antibodies (mAb)[1] [1], are glycosylated. Because glycosylation affects the therapeutic efficacy of mAb, glyco-profiling and characterization of proteins have received increased attention in recent years [2,3]. Liquid chromatography (LC) and mass spectrometry (MS) methods are often employed to investigate glycosylation at the intact protein level [2,4,5], after protein digestion at the glycopeptide level [6–8], or for analysis of glycans released from proteins [1,9–11]. Several reviews and reports on LC and LC–MS IgG glycosylation analysis were published recently and are recommended to readers interested in a comparison of well-established methods and principles of glycoprotein analysis [1–4,12,13].

LC–MS glyco-profiling of intact proteins is a fast and elegant method. However, it has limitations for highly complex proteins, since multiply glycosylated and heterogeneous proteins often do not provide MS spectra of sufficient quality suitable for robust data deconvolution. LC–MS glyco-profiling of mAb with a single glycosylation site has been demonstrated, using on-line sample desalting [2,5].

Another glyco-profiling method relies on the release of glycans from the parent protein [14] and their quantification with high-performance anion exchange chromatography employing pulsed amperometric detection (HPAEC PAD) [15]. The glycans are typically labeled with fluorescent dye and analyzed with reversed-phase liquid chromatography (RP LC) [11], capillary electrophoresis (CE) [16,17], or hydrophilic interaction chromatography (HILIC) methods [1,9,10,14] using sensitive fluorescence detectors.

Neither method noted above (intact protein or glycan profiling) is capable of providing information about protein glycosylation site(s) and their glycan microheterogeneity, with the exception of mono-glycosylated proteins. Characterization of glycosylation sites and their glycan microheterogeneity could be performed after digesting the glycoprotein of interest with proteolytic enzymes. Since glycans remain attached to peptides of unique masses, the microheterogeneity of each site can be investigated by MS [6,7].

RP LC is typically used for analysis of protein digests, but often cannot adequately resolve glycopeptides into their glycoforms. In addition, the glycopeptides coelute with other components in the complex peptide maps. HILIC was found to be more useful for selective analysis of glycopeptides, especially when combined with MS [6,8,10,18–20]. The information about glycosylation sites is in principle available for both N-linked and O-linked glycans [21].

HILIC separation mode was originally used for analysis of highly polar analytes [22] which are inadequately retained in RP LC. The method was later extended to analysis of other types of analytes

* Corresponding author. Fax: +1 508 482 3625.
E-mail address: Martin_Gilar@waters.com (M. Gilar).

[1] *Abbreviatons used:* 2-AB, 2-aminobenzamide; CE, capillary electrophoresis; DTT, dithiothreitol; FA, formic acid; HILIC, hydrophilic interaction chromatography; HPAEC PAD, high-performance anion exchange chromatography employing pulsed amperometric detection; mAb, monoclonal antibodies; RP LC, reversed-phase liquid chromatography; TFA, trifluoroacetic acid.

0003-2697/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.ab.2011.05.028

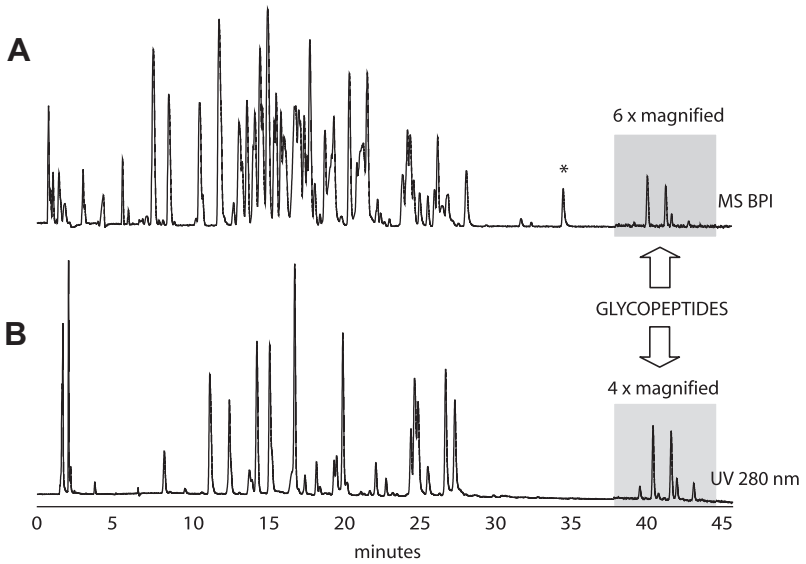

Fig.1. Separation of mAb tryptic digest in HILIC with MS (A) and UV (B) detection. Glycopeptides elute in the highlighted region. The most retained nonglycosylated peptide (VDNALQSGNSQESVTEQDSK, labeled with asterisk) eluted at 34.4 min. BPI stands for base peak intensity; the most intense ions are extracted under each MS peak.

Table 1
Relative quantitation of mAb glycans using glycopeptide UV 280 nm or MS analysis.

Glycan/glycopeptide annotation	Glycan composition	Glycopeptide mass, Da $[MH]^+$ (monoisotopic)	Relative quantity (%) UV 280 nm $N = 3$[a]	Relative quantity (%) XIC MS $N = 3$	Relative quantity (%) glycan FLR[c] $N = 2$[b]
G0	4HexNAc/3Hex	2487.989	6.3 ± 0.3	6.1 ± 0.1	4.8
G0F	4HexNAc/3Hex/1dHex	2634.046	35.7 ± 0.2	38.3 ± 0.8	42.2
G1 a	4HexNAc/4Hex	2650.041	3.1 ± 0.1	2.4 ± 0.2	2.0
G1 b	4HexNAc/4Hex	2650.041	0.8 ± 0.1	1.1 ± 0.1	0.8
Man 5	2HexNAc/5Hex	2405.935	–	1.1 ± 0.0	2.0
G1F a	4HexNAc/4Hex/1dHex	2796.099	34.4 ± 0.1	31.2 ± 0.6	31.1
G1F b	4HexNAc/4Hex/1dHex	2796.099	11.2 ± 0.0	11.6 ± 0.2	8.8
G2F	4HexNAc/5Hex/1dHex	2958.152	8.5 ± 0.2	8.2 ± 0.3	8.4

[a] Man 5 was not resolved from the G1Fa peak; relative quantitation of G1F a peak is slightly biased in UV experiments.
[b] Standard deviation was not calculated for $N = 2$ repetitions.
[c] Results are compared to HILIC analysis of released glycans labeled with 2-AB and detected with a fluorescence detector.

including peptides [23], phosphopeptides [24,25], glycans [9,14], natural compounds [26], and others [27]. The distinct hydrophilic nature of glycan moieties enhances the retention of glycopeptides and enables their separation as a class from unmodified peptides [18].

Recent developments in hydrophilic interaction chromatography columns accelerated the adoption of the method in many analytical laboratories [1,9,18,25,28,29]. Although HILIC received attention as a method for analysis of glycans and glycopeptides [6,8,18–21], and provided some separation benefit compared to RP LC [5,30], only the advent of highly efficient HILIC columns made the challenging separation of peptide glycoforms (glycan microheterogeneity) readily possible.

In this paper we investigate a HILIC column packed with 1.7 μm amide sorbent for separation of glycopeptides and their glycoforms. We propose the HILIC MS method for characterization of N-linked and O-linked glycosylation sites for mAb and for multiply glycosylated proteins. The data-independent MS acquisition (MS^E) enabled detection of unsuspected glycopeptides and analysis of site-specific glycan microheterogeneity.

Materials and methods

Reagents and chemicals

Formic acid (FA), ammonium bicarbonate, dithiothreitol (DTT), iodoacetamide (IAM), trifluoroacetic acid (TFA), and trypsin inhibitor were purchased from Sigma–Aldrich (St. Louis, MO, USA). HPLC grade acetonitrile (MeCN) and ammonium hydroxide were purchased from J.T. Baker (Phillipsburg, NJ, USA). A Milli-Q system (Millipore, Bedford, MA, USA) was used to prepare deionized water (18 MΩ cm) for LC mobile phases. RapiGest SF was obtained from Waters (Milford, MA, USA). Bovine fetuin (P1253), human alpha-1 acid glycoprotein (P02763), human transferrin (P02787), and bovine RNase B (P61823) were purchased from Sigma–Aldrich (St. Louis, MO, USA). Sequencing grade trypsin and PNGase F enzymes were purchased from Promega (Madison, WI, USA). Endoproteinase AspN was obtained from New England BioLabs (Ipswich, MA, USA). Recombinant human CD44/Fc Chimera, CF, was obtained from R&D Systems (Minneapolis, MN, USA).

Protein digestion protocols

Tryptic digestion of proteins was performed as follows: 64 μl of protein (200 μg) dissolved in 50 mM ammonium bicarbonate, pH 7.9, containing 0.1% RapiGest SF was denatured with DTT, alkylated with IAM, and digested for 2–4 h with trypsin (protein:trypsin ratio 20:1, w/w). Tryptic digest of CD44 was further digested with AspN. Trypsin inhibitor was added to the CD44 digest (trypsin:inhibitor ratio 1:1, w/w) followed by the addition of AspN (CD44:AspN ratio 20:1, w/w); digestion was performed overnight. Cleavage of N-linked glycans of fetuin was performed after tryptic digestion by adding trypsin inhibitor (trypsin to inhibitor ratio 1:1,

w/w) followed by addition of 2500 units of PNGase F to 100 μg of digested fetuin.

LC–UV–MS instrumentation, columns, and conditions

LC–UV–MS experiments were carried out using the following instruments: the ACQUITY UPLC system with a tunable ultraviolet detector connected on-line to a Synapt MS Q-Tof instrument (Waters, Milford, MA, USA). The 150 × 2.1 mm, 1.7 μm ACQUITY UPLC BEH glycan column was also obtained from Waters (BEH stands for bridge-ethyl hybrid silica sorbent) [31]. mAb and glycoprotein digests were analyzed using mobile phases A, 10 mM ammonium formate, pH 4.5; and mobile phase B, acetonitrile mixed with 100 mM ammonium formate, pH 4.5, in ratio 9:1 (v/v). The gradient was from 90 to 50% B in 45 min at a flow rate of 0.2 mL/min and a column temperature of 40 °C. For the analysis of bovine fetuin digest we also used MS friendly conditions, mobile phase A, 0.5% FA in water; and mobile phase B, 0.5% FA in acetonitrile. The gradient was 70 to 45% B in 50 min at a flow rate 0.2 ml/min and column temperature of 40 °C. For analysis of CD44 glycoprotein fractions we used mobile phase A, 0.01% TFA in water; and mobile phase B, 0.008% TFA in acetonitrile. The gradient was 70–45% B in 50 min with a flow rate of 0.2 ml/min and column temperature of 40 °C.

CD44 digest fractionation in RP LC was performed using 150 × 2.1 mm, 1.7 μm ACQUITY UPLC BEH C18 column, mobile phase A, 0.02% TFA in water, 0.016% TFA in acetonitrile, flow rate 0.2 ml/min, 40 °C, gradient from 0 to 50% B in 90 min.

HILIC analysis of released 2-aminobenzamide (2-AB)-labeled glycans was performed with a 150 × 2.1 mm, 1.7 μm ACQUITY UPLC BEH glycan column and UPLC system equipped with a fluorescence detector (Waters). Mobile phases consisted of 100 mM ammonium formate buffer and acetonitrile. Chromatographic conditions were described elsewhere [9]. The estimated injected amount of 2-AB-labeled glycans injected on column corresponds to ~5.6 μg of original CD44 protein.

Biopharmalynx v 1.2 software (Waters, Milford, MA, USA) was used to process the peptide maps and extract peptide retention data from the recorded LC–MS chromatograms.

Results and discussion

HILIC–UV–MS analysis of mAb

We experimentally investigated the retention behavior of glycopeptides using an ACQUITY BEH HILIC glycan column packed with efficient 1.7 μm amide sorbent. As reported earlier, HILIC mobile phases containing low concentrations of ammonium formate, FA, or TFA are compatible with on-line MS analysis [9]. This allowed us to collect UV, Fl, or MS data for eluting peaks.

Fig. 1 shows HILIC UV MS analysis of a tryptic digest of a humanized monoclonal antibody (Trastuzumab). The group of peaks exhibiting an enhanced retention (highlighted region in Fig. 1) is glycoforms of glycosylated peptide EEQYNSTYR. The peak identity was confirmed by MS data collected along with a UV chromatogram. The masses of the peaks in the highlighted region correspond to well-known glycoforms of EEQYNSTYR peptide (Table 1).

The extracted mass chromatograms of EEQYNSTYR glycopeptides are shown in Fig. 2. Interestingly, the elution order of peptide glycoforms is similar to HILIC separation of released glycans [9] with the exception of the Man5 peak. Although the HILIC resolution of peptide glycoforms is lower compared to released glycans (Fig. 2D), it is nevertheless suitable for glyco-profiling of mAb glycopeptides. UV- and MS-based relative quantitation of peptide glycoforms is summarized in Table 1. The results are similar to those obtained for released 2-AB-labeled glycans using a sensitive fluorescence detector.

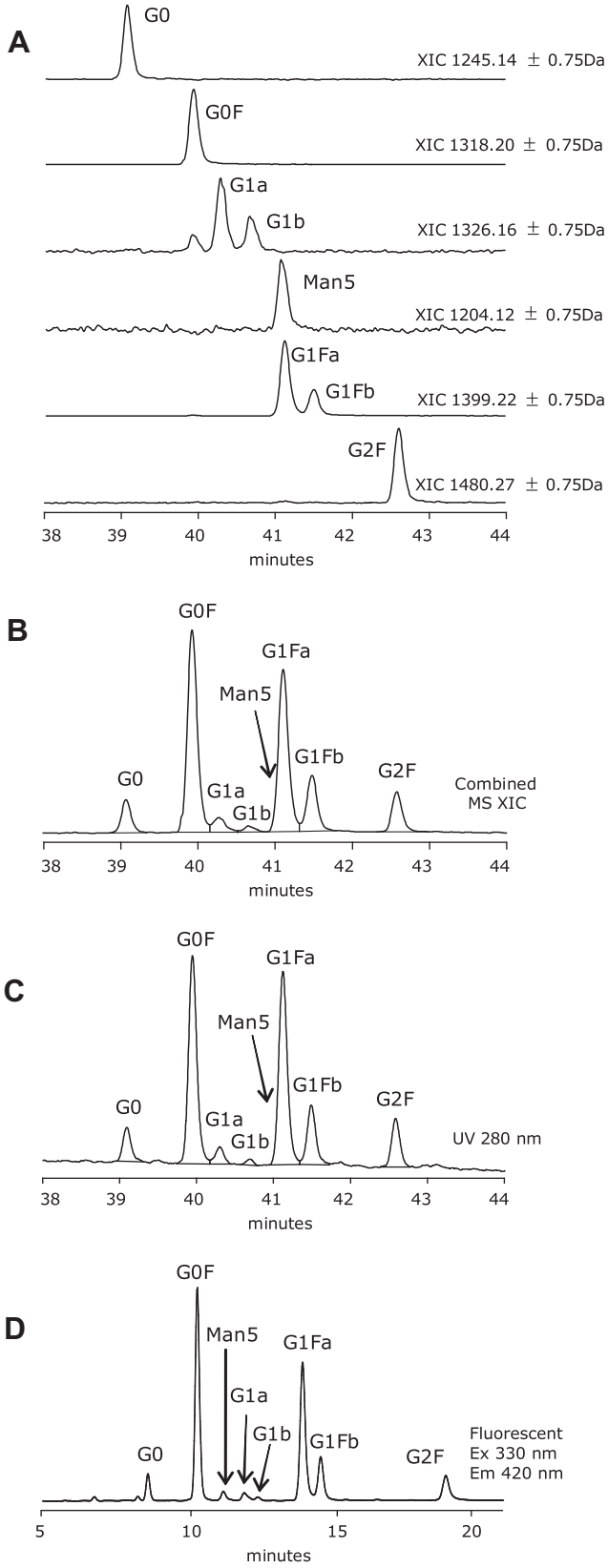

Fig.2. Quantitation of mAb EEQYNSTYR glycoforms. (A) Extracted HILIC MS chromatograms, (B) combined HILIC XIC MS, and (C) HILIC UV chromatogram. Compare with HILIC chromatogram of released 2-AB-labeled glycans; data were acquired with a fluorescent detector. For quantitative data see Table 1.

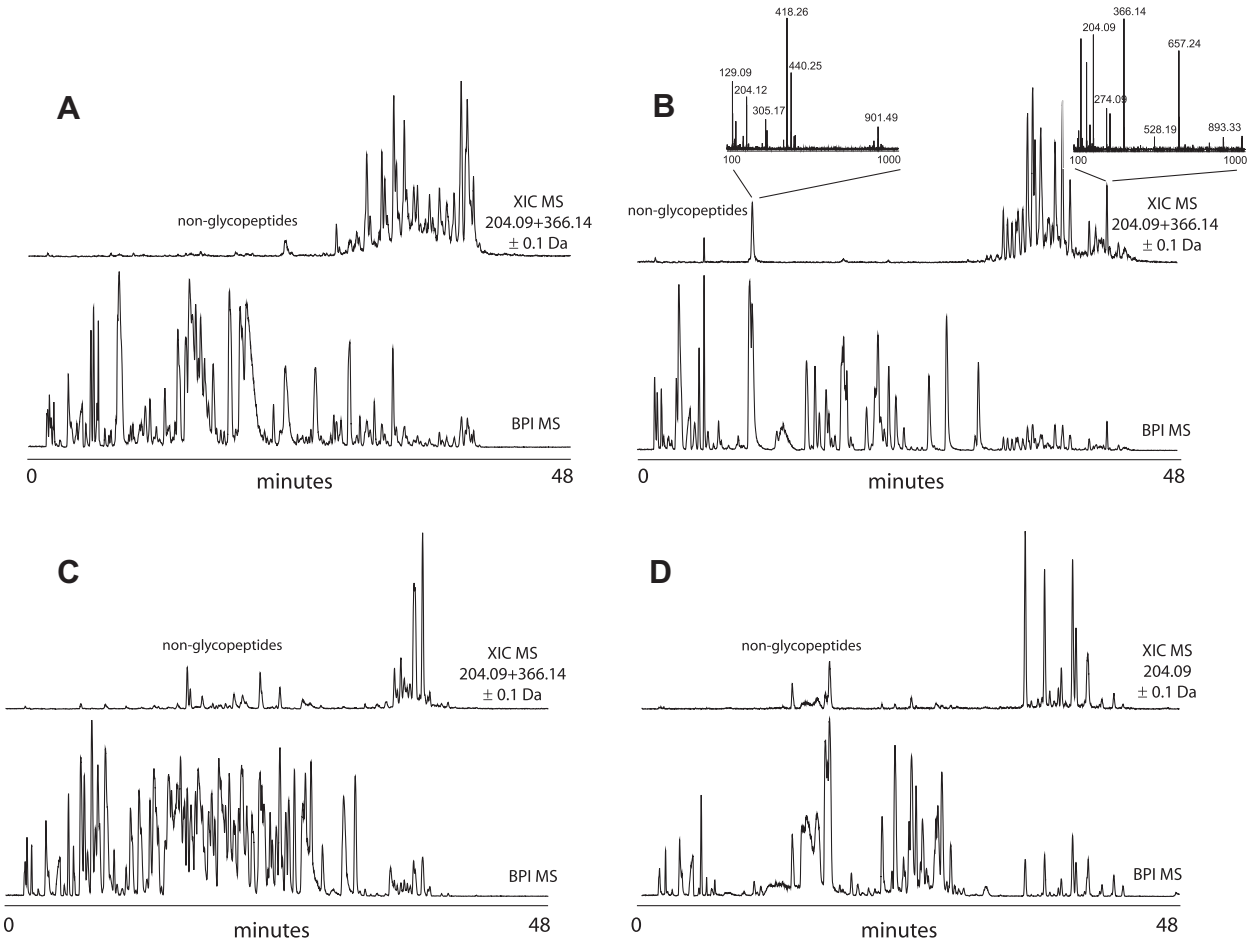

Fig.3. HILIC MS^E analysis of glycoprotein tryptic digests. (A) Bovine fetuin, (B) human alpha-1 acid glycoprotein, (C) human transferrin, and (D) bovine RNase B. The extracted ion chromatograms at 204.09 and 366.14 Da reveal the position of eluting glycopeptides.

Although simple UV detection was sufficient for quantitation of glycoforms (peptide attached to glycoforms serve as equimolar UV tag) the detection wavelength of 280 nm is suitable for glycopeptides containing thyrosines. The allotype of mAb containing phenyl alanine (EEQFNSTFR peptide) does not have a sufficient UV response, and it may be difficult to detect minor glycoforms. Unfortunately, UV detection at shorter wavelengths (e.g., 214 nm) is obscured by an ammonium formate buffer background. Thus, MS quantitation is a potential alternative.

We further investigated HILIC separation of glycopeptides using tryptic digests of simple and complex glycoproteins including bovine fetuin, human alpha-1 acid glycoprotein, human transferrin, and bovine RNase B. Fig. 3 confirms that glycopeptides are generally better retained in HILIC and resolved as a class from nonglycosylated peptides. The analysis was performed using a Synapt MS Q-Tof mass spectrometer operated in data-independent MS acquisition mode (MS^E) described elsewhere [32,33]. Briefly, the MS instrument rapidly alternates between low and elevated collision cell energy settings (duty cycle is typically 0.5–1 s). Two MS chromatograms are collected in this parallel fashion, the first containing peptide precursor masses (low collisional energy setting) and the second includes all fragmentation data acquired with an elevated collisional energy setting. MS chromatograms are linked via a common retention time. We use the MS^E term instead of tandem MS/MS for this data acquisition mode, because in MS^E mode there is no precursor selection; all peptides eluting from the column are fragmented in parallel. Software is required to bin precursors with their corresponding fragments. Only the fragments with identical elution time as precursor are binned together to reconstruct the MS/MS spectrum [32,33].

Fig. 3 illustrates how two collected sets of MS data can be used to investigate glycopeptide elution patterns. Since glycopeptides fragment easily at elevated collisional energy (glycan part), the elevated energy MS data contain glycan-specific signature ions such as 204.09 Da $(HexNAc+H)^+$, 366.14 Da $(HexNAc+Hex+H)^+$, 528.19 Da $(HexNAc+2Hex+H)^+$, and 657.24 Da $(HexNAc+Hex+NeuAc+H)^+$. We extracted the dominant 204.09 and 366.14 Da ions from the elevated energy MS chromatograms (upper traces in Fig. 3) to highlight the positions of glycopeptides in the HILIC chromatograms. All observed glycopeptides elute noticeably later and were well resolved from nonglycosylated peptides. The possible exception of bovine fetuin digest will be discussed in the next section. Extracted MS chromatograms reveal some signal(s) of early eluting peaks, e.g., in Fig. 3B. Manual inspection of reconstructed MS/MS spectra (insets in Fig. 3B) confirms that these are coincidental signals; the spectra do not contain glycan-specific fragment ions.

HILIC–MS^E analysis of complex glycoprotein digest

Fig. 4 shows HILIC MS^E analysis of a bovine fetuin tryptic digest, a protein with several N-linked and O-linked glycosylation sites. In order to maximize the MS signal we utilized MS-compatible mobile phases; both water and acetonitrile mobile phases contained 0.5% FA instead of ammonium formate [9]. We also adjusted the gradient so that the nonglycosylated peptides were weakly retained, while glycosylated ones elute between 10 and 35 min.

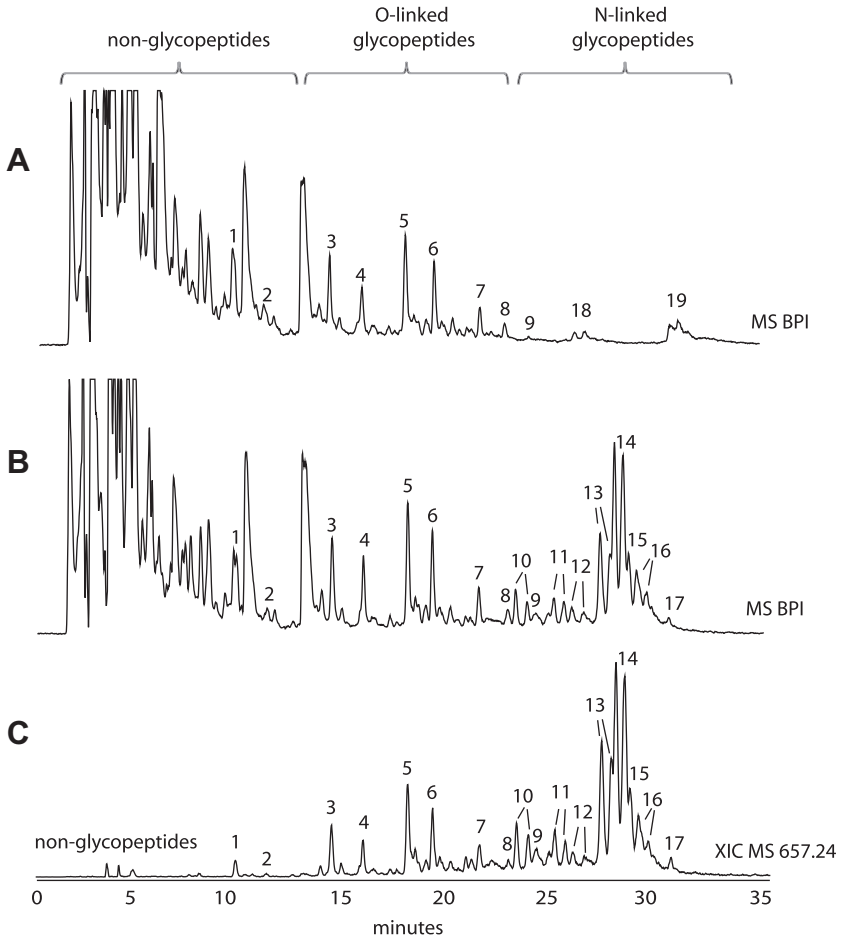

Fig.4. HILIC MS^E analysis of bovine fetuin tryptic digest. Glycopeptides identified by their accurate mass are listed in Table 2. (A) Fetuin tryptic digest exposed to PNGase F, N-linked glycans were removed; only O-linked glycopeptides are present. (B) Fetuin tryptic digest, both N-linked and O-linked glycopeptides are present. (C) Extracted MS ion chromatogram for glycan fragment ion 657.24 Da reveals the position of glycopeptides (signal extracted from chromatogram B). See Table 2 for annotation key.

Similar to the other studied proteins we used MS^E data to identify the position of eluting glycopeptides. All glycopeptides in the sample appear to be sialylated; therefore, we chose to extract 657.24 Da ion $(HexNAc+Hex+NeuAc+H)^+$. Further analysis of accurate MS data combined with literature reports [6] provided the identity of eluting peaks. Annotated glycopeptides from Fig. 4 are listed in Table 2.

Some glycopeptides (peaks 1–9) were only moderately retained in HILIC and partially overlap with a group of early eluting unmodified peptides (Fig. 4). Peaks 1–9 are glycoforms of the large T21 tryptic peptide VTCTLFQTQPVIPQPQPDGAEAEAPSAVPDAAGPTPSAAGPPVASVVVGPSVVAVPLPLHR; the underlined serine and threonine positions indicate O-glycosylated sites. Because of the lack of trypsin-cleavable sites on T21, 4 potential O-linked positions [6] are present in a single T21 peptide. Due to the relative size of the peptide and O-linked glycan moiety, the singly glycosylated form of T21 has chromatographic retention similar to that of nonglycosylated peptides. Additional glycan groups gradually increased the retention of T12 peptide (Fig. 4 and Table 2).

The tryptic digest of bovine fetuin (Fig. 4B) was compared to the same sample analyzed after its further digestion with PNGase F, an enzyme, which cleaved N-linked glycans from the glycopeptides. Fig. 4A illustrates that O-linked glycopeptides 1–9 were unaltered by PNGase F digestion, while the dominant peaks 13–16 (N-linked glycopeptides) were converted into nonglycosylated peptides. Interestingly, the reported tryptic peptide TPIVGQP**S**IPGGPVR O-glycosylated in S_{341} position [6] was not found in our data. It was present only in the nonglycosylated form (eluting at 3.66 min).

Extracted MS chromatograms of O-linked and selected N-linked glycopeptides are shown in Fig. 5. Each extracted mass (peptide glycoform) eluted as 1–4 peaks, suggesting that glycoforms with the same mass are chromatographically resolved. We also observed that in-source fragmentation of glycoforms with higher mass sometimes produces ions with mass corresponding to different glycoform species. For example, when extracting the mass of biantennary T8 glycopeptide, false positive peaks were observed at elution time corresponding to triantennary T8 species. Apparently, in-source fragmentation produced some masses corresponding to biantennary glycopeptides. This observation suggests that in-source fragmentation may complicate MS data analysis and that chromatographic separation of glycoforms remains important for glycopeptide analysis even when using MS detection. In our experience a less extensive in-source fragmentation is observed with mobile phases containing ammonium formate buffers compared to FA.

Thorough analysis of fetuin glycosylation is beyond the scope of this paper. Supplemental Fig. 1 shows XIC MS chromatograms for remaining glycopeptides. Supplemental Fig. 2 shows the HILIC-FLR chromatogram of released N-linked glycans provided for comparison. One may note similarities between the retention order of 2-AB-labeled glycans (Supplemental Fig. 2) and glycopeptides (Fig. 5) in HILIC. For example, biantennary glycopeptides elute prior to triantennary glycoforms, and trisialylated prior to tetrasialylated species; both glycoforms are resolved in two major peaks. This observation confirms that there is a similarity between retention order of glycans and glycopeptides in HILIC, and that

Table 2
Glycopeptides observed in bovine fetuin digest.

Tryptic peptide[a]	Glycan composition	Peak number	Retention time (min)	Theor. mass (monoisotopic, Da)	Dominant charged state; extracted mass (average, Da)[c]
T21 O-linked	1HexNAc/1Hex/1NeuAc	1	9.86	6670.358	5+; 1335.898
T21 O-linked	1HexNAc/1Hex/2NeuAc	2	11.39	6961.453	5+; 1394.150
T21 O-linked	2HexNAc/2Hex/2NeuAc	3	14.56	7326.585	5+; 1467.216
T21 O-linked	2HexNAc/2Hex/3NeuAc	4	16.10	7617.681	5+; 1525.468
T21 O-linked	3HexNAc/3Hex/3NeuAc	5	18.31	7982.813	5+; 1598.534
T21 O-linked	3HexNAc/3Hex/4NeuAc	6	19.51	8273.908	6+; 1380.823
T21 O-linked	4HexNAc/4Hex/4NeuAc	7	21.81	8639.040	5+; 1729.852
T21 O-linked	4HexNAc/4Hex/5NeuAc	8	23.26	8930.136	6+; 1490.254
T21 O-linked	5HexNAc/5Hex/5NeuAc	9	24.51	9295.268	6+; 1551.143
T13N-linked	4HexNAc/5Hex/2NeuAc	10	23.63	3944.602	3+; 1316.674
T13N-linked	4HexNAc/5Hex/2NeuAc	10	24.18	3944.602	3+; 1316.674
T13N-linked	5HexNAc/6Hex/3NeuAc	13	27.81	4600.830	3+; 1535.538
T13N-linked	5HexNAc/6Hex/3NeuAc	13	28.26	4600.830	3+; 1535.538
T13N-linked	5HexNAc/6Hex/4NeuAc	14	28.49	4891.925	4+; 1224.720
T13N-linked	5HexNAc/6Hex/4NeuAc	14	28.89	4891.925	4+; 1224.720
T8N-linked	4HexNAc/5Hex/2NeuAc	11	25.52	5875.532	4+; 1470.777
T8N-linked	4HexNAc/5Hex/2NeuAc	11	26.04	5875.532	4+; 1470.777
T8N-linked	5HexNAc/6Hex/3NeuAc	14	28.51	6531.760	5+; 1308.141
T8N-linked	5HexNAc/6Hex/3NeuAc	14	28.94	6531.760	5+; 1308.141
T8N-linked	5HexNAc/6Hex/4NeuAc	14	28.82	6822.855	5+; 1366.393
T8N-linked	5HexNAc/6Hex/4NeuAc	15	29.94	6822.855	5+; 1366.393
T14N-linked	5HexNAc/6Hex/4NeuAc	–	25.23	6167.662	4+; 1543.813
T14N-linked	5HexNAc/6Hex/4NeuAc	–	25.42	6167.662	4+; 1543.813
T14N-linked	5HexNAc/6Hex/4NeuAc	–	25.98	6167.662	4+; 1543.813
T14N-linked	5HexNAc/6Hex/3NeuAc	9	24.72	5876.567	4+; 1470.998
T12-13N-linked	4HexNAc/5Hex/2NeuAc	11	25.90	4072.701	4+; 1019.802
T12-13N-linked	4HexNAc/5Hex/2NeuAc	12	26.40	4072.701	4+; 1019.802
T12-13N-linked	4HexNAc/5Hex/2NeuAc	–	26.67	4072.701	4+; 1019.802
T12-13N-linked	4HexNAc/5Hex/2NeuAc	–	27.12	4072.701	4+; 1019.802
T12-13N-linked	5HexNAc/6Hex/3NeuAc	15	29.58	4728.928	4+; 1183.950
T12-13N-linked	5HexNAc/6Hex/3NeuAc	–	29.80	4728.928	4+; 1183.950
T12-13N-linked	5HexNAc/6Hex/3NeuAc	–	30.16	4728.928	4+; 1183.950
T12-13N-linked	5HexNAc/6Hex/4NeuAc	16	30.02	5020.024	4+; 1256.765
T12-13N-linked	5HexNAc/6Hex/4NeuAc	–	30.57	5020.024	4+; 1256.765
? N-linked	?	12	26.48	6513.849 [b]	4+; 1630.300
? N-linked	?	12	26.98	6513.849 [b]	4+; 1630.300
? N-linked	?	17	31.19	7188.200 [b]	5+; 1439.367
? O-linked	?	18	~27	2222.831 [b]	2+; 1112.403
? O-linked	?	19	~32	2879.063 [b]	2+; 1441.034

[a] T21: VTCTLFQTQPVIPQPQPDGAEAEAPSAVPDAAGPTPSAAGPPVASVVVGPSVVAVPLPLHR; T13: LCPDCPLLAPLNDSR; T8: RPTGEVYDIEIDTLETTCHVLDPTPLANCSVR; T14: VVHAVEVALATFNAESNGSYLQLVEISR; T12-13: KLCPDCPLLAPLNDSR.
[b] Monoisotopic masses were estimated from experimental MS data.
[c] The masses of dominant isotopic state of dominant charged state were used to extract XMS chromatograms in Fig. 5.

similar separation selectivity observed for mAb glycans and glycopeptides in Fig. 2 is not coincidental.

RP × HILIC two-dimensional (2D) chromatography of glycopeptides

The results presented in previous sections demonstrate that HILIC successfully resolved glycopeptides from a majority of the unmodified peptides and provided useful separation of their glycoforms. However, when analyzing proteins with multiple glycosylation sites, their corresponding glycopeptides are sharing the same HILIC separation space and peak overlap is significant. This makes characterization of complex glycoproteins challenging even when using efficient HILIC columns and high-resolution mass spectrometers.

As more and more glycoproteins are entering clinical trials, greater attention is required for the characterization of glycosylation site-specific microheterogeneity. We used a CD44 fusion protein with five N-linked glycosylation sites to evaluate the feasibility of site-specific glycosylation analysis. The sequence of CD44 is shown in Supplemental Fig. 3.

Tryptic digestion of CD44 was followed by AspN digestion in an attempt to generate specific short peptides with only a single glycosylation site per peptide. Despite the two-step digestion there was still one peptide with two glycosylation sites (Supplemental Fig. 3). This demonstrates a limitation of the chosen approach; because of the lack of a suitable cleavage site some glycosylation sites cannot be physically separated and their microheterogeneity independently evaluated. Less specific enzymes such as chymotrypsin or pronase could potentially solve this problem, but the lack of specificity may generate multiple peptides per glycosylation site and complicate the analysis [7].

CD44 digest was first resolved in RP LC (Fig. 6A), where the peptide retention depends primarily on the amino acid sequence. Different glycoforms of the same parent peptide are typically poorly resolved in RP LC, which in this case is turned into an advantage. The XIC MS chromatogram in Fig. 6B (extracted from elevated energy MS trace; 204.09 + 366.14 Da) shows the position of the eluting glycopeptides, each represented as a single peak or a group of incompletely resolved peaks. It must be noted that the more noticeably resolved group of peaks eluting between 35 and 37 min represents the peptide with two glycosylation sites. In the subsequent RP LC analysis the glycopeptides were collected into fractions as highlighted in Fig. 6.

Fig. 6C and D are from analysis of two selected RP LC fractions in HILIC. The HILIC mobile phase contained 0.01% TFA as modifier to enable sensitive UV and MS analysis. Fig. 6C and D represent UV

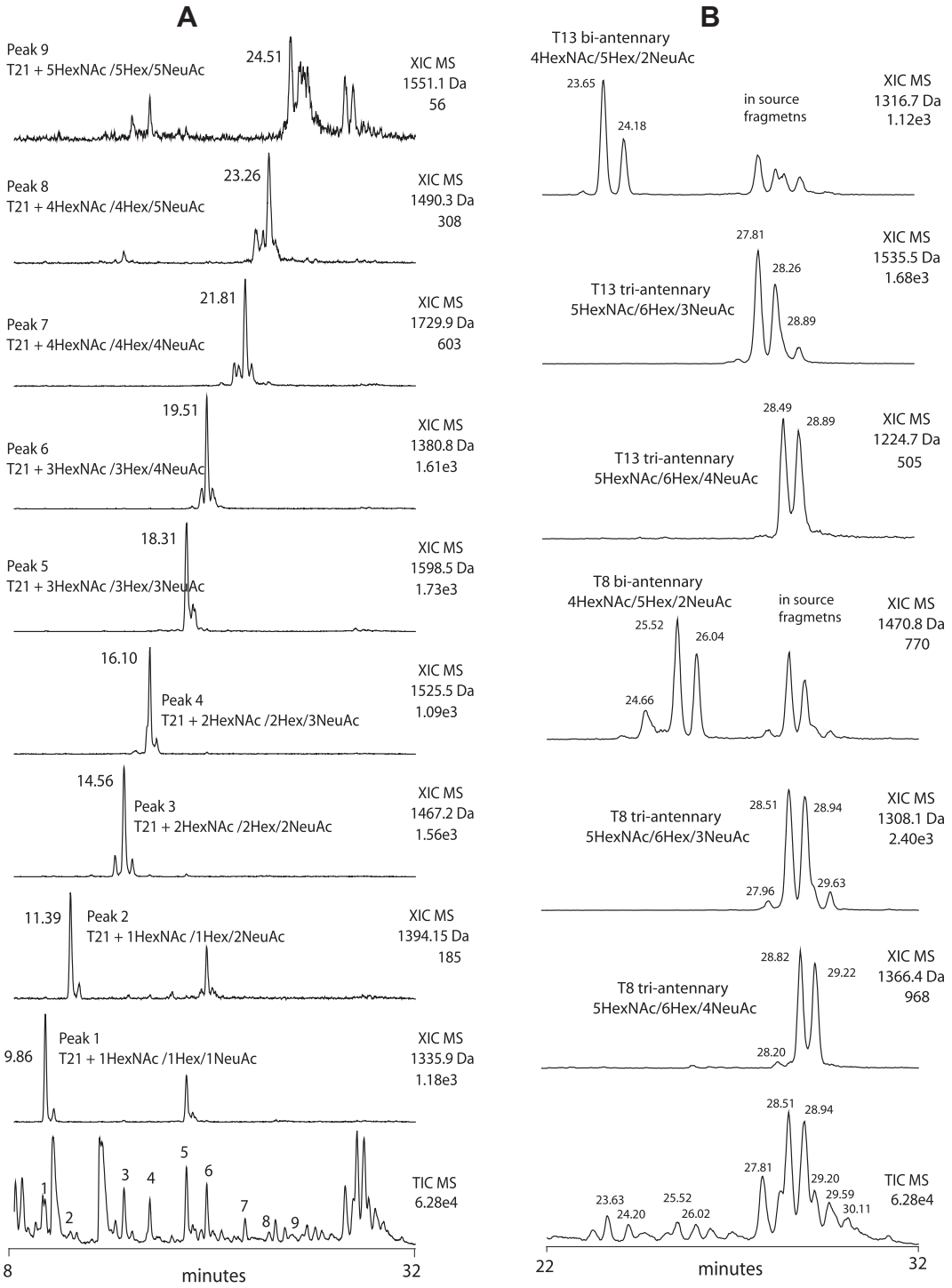

Fig.5. Extracted ion MS chromatograms for bovine fetuin glycopeptides. (A) T21 peptide O-linked glycoforms. (B) Selected N-linked glycopeptide (T13 and T8) glycoforms. For details see Table 2. Target extracted peaks are labeled with elution time.

chromatograms of EEQYNSTYR and DTYCFNASAPPEE peptides, respectively, resolved into their glycoforms. Underlined N represents the glycosylation site. Glycopeptides are labeled with numbers from 1 to 13, corresponding to glycan moieties attached to peptide: 1, 4HexNAc/3Hex/dHex; 2a and 2b, 4HexNAc/4Hex/dHex; 3, 4HexNAc/5Hex/dHex; 4, 5HexNAc/5Hex/dHex; 5, 4HexNAc/6Hex/dHex; 6, 5HexNAc/6Hex/dHex; 7, 4HexNAc/7Hex/dHex; 8, 5HexNAc/7Hex/dHex; 9, 5HexNAc/8Hex/dHex; 10, 5HexNAc/9Hex/dHex; 11, 2HexNAc/6Hex; 12, 2HexNAc/7Hex; and 13, 2HexNAc/8Hex. The peaks in Fig. 6E, presenting HILIC analysis of released 2-AB-labeled glycans, are annotated using the same convention. For example, glycopeptide EEQYNSTYR with attached 4HexNAc/3Hex/dHex glycan labeled 1 in Fig. 6C is correspondingly labeled as peak 1 in Fig. 6E, where it represents released 2-AB-labeled 4HexNAc/3Hex/dHex glycan.

Comparison of the peak heights/ratios in Fig. 6C and D with Fig. 6E illustrates the difficulties of assigning microheterogeneity of specific protein glycosylation sites using the released glycan strategy. When comparing the pattern of glycopeptides related to EEQYNSTYR in Fig. 6C to the corresponding glycans in Fig. 6E, it is

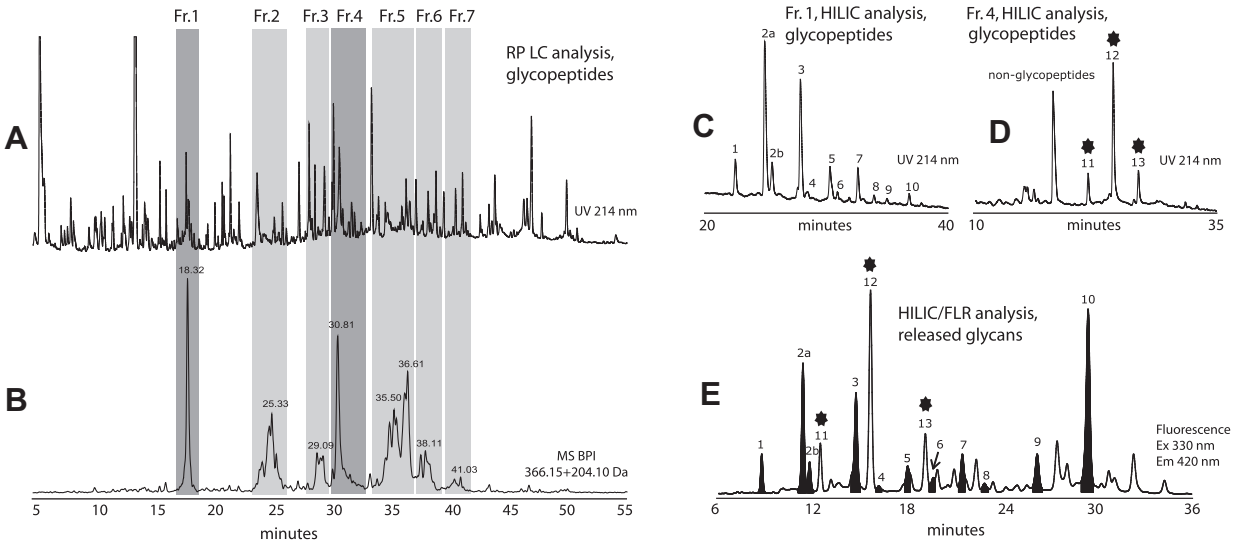

Fig.6. Off-line 2D LC analysis of CD44 tryptic + AspN digest. (A) RP LC isolation of glycopeptides. (B) Extracted glycan fragment ion MS chromatogram reveals the position of glycopeptides in RP LC. Collected fractions were analyzed using an ACQUITY UPLC glycan HILIC column. Examples of HILIC separation of fractions 1 and 4 are shown in panels C and D, respectively. Panel E shows HILIC analysis of N-linked 2-AB-labeled glycans released from CD44 protein.

apparent that peaks 1 to 8 have similar ratios. However, peaks 9 and especially 10 were present in significantly greater amounts in released glycan analysis (Fig. 6E). It is likely that other glycosylation sites in CD44 protein also contained the same type of glycan(s). When released they coelute with the glycans originating from EEQYNSTYR and bias the peak ratios. Consequently, one cannot obtain conclusive information about glycan microheterogeneity of the EEQYNSTYR peptide in CD44 protein using released glycan strategy. This is true for the remaining glycosylation sites as well.

Comparison of Fig. 6D and E reveals that the ratio of glycopeptides 11–13 is similar to that observed for released glycans (Fig. 6E). Conceivably these glycans are unique to this particular CD44 site (DTYCFNASAPPEE) and do not occur elsewhere in the CD44 protein. Therefore the ratio is not skewed by glycans released from other glycosylation sites.

One can note that the pattern of released glycans in Fig. 6E is rather complex, especially when considering minor species. Consequently, the analysis of glycosylation on the peptide level is challenging even when employing a 2D LC approach and MS^E. We believe it is beneficial to first analyze the released glycans and obtain a list of putative glycan identifications (masses) [9]. Armed with this knowledge it is then possible to combine the masses of glycans and peptides (putative glycopeptides) and use this approach to assign the glycopeptides.

Limitations of glycoprotein characterization on a peptide level

While the described method for investigation of the glycan site-specific microheterogeneity is useful, some limitations should be pointed out.

(1) UV Detection of glycopeptides may not be sensitive enough for minor glycoforms. Alternatively, the RP LC isolated glycopeptides could be deglycosylated, liberated glycans could be labeled, and then analyzed by HILIC using a sensitive fluorescence detector [9].
(2) The 2D LC approach is rather complex for routine analysis. Dilution of RP fractions with acetonitrile prior to HILIC is necessary to avoid band broadening or splitting in the HILIC second separation dimension.
(3) The MS^E technique is useful for confirmation and putative assignment of glycopeptides. The peptide identity can be discerned from the molecular weight of the informative fragment corresponding to the mass of peptide plus 203.0794 Da (HexNAc sugar) even in cases when limited peptide fragmentation hinders its sequence confirmation [12,34]. Because MS^E described here was performed with collision-induced dissociation (CID), it often does not provide sufficient fragmentation to sequence the peptide portion of glycopeptides. The electron transfer dissociation (ETD) method is more suitable for this application [35]; it recently became available on a Synapt HDMS instrument in conjunction with MS^E data-independent acquisition.
(4) Caution must be used when using MS data for relative quantitation, as the impact of different glycan forms on MS ionization efficiency is not well understood.

Conclusions

A HILIC column packed with 1.7 μm amide sorbent was suitable for characterization of glycoproteins after their digestion into glycopeptides. The presence of a hydrophilic glycan moiety significantly enhances peptide retention in HILIC and facilitates separation of peptides from glycopeptides. Glycan microheterogeneity (same peptide, various attached glycans) separation was achieved and used to study glycan site occupancy and heterogeneity for singly glycosylated proteins such as mAbs. The analysis of multiply glycosylated proteins is a more complex task; various glycopeptides and their glycoforms occupy the same separation space and tend to overlap. We demonstrated the utility of two-dimensional chromatography, where RP LC was employed for isolation of glycopeptides according to their sequence (hydrophobicity), and HILIC dimension for resolution of isolated glycopeptides into their glycoforms. MS^E-independent data acquisition was useful for distinguishing glycopeptides in the complex sample. The proposed method is potentially applicable for discovery of novel glycosylation sites and useful for comprehensive characterization of glycoproteins.

Acknowledgments

We thank Jeff Mazzeo, John C. Gebler, and Gregory T. Roman for comments on the manuscript. Dr. Xiaohong Qian acknowledges that this work was in part supported by China National Key Basic

Research Program (Grant No. 2011CB910603) and National Natural Science Foundation of China (Grant No. 20735005).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ab.2011.05.028.

References

[1] L. Mauko, A. Nordborg, J.P. Hutchinson, N.A. Lacher, E.F. Hilder, P.R. Haddad, Glycan profiling of monoclonal antibodies using zwitterionic-type hydrophilic interaction chromatography coupled with electrospray ionization mass spectrometry detection, Anal. Biochem. 408 (2011) 235–241.
[2] H. Xie, A. Chakraborty, J. Ahn, Y.Q. Yu, D.P. Dakshinamoorthy, M. Gilar, W. Chen, S.J. Skilton, J.R. Mazzeo, Rapid comparison of a candidate biosimilar to an innovator monoclonal antibody with advanced liquid chromatography and mass spectrometry technologies, MAbs 2 (2010) 4.
[3] C. Huhn, M.H. Selman, L.R. Ruhaak, A.M. Deelder, M. Wuhrer, IgG glycosylation analysis, Proteomics 9 (2009) 882–913.
[4] S. Sinha, G. Pipes, E.M. Topp, P.V. Bondarenko, M.J. Treuheit, H.S. Gadgil, Comparison of LC and LC/MS methods for quantifying N-glycosylation in recombinant IgGs, J. Am. Soc. Mass Spectrom. 19 (2008) 1643–1654.
[5] P. Olivova, W. Chen, A.B. Chakraborty, J.C. Gebler, Determination of N-glycosylation sites and site heterogeneity in a monoclonal antibody by electrospray quadrupole ion-mobility time-of-flight mass spectrometry, Rapid Commun. Mass Spectrom. 22 (2008) 29–40.
[6] G. Zauner, C.A. Koeleman, A.M. Deelder, M. Wuhrer, Protein glycosylation analysis by HILIC–LC–MS of proteinase K-generated N- and O-glycopeptides, J. Sep. Sci. 33 (2010) 903–910.
[7] Y.Q. Yu, J. Fournier, M. Gilar, J.C. Gebler, Identification of N-linked glycosylation sites using glycoprotein digestion with pronase prior to MALDI tandem time-of-flight mass spectrometry, Anal. Chem. 79 (2007) 1731–1738.
[8] Y. Takegawa, K. Deguchi, T. Keira, H. Ito, H. Nakagawa, S. 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.
[9] J. Ahn, J. Bones, Y.Q. Yu, P.M. Rudd, M. Gilar, Separation of 2-aminobenzamide labeled glycans using hydrophilic interaction chromatography columns packed with 1.7 micron sorbent, J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 878 (2009) 403–408.
[10] K. Deguchi, T. Keira, K. Yamada, H. Ito, Y. Takegawa, H. Nakagawa, S. Nishimura, Two-dimensional hydrophilic interaction chromatography coupling anion-exchange and hydrophilic interaction columns for separation of 2-pyridylamino derivatives of neutral and sialylated N-glycans, J. Chromatogr. A 1189 (2008) 169–174.
[11] X. Chen, G.C. Flynn, Analysis of N-glycans from recombinant immunoglobulin G by on-line reversed-phase high-performance liquid chromatography/mass spectrometry, Anal. Biochem. 370 (2007) 147–161.
[12] E. Wagner-Rousset, A. Bednarczyk, M.C. Bussat, O. Colas, N. Corvaia, C. Schaeffer, A. Van Dorsselaer, A. Beck, The way forward, enhanced characterization of therapeutic antibody glycosylation: comparison of three level mass spectrometry-based strategies, J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 872 (2008) 23–37.
[13] C.W. Damen, W. Chen, A.B. Chakraborty, M. van Oosterhout, J.R. Mazzeo, J.C. Gebler, J.H. Schellens, H. Rosing, J.H. Beijnen, Electrospray ionization quadrupole ion-mobility time-of-flight mass spectrometry as a tool to distinguish the lot-to-lot heterogeneity in N-glycosylation profile of the therapeutic monoclonal antibody trastuzumab, J. Am. Soc. Mass. Spectrom. 20 (2009) 2021–2033.
[14] L. Royle, M.P. Campbell, C.M. Radcliffe, D.M. White, D.J. Harvey, J.L. Abrahams, Y.G. Kim, G.W. Henry, N.A. Shadick, M.E. Weinblatt, D.M. Lee, P.M. Rudd, R.A. Dwek, HPLC-based analysis of serum N-glycans on a 96-well plate platform with dedicated database software, Anal. Biochem. 376 (2008) 1–12.
[15] T.R. Cataldi, C. Campa, G.E. De Benedetto, Carbohydrate analysis by high-performance anion-exchange chromatography with pulsed amperometric detection: the potential is still growing, Fresenius J. Anal. Chem. 368 (2000) 739–758.
[16] M. Nakano, D. Higo, E. Arai, T. Nakagawa, K. Kakehi, N. Taniguchi, A. Kondo, Capillary electrophoresis–electrospray ionization mass spectrometry for rapid and sensitive N-glycan analysis of glycoproteins as 9-fluorenylmethyl derivatives, Glycobiology 19 (2009) 135–143.
[17] Z. Szabo, A. Guttman, T. Rejtar, B.L. Karger, Improved sample preparation method for glycan analysis of glycoproteins by CE–LIF and CE–MS, Electrophoresis 31 (2010) 1389–1395.
[18] Y. Takegawa, K. Deguchi, H. Ito, T. Keira, H. Nakagawa, S. Nishimura, Simple separation of isomeric sialylated N-glycopeptides by a zwitterionic type of hydrophilic interaction chromatography, J. Sep. Sci. 29 (2006) 2533–2540.
[19] J. Wohlgemuth, M. Karas, W. Jiang, R. Hendriks, S. Andrecht, Enhanced glyco-profiling by specific glycopeptide enrichment and complementary monolithic nano-LC (ZIC-HILIC/RP18e)/ESI-MS analysis, J. Sep. Sci. 33 (2010) 880–890.
[20] M. Wuhrer, C.A. 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.
[21] Y. Takegawa, H. Ito, T. Keira, K. Deguchi, H. Nakagawa, S. Nishimura, Profiling of N- and O-glycopeptides of erythropoietin by capillary zwitterionic type of hydrophilic interaction chromatography/electrospray ionization mass spectrometry, J. Sep. Sci. 31 (2008) 1585–1593.
[22] A.J. Alpert, Hydrophilic-interaction chromatography for the separation of peptides, nucleic acids and other polar compounds, J. Chromatogr. 499 (1990) 177–196.
[23] T. Yoshida, Peptide separation in normal phase liquid chromatography, Anal. Chem. 69 (1997) 3038–3043.
[24] D.E. McNulty, R.S. Annan, Hydrophilic interaction chromatography reduces the complexity of the phosphoproteome and improves global phosphopeptide isolation and detection, Mol. Cell. Proteomics 7 (2008) 971–980.
[25] A.J. Alpert, Electrostatic repulsion hydrophilic interaction chromatography for isocratic separation of charged solutes and selective isolation of phosphopeptides, Anal. Chem. 80 (2008) 62–76.
[26] Y. Liu, X. Xue, Z. Guo, Q. Xu, F. Zhang, X. Liang, Novel two-dimensional reversed-phase liquid chromatography/hydrophilic interaction chromatography, an excellent orthogonal system for practical analysis, J. Chromatogr. A 1208 (2008) 133–140.
[27] P. Jandera, Stationary phases for hydrophilic interaction chromatography, their characterization and implementation into multidimensional chromatography concepts, J. Sep. Sci. 31 (2008) 1421–1437.
[28] D.V. McCalley, Is hydrophilic interaction chromatography with silica columns a viable alternative to reversed-phase liquid chromatography for the analysis of ionisable compounds?, J Chromatogr. A 1171 (2007) 46–55.
[29] P. Jandera, T. Hajek, V. Skerikova, J. Soukup, Dual hydrophilic interaction-RP retention mechanism on polar columns: structural correlations and implementation for 2-D separations on a single column, J. Sep. Sci. 33 (2010) 841–852.
[30] T. Imre, G. Schlosser, G. Pocsfalvi, R. Siciliano, E. Molnar-Szollosi, T. Kremmer, A. Malorni, K. Vekey, Glycosylation site analysis of human alpha-1-acid glycoprotein (AGP) by capillary liquid chromatography-electrospray mass spectrometry, J. Mass. Spectrom. 40 (2005) 1472–1483.
[31] K.D. Wyndham, J.E. O'Gara, T.H. Walter, K.H. Glose, N.L. Lawrence, B.A. Alden, G.S. Izzo, C.J. Hudalla, P.C. Iraneta, Characterization and evaluation of C18 HPLC stationary phases based on ethyl-bridged hybrid organic/inorganic particles, Anal. Chem. 75 (2003) 6781–6788.
[32] S.J. Geromanos, J.P. Vissers, J.C. Silva, C.A. Dorschel, G.Z. Li, M.V. Gorenstein, R.H. Bateman, J.I. Langridge, The detection, correlation, and comparison of peptide precursor and product ions from data independent LC–MS with data dependant LC–MS/MS, Proteomics 9 (2009) 1683–1695.
[33] J.C. Silva, M.V. Gorenstein, G.Z. Li, J.P. Vissers, S.J. Geromanos, Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition, Mol. Cell. Proteomics 5 (2006) 144–156.
[34] S.M. Peterman, J.J. Mulholland, A novel approach for identification and characterization of glycoproteins using a hybrid linear ion trap/FT-ICR mass spectrometer, J. Am. Soc. Mass Spectrom. 17 (2006) 168–179.
[35] I. Perdivara, R. Petrovich, B. Allinquant, L.J. Deterding, K.B. Tomer, M. Przybylski, Elucidation of O-glycosylation structures of the beta-amyloid precursor protein by liquid chromatography-mass spectrometry using electron transfer dissociation and collision induced dissociation, J. Proteome Res. 8 (2009) 631–642.