PosterREPRINT

Waters

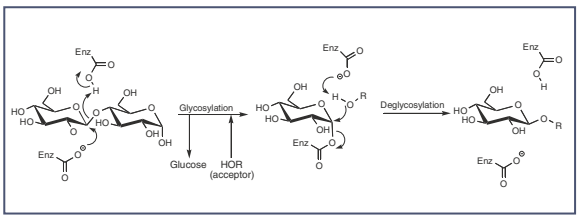
THE USE OF Q-TOF FOR THE CHARACTERISATION OF AN N-LINKED GLYCOSYLATION SITE IN NATIVE CAULIFLOWER XYLOGLUCAN ENDOTRANSGLYCOSYLASE

[1]Iain Campuzano, [1]Mark Richie, [1]James Langridge, [2]Harry Brumer, [2]Hongbin Henriksson and [2]Tuula Teeri
[1]Waters Corporation, Manchester, M22 5PP, UK [2]Dept. of Biotechnology, Royal Institute of Technology, SCFAB, S-106 91 Stockholm, Sweden.

Presented at BMSS 2001, 9th-12th September, 2001

INTRODUCTION

Xyloglucan endotransglycosylases (XETs) are a class of enzymes closely related to the well-studied glycosyl hydrolases (GH) which have been implicated in playing a role in plant cell wall expansion during growth and development. XET catalyses the cleavage and re-ligation of high molecular weight xyloglucan, which acts as a "glue" to hold cellulose microfibrils together in a composite matrix thorough hydrogen bonding interactions [1]. On the basis of sequence similarity, all known xyloglucan endotransglycosylases (XETs) have been classified as members of glycosyl hydrolase family 16 [http://afmb.cnrs-mrs.fr/~pedro/CAZY/db.html]. By analogy with other well-studied members of this family, the β1-3, β1-4 endoglucanases, utilise a two-step configuration retaining mechanism like that shown in **Scheme 1**.



Scheme 1. Two-step retaining glycosyl transfer mechanism. R = H for hydrolytic enzymes, R = sugar moiety for transglycosylating enzymes.

Early studies have indicated that protein glycosylation may be important for the catalytic function of heterologously expressed *Arabidopsis thaliana* XET [2], although this was not explored in any mechanistic detail. Interestingly, a conserved N-glycosylation site is found in Family 16 XETs proximal to the active site glutamic acid residues (**Figure 1**).

```
ArabidopsisTCH4   SAGTVTTLYLKSPGTTWDEIDFEFLGNSSGEPYTLHTNVYTQGKGDKEQQFKLWFDPTANFH 140
Kiwi              SAGTVTAFYLSSQNSEHDEIDFEFLGNRTGQPYILQTNVFTGGKGDREQRIYLWFDPTKDYH 142
Tomato            SAGVVTAFYLSSNNAEHDEIDFEFLGNRTGQPYILQTNVFTGGKGNREQRIYLWFDPTKGYH 150
Tobacco           SAGVVTAFYLSSNNAEHDEIDFEFLGNRTGQPYILQTNVFTGGKGDREQRIYLWFDPTKGYH 149
Soybean           SAGTVTAFYLSSQNAEHDEIDFEFLGNRTGQPYILQTNVFTGGKGDREQRIYLWFDPTKEYH 147
MS/MS seq.                 YLSSTNNEHDELDFEFLGDRTGQPVLLQTNVFTGGK
```

Figure 1. Sequence alignments of some GH family 16 XETs, showing the conserved catalytic residues and potential N-glycosylation site. The MS-MS sequence obtained from two tryptic peptides (m/z 1201.0 and m/z 830.5) are shown.

The ability to purify cauliflower XET from the native source as a single glycoform has promted us to re-examine the role of protein glycosylation regarding the ability of XET to favour transglycosylation over hydrolysis.
We report here the characterisation of this glycosylation by a combined LC-MS and MS/MS approach on native and enzymatically deglycosylated samples.

EXPERIMENTAL

Tryptically digested samples were introduced using the Micromass CapLC (Micromass UK Ltd) system, a low liquid chromatograph. The stream select module, attached directly to the nano Z spray source, was configured with a trapping column (0.3 x 5mm, C18) where samples were first loaded via the autosampler and desalted. The 10 port valve was switched after ca. 3 mins and a gradient was run to elute the sample from the trap to the analytical column (LC Packings, 15 x 0.075mm, C18 PepMap). A precolumn split gave a resultant flow though the column of 250 nL/min, with the pump delivering a flow of 2.5 uL/min.

Intact (undigested) proteins of interest were infused into the mass spectrometer via a borosilicate nanovial. A voltage of 850V was applied to the tip of the needle to generate electrostatic nebulisation.

MICROMASS MS TECHNOLOGIES

The proteins were dissolved into a solution of 50% acetonitrile, 0.5% formic acid and to a protein concentration of 1 pmol/uL.

MASS SPECTROMETRY

Mass spectroscopic analysis was carried out on a Q-Tof 2 (Micromass UK Ltd.) mass spectrometer fitted with a nano Z spray source. The instrument was calibrated over the mass range *m/z* 50-2500, using a sodium iodide solution. Operating the Q-Tof 2 in a wide band pass mode (quadrupole set in RF only) was carried out for intact (undigested) protein analysis. Data was acquired over the mass range *m/z* 400-2500.

Data dependent acquisition (DDA) was performed on the tryptic digest samples. An initial TOF-MS survey scan was acquired over the mass range *m/z* 350-1600, with the switching criteria for MS to MS-MS switching including ion intensity and charge state. The Q-Tof 2 was programmed to ignore singly charged ions and perform MS-MS on up to four co-eluting species, with 1-second integration times. Switch back into MS survey mode was triggered after a time period of 4 seconds.

The collision energy used to perform MS-MS was varied according to the mass and the charge state of the eluting peptide. Data for MS-MS was acquired over the mass range *m/z* 50-2000.

DATA PROCESSING

The raw combined spectral data obtained from the undigested proteins was background subtracted and subjected to Maximum Entropy 1 deconvolution, to obtain a zero charge spectrum. The raw combined spectral data obtained from the MS-MS fragmentation of a selected precursor ion was subject to Maximum Entropy 3 deconvolution. *De novo* sequencing was performed using the software program PepSeq (MassLynx, Micromass UK Ltd).

RESULTS AND DISCUSSION

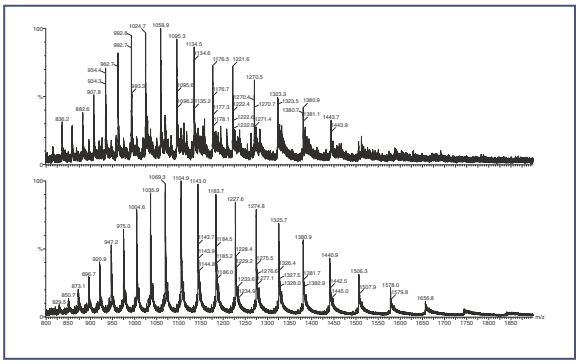


Figure 2. Nanospray TOF-MS spectra of native (lower) and N-glycosidase F-treated (upper) cauliflower xyloglucan endotransglycosylase

A background subtraction was applied to the raw data (**Figure 2**). Maximum Entropy 1 deconvolution was then carried out on the background subtracted spectra (data not shown). The deconvoluted mass for the N-glycosidase F treated XET (upper spectra) was 31736.8 Da. The deconvoluted mass for the native (lower spectra) was 33117.9 Da. The mass difference between the native XET and the N-glycosidase F-treated XET was 1381.0 Da. This mass is consistent with an N-linked $GlcNAc_2Hex_6$ oligosaccharide.

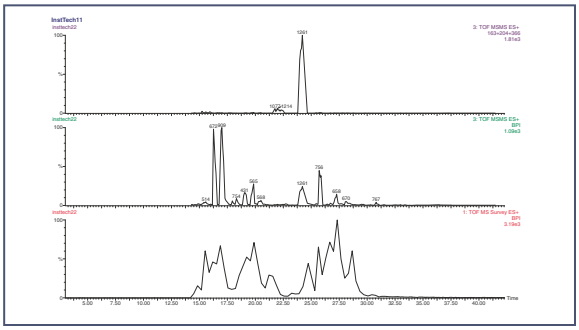


Figure 3. The LC-MS-MS chromatography of a tryptic digest of native cauliflower xyloglucan endotransglycosylase. The lower chromatogram represents the TOF-MS survey. The central chromatogram represents one of the TOF-MS-MS components. The top chromatogram is a reconstructed ion chromatogram of the ions m/z 163, 204 and 366

Through the use of marker ions such as *m/z* 163.0 (protonated hexose residue [Hex]), *m/z* 204.0 (protonated N-acetylhexosamine residue [HexNAc]) or *m/z* 366.0 (HexHexNAc), it is easily possible to identify glycopeptides. Upon MS-MS analysis of glycopeptides, the above ions are frequently observed due to the decomposition of the oligosaccharide function within the collision cell of the mass spectrometer. *Figure 3* shows the TOF-MS and a selected TOF MS-MS chromatogram from the DDA experiment carried out on the tryptic digest of native XET. A reconstructed ion chromatogram for the ions of *m/z* 163.0, 204.0 and 366.0 was performed on the selected TOF-MS-MS chromatogram. As can be observed from the upper chromatogram there is a large chromatographic peak produced corresponding to the MS-MS of the ion *m/z* 1260.5. This suggest that there is an oligosaccharide moiety attached to the ion *m/z* 1260.5.

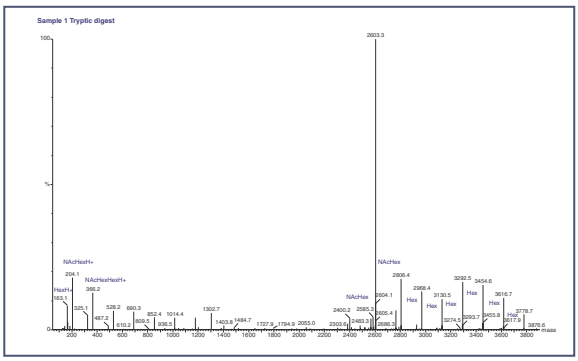


Figure 4. Maximum Entropy 3 deconvoluted spectrum obtained from MS-MS of the ion m/z 1260.5. The Progressive neutral losses of hexose and N-acetyl hexoseamine have been annotated by Hex and NAcHex respectively

The scans constituting MS-MS of the triply charged ion of *m/z* 1260.5 were combined. The combined spectra was then subject to Maximum Entropy 3 deconvolution. The deconvoluted spectrum is shown in **figure 4**. What is evident from this spectra is the multiple neutral losses of 162.0 amu and 203.0 amu at the upper end of the spectrum, which correspond to the loss of hexosamine and N-acetylhexosamine respectively from the peptide bound oligosaccharide chain. At the lower end of the spectra the ions of *m/z* 163.1, 204.1 and 366.2 which correspond to protonated hexose, protonated N-acetyhexosamine and a hexose N-acetylhexosamine respectively. From the spectra it can be clearly deduced that there are six neutral losses of 162.0 amu and two neutral losses of 203.0 amu.

This value corresponds with the postulated composition of an N-linked $GlcNAc_2Hex_6$ oligosaccharide, determined by the mass difference between the N-glycosidase F-treated and untreated cauliflower xyloglucan endotransglycosylase.

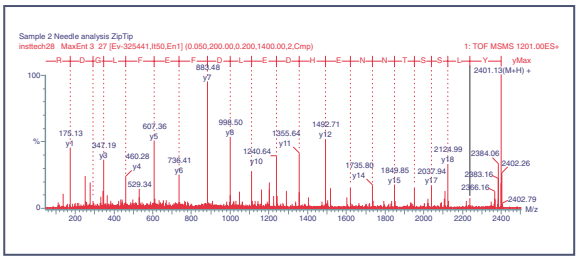


Figure 5a. Full sequence coverage obtained from nanospray MS-MS of the doubly charged precursor ion m/z 1201.0

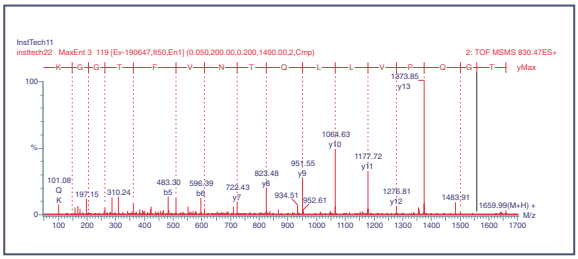
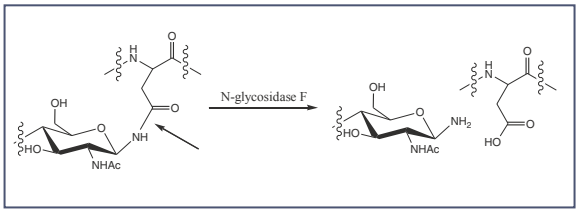


Figure 5b. Full sequence coverage obtained from nanospray MS-MS of the doubly charged precursor ion m/z 830.5

The tryptic digest of the N-glycosidase F-treated XET was desalted into 50% acetonitrile, 0.5% formic acid, using a C18 ZipTip (Millipore) and loaded into a borosilicate nanovial. This enabled an extended period of MS-MS to be performed on an ion of interest.

The ion corresponding to the peptide on which the oligosaccharide moiety had been enzymatically removed (N-glycosidase F-treated) prior to tryptic digestion was the doubly charged ion *m/z* 1201.0. Three minutes of MS-MS was performed on this doubly charged ion *m/z* 1201.0. The data was combined and subject to Maximum Entropy 3 deconvolution and full sequence coverage was obtained (**Figure 5a**). The adjacent C-terminal peptide within the sequence of XET was also chosen for MS-MS analysis. Full sequence coverage was also obtained for this peptide (**Figure 5b**, *m/z* 830.5). The two C-terminal residues of peptide *m/z* 1201.0 are aspartic acid and arginine (D & R), and the N-terminal residue of peptide *m/z* 830.5 is threonine (T). The conserved glycosylation site within the xyloglucan endotransglycosylases is asparagine-arginine-threonine (NRT, **Figure 1**).
The conserved sequence obtained by MS-MS was aspartic acid-arginine-threonine (DRT). Upon removal of the oligosaccaharide by N-glycosidase F-treatment, the residue asparagine (N) is hydrolysed to aspartate (D, see **Scheme 2**). The one mass unit shift (asparagine to aspartic acid) was identified by MS-MS and subsequent *de novo* sequencing of the fragmented peptide, thus confirming the presence and position of the conserved glycosylation site within cauliflower xyloglucan endotransglycosylase.



Scheme 2. Schematic representation of N-glycosidase F hydrolysis of an N-linked oligosaccharide. Prior to hydrolysis the oligosaccahride is attached to an asparagine residue. Upon release of the oligosaccharide the asparagine residue is hydrolysed to an aspartic acid residue

CONCLUSION

Mass spectral analysis of the undigested N-glycosidase F treated and untreated cauliflower XET produced evidence of a post translational modification in the form of glycosylation, consistent with an N-linked $GlcNAc_2Hex_6$ oligosaccharide. Following trypsinolysis of N-glycosidase F treated and untreated cauliflower XET, MS-MS was performed on the peptide on which the oligosaccharide resided. Full oligosaccharide coverage and sequence was obtained, confirming the hypothesised $GlcNAc_2Hex_6$ sequence. The conserved sequence of oligosaccharide attachment was mapped and confirmed to be NRT.

REFERENCES

1. P. Campbell and J. Braam (1999) *Trends in Plant Sci.*, 4, 361-366
2. P. Campbell and J. Braam (1998) *Plant Journal*, 15, 553-561

PosterREPRINT

Author to whom all correspondence should be addressed:
Mark McDowall
Waters Corporation
(Micromass UK Limited)
Floats Road, Wythenshawe
Manchester, M23 9LZ
Tel: + 44 (0) 161 946 2400
Fax: + 44 (0) 161 946 2480
e-mail: mark.mcdowall@micromass.co.uk

WATERS CORPORATION
34 Maple St.
Milford, MA 01757 U.S.A.
T: 508 478 2000
F: 508 872 1990
www.waters.com

Made in the United Kingdom

Waters
RIGHT ON TIME.

MICROMASS®
MS TECHNOLOGIES

Certificate No: 951387