Challenges of Determining O-Glycopeptide Heterogeneity: A Fungal Glucanase Model System

ARTICLE *in* ANALYTICAL CHEMISTRY · APRIL 2010

Impact Factor: 5.64 · DOI: 10.1021/ac901717n · Source: PubMed

5 AUTHORS, INCLUDING:

Daniel Kolarich
Max Planck Institute of Colloids and Interfa…
95 PUBLICATIONS **3,134** CITATIONS
SEE PROFILE

Helena Nevalainen
Macquarie University
150 PUBLICATIONS **4,222** CITATIONS
SEE PROFILE

Nicolle Packer
Macquarie University
182 PUBLICATIONS **5,495** CITATIONS
SEE PROFILE

Pia Hønnerup Jensen
University of Southern Denmark
16 PUBLICATIONS **611** CITATIONS
SEE PROFILE

Anal. Chem. **2010,** *82,* 3500–3509

Challenges of Determining O-Glycopeptide Heterogeneity: A Fungal Glucanase Model System

Maja N. Christiansen,[†,‡,§] Daniel Kolarich,[†] Helena Nevalainen,[†] Nicolle H. Packer,[†] and Pia Hønnerup Jensen*[,†,‡]

Biomolecular Frontiers Research Centre, Department of Chemistry and Bimolecular Sciences, Faculty of Science, Macquarie University, Sydney, NSW, 2109, Australia, Protein Research Group, Department of Biochemistry and Molecular Biology, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark, and Protein Technology Department, Novozymes A/S, Krogshøjvej 36, 2880 Bagsværd, Denmark

O-Linked glycosylation often occurs in mucin-type domains that are heavily and heterogeneously glycosylated and are challenging to analyze. The analysis of these domains is often overlooked because of these difficulties, but changes in mucinlike domain glycosylation are implicated in many diseases. Here we have explored several strategies to determine the heterogeneity of mucinlike O-glycosylated domains. Four glucanases secreted in large quantities from *Trichoderma reesei*, all containing heavily O-glycosylated mucinlike linker regions, were used as a model system. The strategies involved monosaccharide compositional analysis and identification of the released glycans by HPAEC-PAD and carbon-LC ESI-MS/MS. Glycosylated peptides were generated by different protease digestions (trypsin, papain, Asp-N, PreTAQ) and enriched by HILIC microcolumns, to determine the glycopeptide heterogeneity and glycosylation sites. The complex O-glycan heterogeneity on the intact glycoproteins and the enriched mucin-type domains was determined by MALDI-MS and ESI-MS, but the dense O-glycosylation in the mucin-type domains conferred high resistance to protease cleavage. ETD-MS/MS of the glycopeptide-enriched protease digests was unsuccessful for the de novo assignment of O-glycosylation at individual sites within the mucin-type domains but allowed several previously unknown O-linked sites outside the defined linker region to be found on two of the four glucanases. The protease digests produced many glycopeptides as determined by CID-MS/MS, but ETD fragmentation of these resulted in only a few interpretable spectra, suggesting that the use of ETD for determining the heterogeneous O-glycosylation at specific sites in regions of multiple occupancy is still in its infancy.

The glycosylation of mucins has been implicated in many different functions, and methods for their detailed characterization are essential for determining their role. Altered glycosylation has been associated with many different human diseases; for instance, the role of mucin O-glycosylation is well established in tumor formation and metastasis,[1−4] and plays an important role in altered immune response[5] as well as being essential to successful development.[6] A number of features make the analysis of mucin-type domains very challenging. First, no known O-glycosylation amino acid consensus sequence exists; second, there is no universal enzyme for the general release of O-glycans; and third, the heterogeneity of these glycoprotein domains is huge, due both to the number and diverse occupancy of glycosylation sites. Several different strategies have previously been employed to enrich for and characterize mucin type O-glycosylation (reviewed in refs 7 and 8). Lectins and antibodies are widely used to enrich for glycoproteins from a complex mixture,[9] but the specificity of lectins and antibodies does not help in the characterization of densely packed oligosaccharides. The most favored technique for analysis of these domains is to release the O-glycans from the peptide,[10] most commonly by reductive β-elimination. This gives information on the type and structure of the attached glycans but not on their site occupancy. If information on site occupancy is to be gained, one has to look at the peptide level. Materials such as graphitized carbon or hydrophilic interaction liquid chromatography (HILIC) material have been used to enrich glycopeptides from digest mixtures.[11,12] Their success at enrichment however depends on the hydrophilicity of the glycopeptides, with very hydrophilic glycopeptides containing sialic acid being best retained

* Corresponding author. E-mail: pia.honnerup.jensen@glycoproteomics.eu. Fax: +61 (0)2 9850 8313.
† Macquarie University.
‡ University of Southern Denmark.
§ Novozymes A/S.

(1) Kim, Y. J.; Varki, A. *Glycoconj. J.* **1997**, *14*, 569–76.
(2) Ono, M.; Hakomori, S. *Glycoconj. J.* **2004**, *20*, 71–78.
(3) Springer, G. F. *J. Mol. Med.* **1997**, *75*, 594–602.
(4) Rhodes, J. M.; Campbell, B. J.; Yu, L. G. *Biochem. Soc. Trans.* **2008**, *36*, 1482–86.
(5) Broide, D. H.; Miller, M.; Castaneda, D.; Nayar, J.; Cho, J. Y.; Roman, M.; Ellies, L. G.; Sriramarao, P. *Am. J. Physiol Lung Cell Mol. Physiol.* **2002**, *282*, L259–L266.
(6) Tian, E.; Ten Hagen, K. G. *Glycoconj. J.* **2009**, *26*, 325–34.
(7) Jensen, P. H.; Kolarich, D.; Packer, N. H. *FEBS J.* **2009**, *277*, 81–94.
(8) Robbe, C.; Michalski, J. C.; Capon, C. *Methods Mol. Biol.* **2006**, *347*, 109–23.
(9) Wu, Y. M.; Nowack, D.; Omenn, G.; Haab, B. *J. Proteome Res.* **2009**, *8*, 1876−1886.
(10) Lee, Y. C.; Rice, K. G. *Glycobiology a practical approach*; Fukuda, M.; Kobata, A., Eds.; Oxford University Press: Cary, NC, 1993.
(11) Hagglund, P.; Bunkenborg, J.; Elortza, F.; Jensen, O. N.; Roepstorff, P. *J. Proteome Res.* **2004**, *3*, 556–66.
(12) Larsen, M. R.; Cordwell, S. J.; Roepstorff, P. *Proteomics.* **2002**, *2*, 1277–87.

10.1021/ac901717n
Published on Web 04/13/2010

by titanium dioxide.[13] Edman sequencing has been used to detect glycosylation sites,[14] but mass spectrometry is now the preferred technique. The specific fragment ions produced by collision-induced dissociation (CID) of an oligosaccharide can verify that the peptide has attached glycans and can give information on size and monosaccharide composition.[15] Newer fragmentation techniques such as electron capture dissociation (ECD)[16] and electron transfer dissociation (ETD)[17] have the capacity to fragment the peptide while retaining the sugar modification.[18] A few studies have identified some O-glycosylation sites[15,18,19] on lightly glycosylated defined peptides. The use of nonspecific proteases, that are required to penetrate the highly glycosylated mucinlike peptides, requires de novo sequencing of the peptide for any site glycosylation information to be obtained.

The cellulolytic enzymes of the fungus *Trichoderma reesei*, the endoglucanases (EGI and EGII), and the exoglucanase cellobiohydrolases (CBHI and CBHII), all contain a heavily O-glycosylated linker region. These linker regions of EGI, EGII, CBHI, and CBHII are often called mucinlike domains, but, as in mucins, the structure and function of the glycosylation of these regions are largely unknown. The linker is positioned between the carbohydrate binding and catalytic domains and is characterized by a high prevalence of serine (Ser), threonine (Thr), and proline (Pro) amino acids, much like the highly O-glycosylated repeats known in mucins. The linker region has proven useful in the expression of heterologous recombinant proteins in *Trichoderma reesei* where the protein of interest was fused to the CBHI linker region. The linker region physically separates the endogenous catalytic region of the cellulase from the recombinant protein construct and also facilitates secretion.[20] In contrast to the mammalian O-glycosylation, that comprises the addition of *N*-acetylgalactosamine, *N*-acetylglucosamine, galactose, and sialic acid, the fungal linker regions are heavily glycosylated by short chains of mannose residues with some evidence of sulfation.[21] The fungal cellulases are secreted in gram quantities in culture,[22] are easily purified, and provide a good model system for developing methods to analyze these heterogeneously O-glycosylated domains.

We have used these four fungal glucanases as model glycoproteins to develop complementary strategies to obtain structural information on the glycosylation of these difficult to analyze glycopeptide domains.

(13) Larsen, M. R.; Jensen, S. S.; Jakobsen, L. A.; Heegaard, N. H. *Mol. Cell Proteomics* **2007**, *6*, 1778–87.
(14) Sparrow, L. G.; Gorman, J. J.; Strike, P. M.; Robinson, C. P.; McKern, N. M.; Epa, V. C.; Ward, C. W. *Proteins* **2007**, *66*, 261–65.
(15) Mirgorodskaya, E.; Roepstorff, P.; Zubarev, R. A. *Anal. Chem.* **1999**, *71*, 4431–36.
(16) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265–66.
(17) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 9528–33.
(18) Kjeldsen, F.; Haselmann, K. F.; Budnik, B. A.; Sorensen, E. S.; Zubarev, R. A. *Anal. Chem.* **2003**, *75*, 2355–61.
(19) Wu, S. L.; Huhmer, A. F.; Hao, Z.; Karger, B. L. *J. Proteome Res.* **2007**, *6*, 4230–44.
(20) Kubicek, C. P.; Panda, T.; Schreferlkunar, G.; Gruber, F.; Messner, R. *Can. J. Microbiol.* **1987**, *33*, 698–703.
(21) Harrison, M. J.; Nouwens, A. S.; Jardine, D. R.; Zachara, N. E.; Gooley, A. A.; Nevalainen, H.; Packer, N. H. *Eur. J. Biochem.* **1998**, *256*, 119–27.
(22) Salovuori, I.; Makarow, M.; Rauvala, H.; Knowles, J.; Kaariainen, L. *Nat. Biotechnol.* **1987**, *5*, 152–56.

EXPERIMENTAL PROCEDURES

Chemicals. The proteins used were cellobiohydrolase I (Swiss-Prot accession number P62694 missing $_{440}$GNRG$_{443}$) and II (P07987) and endoglucanase I (P07981) and II (P07982) from *Trichoderma reesei*. AspN, Proteinase K, α-cyano-4-hydroxycinnamic acid (HCCA), 3-nitrobenzyl alcohol (NBA), dithiothreitol (DTT), iodoacetic acid (IAA), and sodium borohydride were obtained from Sigma, and papain was sourced from Roche Applied Science. Trypsin was obtained from Promega and PreTAQ from Gibco BRL. All solvents used for LC-MS were LC-MS grade, and all other chemicals were sequence grade.

Release and Purification of O-Glycans. The O-linked glycans from the glucanases were released by reductive β-elimination. A 200 μg of sample was incubated with 50 μL of 0.5 M $NaBH_4$ and 50 mM KOH for 16 h at 50 °C, after which the reaction was stopped by adding 5 μL of acetic acid. The glycans were desalted on custom-made cation exchange columns. Methanol-washed AG50W-X8 resin (100 μL) (BioRAD) was packed on top of a 1 mL Strata-X 33u SPE cartridge (Phenomenex) prewetted with methanol. The columns were prewashed three times with 200 μL of 1 M HCl, three times with 200 μL of methanol, and three times with 200 μL of water. The β-eliminated samples were added to the columns, and the reduced O-glycans were eluted with 200 μL of water and lyophilized in a SpeedVac. The samples were redissolved in 100 μL of methanol containing 1% acetic acid, dried down again, and then washed three times with 100 μL of methanol to remove the residual borate. The glycans were taken up in water before analysis.

Oligosaccharide and Mass Spectrometric Glycan Analysis. The composition and relative ratio of the released, reduced mannose oligosaccharides were determined by high performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) analysis. The instrumentation was as follows: AS50 autosampler and thermal unit, EG50 eluent generator, ED50 detector, and GD50 gradient pump (Dionex). The β-eliminated and purified glycans were eluted isocratically with a flow rate of 0.5 mL/min from a CarboPac PA20 column (3 × 150 mm, Dionex) with 5 mM NaOH over 20 min, and the column was washed with 100 mM NaOH for 15 min and re-equilibrated for 15 min with 5 mM NaOH. A mixture of 100 pmol each of mono-, di-, tri-, and pentamannitol was used as a standard. The area under the chromatogram peaks of the sample relative to the standard was used to calculate the amount of each oligosaccharide. The eluted oligosaccharides were collected and neutralized individually into tubes containing 3 μL of glacial acetic acid, and the pH was maintained at <7 to avoid peeling of the glycans. Each fraction was desalted over custom-made cation exchange columns as described above, and the composition of each peak was verified by carbon LC-ESI MS/MS as described below.

The β-eliminated, reduced O-glycans were also analyzed directly by carbon LC-ESI MS/MS either using an Agilent LC/MSD Trap XCT Ultra or a Bruker HCT Ultra ion trap, using a homemade graphitized carbon column (Hypercarb, 5 μm material, 300 μm ID, 100 mm length), with detection in the negative ion mode. The O-linked glycans were eluted using the following gradient of acetonitrile in 10 mM NH_4HCO_3 with a flow of 5 μL/min: 0−25% in 25 min, 25−45% in 4 min, and 45−90% in 0.5

min. The column was then washed with 90% acetonitrile in 10 mM NH_4HCO_3 for 4.5 min and equilibrated in 10 mM NH_4HCO_3 for 7 min. A full scan of the mass range m/z 100−2500 was acquired with data-dependent MS/MS scans of the three most intense ions. The MS/MS fragmentation amplitude was 1.00 V. Dynamic exclusion of ions selected for MS/MS was introduced after two spectra and released after 1 min.

Intact Protein Mass Spectrometry. The intact proteins were analyses by LC-ESI MS on a Bruker micrOTOF using a BioBasic-4 column from Thermo Scientific (5 μm, 300 Å, 2.1 mm ID × 100 mm). A 50 μg amount of protein was used for each analysis. The 32 min gradient of acetonitrile in 0.1% formic acid with a flow rate of 200 μL/min was equilibration with 6% acetonitrile for 2.5 min and then 6−100% acetonitrile over 20 min. The column was washed with 100% acetonitrile/0.1% formic acid for 5 min and equilibrated in 6% acetonitrile/0.1% formic acid for 4.5 min. Data were analyzed with Data Analysis vers. 3.4 (Bruker Daltonics) and deconvoluted using MaxEnt.

Proteolytic Digestions and HILIC Enrichment. All proteolytic digestions were performed in 50 mM NH_4HCO_3 buffer. For AspN digestions, 1 μg of protease was added per 200 μg of glucanase and incubated at 37 °C for 4 h. For papain digestions, 2 μL of papain solution (20 μg) was added per 200 μg of glucanase and incubated at 37 °C O/N. The peptides containing the glucanase linker regions were HILIC enriched from the AspN and papain digests. HILIC material, PolyHYDROXYETHYL A (PolyLC INC., 12 μm, 300 Å), was packed into an emptied 1 mL strata-X cartridge (Phenomenex). The column size and volumes of buffers were adjusted according to the amount of sample to be purified; for a ∼0.5 cm column the buffer volume was 100 μL. The samples were loaded in 80% acetonitrile/5% formic acid, and the columns were washed with the same buffer before eluting the glycopeptides with 0.1−0.5% formic acid.[23] Further digestions of these HILIC-enriched linker regions were carried out. For trypsin cleavage, 1 μg of trypsin was added to the amount of sample enriched from 10 μg of original glucanase sample and incubated overnight at 37 °C. For proteinase K, 0.4 U was added and incubated at 37 °C over a time course of 5 min to 18 h. For the PreTAQ digest, 0.15 U was added and incubated at 70 °C over a time course of 1 min to 18 h. All samples were dissolved in 0.1% formic acid prior to C18 LC-ESI MS/MS analysis. In addition, an aliquot of the glycopeptide samples from the PreTAQ digest (∼100 pmol) was reduced with DTT (10 mM final concentration in sample, 30 min at 50 °C) before adding IAA to a final concentration of 50 mM and incubation in the dark for 2 h at room temperature. The reaction was quenched with DTT and directly subjected to C18 LC-ESI MS/MS.

Mass Spectrometry of Glycopeptides. MALDI MS spectra were obtained on a Bruker Microflex mass spectrometer in linear mode, using 10 μg/μL α-cyano-4-hydroxycinnamic acid (HCCA) in 70% acetonitrile/0.1% TFA as matrix solution.

C18 LC-ESI MS/MS was performed on a Bruker HTC ultra equipped with ETD fragmentation. The glycopeptides were separated over either a 45 or 135 min gradient with a C18 capillary column (ProteCol Exclusive 303, 5 μm, 300 Å, 300 μm ID × 100 mm. SGE Analytical Science, Ringwood, Australia). The 45 min gradient of acetonitrile in 0.1% formic acid with a flow rate of 5 μL/min was 0−50% acetonitrile over 18 min and 50−80% acetonitrile over 1 min. The column was then washed with 80% acetonitrile/0.1% formic acid for 10 min and equilibrated in 0.1% formic acid for 7 min. The 135 min gradient of acetonitrile in 0.1% formic acid with a flow rate of 5 μL/min was 0−50% acetonitrile over 112 min and 50−80% acetonitrile over 1 min. The column was then washed with 80% acetonitrile/0.1% formic acid for 4 min and equilibrated in 0.1% formic acid for 9 min. For experiments with 3-nitrobenzyl alcohol (NBA), 0.1% (w/v) NBA was added to both solvents.[24] 4-Sulfophenyl isothiocyanate (SPITC) derivatization of peptides was performed according to Wang et al.[25] The eluted glycopeptides were detected with the ion trap mass spectrometer in positive ion mode with acquisition of a full scan with mass range m/z 300−2000 and then data-dependent MS/MS scans of the two most intense ions. The MS/MS fragmentation amplitude was 1.00 V. Dynamic exclusion of ions selected for MS/MS was introduced after three spectra and released after 1.5 min. For ETD experiments, the ICC target for the fluoranthene was set to 600 000 ions and the reaction time to 150 ms. Alternating CID and ETD was acquired of the same parent ions, and the smart decomposition was set to auto. Several runs used manual ETD of a specific precursor mass throughout the whole chromatographic separation. The same ETD settings were applied to these experiments. All data were interpreted manually, and spectra are annotated with amino acid numbers corresponding to the expressed proteins, that is, without the signal peptide sequence.

RESULTS AND DISCUSSION

The glucanases are cellobiohydrolase and endoglucanase enzymes purified in gram quantities from the culture supernatant of *T. reesei* grown on lactose medium.[26] Figure 1 shows our strategy of analyzing the released glycans to obtain detail on glycan composition and structure and combining this knowledge to assist in the analysis of the highly heterogeneous glycopeptide linker region to obtain site occupancy data. The identification of glycopeptides was facilitated by use of an in-house-developed bioinformatic tool, "GlycoSpectrumScan" (www.glycospectrumscan.org[27]). GPMAW software[28] was used to verify the manually assigned sites of glycosylation by in silico digestion of the proposed glycopeptide. Examples from the characterization of the different glucanases are given to highlight the protein-dependent success of the various analytical tools.

O-Glycan Analysis. O-Glycans were released by reductive β-elimination from all four glucanases and were analyzed by HPAEC PAD and carbon LC-ESI MS/MS to gain information on the O-glycans present on each protein. HPAEC-PAD was used to identify and quantify the mannose (Man) oligosaccharides present on the proteins. Monosaccharides (Man1), disaccharides (Man2),

(23) Hagglund, P.; Matthiesen, R.; Elortza, F.; Hojrup, P.; Roepstorff, P.; Jensen, O. N.; Bunkenborg, J. *J. Proteome Res.* **2007**, *6*, 3021–31.

(24) Kjeldsen, F.; Giessing, A. M.; Ingrell, C. R.; Jensen, O. N. *Anal. Chem.* **2007**, *79*, 9243–52.

(25) Wang, D.; Kalb, S. R.; Cotter, R. J. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 96–102.

(26) Henrissat, B.; Driguez, H.; Viet, C.; Schulein, M. *Nat. Biotechnol.* **1985**, *3*, 722–26.

(27) Deshpande, N.; Jensen, P. H.; Packer, N. H.; Kolarich, D. *J. Proteome Res.* **2009**, *9*, 1063–75.

(28) Peri, S.; Steen, H.; Pandey, A. *Trends Biochem. Sci.* **2001**, *26*, 687–89.

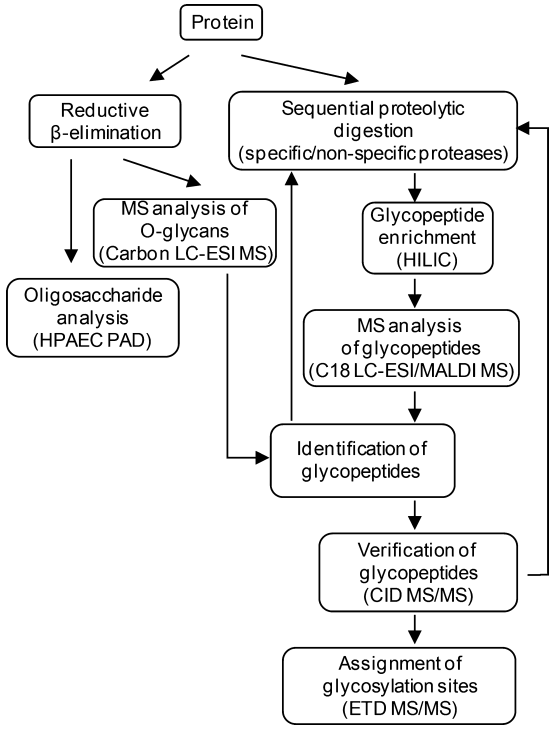

Figure 1. Overview of the experimental strategy developed for comprehensive analysis of mucinlike O-glycosylation. The flowchart illustrates the workflow and techniques employed to characterize the linker regions of the four fungal glucanases.

Table 1. Relative Ratios of Glycans Normalized to Integers, Present on the Four Fungal Glucanases Determined by HPAEC PAD Analysis[a]

composition	CBH I	CBH II	EG I	EG II
Man1	4	2	4	5
Man2	3	1	3	4
Man3	2	2	5	4
Man4	1/8	–	–	–

[a] For simplicity the Man4 for CBHI was set to the fraction 1/8, illustrating that the amount of this oligosaccharide composition is very low.

and trisaccharides (Man3) were found to be present on all four proteins. Additionally CBHI was shown to also contain a small amount of tetrasaccharide (Man4). Table 1 shows the relative ratios of these oligosaccharides, normalized to integers, on the four glucanases. This shows that there is different heterogeneity of attached O-mannose structures on the different enzymes.

Additional information on the released glycans was obtained by carbon LC-ESI MS and MS/MS analysis. The MS analysis, being more sensitive than the HPAEC-PAD monosaccharide analysis, revealed low amounts of Man4 also in both EGI and EGII (data not shown). The carbon chromatography can separate isobaric isomers of identical monosaccharide composition based on differences in linkage and branching,[29] and CID fragmentation of the isomers yields partial information about sequence and branching, so structural information on the mannose oligosaccharides was acquired by this approach. This is illustrated in Figure 2, using CBHI as an example. The summed mass spectrum

(29) Packer, N. H.; Lawson, M. A.; Jardine, D. R.; Redmond, J. W. *Glycoconj. J.* **1998**, *15*, 737–47.

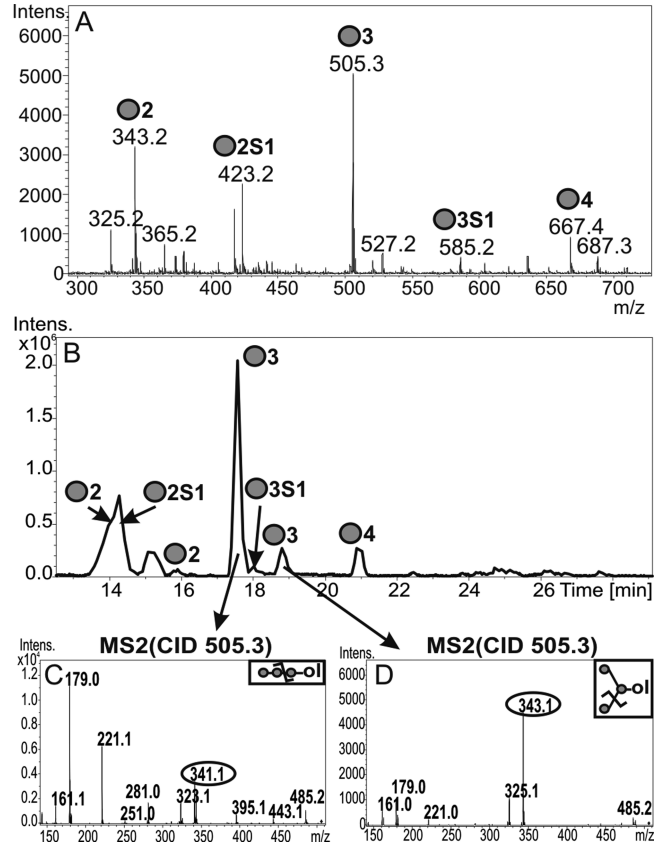

Figure 2. Carbon LC-ESI MS and MS/MS analysis of released oligosaccharides of CBHI. A: summed MS of the eluting glycans over the retention time frame 13−22 min. Masses corresponding to the compositions of Man2 (2), sulfated Man2 (2S1), Man3 (3), sulfated Man3 (3S1) and Man4 (4), and the Na^+ adducts (+22 Da) of the neutral glycans are seen. B: base peak intensity (BPI) of the eluting glycans. The peaks are labeled with the composition identified with clear separation of two isomers of Man2 and Man3. C and D: CID MS/MS of the differentially eluting isomers of Man3. The fragmentation yields different ions that lead to partial structural determination. The *m*/*z* 341.1 in C corresponds to the nonreduced dimannose fragment, which can only arise from the fragmentation of a linear trisaccharide. D: *m*/*z* 343.1 is the mass of a dimannitol fragment from the fragmentation of the reduced branched trisaccharide, as illustrated by the inset.

of all the O-glycans (Figure 2A) eluting from the column (t_R 13−22 min) shows the hexose masses corresponding to the mannose oligosaccharides (Man2, Man3, and Man4) already identified by HPAEC PAD analysis. Sodium adducts in negative ion ionization (+22 Da) of these masses are also seen. Man1 is not detected in the LC-ESI MS spectrum, as it is not retained by the graphitized carbon column[29] and emphasizes the need to use alternate analysis methods. The dihexose and trihexose are also seen to be sulfated, which was not detected in the neutral sugar HPAEC PAD analysis. Sulfation of the O-linker of CBHI has been previously reported,[21] and this analysis has identified the sulfate as located on the di- and trimannose oligosaccharides. The chromatographic separation of the oligosaccharides on the carbon column, illustrated by the base peak intensity (BPI) (Figure 2B), clearly shows the separation of two isomers of Man3 with some Man2 isomerization. The CID fragmentation of the two chromatographically separated isomers

of Man3 (Figure 2C) show different MS/MS spectra that contain diagnostic masses able to differentiate the linear trimannose (eluting at 17.5 min) from a branched trimannose (eluting at 18.5 min). The diagnostic fragment m/z 343.1 (Figure 2D) corresponds to the cleaved dimannitol terminal fragment whereas the m/z 341.1 (Figure 2C) fragment is the mass of the dimannose from the cleavage of the linear trisaccharide. The m/z 325.1 and m/z 323.1 seen are the result of loss of H_2O from these fragments. The use of software able to theoretically fragment different glycan structures to compare with the experimental MS/MS spectrum, such as GlycoWorkbench available from EurocarbDB's homepage (www.eurocarbdb.org/), facilitates this type of interpretation.

Intact Protein Mass Analysis. The four glucanases were subjected to intact protein mass analyses in order to estimate the overall extent of glycosylation. The LC-ESI MS spectra had a low signal-to-noise ratio, which is often the case when analyzing proteins of this size and heterogeneity, but it was possible to approximate the total number of attached mannoses on each glucanase to 20−31 for CBHI, 47−66 for CBHII, 29−42 for EGI, and 24−45 for EGII (Supporting Information Figure S1).

Linker Glycopeptide Heterogeneity. Working with mucinlike glycopeptide domains poses an analytical mass spectrometric challenge. The domains have a high frequency of Ser, Thr, and Pro with heterogeneous glycosylation at each Ser and Thr, which are not usually encountered by regular proteomic analysis. First, it is rarely possible to derive mucinlike glycopeptides of a measurable size using specific proteases such as trypsin, AspN, or chymotrypsin, as the specific target amino acids (K, R, D, W, Y, F) are not commonly found within these domains. The use of nonspecific proteases makes the fragmentation data more difficult to interpret. Second, the Ser and Thr residues each serve as potential O-glycosylation sites that may be modified by a number of different possible glycoforms thus increasing the complexity of the mass spectrum. Third the glycopeptide needs to be separated from the other, usually more abundant, nonglycosylated peptides in the protease digest in order to increase the mass spectral ion signal. The resultant mass spectrum of such a mucinlike domain glycopeptide resembles a "porcupine" (Figure 3A and 3B). The glycopeptide is more readily seen by MALDI MS than ESI MS because of the multiple charging of the latter ionization as evidenced by the MALDI MS spectrum (Figure 3A(i)) and a summed ESI MS spectrum (3A(ii)) of the AspN-digested, HILIC-enriched linker region from EGII. The papain-digested, HILIC-purified linker region from CBHI (Figure 3B) is even more complex due to the sulfation of the attached oligosaccharides (shown as asterisked peaks). Due to the use of proteases with known cleavage sites, both have a well-defined peptide moiety of $M = 5830.6$ Da and $M = 8180.9$ Da, respectively. This makes it possible to assign the number of mannoses (and sulfates) to each mass in the spectra. The glycopeptide data together with the intact protein mass analysis also gives an estimate of the number of oligosaccharides located outside the linker region. In the case of EGII, at least 45 mannoses are attached to the intact protein, but only 37 mannose residues can be detected on the purified linker region (Figure 3A(ii) and Supporting Information Figure S1D), which indicates that at least eight mannoses are located outside the purified linker region. Interpretation of such complex glycopeptide spectra was facilitated by using an in-house-developed tool, GlycoSpectrumScan,[27] which finds the possible glycopeptide masses in complex spectra if the compositions of the attached oligosaccharides are known (freely available under www.glycospectrumscan.org). Alternatively, data interpretation needs to be performed manually, which is very time-consuming. For example, if we take the EGII linker region peptide, there are 27 possible sites of glycosylation with 4 possible mannose compositions at each site which theoretically gives 4060 possible glycoforms of the one peptide. Experimentally, we actually see 19 linker glycopeptides of EGII (Figure 3A(i)).

The complexity and size of these glycopeptides need to be reduced to more easily allow for detailed analysis of the heterogeneous site glycosylation, so we attempted to further digest the HILIC-enriched glycopeptide linker regions of CBHII and EGII. For EGII it was possible to get a little closer to achieving this goal while still retaining a well-defined peptide. There are two arginine (R) tryptic cleavage sites in the EGII linker region (Figure 7B), so the AspN-derived linker region was further cleaved by trypsin, resulting in peptides of approximately half the linker region with different numbers of glycans attached. Three different glycopeptides were found with 27−40 hexoses attached. The use of specific proteases allows the calculation and assignment of peptides and the attached number of glycans with reasonable certainty. For example (Figure 4A), the triply charged tryptic peptide $_{58}$ATSTSSSTPPTSSGVR$_{73}$ was found as 8 glycoforms with between 9 and 16 hexoses attached. The glycopeptide nature of the peptides was verified by CID spectra obtained for each peak which clearly showed fragmentation of the glycosidic bonds. For example, the CID spectrum derived from the m/z 1210.5 peak shows triply charged glycopeptide masses corresponding to the loss of each of 13 hexoses (Figure 4B).

Additional glycosylation sites were also assigned in this way on CBHII based on a singly charged MS ladder of peptides resulting from a PreTAQ subdigest of the AspN-digested linker region. Unambiguous assignment of a Man1 on S_{106} and 2 Man1 or 1 Man2 on $_{108}$VSSL$_{111}$ was determined (Supporting Information Figure S2, summarized in Figure 7A).

This type of data defines the total number of sugars attached to a peptide; however, no, or limited, information about the number of occupied sites or the particular oligosaccharides at each site is gained.

Site-Specific Glycosylation of Mucinlike Glycopeptides. The final and most difficult piece of information to obtain is the occupation and heterogeneity of oligosaccharide structures at each Ser and Thr site. A range of strategies have previously been attempted. For instance, the saccharides can be trimmed down or removed, leaving only a mass tag by which MS/MS fragmentation can identify the glycosylation site,[30,31] and an Edman sequencing blank cycle can indicate a glycosylation site,[14] but these methods do not reveal the glycan heterogeneity at each site. CID mass spectrometry of glycopeptides leads mainly to glycosidic bond fragmentation, which can verify that the parent mass is that of a glycopeptide but yields little information on the attached peptide. A relatively new fragmentation technique, electron transfer dissociation (ETD) in ion trap ESI-MS, has shown great promise for the localization of O-phosphorylation(s) on the

(30) Wiesner, J.; Premsler, T.; Sickmann, A. *Proteomics* **2008**, *8*, 4466–83.
(31) Mikesh, L. M.; Ueberheide, B.; Chi, A.; Coon, J. J.; Syka, J. E.; Shabanowitz, J.; Hunt, D. F. *Biochim. Biophys. Acta* **2006**, *1764*, 1811–22.

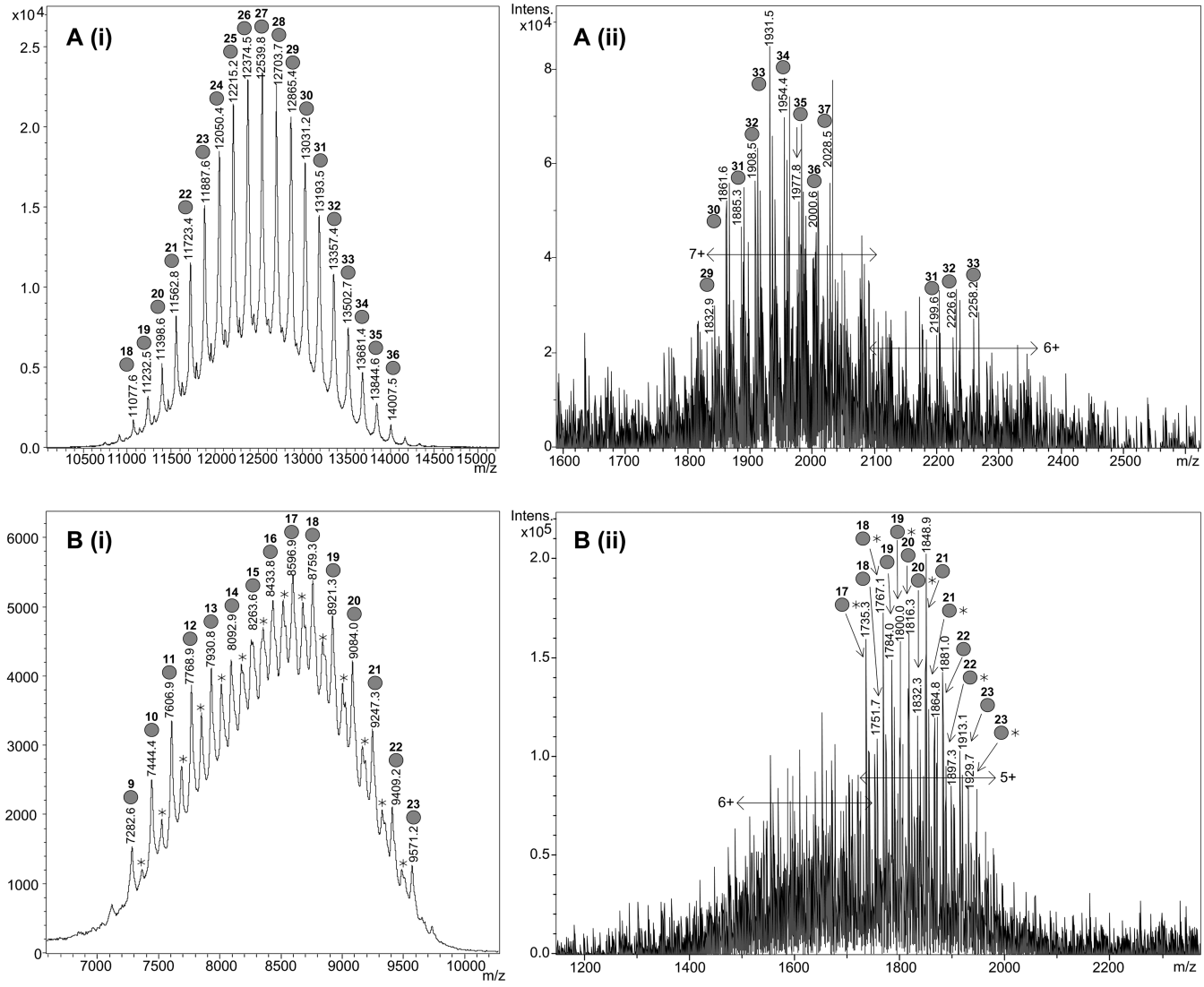

Figure 3. MALDI MS and ESI MS spectra of the linker regions of fungal glucanases. The different glycoforms and charge states (denoted as arrows on the ESI MS spectra) are shown on the figure. The asterisks denote sulfated glycopeptides. A: (i) MALDI MS and (ii) ESI MS of the AspN-digested HILIC-enriched linker region of EGII. B: (i) MALDI MS and the (ii) ESI MS of the papain-digested, HILIC-enriched linker region of CBHI.

same amino acids (Ser and Thr) in a peptide,[31,32] as mostly peptide bond fragmentation occurs, with the retention of the phosphate modification on the amino acid. Although ETD fragmentation has been shown to have some potential for identifying single glycosylation sites on a peptide,[18,19,33] there are several issues associated when attempting to obtain ETD fragmentation from samples with the density and heterogeneity of glycosylation shown on mucinlike glycopeptides (Figure 3). The huge heterogeneity lowers the signal intensity of each particular glycopeptide parent ion mass compared to unmodified peptides. Enough of one particular ion is not accumulated during the MS scan to trigger a MS/MS experiment. For example, both of the CBHII and EGII linker peptides are too big and too heterogeneous to provide useful data by any CID and/or ETD fragmentation experiments from the $[M + H5]^{5+}$, $[M + H6]^{6+}$, or $[M + H7]^{7+}$ ions obtained by ES ionization (Figure 3). The manual interpretation of an ETD spectrum triggered from the fragmentation of a $[M + H5-7]^{5-7+}$ ion from a low resolution instrument, such as the ion trap, will be very challenging, bordering on impossible. The fact that each parent ion mass could correspond to different glycoforms of the same peptide only means that the abundance of each particular glycopeptide is even lower and that each MS/MS spectrum may contain fragment ions from all these different glycoforms.

To obtain glycopeptides in the suggested optimal size and charge state for successful ETD fragmentation (m/z 500−1000 with charge states $[M + 2H]^{2+}$ or $[M + 3H]^{3+}$, personal communication, Bruker Daltonics, based on phosphopeptide fragmentation properties), different specific and nonspecific proteases were used to try to reduce the size of the linker region of all four glucanases. The HILIC-enriched linker regions generated by papain (CBHI and II) or AspN (EGI and II) were further digested with trypsin, Pronase, proteinase K, or PreTAQ, over a time course. These smaller glycopeptides

(32) Chi, A.; Huttenhower, C.; Geer, L. Y.; Coon, J. J.; Syka, J. E.; Bai, D. L.; Shabanowitz, J.; Burke, D. J.; Troyanskaya, O. G.; Hunt, D. F. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 2193–98.

(33) Perdivara, I.; Petrovich, R.; Allinquant, B.; Deterding, L. J.; Tomer, K. B.; Przybylski, M. *J. Proteome Res.* **2009**, *8*, 631–42.

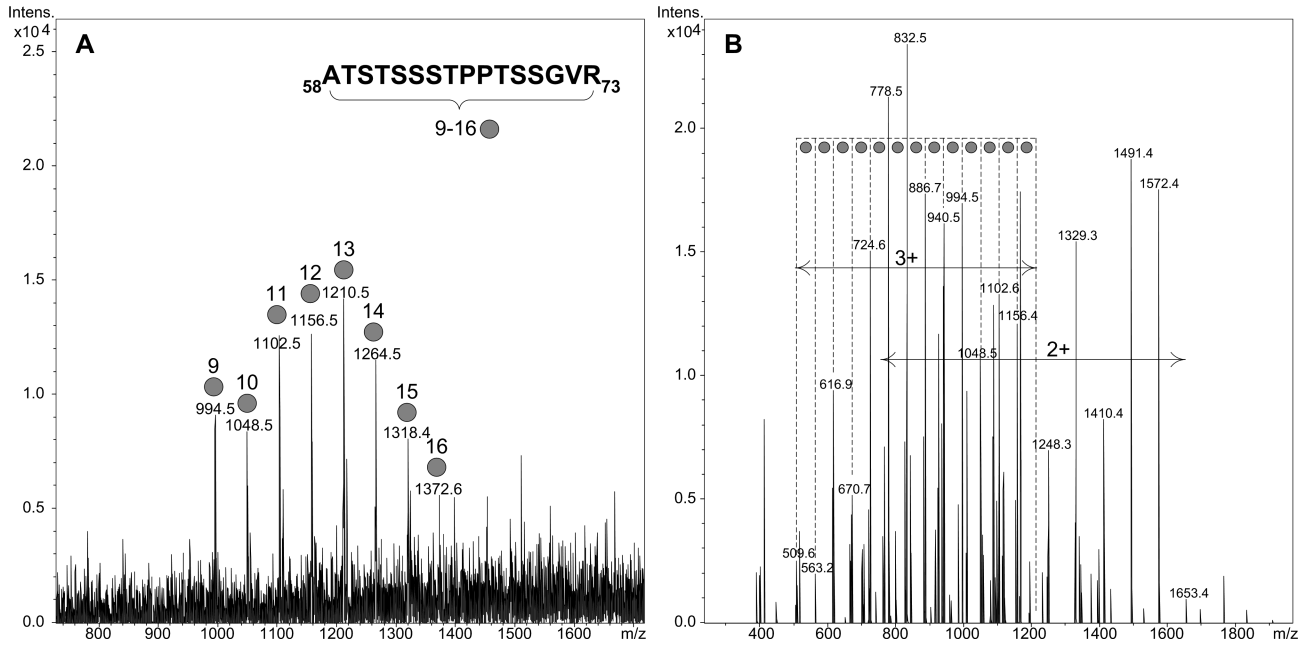

Figure 4. ESI MS and MS/MS analysis of a tryptic glycopeptide from digestion of the AspN-digested linker region of EGII. A: summed MS spectrum of $_{58}$ATSTSSSTPPTSSGVR$_{73}$ (retention time 16.4−18.2 min) with 9−16 mannoses attached. B: CID MS/MS spectrum resulting from fragmentation of the *m*/*z* 1210.5 glycopeptide peak in A. Both triply and doubly charged states are visible with the masses corresponding to loss of all 13 mannoses seen in the triply charged state, with loss of 11 seen in the doubly charged state.

were not hydrophilic enough for reproducible binding to either HILIC or graphite μ-columns for glycopeptide pre-enrichment, so shallow, reversed phase C18 LC gradients were used to achieve separation of glycopeptides from nonglycosylated peptides online to ESI-MS. This approach produced a long list of promising ETD candidates based on the size and charge criteria, and all were subjected to ETD fragmentation for attempted glycosylation site assignment (a list of the properties of all the glycopeptide candidates from CBHII and EGII can be seen in Supporting Information S2). For example, from a single chromatographic run of PreTAQ-digested EGII linker region, more than 30 potential candidate glycopeptides were found. Surprisingly, interpretable ETD fragmentation was achieved on only a few of these glycopeptides and these are detailed below.

ETD fragmentation of the tryptic glycopeptide of the AspN-derived linker region of EGII (shown in Figure 4), the triply charged tryptic peptide $_{58}$ATSTSSSTPPTSSGVR$_{73}$ (*m*/*z* range 995−1372), did not allow the assignment of the glycosylation at specific sites. This is despite the fact that at least one or two of the glycoforms were within the range and charge state suggested for successful ETD fragmentation. The resultant ETD spectrum showed, by the presence of the charge reduced species, that the reaction with fluoranthene ion had taken place (data not shown), but the peptide did not fragment to an extent that allowed any reliable assignment of sequence. Increasing the charge state of the peptide with NBA[24] did not improve the ETD fragmentation.

A successful example of the glycosylation site assignment by ETD fragmentation was from a papain-digested, HILIC-enriched glycopeptide fraction of CBHII which was further digested by trypsin. Figure 5 (A and B) shows the ETD spectra of two glycopeptides (*m*/*z* 678.3, 2+ and 667.3, 2+, respectively) only differing by the mass of 22 Da. The CID of both peptides clearly showed the loss of four hexoses, leaving a peptide mass of 707.3 and 685.3 Da, respectively (data not shown). This did not match any tryptic peptides in CBHII, so de novo sequencing was necessary. Figure 5A shows the assignment of the glycosylation on a peptide $_{83}$SGTATYS$_{89}$ with a Na^+ mass adduct (+22 Da) being a Man1 on S_{89}, and a Man3 on T_{87}. Figure 5B shows the assignment to the same peptide sequence of a Man1 on S_{89}, a Man2 on T_{87}, and the last Man1 on $_{83}$SGTA$_{86}$. The lack of ions in the lower mass region did not allow for the exact assignment of the last mannose in this glycoform as being on S_{83} or T_{85}. A theoretical fragmentation of this peptide with the last Man in either of these two possible positions, S_{83} and T_{85}, was performed using the software GPMAW (Lighthouse Data, Odense, Denmark). The resulting *z* ion series shown in the inset in Figure 5B demonstrates that the presence of the diagnostic ion of 1174.0 Da, within the mass accuracy limits of the ion trap, leads to the assignment of T_{85} as the remaining glycosylation site occupied by a single mannose, since the theoretical fragmentation by GPMAW of the alternative peptide with mannose on S_{83} (data not shown) did not result in this mass ion.

Other successful examples of ETD sequencing of glycopeptides with specific glycosylation site determination are summarized in Figure 6; the peptides are from a PreTAQ digest of the AspN-digested, HILIC-enriched peptides of EGII. Figure 6A shows the ETD fragmentation of a doubly charged *m*/*z* 919.4 ion obtained after reduction and carboxymethylation of the glycopeptide. CID data indicated the attachment of two hexose residues to this peptide (data not shown). The mass of the "deglycosylated" peptide was then submitted to FindPept Tool[34] (freely available on the SwissProt ExPASy server), and a list of potential candidate peptides in EGII was obtained. Theoretical fragments for each potential peptide were calculated using GPMAW which allowed

(34) Gattiker, A.; Bienvenut, W. V.; Bairoch, A.; Gasteiger, E. *Proteomics* **2002**, *2*, 1435–44.

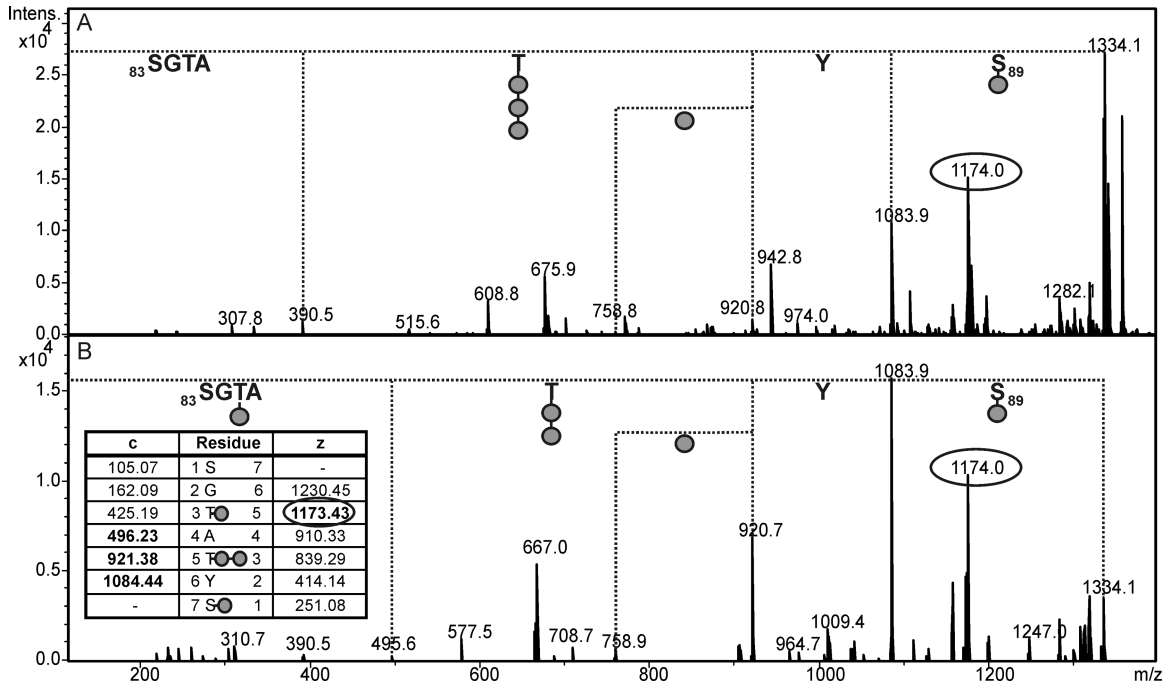

Figure 5. ETD MS/MS analysis of two glycoforms of a CBHII glycopeptide. A: ETD MS/MS of *m/z* 678.3 [2+], retention time 7.2−7.4 min, from an overnight tryptic digest, $[M + HNa]^{2+}$ ion. The ETD fragmentation led to the identification of the peptide $_{83}$SGTATYS$_{89}$ with Man3 on T_{87} and Man on S_{89}. This is illustrated by the *c* ion series shown in the figure. B: ETD MS/MS of *m/z* 667.3 [2+], retention time 7.3 min from the same overnight tryptic digest, $[M + 2H]^{2+}$ ion. The ETD fragmentation led to identification of the same peptide (of same mass minus Na^+) but with Man2 on T_{87} and Man on S_{89} and the last Man on $_{83}$SGTA$_{86}$ as shown by the *c* ion series. Theoretical fragmentation using GPMAW of the peptide with the last Man on T_{85} (inset) led to the assignment of the last Man residue to T_{85} based on the circled *m/z* 1174.0 *z5* ion.

a clear assignment of the *c*-ion series, proving the attachment of a single hexose to S_{26} and T_{27} in the peptide NCAPGSACSTLNPY outside the linker region (Figure 6A). In the same sample before reduction and carboxymethylation, a glycosylated peptide STLNPY was detected containing the same glycosylation sites but cleaved between C_{25} and S_{26} ($[M + 2H]^{2+} = 509.7$ Da, Figure 6B1 and B2). The CID data showed the loss of two hexoses from both the singly and doubly charged parent ions; however, in the ETD spectrum distinct signals corresponding to the loss of hexoses were surprisingly detected as well (Figure 6 B1 and B2, respectively). ETD fragmentation on the ion trap MS uses a "smart decomposition" function of reduced energy CID on reacted ion species. Deactivating this feature produced spectra from this glycopeptide that did not contain these CID peaks, however the intensity of the resultant ETD fragment peaks was very low and would not have allowed unambiguous identification of the sequence shown (this glycan loss as a result of using supplemental activation energy has been noted previously[33]). Thus this "smart decomposition" function increases spectral complexity but is necessary for the fragmentation of small glycopeptides. The loss of sugars from the glycopeptides by the smart decomposition function was less obvious in larger glycopeptides. As a result of this observation, we recommend that both CID and ETD spectra are used together to filter out "contaminating" peaks that result from low CID-like fragmentation in ETD spectra. The *c*-ion fragment series is observed to be more abundant than the *z*-ions in the ETD spectra, which suggests the positive charge is located mainly on the N-terminus by the use of nonspecific proteases, in contrast to the positively charged amino acid on the C-terminus obtained by trypsin digestion. Taouatas et al. reported similar spectra where the *c*-ion series dominates the ETD fragmentation after the use of LysN. LysN cleaves on the N-terminal side of a lysine residue, thus concentrating the positive charge of the Lys and the N-terminus on the same amino acid.[35]

ETD fragmentation was also helpful in mapping one disulfide bond in this glycopeptide of EGII. A doubly charged signal at *m/z* 713.3 in the spectrum of the same nonreduced PreTAQ digest of the AspN-digested, HILIC-enriched EGII linker region showed the typical loss of two hexoses in the CID spectrum (Supporting Information Figure S3). However, the overall mass could not be assigned to any glycopeptide in the EGII protein. Some prominent *y*2 and *y*3 ions indicated the peptide sequence NPY as part of this peptide. After reduction and alkylation with IAA, the mass of *m/z* 713.3 disappeared and a doubly charged peptide with *m/z* of 590.2 was detected and identified as *Cm*STLNPY, with Cys modified by IAA and S and T each having one hexose attached. (Supporting Information Figure S3, B1, and B2). The UniProt EGII entry indicates a possible disulfide bridge between C_8 and C_{25}. The *m/z* 713.3 in the spectrum of the nonreduced sample can thus be interpreted as the tripeptide $_7$QCG$_9$ (loss seen in ETD spectrum) linked by a disulfide bridge to peptide $_{25}$CSTLNPY$_{31}$ (sequence obtained from ETD spectrum) (Supporting Information Figure S3, A1, and A2). ETD has already been described to cleave disulfide bonds,[36] where the sugars remain on the respective amino acids.

The two peptides of EGII characterized in Figure 6 are located outside the defined linker region (Figure 7B), as is the glycopep-

(35) Taouatas, N.; Drugan, M. M.; Heck, A. J.; Mohammed, S. *Nat. Methods* **2008**, *5*, 405–07.
(36) Wu, S. L.; Jiang, H.; Lu, Q.; Dai, S.; Hancock, W. S.; Karger, B. L. *Anal. Chem.* **2009**, *81*, 112–22.

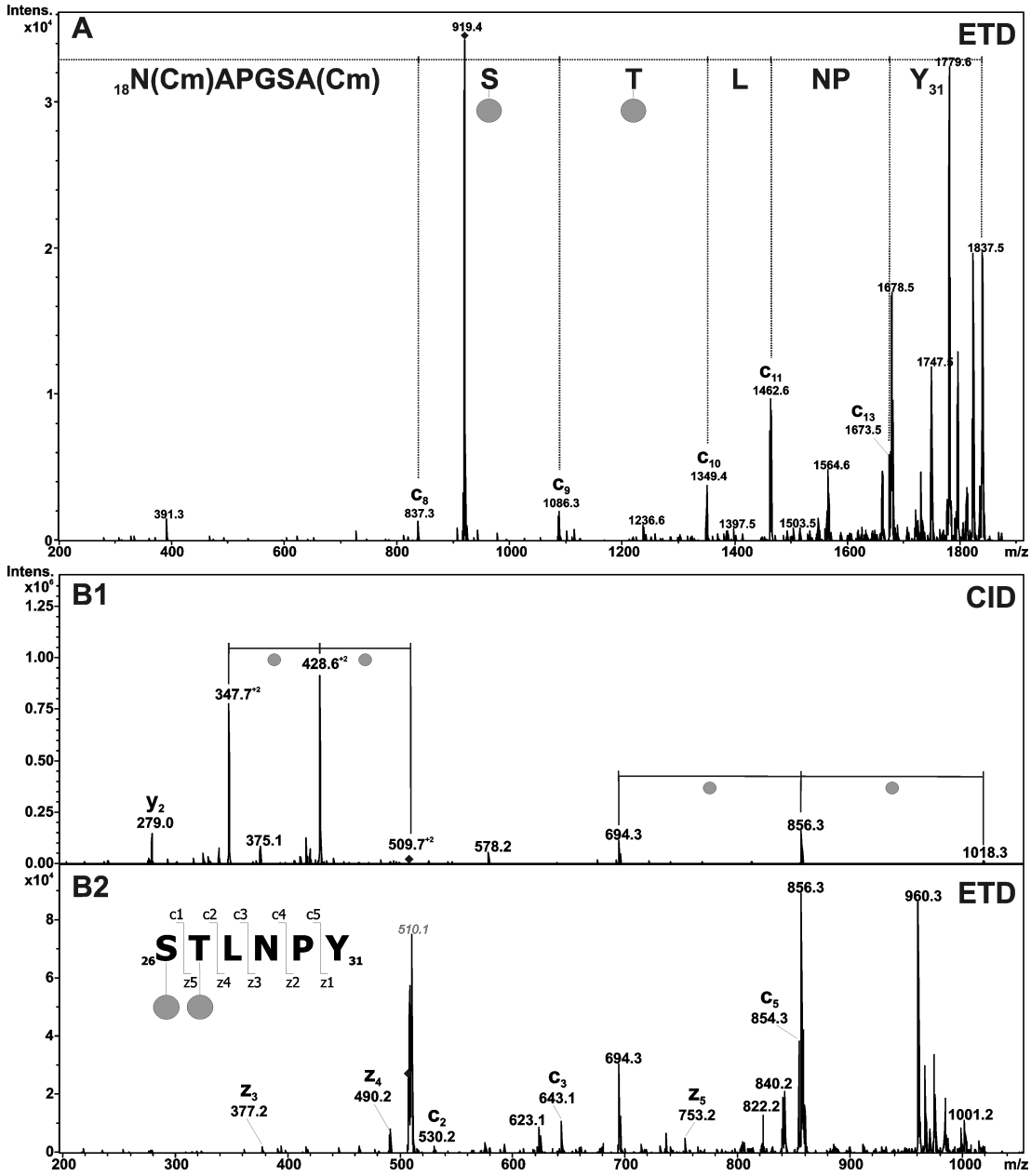

Figure 6. ETD MS/MS analysis of EGII glycopeptides. A: ETD MS/MS of *m*/*z* 919.4[2+], retention time 37.8−38.3 min, from a 2 h PreTAQ digest after reduction and alkylation with IAA. The ETD fragmentation identified the peptide as $_{18}$NCAPGSACSTLNPY$_{31}$ with Man at S_{26} and at T_{27}, illustrated by the *c*-ion series. B: ETD MS/MS of *m*/*z* 509.7[2+], retention time 27.1 min, from a nonreduced digest. B1: The CID spectrum shows the loss of two hexose residues from both the singly and doubly charged parent ion. B2: The ETD spectrum shows a more intense *c*- and part of the *z*-ion series of amino acid sequence $_{26}$STLNPY$_{31}$; however, prominent CID peaks corresponding to the loss of one and two hexoses are found by ETD fragmentation of this low molecular weight glycopeptide, a result of the use of the "smart decomposition" function. The singly charged peak of 510.1 Da is a result of an unidentified component that is partially coeluting with the doubly charged glycopeptide.

tide characterized above in CBHII (Figure 5, Figure 7A). O-Glycosylation has not previously been described on these peptides. No glycopeptide from the defined linker region was found after the many nonspecific proteases used on all the enriched linker region glycopeptides from the four glucanases.

The strategy to analyze mucinlike domains presented here and tested on the heavily O-glycosylated linker regions of four fungal glucanases resulted in a surprisingly small number of assigned glycosylation sites despite the exhaustive analysis. Figure 7 summarizes the O-glycosylation sites that were successfully assigned in the CBHII and EGII sequences, of which all were outside the linker region; no glycosylation sites could be assigned in CBHI and EGI. The sites were assigned using a range of approaches (Figure 1), including oligosaccharide analysis by HPAEC-PAD and graphitized carbon LC-ESI MS, sequential specific and nonspecific protease digestions, enrichment of the glycopeptides by HILIC, reversed phase LC-ESI MS/MS with CID, and ETD fragmentation, with and without NBA and using some specifically designed software. Despite all these techniques, which are usually successful on lightly glycosylated peptides, only six glycosylation sites on O-glycopeptides from two out of the four glucanases could be assigned with the structure of the intact

Figure 7. Sequences of fungal glucanases CBHII (A), EGII (B), CBHI (C), and EGI (D) summarizing the assigned glycosylation sites found in this study. The linker regions are marked in bold, the assigned glycopeptides are boxed, and the characterized glycosylation sites are marked with circles and italics. Circling of more than one amino acid indicates that the attachment of the glycan could not be assigned. No sites could be identified by ETD in CBHI and EGI.

attached oligosaccharide. The O-glycosylation prediction tool NetOGlyc www.cbs.dtu.dk/services/NetOGlyc/[37] predicted only two out of the six sites assigned here. The failure to detect glycopeptides from the linker region reinforces the notion[38] that a function of these heavily glycosylated regions is to render these proteins resistant to protease digestion in vivo. The data also determined that O-glycosylation is present outside of the linker region, which has not been shown before in these enzymes.

This study shows that ETD can successfully, but not routinely, assign multiple O-glycosylation sites on a glycopeptide with their heterogeneous oligosaccharides still attached. The assignment of O-linked glycosylation sites by ETD MS is clearly peptide dependent and may depend on a range of chemical and physical properties. Specific proteases for digesting these regions are not generally useful, as these domains rarely seem to contain lysine, arginine, or aspartic acid. The effect of the introduction of charge to the termini of the peptide by protease digestion[39,40] in ETD fragmentation is unknown, and nonspecific protease digestion makes data interpretation more difficult because of the undefined cleavage sites. The MS fragmentation properties that are conferred by glycosylation site occupancy appear to differ from those of phosphosite occupancy since multiple sites of phosphorylation of peptides have been shown to be very amenable to ETD sequencing.[32] Methods shown to improve the ETD sequencing of heavily O-phosphorylated peptides were attempted in this study to improve the success in assigning heavily O-glycosylated peptides. Neither (i) introduction of a charged group to the glycopeptide prior to MS/MS analysis by derivatization with 4-sulfophenyl isothiocyanate (SPITC)[41] nor (ii) increasing the charge state of singly charged peptides using NBA[24] resulted in routine interpretable ETD spectra of glycopeptides. Clearly, the use of ETD MS/MS as a means of obtaining occupancy site information for glycopeptides requires different conditions to that currently available and optimized in phosphosite determination. This will no doubt improve with time as new mass spectrometers are developed and the understanding of the mechanisms of ETD, and other molecular fragmentation, increases.

CONCLUSION

No one method is currently effective for the analysis of site heterogeneity in highly O-glycosylated mucinlike domains. However, by using different types of analytical approaches and putting the information together, we can determine many of the structural details of the glycan heterogeneity on these heavily glycosylated O-linked glycopeptides. Analysis of the released glycans yields information on composition, structure, and relative amounts of oligosaccharide on the glycopeptides. Mass spectrometry of the protease-digested and HILIC-enriched highly glycosylated domains allows the overall oligosaccharide heterogeneity to be determined. However, the O-glycan site occupancy of the mucinlike domains, with the modification intact, requires de novo peptide sequencing. ETD MS/MS has the potential to assign O-glycosylation sites but not on a routine basis and not on the protease-resistant mucinlike domains.

ACKNOWLEDGMENT

M.N.C. and P.H.J. contributed equally to this work. M.N.C. was supported by the Danish Ministry of Science, Technology and Innovation (Industrial PhD program). P.H.J. was supported by the Danish Agency for Science, Technology and Innovation (grant 272-07-0066). D.K. was supported by an Erwin Schrödinger Fellowship from the Austrian Science Fund (grant J2661) and Macquarie University. The authors thank Peter Højrup for the use of GPMAW.

SUPPORTING INFORMATION AVAILABLE

Additional information as noted in the text. This material is available free of charge via the Internet at http://pubs.acs.org.

Received for review July 31, 2009. Accepted March 13, 2010.

AC901717N

(37) Julenius, K.; Molgaard, A.; Gupta, R.; Brunak, S. *Glycobiology* **2005**, *15*, 153–64.
(38) Langsford, M. L.; Gilkes, N. R.; Singh, B.; Moser, B.; Miller, R. C., Jr.; Warren, R. A.; Kilburn, D. G. *FEBS Lett.* **1987**, *225*, 163–67.
(39) Hennrich, M. L.; Boersema, P. J.; van den Toorn, H.; Mischerikow, N.; Heck, A. J. R.; Mohammed, S. *Anal. Chem.* **2009**, *81*, 7814–22.
(40) Gauci, S.; Helbig, A. O.; Slijper, M.; Krijgsveld, J.; Heck, A. J.; Mohammed, S. *Anal. Chem.* **2009**, *81*, 4493–501.
(41) Madsen, J. A.; Brodbelt, J. S. *Anal. Chem.* **2009**, *81*, 3645–53.