



Separation and Characterization of N-linked Glycopeptides on Hemagglutinins in a Recombinant Influenza Vaccine

Hongwei Xie, Weibin Chen, Martin Gilar, StJohn Skilton, Jeffrey R. Mazzeo

Waters Corporation



Abstract

Influenza vaccines provide protection by neutralizing antibodies to viral hemagglutinin (HA), a protein that plays a critical role in influenza infection. In this application note, we demonstrate that UPLC-MS^E is capable of separating and characterizing N-linked glycosylations of HA proteins in a recombinant influenza vaccine candidate expressed from insect cell-baculovirus expression vector system (BEVS). Glycopeptides and glycoforms are separated by an ACQUITY UPLC System at the peptide level, and are detected online by a SYNAPT MS System. The UPLC-MS^E data are processed by BiopharmaLynx Software to report N-linked glycosylation information.

Introduction

Influenza is a viral infection of the respiratory tract. It is one of the leading causes of death in the U.S., killing more than 50,000 people per year.¹ Influenza vaccination is a primary prophylactic method and the principal strategy for reducing morbidity and mortality due to seasonal influenza. Vaccines provide protection by neutralizing antibodies to viral hemagglutinin (HA), a protein that plays a critical role in influenza infection.

Licensed, inactivated vaccines for seasonal influenza usually contain a preset amount of HAs – a mixture of H1, H3, and B, the corresponding HA proteins of the three most common viruses – influenza A subtypes H1N1 and H3N2, and influenza B. These HA proteins are glycoproteins, with multiple N-linked glycosylation motifs and multiple glycoforms for each glycosylation site. Detailed characterization and monitoring of glycosylation in HAs is important for both vaccine development and production because of their role in determining the function of influenza binding onto host cells and therefore infection.

Currently, the methods for characterization of N-linked glycosylations include released free glycan analysis²⁻⁴ and intact mass analysis.⁵⁻⁷ These methods are useful for analyzing glycoproteins with defined glycosylation sites, such as monoclonal antibodies. Glycan profiling can be performed at the intact protein level (intact mass analysis) or as carbohydrates (free glycan profiling). The glycosylation site can usually be determined by peptide mapping⁸ after the glycan moieties are enzymatically removed (mostly by PNGase F) because the mass of asparagine (N) residues with glycosylation increase by 1 Da upon deglycosylation.

However, it is a challenge for these methods to characterize glycoproteins with multiple glycosylation sites, such as HAs, because it is difficult to distinguish between glycan moieties that link to different sites in a

protein. Furthermore, complex samples like vaccines have multiple N sites with an -NXS/T- motif. The determination of N-linked glycosylation sites by peptide mapping may be extremely challenging because modifications of N sites with a 1 Da mass increase could be due to either glycosylation or deamidation.

Using improved resolution of UPLC, four major N-linked glycoforms of a tryptic peptide prepared from a monoclonal mouse IgG1 antibody tryptic digestion were analyzed by LC/UV-MS.⁹ The experiment demonstrated that glycosylations could be detected and quantified. Both glycosylated site and glycan moiety could be characterized simultaneously.

In previous studies,¹⁰⁻¹¹ we demonstrated that tryptic peptide mapping with UPLC-MS^E was capable of separating and characterizing site-specific modifications such as N-deamidation and M-oxidation in an unbiased manner.

In this application note, we demonstrate that UPLC-MS^E is capable of separating and characterizing N-linked glycosylations of HA proteins in a recombinant influenza vaccine candidate expressed from insect cell-baculovirus expression vector system (BEVS). Glycopeptides and glycoforms are separated by an ACQUITY UPLC System at the peptide level, and are detected online by a SYNAPT MS System. The UPLC-MS^E data are processed by BiopharmaLynx Software to report N-linked glycosylation information. The method offers a way to improve the characterization as well as reduce the amount of time spent on data processing. Furthermore, having a general technique that can be applied to such problems opens this type of analysis to non-experts and could benefit an organization by streamlining their work.

Experimental

A tryptic digest was prepared for the influenza vaccine candidate sample containing HA proteins H1, H3, and B that were expressed from insect cell-BEVS system. The peptide mixture contained N-linked glycopeptides together with other peptides from the proteins. The preparation procedure included:

1. Protein denaturing by heating at 80 °C for 10 min in the presence of 0.05% RapiGest SF and at pH 7.4
2. Reduction with dithiothreitol for 30 min at 56 °C
3. Alkylation with iodoacetamide for 30 min at room temperature in the dark
4. Digestion with trypsin for 4 hrs at 37 °C and pH 7.4

5. Quenching and de-activation of trypsin by adding formic acid (FA, at 0.1%)

The digest was diluted to 0.2 $\mu\text{g}/\mu\text{L}$ with 0.1% FA in 5% acetonitrile (ACN) for UPLC-MS^E analysis.

UPLC-MS^E experiments were performed using a SYNAPT MS System coupled with an ACQUITY UPLC System. The UPLC system was configured with a 2.1 x 150 mm, BEH300 C₁₈ 1.7- μm Peptide Separation Technology column. About 4 μg of the peptide mixture in a 20- μL volume was injected, and eluted using a 120-min gradient (1 to 40% ACN in 0.1% FA) at a flow rate of 0.2 mL/min and column temperature of 60 °C. Four injections were repeated.

MS^E data were acquired in ESI positive ion mode, with collision cell energy alternating between low energy (5 V) to collect peptide precursor (MS) data and elevated energy (ramping from 20 to 40 V) to obtain peptide fragmentation (MS^E) data. The scan time was 0.5 sec (1 sec total duty cycle). Capillary voltage of 3.0 kV, source temperature of 100 °C, cone voltage of 37 V, and cone gas flow of 10 L/h were maintained during the analyses. The system was tuned for a minimum resolution of 10,000 (V-mode) and calibrated using a 100 fmol/ μL Glu1-fibrinopeptide B (GFP) infusion. Sampling of the lock spray channel (100 fmol/ μL GFP in 50:50 ACN/water containing 0.1% FA) was performed every 1 min to ensure high mass accuracy.

The collected data were processed by BiopharmaLynx, v.1.2, an application manager for MassLynx Software, using a strict tryptic cleavage rule, and setting cysteine carbamidomethylation as a fixed modification and N-linked glycosylations as variable modifications. Additional BiopharmaLynx method settings were detailed in a previous publication.¹²

Results and Discussion

It has been reported¹³ that glycoproteins expressed from insect cell lines present two major types of N-linked glycans: paucimannosidic structures (Man(1-3)GlcNAc₂ or Man(1-3)GlcNAc[Fuc]GlcNAc) and oligomannose structures (Man(5-9)GlcNAc₂), where Man represents mannose, GlcNAc is N-acetylglucosamine, and Fuc is fucose. All 11 possible glycosylation forms were entered as variable N-linked glycosylation modifications for BiopharmaLynx data processing. The number of N sites available for N-linked glycosylation (with -NXS/Tmotif) are 9, 12, and 10 in HA proteins H1, H3, and B, respectively. After BiopharmaLynx processing of the collected UPLC-MS^E data, we were able to identify 7, 4, and 6 N sites with glycosylation for H1, H3, and B, respectively. Each N-linked glycosylation site may have multiple attached glycoforms. The following are three typical examples (one from each HA protein in the vaccine sample) selected to demonstrate how

UPLC-MS^E can separate and characterize glycopeptides and glycoforms.

Figure 1 shows the separation and identification of unmodified and glycosylated tryptic peptide T8 in HA protein B (named B_T8 and B_T8*, respectively; in the following text, all glycosylated peptides are marked with an asterisk (*) to signify a corresponding modified form).

B_T8* eluted earlier than B_T8 because the attached sugar group increases hydrophilicity (Figure 1A). The tryptic peptide sequence was confirmed by MS^E spectra (Figures 1B and 1C). The glycosylation of B_T8* was indicated by a series of characteristic sugar ions at m/z 138.05, m/z 204.09, m/z 366.14, and m/z 528.12 in the low m/z range and confirmed by ion y28. The 1038.36 mass difference of y28 before and after glycosylation corresponds to a glycan moiety (-Man3NAcGlc[Fuc]NAcGlc). The fragment ions with sugar groups at high m/z range give further structural information about the attached glycan moiety. Based on the MS signal intensities processed by BiopharmaLynx and extracted ion chromatographic areas of precursor masses, the relative concentration of B_T8* is 20%.

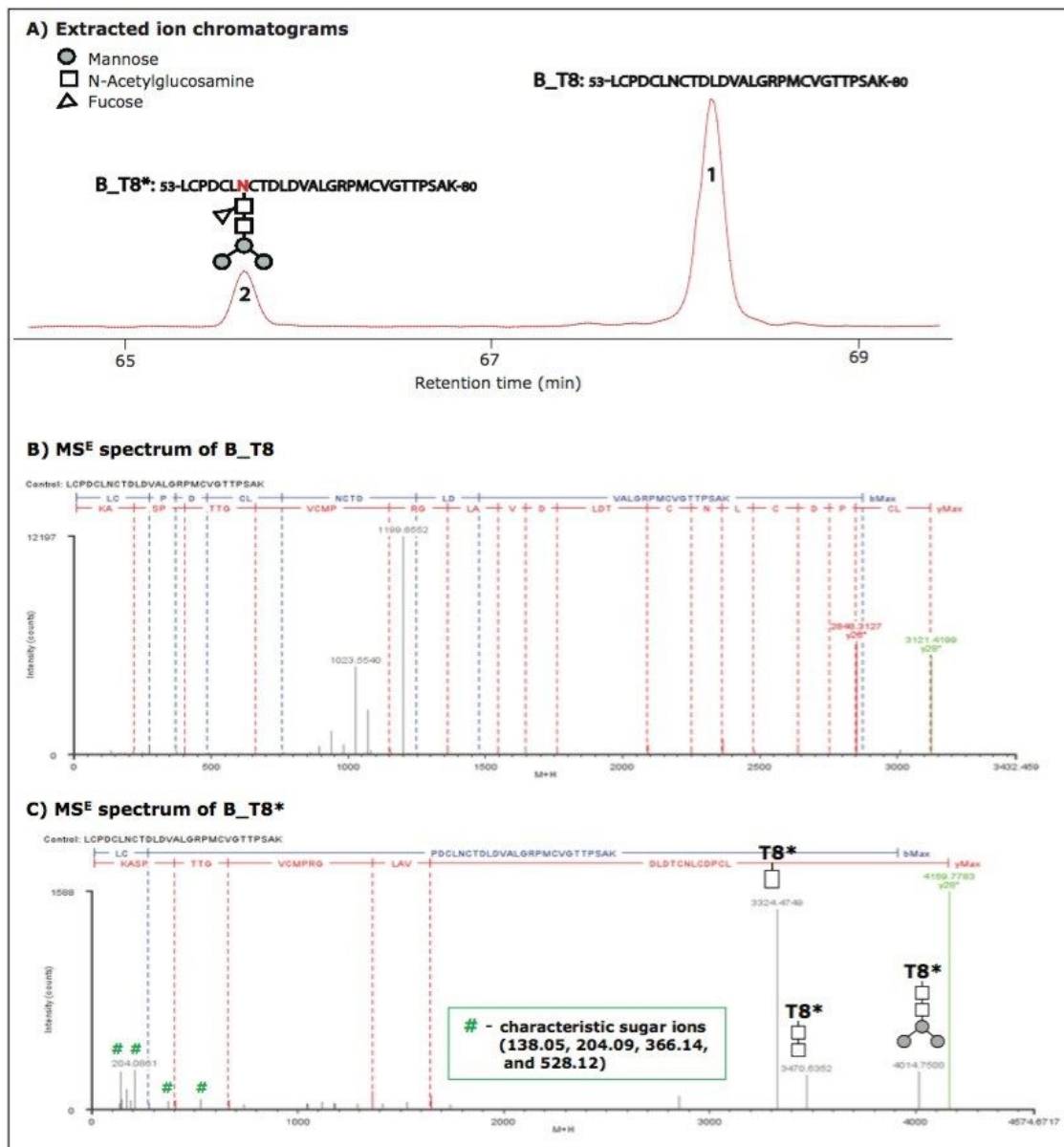


Figure 1. Separation and characterization of unmodified and glycosylated tryptic peptide T8 in HA protein B.

A) Extracted ion chromatograms of precursors

B) MS^E spectrum of B_T8 (The unassigned ions m/z 1199.65 and m/z 1023.55 are fragments of coeluted Twin-20 that was added during vaccine formulation)

C) MS^E spectrum of B_T8*

Unlike B_T8*, most identified glycopeptides were fully glycosylated with multiple glycoforms. For example, peptide T24 of HA protein H1. No unmodified H1_T24 was identified, but four glycoforms of H1_T24* were

chromatographically resolved and identified (Figure 2A). The glycoform elution order correlated to the size of glycan moiety. The heavier the glycan moiety, the earlier the glycoform eluted for this glycosylated peptide. The peptide has two N sites. MS^E spectra confirmed that the glycosylation only occurred on N286 which has the -NSS- motif, and that no glycans attached to N293 with the -NVH- motif. This is consistent with the rule that N-linked glycosylations only occur on N sites with -NXS/T- motif. The MSE spectrum of H1_T24*-Man3NAcGlc2 is shown in Figure 2B. Again, the characteristic sugar ions in the low *m/z* range and the fragment ions with sugar groups in the high *m/z* range further confirm the glycosylation and provide structural information of the glycan moiety.

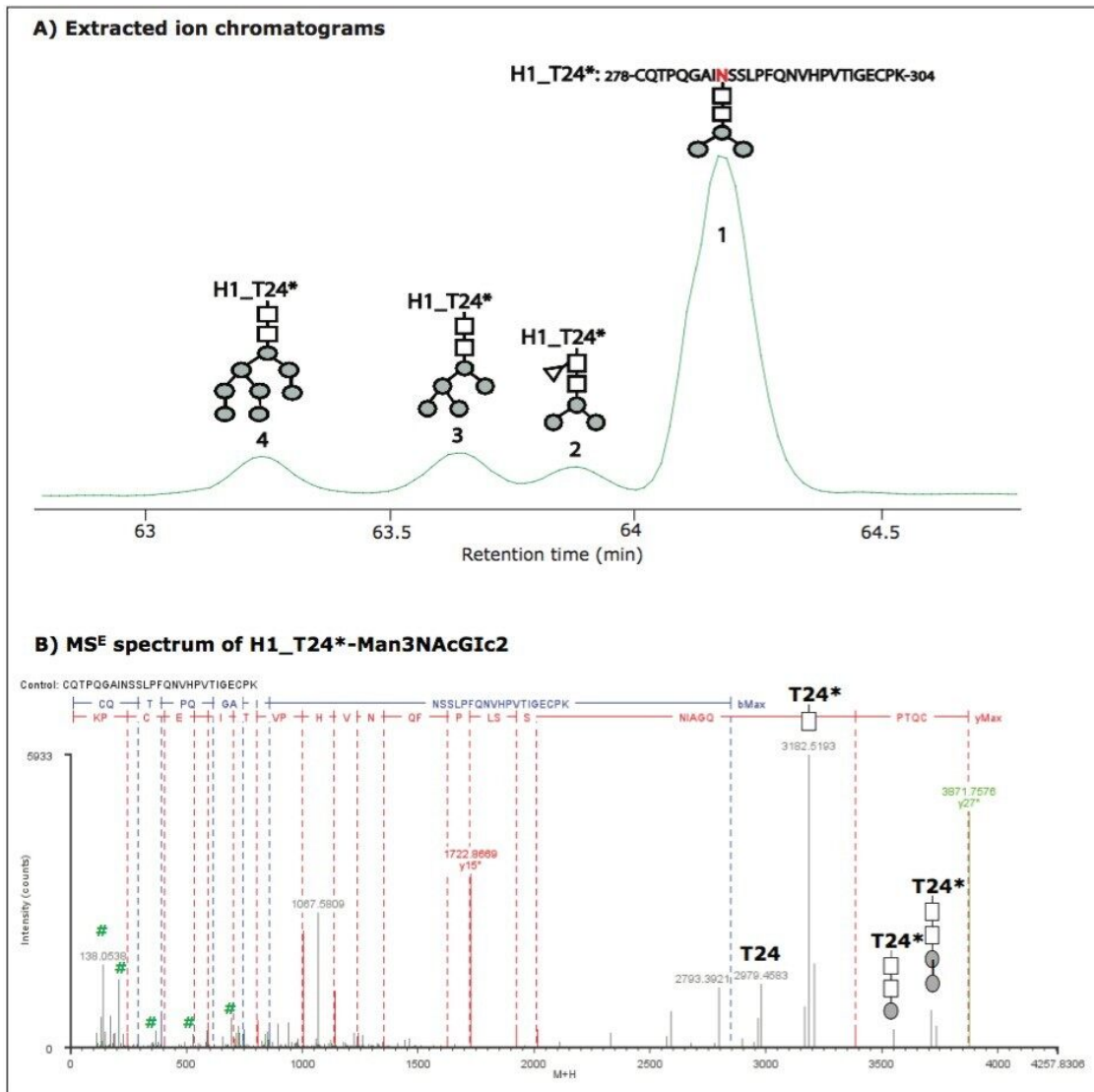


Figure 2. Separation and characterization of glycoforms of glycopeptide T24* in HA protein H1.

A) Extracted ion chromatograms of precursors

B) MS^E spectrum of H1_T24*-Man3NAcGlc2.

- characteristic sugar ions

However, not all identified glycoforms were chromatographically resolved. For example, five glycoforms were identified on tryptic peptide T21 from HA protein H3 (see Figure 3A), but these glycoforms were eluted in one LC peak at 91.65 min. Their retention times show small differences (~ 2 seconds) in the extracted ion chromatograms (data not shown). This evidence and the example described above indicate that the chromatographic behavior of glycoforms is related to the nature of the peptide sequence to which the glycan is attached. The peptide sequence of H3_T21 and glycosylations could be confirmed by MS^E spectra. An

example MS^E spectrum of H3_T21*-Man9NAcGlc2 was plotted in Figure 3B.

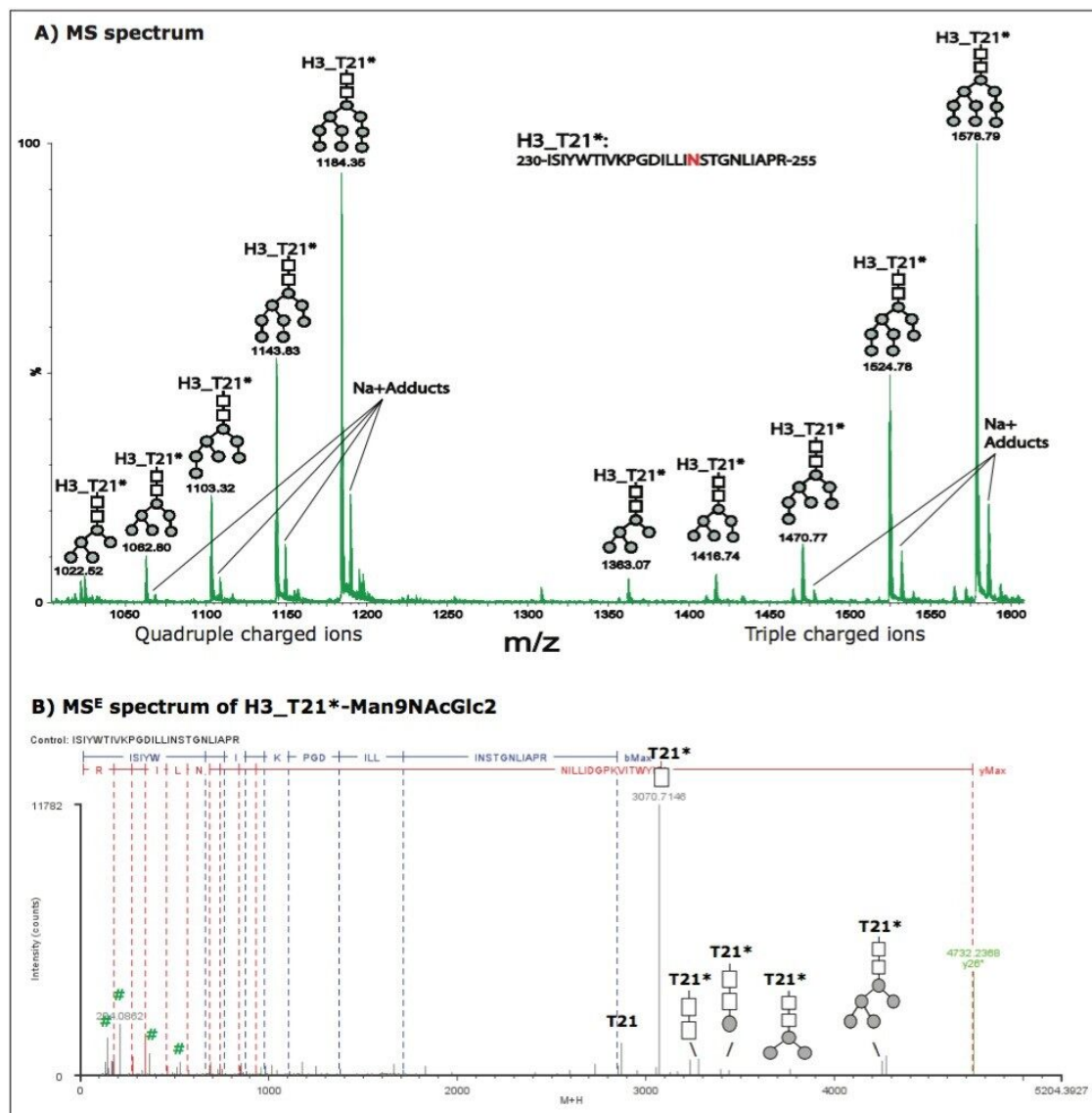


Figure 3. Characterization of glycoforms of glycopeptide T21* in HA protein H3.

A) MS spectrum

B) MS^E spectrum of H3_T21*-Man9NAcGlc2.

- characteristic sugar ions

Conclusion

The results presented here demonstrate that UPLC-MS^E can separate and characterize multiple glycopeptides and multiple glycoforms on hemagglutinin glycoproteins in an influenza vaccine sample. Because HA proteins have multiple glycosylation sites and multiple glycoforms, characterization of glycosylations in influenza vaccine samples would be a challenging task for traditional methods such as glycan analysis and intact mass analysis.

Unlike traditional glycosylation characterization methods, in this work the glycosylation sites were unambiguously identified utilizing the MS^E technique. MS^E also provided useful structural information of glycan moieties from the glycan fragments. In addition, the methods do not involve any additional enrichment or purification procedure, and the data acquired can be automatically processed by BiopharmaLynx.

Further optimization could make this method applicable not only to the characterization of glycosylations in complex samples such as an influenza vaccine, but also as a rapid and routine method for characterizing glycoproteins.

References

1. Stevens J, Corper AL, Basler CF, Taubenberger JK, Palese P, Wilson IA. *Science*. 2004; 303:1866-1870.
2. Yu YQ, Gilar M, Kaska J, Gebler JC. Rapid Commun. *Mass Spectr.* 2005; 19:2331-6.
3. Ahn J, Yu YQ, Gilar M. Rapid sample cleanup method development for 2-AB labeled N-linked glycans released from glycoproteins. Waters Application Note. 2008; 720002622en.
4. Gillece-Castro B, van Tran K, Turner JE, Wheat TE, Diehl DM. N-linked glycans of glycoproteins: a new column for completed resolution. Waters Application Note. 2009; 720003112en.
5. Oliviva P, Chen W, Chakraborty AB, Gebler JC. Rapid Commun. *Mass Spectr.* 2008; 22:29-40.
6. Chakraborty AB, Berger SJ, Gebler JC. Characterization of an IgG1 monoclonal antibody and related sub-structures by LC/ESI-TOF-MS. Waters Application Note. 2008; 720002107en.
7. Chakraborty AB, Chen W, Gebler JC. Characterization of reduced monoclonal antibody by online UPLC-UV/ESI-TOF MS. Waters Application Note. 2009; 720002919en.
8. Orlando R. Rapid Commun. *Mass Spectr.* 2007; 21:674-682.

9. Gillece-Castro BL, Wheat TE, Mazzeo JR, Diehl D M. UPLC technology for the analysis of antibody glycopeptides. Waters Application Note. 2008; 720002382en.
10. Xie HW, Gilar M, Gebler JC. Analysis of deamidation and oxidation in monoclonal antibody using peptide mapping with UPLC/MS^E. Waters Application Note. 2009; 720002897en.
11. Xie HW, Gilar M, Gebler JC. *Anal. Chem.* 2009; 81:5699-5708.
12. Ahn J, Gillece-Castro BL, Berger S. BiopharmLynx: A new bioinformatics tool for automated LC/MS peptide mapping assignment. Waters Application Note. 2008; 720002754en.
13. Tomiya NT, Narang S, Lee YC, Betenbaugh MJ. *Glycoconjugate J.* 2004; 21:343-360.

Featured Products

ACQUITY UPLC System <<https://www.waters.com/514207>>

BiopharmaLynx <<https://www.waters.com/513793>>

720003173, August 2009