ACCURATE MASS ANALYSIS OF GLYCOPROTEIN ISOFORMS BY ELECTROSPRAY IONIZATION, ORTHOGONAL ACCELERATION TIME-OF-FLIGHT MASS SPECTROMETRY AND MAXIMUM ENTROPY

Iain Campuzano[1], Ashley Sage[1], Therese McKenna[1], Jim Langridge[1], Harry Brumer[2] and Kathleen Piens[2]
[1]Waters Corporation, Manchester, UK
[2]Department of Biotechnology, Royal Institute of Technology (KTH), AlbaNova University Centre, Stockholm, Sweden

OVERVIEW

We demonstrate accurate mass measurement of the isoforms of a glycoprotein (34 kDa) using an internal calibrant and Maximum Entropy (MaxEnt) deconvolution of the data obtained on an orthogonal acceleration time-of-flight (oa-TOF) instrument.

INTRODUCTION

Over the past 15 years, electrospray ionization (ESI) combined with mass spectrometry has proven to be a powerful combination for determination of the molecular weights (masses) of proteins. The differences between the observed and expected masses can provide insight into the presence of post-translational modifications or errors within publicly available protein sequences.

The mass measurement of low molecular weight molecules (<1000 Da) to better than 5 ppm accuracy is currently a routine procedure, provided an internal calibrant or lock mass is introduced with the sample. However, it is only recently that this level of accuracy was extended to the measurement of intact proteins. The mass of the human hemoglobin β-chain (15867.2 Da) was determined with a standard deviation of ±0.05 Da (±3.2 ppm) using the α-chain (15126.4 Da) for internal calibration of the mass scale.[1]

Within an electrospray spectrum of a pure protein, there are a series of multiply-charged peaks whose mass-to-charge (m/z) ratio is given by (Mr + nH)/n, where Mr is the molecular weight of the protein, H is the mass of a proton, and n is a series of integers representing the number of charges associated with each species.

In order to simplify interpretation, particularly from mixtures of proteins, various algorithms have been developed to deconvolute these multiply-charged spectra and condense the series representing each protein into a single peak on a true molecular weight scale. Of these algorithms, the one that uses a maximum entropy-based approach (MaxEnt) is the most powerful because it is automatic, it enhances the resolution, improves the signal-to-noise ratio, and can deconvolute complex mixtures of proteins.[2-3]

Xyloglucan endotransglycosylases (XET, EC 2.4.1.207) are unique enzymes that perform an endolytic cleavage of a xyloglucan chain with subsequent transfer of the newly-created chain end to the non-reducing end of a different xyloglucan molecule. This activity has been proposed to play a major role in the transient cell wall loosening required for cell wall expansion. Recent evidence suggests that XET activity may also contribute to reinforcing the connections between primary and secondary cell walls in wood-forming tissues.

PttXET16A contains a conserved N-glycosylation site situated proximal to the predicted catalytic residues in the sequence EIDFEFLGNRT (glycosylation site underlined). When expressed in P.pastoris, mass spectrometric analysis indicated that PttXET16A is heterogeneous due to the presence of variable N-glycosylation and incomplete cleavage of the α -factor secretion signal peptide.

EXPERIMENTAL

Waters LCT Premier Mass Spectrometer

Figure 1 shows a schematic diagram of the Waters LCT Premier™ oa-TOF mass spectrometer. Samples under investigation are introduced into the instrument via the ZSpray™ electrospray ion source, which operates at atmospheric pressure. Analytes of interest are ionized into the gas phase through the electrospray process and enter the instrument through the sampling cone orifice. Three RF-only ion guides provide high-efficiency transfer of the gas phase ions from the ion source to the oa-TOF pusher optics of the analyzer. Within this region, the ions are spatially and time-focused and then orthogonally pulsed into the TOF analyzer at a pusher repetition rate up to 30 kHz.

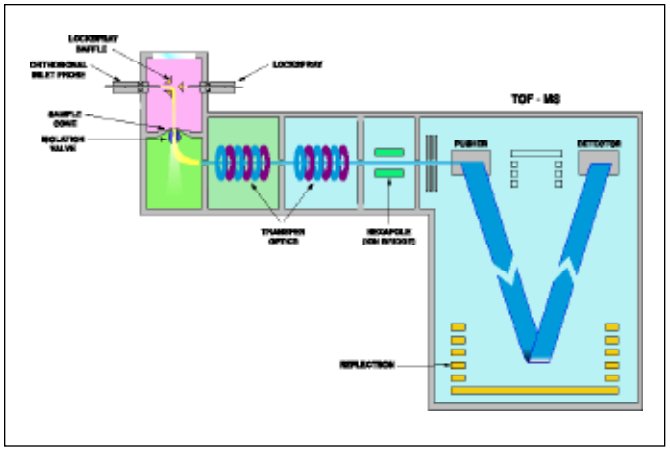

Figure 1. Instrument configuration of the LCT Premier Mass Spectrometer.

To allow time-of-flight mass spectrometry to be performed with a benchtop instrument, the instrument's design incorporates a high precision reflectron mirror at the opposing end of the field free drift region of the TOF analyzer. The flight path of the ions are reversed by the reflectron and sent back towards the detector where the ion arrivals are recorded and outputted to a computer display via a time-to-digital converter (TDC). A single pass of the TOF analyzer results in an effective flight length of 0.8 m.

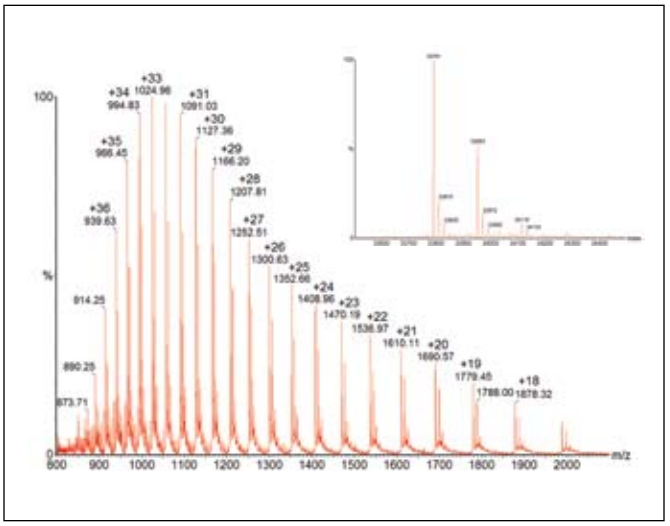

Figure 2. Background subtracted and smoothed electrospray spectrum from PttXET16A (E85A) infusion. The spectrum represents four minutes of combined continuum data. The inset shows m/z 1200 to1950 after deconvolution by MaxEnt1.

Initially a protein sample of ***Populus tremula x tremuloides*** Xyloglucan Endotransglycosylase 16A (PttXET16A E85A) was introduced into an LCT Premier, at 5 µL/min from a 250 µL gas-tight syringe, at 10 µM concentration, in an aqueous solution of 50% acetonitrile (v/v), 0.1% formic acid (v/v). Typical source operating conditions were: 3 kV capillary voltage; 100 V cone voltage; 50 V ion guide 1; 25 V aperture 1. Four minutes of continuum data were acquired over the m/z range 600 to 4000. The m/z scale was externally calibrated with a 2 µg/µL cesium iodide (50% aqueous solution of isopropanol v/v). The instrument was operated at a resolution of 5,500 FWHM. In a second series of analysis, horse heart myoglobin (Sigma M-1882, sequence mass 16951.499 Da) was introduced with the sample to act as an internal calibrant.

Mass calibration

Using standard routines provided in MassLynx™ Software, the raw combined spectral data was baseline-subtracted with a 25-order polynomial curve, such that 5% of the acquired data fell below the new baseline. The subtracted data were then smoothed (2 x 4 channels, Savitzky-Golay smooth). The mass centroids of the smoothed peaks were then determined using the top 20% of each peak and used to recalibrate the m/z scale of the spectrum.

Maximum entropy deconvolution

Three multiply-charged peaks belonging to the enzyme PttXET16A (E85A) were chosen for deconvolution from the internally recalibrated data. The mass output parameters ranged from 33,600 to 34,400 Da and resolution of 0.02 Da/channel. The Uniform Gausian peak width at half height, for the Damage Model, was set up 0.44 Da for deconvolution. Left and right intensity values were set to 40%, 40% respectively. A centroid spectrum was created to obtain accurate average protein molecular mass. Five data sets were acquired and standard deviations were calculated for mDa and ppm errors.

RESULTS AND DISCUSSION

Figure 2 shows a typical multiply charged spectrum obtained by infusing a 10 µM solution of the mutant enzyme PttXET16A (E85A). After deconvolution by MaxEnt1 (m/z 1200 to 1950), several glycoforms were revealed (Figure 2 inset). The measured mass of the most abundant species (33791.32 Da) is within experimental error of the sequence mass of the enzyme plus a glycan composition of $HexNAc_2Hex_8$ (calculated mass 33791.80 Da), i.e. is 0.48 Da (14.2 ppm) lower than calculated. Such an error is typical for externally calibrated data. Although the mass accuracy in this instance is probably adequate to assign the glycan composition, a more accurate assessment was made by repeating the analysis of the enzyme, PttXET16A (E85A), after adding myoglobin to act as an internal calibrant.

Figure 3 shows a typical multiply charged spectrum obtained by infusing a 10 µM solution of the mutant enzyme PttXET16A (E85A) in the presence of horse heart myoglobin (final concentration of 500 fmol/µL). Myoglobin can now be used as a calibrant, with which a spectrum calibration was carried out. Upon spectral recalibration, Maximum Entropy 1 deconvolution can be used to generate an accurate zero charge mass (Figure 4) for the protein under investigation.

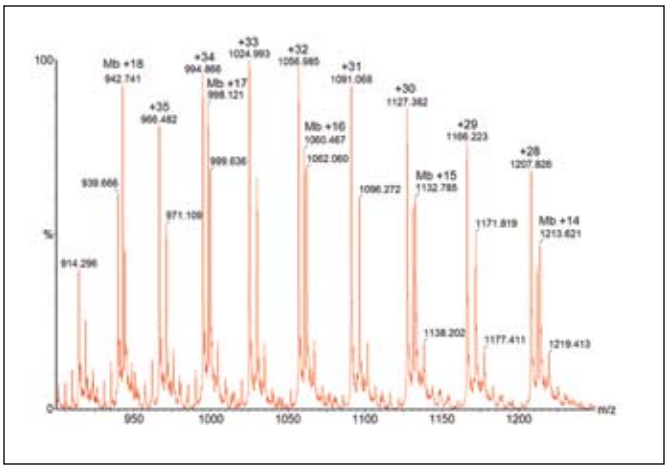

Figure 3. Background subtracted and smoothed electrospray spectrum from PttXET16A (E85A) infusion in the presence of a 500 fmol/µL (final concentration) of horse heart myoglobin. The spectrum represents four minutes of combined continuum data.

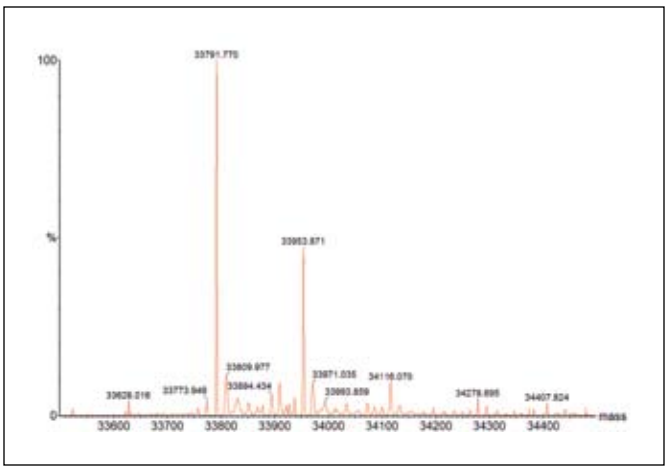

Figure 4. A typical Maximum Entropy deconvoluted spectra for PttXET16A (E85A).

Protein Analyzed	Expected Molecular Weight (Da)	Mean Measured Molecular Weight ±SD (DA)	Error mDA/pp m
PttXET16AE85A HexNAc2Hex8	33791.800	33791.792 ±0.034 ±1.0 ppm	-8.0 / -0.3
PttXET16AE85A HexNAc2Hex9	33953.943	33953.854 ±0.043 ±1.3 ppm	89.0 / -2.6
PttXET16AE85A HexNAc2Hex10	34116.085	34116.081 ±0.024 ±0.7 ppm	-4.0 / -0.1
PttXET16AE85A HexNAc2Hex11	34278.228	34278.620 ±0.137 ±3.9 ppm	392.0 / 11.4

Table 1. A comparison of the accurately-measured masses against the theoretically-calculated masses for the mutant enzyme PttXET16A (E85A). HexNAc; N-acetyl hexosamine. Hex; Mannose represent the post-translational modifications present on the enzyme.

The results summarized in Table 1 indicate that the PttXET16A (E85A) glycoforms were measured with a standard deviation of ±43 mDa (±1.3 ppm), or better on the 8, 9 and 10-mannose glycoforms. The mass standard deviation of the 11-mannose variant was found to be ±137 mDa (±3.9 ppm).

Over the four minutes of continuum data acquisition, the combined number of ion counts for the 11-mannose glycoform is considerably lower than those of the 8, 9 and 10-mannose glycoforms. This is reflected in the poorer accuracy and precision. Another possibility is that there could be low-level background interference affecting the accuracy. The fourth column within Table 1 represents the accurately-measured glycoform masses compared with the theoretically-calculated glycoform masses.

The data presented in the table above are comparable to the errors obtained during the accurate mass analysis of compounds whose molecular weight is below 1000 Da. With such high-mass accuracies, one can, therefore, determine amino acid substitutions that may only differ in 1 amu, such as leucine to asparagine. Also, the presence of post-translational modifications, such as phosphorylation, sulphation, oxidation, and disulphide bonds, can be determined.

CONCLUSIONS

Using the combination of electrospray ionization, oa-TOF, and Maximum Entropy, one can accurately mass measure intact proteins.

On three out of the four glycoforms analyzed, a standard deviation of ±43 mDa (±1.3 ppm) or better was obtained.

On three out of the four glycoforms analyzed, the errors obtained in comparison to the theoretically-calculated masses of the glycoforms were sub-3 ppm. Two out of the three glycoforms were sub-0.5 ppm.

References

1. DK Rai, WJ Griffiths, B Landin, BJ Wild, G Alvelius & BN Green, Anal. Chem. 2003, 75, 1978- 1982.

2. AG Ferridge, MJ Seddon, BN Green, SA Jarvis & J Skilling, Rapid Comms. Mass Spec., 1992, 6, 707-711.

3. Waters Application Note, number 212 (720000553EN) MaxEnt: An Essential Maximum Entropy Based Tool for Interpreting Multiply-Charged Electrospray Data.

Waters Corporation
34 Maple Street
Milford, MA 01757 U.S.A.
T: 1 508 478 2000
F: 1 508 872 1990
www.waters.com