International Journal of Mass Spectrometry xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

International Journal of Mass Spectrometry

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

Comparison of glycopeptide fragmentation by collision induced dissociation and ultraviolet photodissociation

Byoung Joon Ko, Jennifer S. Brodbelt *

Department of Chemistry, University of Texas at Austin, 1 University Station A5300, Austin, TX 78712-1167, United States

ARTICLE INFO

Article history:
Received 15 April 2014
Received in revised form 12 July 2014
Accepted 17 July 2014
Available online xxx

Keywords:
Ultraviolet photodissociation
Collision induced dissociation
Glycopeptide

ABSTRACT

A comparison of the fragmentation pathways of both protonated and deprotonated O-linked glycopeptides from fetuin and κ-casein obtained upon collision induced dissociation (CID) and 193 nm ultraviolet photodissociation (UVPD) in a linear ion trap is presented. A strategy using non-specific pronase digestion, zwitterionic hydrophilic interaction liquid chromatography (ZIC–HILIC) solid phase extraction (SPE) enrichment, and nano-liquid chromatography (nano-LC) is employed. UVPD of deprotonated glycopeptides generally produced the greatest array of fragment ions, thus affording the most diagnostic information about both glycan structure and peptide sequence. In addition, UVPD generated unique fragment ion such as *Y*-type ions arising from cleavage at the N-terminus of proline. CID and UVPD of protonated glycopeptides produced fragment ions solely from glycan cleavages.

1. Introduction

Glycosylation is the most abundant post-translational modification and contributes to the functional roles of proteins that are related to cell adhesion, signaling, immune response, and inflammation [1–3]. In addition, there have been significant inroads in understanding the impact of the glycosylation status of proteins on human diseases such as cancer and inflammatory diseases [4–7] Despite the prevalence of glycosylation, it is estimated that only ~10% of glycoproteins have been characterized [8]. In order to understand the role of glycosylation, systematic characterization of this modification is required. Tandem mass spectrometry has become one of the most versatile approaches for elucidating the glycoproteome, typically via sequencing the glycan and protein portions via bottom-up strategies [9,10].

Glycosylation analysis is especially challenging due to the complexity that arises from the heterogeneity of glycosylation, including the formation of branched glycan structures (unlike the linear sequences of proteins and nucleic acids). Moreover, variations in glycosylation sites and occupancy, as well as differences stemming from two dominant types of glycosylation (comprising N- and O-linkages), further complicate glycosylation characterization [2,11]. Although cells may produce the same protein sequences with the same glycosylation sites, even then the glycans can be attached via different branching patterns and intersaccharide linkages [3,11]. With respect to the glycan attachment sites, N-linked glycosylation occurs at specific sequences: Asp-Xxx-Ser/Thr (where the Asp side chain is the glycosylation site, and Xxx can be any amino acid except proline). In contrast, O-linked glycans may be attached to serine or threonine without specific sequon to help identify their locations in proteins. The differences in the attachment sites and constituent complexity of N-glycans versus O-glycans has motivated the development of different tandem mass spectrometric methods for their analysis [3]. For example, N-linked glycans are typically substantially larger than O-linked glycans and can be selectively cleaved from proteins with peptide-N-glycosidase F (PNGase F) [3]. In addition, the known N–X–S/T amino acid sequon streamlines the mapping of the N-linked glycans because they are restricted to attachment at a well-defined amino acid sequence. For these reasons, N-glycosylation analysis has generally focused more heavily on elucidating the glycan structures, not the glycosylation sites. In contrast, the smaller size of O-glycans and less specific sites of attachment (in fact, conceivably over 10% of residues can be O-glycosylated based on the frequency of Thr and Ser amino acids in proteins) means that identifying the O-glycosylation sites is usually the greater challenge.

Collision-induced dissociation (CID) has been the most popular MS/MS technique for glycopeptide analysis [11–17], but it typically causes only glycosidic cleavages that are useful for assigning the glycan portions but not for sequencing the peptide portions or pinpointing the glycosylation sites. CID also produces some reporter ions such as *m/z* 204 ($[HexNAc+H]^+$), 292 ($[NeuNAc+H]^+$), and 366 ($[Hex\text{-}HexNAc+H]^+$) which are diagnostic for the

* Corresponding author. Tel.: +1 512 471 0028; fax: +1 512 471 8696.
E-mail address: jbrodbelt@cm.utexas.edu (J.S. Brodbelt).

http://dx.doi.org/10.1016/j.ijms.2014.07.032
1387-3806/

presence of glycan modifications. These fragments are dominant because CID preferentially produces *Y*- and *B*-type glycosidic cleavage ions for glycopeptides [10] in which *B*-type ions contain the non-reducing end of the glycan and *Y*-type ions contain the reducing end. The peptide sequences of the glycopeptides may be deduced after CID cleavage of the glycan portion, albeit requiring more elaborate multi-step MS^3 strategies [18]. Electron-based dissociation techniques including electron transfer dissociation (ETD) [13,19–25] and electron capture dissociation (ECD) [26–30] have also been employed to analyze glycopeptides. Interestingly, the electron-based methods predominantly result in formation of peptide sequence ions, not the glycan ions created by CID. In addition, retention of the glycan moieties by the *c*-type ions allows facile determination of the glycosylation sites of the peptides [30,31]. Although more data intensive, pairing CID with electron-based MS/MS methods provides a compelling route for obtaining both glycan and peptide information [23,28,32]. Photon-based activation methods including infrared multiphoton dissociation (IRMPD) [12,26–28] and ultraviolet photodissociation (UVPD) [33] also have shown promising results for characterization of glycopeptides. For example, Lebrilla et al. reported the use of IRMPD for analysis of N-linked glycopeptides from mucin-type glycoproteins, generating spectra that displayed peptide fragments in addition to products from glycosidic bond cleavages [32]. Reilly and Zhang utilized UVPD using 157 nm photons to obtain both glycan and peptide sequence information from rich spectra containing *x*-, *v*-, *w*-, and *y*-type peptide fragment ions that retained the glycan as well as both glycosidic fragments and cross-ring cleavage products from N-linked glycopeptides [33]. We demonstrated 193 nm UVPD for the characterization of acidic O-linked glycopeptide anions produced upon trypsin/Glu-C digestion of O-glycoproteins [31].

The MS/MS methods used for protein identification have most commonly been applied for positively charged ions because trypsin is the most frequently used protease and yields Arg- or Lys-terminated peptides amenable to protonation. However, trypsin digestion is not as favorable for glycoproteins in general because the bulky glycans may hinder proteolysis and the low distribution of basic residues around the glycan site may lead to production of overly large peptides, impeding bottom-up strategies. Moreover, acidic glycans, such as those containing sialic acid, may suppress protonation. For these reasons, several groups have explored the characterization of glycopeptides in the negative mode [12,14,18,31,34–36]. For example, Lebrilla and coworkers reported that IRMPD of deprotonated O-glycopeptide ions resulted in Ser- and Thr-specific side-chain losses that proved useful for pinpointing glycosylation sites [12,34]. Another effort to overcome the shortcomings of trypsin entails the use of pronase which is a non-specific protease [12,32,34,36,37]. Pronase is a mixture of several proteases that leads to more extensive digestion of proteins into smaller peptides containing fewer amino acids. Because of the non-specificity of proteolytic cleavage using pronase, there is no expected amino acid terminus or length of peptide after pronase digestion. The Lebrilla group has reported the use of immobilized pronase to decrease the sizes of the glycopeptide sequences and correspondingly increase their dissociation efficiencies while reducing the array of confounding non-glycosylated peptides [12,32,34].

Our group has developed 193 nm UVPD for structural characterization of a variety of biologically-relevant molecules, including proteins [38–40], N-linked glycans [41], lipids [42], and nucleic acids [43]. Several other groups have explored the development and application of UVPD (either using 157 nm or 193 nm photons) as an activation method [33,44–50], and the high energy deposited by UV photons makes it an attractive option for biological molecules. UVPD produces rich fragmentation patterns that enhance the confident application of database search algorithms [39] and provide diagnostic ions that allow differentiation of subtle modifications of biopolymers. The recent promising results obtained upon the elucidation of glycan anions by UVPD (193 nm) [41], N-linked glycopeptides by UVPD (157 nm) [33], and O-linked glycopeptides by UVPD (193 nm) [31] motivated our interest in extending this methodology to O-linked glycopeptides created upon pronase digestion of glycoproteins. We report a systematic comparison of CID and UVPD for both positively- and negatively-charged O-linked glycopeptides from fetuin and κ-casein in the present study.

2. Experimental

2.1. Reagents

Fetuin, κ-casein, and pronase were purchased from Sigma (St. Louis, MO). ZIC–HILIC SPE column material was obtained from SeQuant (Southboro, MA). All other chemicals were purchased from Fisher Scientific (Pittsburgh, PA) and used without any further purification.

2.2. Pronase digestion

The general experimental strategy is illustrated in Fig. 1. Each glycoprotein was digested with pronase to cause non-specific digestion and creating a greater array of sizes of peptides. The usage of pronase instead of more specific proteases such as trypsin and Glu-C was justified based on the frequencies and locations of basic (R,K) and acidic (D,E) residues in each protein. For example, bovine κ-casein has many acidic residues around the O-linked glycosylation sites. In contrast, bovine fetuin has only a limited number of basic and acidic residues around the O-glycosylation sites. The low frequency of cleavable residues result in overly large glycopeptides, hampering their efficient dissociation by conventional ion activation methods. In order to overcome this factor, pronase was used as an alternative to trypsin or Glu-C.

Fetuin and κ-casein were reduced by adding 200 mM dithiothreitol (DTT) in 100 mM of NH_4HCO_3. These solutions were incubated for 10 min in boiling water for denaturation. Then the proteins were alkylated with 300 mM of iodoacetamide solution for 1 h at room temperature in the dark. For pronase digestion, pronase (1 mg/mL) was added at ratio of 1:50 (pronase/protein w/w) followed by incubation at 37 °C for overnight.

3. ZIC–HILIC SPE

The resulting glycopeptides were enriched using a ZIC–HILIC SPE column prior to separation via C_{18} reversed phase nanoHPLC and then analysis by CID and UVPD in both the positive and negative modes. A self-packed zwitterionic hydrophilic interaction liquid chromatography (ZIC–HILIC) SPE column (25 mg of packing material) was used, following the manufacturer's protocol was followed. The column was washed with 2× 1 mL of water, and the washed column was primed with 1 mL of 80% acetonitrile/15% water/5% formic acid. 20 μg of dried glycopeptides were dissolved in 1 mL of 80% acetonitrile/15% water/5% formic acid, and loaded onto the primed column. The column was washed with 2× 1 mL of 80% acetonitrile/15% water/5% formic acid. Finally, the glycopeptides were eluted with 500 μL of water. The enriched glycopeptides were dried prior to LC–MS analysis.

3.1. Liquid chromatography and mass spectrometry

The glycopeptides were separated by using a Dionex Acclaim PepMap RSLC C_{18} analytical column (75 μm × 15 cm, 2 μm particle

Please cite this article in press as: B.J. Ko, J.S. Brodbelt, Comparison of glycopeptide fragmentation by collision induced dissociation and ultraviolet photodissociation, Int. J. Mass Spectrom. (2014), http://dx.doi.org/10.1016/j.ijms.2014.07.032