

UCSF

UC San Francisco Previously Published Works

Title

N-Glycopeptide Profiling in Arabidopsis Inflorescence*

Permalink

<https://escholarship.org/uc/item/1zq968k6>

Journal

Molecular & Cellular Proteomics, 15(6)

ISSN

1535-9476

Authors

Xu, Shou-Ling

Medzihradzsky, Katalin F

Wang, Zhi-Yong

et al.

Publication Date

2016-06-01

DOI

10.1074/mcp.m115.056101

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

N-Glycopeptide Profiling in *Arabidopsis* Inflorescence*[§]

Shou-Ling Xu^{‡§}, Katalin F. Medzihradzsky[§], Zhi-Yong Wang[‡], Alma L. Burlingame[§], and Robert J. Chalkley^{§¶}

This study presents the first large-scale analysis of plant intact glycopeptides. Using wheat germ agglutinin lectin weak affinity chromatography to enrich modified peptides, followed by electron transfer dissociation (ETD)¹ fragmentation tandem mass spectrometry, glycan compositions on over 1100 glycopeptides from 270 proteins found in *Arabidopsis* inflorescence tissue were characterized. While some sites were only detected with a single glycan attached, others displayed up to 16 different glycoforms. Among the identified glycopeptides were four modified in nonconsensus glycosylation motifs. While most of the modified proteins are secreted, membrane, endoplasmic reticulum (ER), or Golgi-localized proteins, surprisingly, N-linked sugars were detected on a protein predicted to be cytosolic or nuclear. *Molecular & Cellular Proteomics* 15: 10.1074/mcp.M115.056101, 2048–2054, 2016.

Protein glycosylation is the most abundant and diverse posttranslational modification of proteins in both plants and animals. Most membrane proteins and the majority of secreted proteins are glycosylated (1).

The different types of glycosylation are grouped according to the heteroatom involved in the peptide–glycan bond. Thus, the modification of Asn side chains is termed N-glycosylation. The general features of N-linked glycosylation are the same in both plant and animal kingdoms. Asparagines within the consensus motif Asn-Xxx-Ser/Thr/Cys, where Xxx is any residue other than proline, are modified cotranslationally with the same initial structure (with a few exceptions (2)) in the ER (GlcNAc₂Man₉Glc₃) before being trimmed and built up in the Golgi. However, there are tissue- and kingdom-specific

differences in the Golgi processing (3). The modifications' functions have been best characterized in mammalian species, where in addition to being important in controlling protein folding in the ER, they are involved in cell adhesion (4), cell–cell recognition, including host–pathogen recognition in immune response (5), and regulation of cellular signaling through modification of ligand receptors (6, 7). Glycosylation affects protein solubility, stability, and activity, and as the majority of protein therapeutics are glycoproteins, there is a lot of pharmaceutical interest in characterizing glycosylation patterns on recombinant proteins (8).

Much less attention has been paid to plant glycosylation, and its functions beyond quality control in protein folding are largely unknown. Genetic data have shown mutations in genes involved in protein N-glycosylation can cause severe phenotypes. Deficiencies in the oligosaccharyltransferase complex can be either gametophytic lethal (9) or embryo lethal (10). Deficiencies in enzymes involved in the subsequent processing occurring in the Golgi apparatus display varied phenotypes. For example, in *Arabidopsis*, disruption of N-acetylglucosaminyltransferase I (GnT1) results in higher sensitivity to salt stress, but the mutant is able to develop normally under normal growth condition (11, 12) despite the fact that the mutant plant is unable to produce complex glycans. On the other hand, a rice *gnt1* mutant exhibits severe developmental abnormalities resulting in early lethality (13).

Genetically modified plants could be cultivated to produce drugs, including protein pharmaceuticals. Unfortunately, in plants, the fucose attached to the core GlcNAc of N-glycans is in an α 1–3 linkage, whereas in mammals, including humans, there is an α 1–6 linkage instead (3). Also, plants employ a unique sugar residue, xylose, which is often attached to the central mannose residue in the GlcNAc₂Man₃ core (3). Both of these plant-specific structures are immunogenic in humans (14). In fact, a lot of the recent glycosylation research in plants has been focused on how to mutate plants to produce animal-like N-glycans so that they could be used for protein therapeutics production (2).

The glycosylation of Ser, Thr, and less frequently Tyr, Hyl, and Hyp residues is classified as O-glycosylation. A specific class of these modifications is O-GlcNAcylation: Nuclear and cytoplasmic proteins can be modified with single N-acetylglucosamine residues in a dynamic fashion to regulate protein

From the [‡]Department of Plant Biology, Carnegie Institution for Science, Stanford, California 94305; [§]Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, California 94143

Received February 12, 2016, and in revised form, February 12, 2016

Published, MCP Papers in Press, April 11, 2016, DOI 10.1074/mcp.M115.056101

Author contributions: S.X., Z.W., A.L.B., and R.J.C. designed the research; S.X. performed research; S.X., K.F.M., and R.J.C. analyzed data; S.X., K.F.M., and R.J.C. wrote the paper.

¹ The abbreviations used are: ETD, electron transfer dissociation; HCD, higher energy C-trap dissociation; LWAC, lectin weak affinity chromatography.

activity (15). This modification is receiving increasing interest in mammalian systems, where it is intricately involved in etiologies such as diabetes and neurodegenerative disorders (16). Studies of this modification in plant systems are very limited; only one site of modification has been reported yet (17). However, the first global study of this modification in plants is about to be published from data acquired as part of the same experiment as described in this manuscript (manuscript in preparation).

Extracellular O-glycosylation also occurs in both plants and animals, but there is very little similarity between the types of modifications employed. In mammals, the most common O-glycans are extended structures starting with an N-acetylgalactosamine attached to serine, threonine, and tyrosine residues (18). In plants, the most common O-glycosylation is through hydroxyproline residues, although serine glycosylation is also employed, and the sugars attached are either galactose or arabinose, often in very long chains (3).

Glycoprotein analysis has traditionally followed a divergent strategy. Due to the difficulty and complexity of glycopeptide analysis, most researchers have chosen to focus on N-linked glycosylation, where enzymes (most notably PNGase F or A) can cleave off sugar structures. Researchers then either study the released glycans for sugar heterogeneity or study the deglycosylated peptides to identify former sites of glycosylation, making use of the fact that PNGase F/A convert the formerly glycosylated asparagine into an aspartate: a +1 Da mass shift that can be detected using mass spectrometry.

A couple of broad studies of deglycosylated peptides from *Arabidopsis* have been published (19, 20). However, in this work, we present the first large analysis of *Arabidopsis* intact glycopeptides. Using wheat germ agglutinin lectin weak affinity chromatography (LWAC) (21, 22) to enrich modified peptides then ETD fragmentation tandem mass spectrometry, 1152 unique N-glycopeptides were identified on 348 different sites in 270 proteins from *Arabidopsis* inflorescence tissue.

MATERIALS AND METHODS

Experimental Design—This first large-scale analysis allowing simultaneous site and glycan characterization in a plant tissue was performed from a single biological preparation and analysis of inflorescence tissue from many plants.

Sample Preparation—Proteins from three grams of *Arabidopsis Columbia* inflorescence tissues (5–6 weeks old, grown in a greenhouse) were extracted as previously described in (17) with a slight modification in extraction buffer (0.1 M Tris-HCl, pH 8.0, 2% SDS, 20 mM EGTA, 20 mM EDTA, 1.2% Triton X-100, PhosStop, Protease inhibitor, 20 μ M PUGNAc (Sigma)), followed by reduction using dithiothreitol and alkylation using iodoacetamide, tryptic digestion (Thermo, Waltham, MA) and then reverse phase desalting using a Sep-Pak C18 cartridge (Millipore, Billerica, MA).

Lectin Weak Affinity Chromatography—The wheat germ agglutinin-poros column was packed as previously described

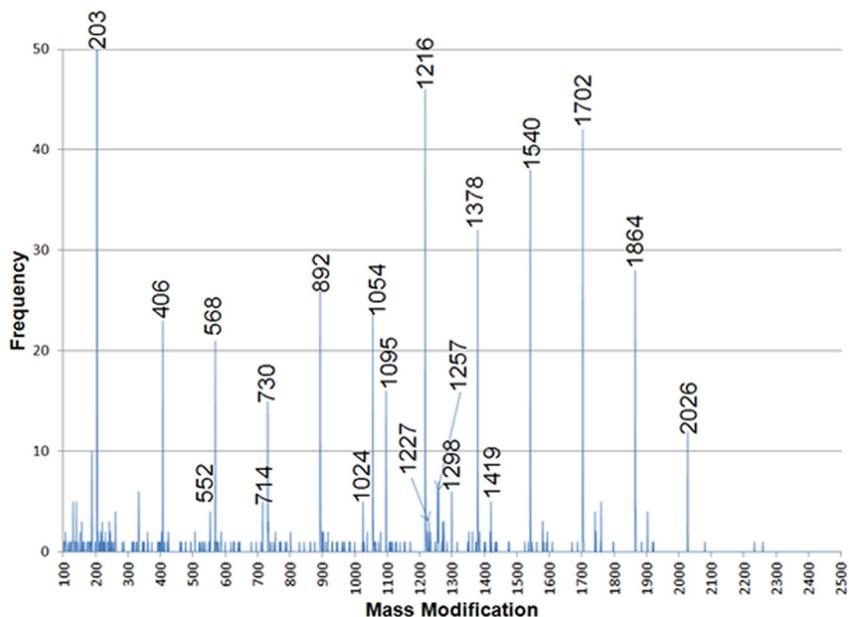
(23). The enrichment of glycosylated peptides was slightly modified from described in (23). Briefly, peptides were resuspended in 100 μ l LWAC buffer (100 mM Tris, pH 7.5, 150 mM NaCl, 2 mM MgCl₂, 2 mM CaCl₂, 5% acetonitrile). Chromatography was performed at a flow rate of 100 μ l/min. After 3.0 ml of elution, 100 μ l of 40 mM GlcNAc in LWAC buffer was injected to elute any bound glycopeptides. A single glycopeptide-enriched fraction was collected between 1.3 and 6.7 ml. To decrease the chance of overloading the column, 20 mg of starting peptides were split into 12 aliquots; each portion was run separately, and the glycopeptide-enriched fractions were then pooled. For subsequent rounds of LWAC enrichment, the pooled fractions were run as before, each time collecting the same glycopeptide-enriched tail.

High pH Reverse Phase Chromatography—High pH reverse phase chromatography was performed using an AKTA purifier (GE Healthcare, Little Chalfont, UK) equipped with a 1 \times 100 mm Gemini 3 μ C18 column (Phenomenex, Torrance, CA). The glycopeptide-enriched fraction was loaded onto the column in 240 μ l buffer A (20 mM ammonium formate, pH 10). Buffer B consisted of buffer A with 50% acetonitrile. The gradient was from 1% B to 21% B over 1.1 ml, to 62% B over 5.4 ml, then directly to 100% B. The flow rate was 80 μ l/min. Fractions from 1.4 ml to 7.3 ml were collected and dried down using a SpeedVac concentrator.

Mass Spectrometry—High pH reverse phase fractions were analyzed on an LTQ-Orbitrap Velos (Thermo Fisher, San Jose, CA) mass spectrometer equipped with a nano-Acquity UPLC (Waters, Milford, MA) system. Electron transfer dissociation (ETD), sequential HCD and ETD, or HCD triggered ETD data were acquired (see Supplement 1 for details of individual runs and acquisition parameters). Peptides were analyzed using either 1 h or 2 h reverse phase gradients based on the expected complexity of the sample according to UV absorbance from the high pH reverse phase chromatography. Raw data are available through proteomeXchange (24) (accession PXD003008) by submission to MassIVE (MassIVE Accession MSV000079345).

Database Searching—Tandem mass spectrometry data were converted into mgf peak lists using Proteome Discoverer. Database searching was performed using Protein Prospector version 5.14.20. Searches were performed against the The Arabidopsis Information Resource (TAIR) database (<https://www.arabidopsis.org/>) from December 2010, concatenated with sequence randomized versions of each protein (a total of 35,386 entries). All peptides were assumed to be fully tryptic. Precursor and fragment mass tolerances of 10 ppm and 0.6 Da were considered for ETD data. ETD data were initially searched using a mass modification strategy similar to that previously published (22, 25), where any mass modification between 100 and 2500 Da on serine, threonines, or asparagine residues was considered. For subsequent searches, modifications were only considered on asparagine residues located within the glycosylation motif N-IP-S/T, a

FIG. 1. Histogram of N-linked mass modifications observed in LWAC-enriched *Arabidopsis* flower proteins. Repeat identifications of the same mass modification on the same peptide are collapsed to a single count. The y axis range has been curtailed at 50 to facilitate viewing of low count peaks; the 203 mass modification peak should extend to 172 counts. Masses labeled correspond to glycan compositions and these are summarized in Table I.



capability newly introduced in Protein Prospector version 5.14.20. In all searches, additional modifications considered were methionine or tryptophan oxidation, pyroglutamate formation from peptide N-terminal glutamines, pyrocarbamidomethyl cysteine formation from peptide N-terminal cysteines, and protein N-terminal methionine removal, acetylation, or combinations thereof. A total of two variable modifications and one mass modification were allowed per peptide. The second search considered all proteins in the TAIR database; a third search used the same parameters except only considering proteins identified in the second search. From these results a histogram of mass modifications was produced (Fig. 1), and masses were translated into sugar compositions (Table I) that were subsequently specified as defined variable rare modifications (only one glycosylation allowed per peptide) in a final search. A single GlcNAc was an exception to this rule and was considered up to twice per peptide. Results were sorted by expectation value and thresholded to an estimated 0.2% false discovery rate at the unique glycopeptide level according to target: decoy database searching. The glycopeptide ETD results from this search are presented in Supplemental Table 2 and are also uploaded to MS-Viewer (prospector2.ucsf.edu, 26)) with the Searchkey 6mvcoan57m, which allows viewing annotated spectra of all results.

HCD data, when available, were used for manual verification of some assignments based on the presence of diagnostic low mass oxonium ions, Y1 ions, and peptide fragments (27).

RESULTS

High pH reverse phase fractions after LWAC glycopeptide enrichment were analyzed by either HCD and ETD on every precursor, HCD-triggered ETD based on the observation of

the HexNAc oxonium ion at m/z 204.087 or ETD only. To identify complex glycan-modified peptides, data were searched using Protein Prospector, allowing for unspecified mass modifications within the mass range of 100 to 2500 on serine, threonine, or asparagine residues. Results from this initial search showed no evidence of mass modifications on serines or threonines other than a single HexNAc, so future searches only allowed mass modifications on asparagine residues within the Asn-Xxx-Ser/Thr motif. The list of accession numbers of proteins identified in this second search was then used for a further discovery search to identify as many sugar mass modifications as possible.

Figure 1 shows the histogram of unique peptide + mass modification combinations observed, and Table I provides a translation of these masses into modification compositions. Many of the unlabeled peaks in Fig. 1 correspond to combinations of sugar and other modifications. For example, potassium and iron adducts of glycans are the explanations for several unlabeled peaks. These results show that high mannose structures dominated the glycans detected, but both hybrid and complex glycans were also observed.

Finally, the list of identified modifications discovered and summarized in Table I were specified as defined variable modifications for a search of the whole *Arabidopsis* proteome to maximize glycopeptide discovery. This resulted in the identification of 1148 unique N-linked glycopeptides (peptide + modification state combinations).

In addition, the initial mass modification search where the consensus glycosylation motif was not enforced allowed the discovery of a further four glycopeptides that are modified in a nonconsensus motif. An example of one of these is shown in Fig. 2, which identifies a peptide from O-glycosyl hydrolases family 17 protein, where a HexNAc (GlcNAc) is at-

tached to an asparagine in a reversed consensus motif (Thr-Xxx-Asn). So, in total, 1152 unique N-linked glycopeptides, representing 348 sites in 270 different proteins were identified (Supplemental Table 2, annotated spectra of each can be viewed using MS-Viewer at prospector2.ucsf.edu with the Searchkey 6mvc0an57m).

861 unique glycoforms are in this list (Supplemental Table 2). Almost 60% of the glycosylation sites (198 sites) were identified from a single glycoform, and more than half of these glycopeptides (110 unique sites) featured the core GlcNAc only. Interestingly, one of these glycopeptides from putative fascilin-like arabinogalactan protein 20 contained two modified consensus glycosylation motifs. At the other end of the spectrum, 15 glycosylation sites displayed at least 10 different glycans (16 different glycoforms were the most detected on a single site), and these glycosylation sites are all found in different proteins. There seems to be a significant variation in the site-specific heterogeneity depending on the individual proteins. For example, putative cysteine-rich repeat secretory protein 13 (AT3G29040.1 = Q9LJW2 in the SwissProt database) contains four N-glycosylation motifs, Asn-34, 56, 147, and 153; the latter three of which were detected bearing truncated structures: GlcNAc₂Man; GlcNAc₂Man₃ and GlcNAc₂Fuc, respectively. Five glycosylation sites were observed modified from peroxidase 34 (AT3G49120.1). Asn-43, 184, and 244 were each identified with a single glycoform; while Asn-43 and 244 featured only the core GlcNAc, Asn-184

was modified with a Man₅ structure. Asn-285 displayed two glycoforms, bearing either a single GlcNAc or Man₅. Asn-316 was represented with four glycoforms, bearing GlcNAc₁₋₂, Man₅, and a complex glycan, GlcNAc₃Man₃Hex. Asn-59 and 154 from calreticulin 1 (AT1G56340.1) were detected modified with 12 and 8 different glycoforms, respectively. In addition to a series of oligomannose structures, each was also observed bearing only the core GlcNAc.

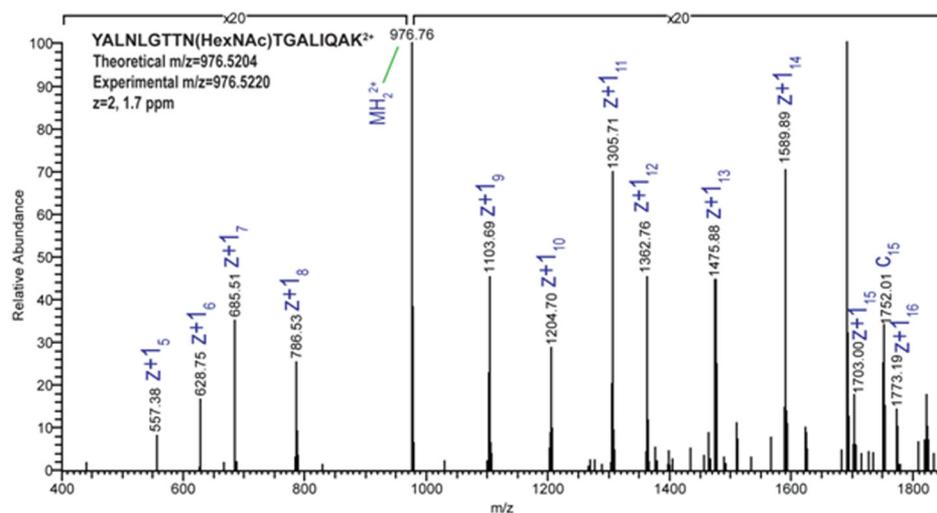
The glycan distribution among the unique glycoforms is presented in Fig. 3. Although the high mannose structures were the dominant glycans, surprisingly, the truncated structures, starting from core GlcNAc and paucimannose glycans added, were present at approximately the same level. We also detected some structures where the processing had not been completed, *i.e.* glycopeptides featuring GlcNAc₂Man₉Glc (23 instances) and a limited number of complex or more likely hybrid glycans. Sixty-one sites were detected with sugar structures containing three GlcNAcs; their hexose (most likely mannose) content varied from three to seven. 75% of these glycans might be considered paucimannose except for the presence of the extra GlcNAc. The plant-specific, xylose-containing structures were detected in very low numbers and always on truncated structures. The presence of more fucosyl structures (30) are indicated. However, because of limited ETD fragmentation, in six cases, it could not be determined whether the glycan contained a fucose, and a Met or Trp residue nearby in the sequence was oxidized or the unoxidized peptide featured a glycan with a hexose.

The glycoproteins identified are predicted to be secreted or transmembrane proteins, quite a few of them residing in the ER or the Golgi, as expected. However, we also identified a protein N-glycosylated that is predicted to be located either in the cytosol or the nucleus. This protein, ubiquitin E2 variant 1D-4 (AT3G52560.2), which contains the UBCC (ubiquitin-conjugating enzyme E2), was found modified by high mannose glycans. Two ETD spectra identified a peptide from this E2

TABLE I
Sugar compositions of N-linked mass modifications identified in Fig. 1

Mass	Composition	Mass	Composition	Mass	Composition
203	HexNAc	1024	HexNAc2Hex3Xyl	1378	HexNAc2Hex6
406	HexNAc2	1054	HexNAc2Hex4	1419	HexNAc3Hex5
552	HexNAc2Fuc	1095	HexNAc3Hex3	1540	HexNAc2Hex7
568	HexNAc2Hex	1216	HexNAc2Hex5	1702	HexNAc2Hex8
714	HexNAc2HexFuc	1227	HexNAc3Hex3Xyl	1864	HexNAc2Hex9
730	HexNAc2Hex2	1257	HexNAc3Hex4	2026	HexNAc2Hex10
892	HexNAc2Hex3	1298	HexNAc4Hex3		

FIG. 2. **Identification of a nonconsensus N-glycosylation.** The ETD spectrum of m/z 976.5220 (2+) identifies a peptide from O-glycosyl hydrolases family 17 protein (At5g58090) spanning residues 337–353 with a single GlcNAc residue attached to Asn-345. The mass difference between $z+1_8$ and $z+1_9$ unambiguously identifies the site of modification.



enzyme with HexNAc₂Hex₈ and HexNAc₂Hex₉ attached (Fig. 4, Supplemental Table 2).

DISCUSSION

This first attempt at a large-scale study of plant glycopeptides suggests that high-mannose glycan structures dominate in *Arabidopsis* flower tissue. Many hybrid glycopeptides were identified, but only a handful of complex glycans were observed. Interestingly, relatively few fucosylated (29/1152) and even fewer xylose-containing glycopeptides (14/1152) were observed, and no occurrence of both fucose and xylose on the same glycan was detected. A previous N-glycan analysis reported that ~30% of *Arabidopsis* glycans contained a xylose (28). However, the reference study was performed on whole plant tissue, so one explanation for the difference could be a different distribution in inflorescence tissue. Also, even though a relatively limited number of proteins and sites in the present study contained xylose, these included the most heavily glycosylated protein observed (TGG1), so as a percentage of the total released glycan pool, the value may be

higher. It is also possible that xylose interferes with the ability of wheat germ agglutinin to enrich this population of glycopeptides. This highlights a difference between glycopeptide and released glycan analysis and how they could provide different information.

The proposed glycan processing pathway predicts xylose is added first to complex glycans, and then α 1-3 fucose is subsequently added (2, 20). We detected xylose on GlcNAc₂Man₃GlcNAc and paucimannose glycans, whereas fucose was only observed on truncated structures. The expected mature complex glycan version in the plant vacuole is a xylose and fucose modified paucimannose glycan, so it is unclear why fucose was missing in our observed structures.

This is the first global intact glycopeptide study performed in any plant species. However, previous researchers have analyzed enzymatically deglycosylated peptides, so it is possible to investigate the overlap in glycosylation site identifications. By far, the largest of the previous studies reported the identification of 2188 formerly glycosylated sites (20). Comparing to these results, 84 of our 346 glycosylation sites are novel.

In this study, we identified three nonconsensus glycosylation sites. While we, and others, have observed nonconsensus glycosylation in other species (20, 25, 29, 30), we believe this is the first confirmation that it also occurs naturally in plant proteins.

This study identified a large number of glycopeptides bearing heavily truncated structures, and the core GlcNAc alone seems to be the dominant individual modification. This large number of truncated structures is against conventional wisdom but is consistent with glycopeptide studies carried out by us and others (22, 25, 31). We believe that released glycan analysis misses these structures for a few reasons: 1) They mostly analyze glycans released from the plasma membrane, whereas our results are from the whole cell including compartments such as the ER and Golgi; 2) PNGase F is probably less efficient at cleaving truncated structures (it is known to

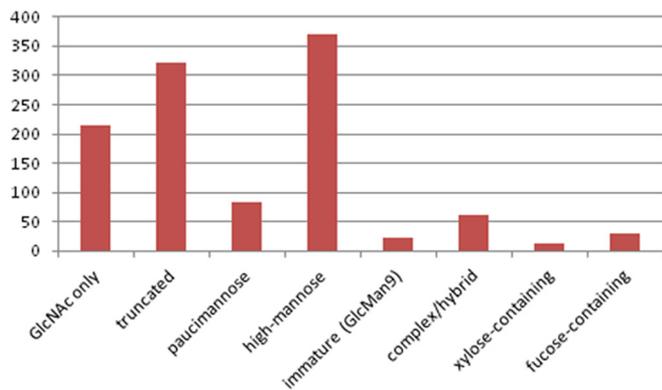
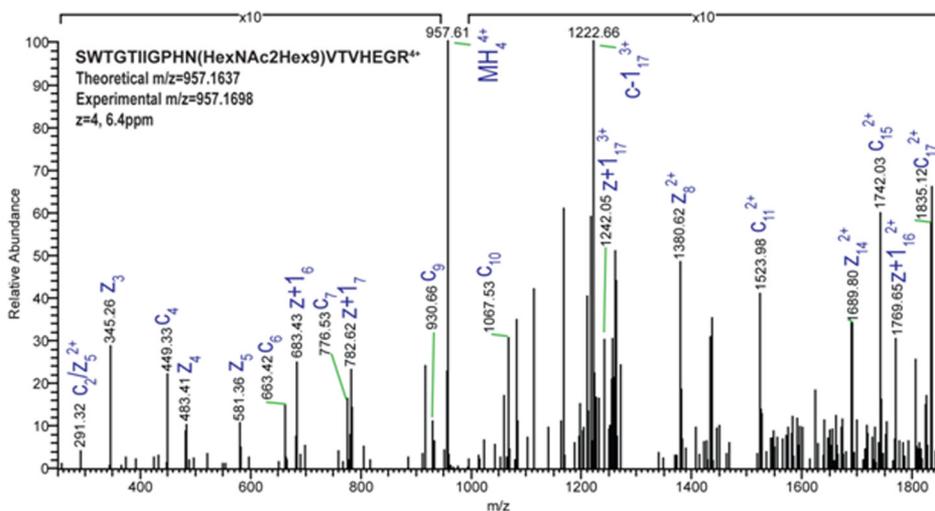


FIG. 3. Glycan distribution among the unique glycoforms identified (Supplemental Table 2). The “GlcNAc only” category is a subset of the truncated structures, and the xylose- and fucose-containing are also subsets of other categories.

FIG. 4. Identification of N-glycosylation on a putative cytosolic or nuclear protein. ETD spectrum of m/z 957.1698 (4+). From these data, a tryptic peptide (47–64) from ubiquitin E2 variant 1D-4 was identified with GlcNAc₂Man₉ attached to Asn-58.



not cleave off a single GlcNAc); and 3) it is difficult to retain short glycan structures in released glycan analysis. It also has been observed that recombinant proteins expressed in plants can have glycans trimmed to a single sugar residue (32, 33). The sample preparation for this study did not make a specific attempt to solubilize transmembrane proteins, so the results here may be biased toward more soluble proteins, with many Golgi-resident proteins identified.

It is interesting that such a high number of enzymes involved in carbohydrate metabolism were identified as glycoproteins themselves. The protein for which the most glycopeptides were identified, beta glucosidase 38 (TGG1), converts glucosinates into reactive thiocyanates that act as part of the plant defense against pests and disease. This protein has previously been reported to only bear high mannose glycan structures (34). However, in our results, while residues 108, 236, and 493 were only detected with high mannose modifications, Asn-379 was also detected with hybrid glycan structures and complex structures with xylose attached.

Unexpectedly, we identified ubiquitin conjugating E2 enzyme modified with high-mannose N-glycans. Ubiquitin E2 enzymes are usually located either in cytosol or nucleus. The N-glycosylation of this enzyme suggests that perhaps this protein is also targeted to the ER or the Golgi.

The lack of observation of extended O-linked structures could be due to several factors. Our lectin enrichment strategy using wheat germ agglutinin is supposed to have affinity for terminal GlcNAc residues, but this and previous studies show it has much wider glycan enrichment specificity. Nevertheless, it may have no affinity for galactose or arabinose, employed by plant O-glycans. Secondly, most plant O-glycosylation features extended sugar chains and occurs in protein domains rich in hydroxyprolines and with very few tryptic cleavage sites, so these glycopeptides may not be accessible to our analysis strategy.

CONCLUSIONS

This first *Arabidopsis* plant intact glycopeptide study found that xylose and fucose modification of plant proteins in inflorescence tissue is much lower than reported in previous global glycan studies. It also highlights that unique information can be achieved from glycopeptide analysis compared with deglycosylated peptide and glycan analysis. For example, the single GlcNAc-modified sites will be most likely overlooked since PNGase F does not remove these sugars and the hydrolysis speed of PNGase A for such structures is very low. The LWAC strategy employed here seems to have broad specificity for N-linked glycosylation analysis and should be a powerful approach to help unravel the roles of this modification in planta. However, different strategies are going to be required for studying plant O-glycosylation.

* This work is supported by the Chemical Sciences, Geosciences and Biosciences Division, Office of Basic Energy Science, Office of Science, U.S., Department of Energy DOE (DEFG02-08ER15973),

NIH Grant NIGMS 8P41GM103481, and the Howard Hughes Medical Institute. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

§ This article contains supplemental material.

¶ To whom correspondence should be addressed: Department of Pharmaceutical Chemistry, University of California San Francisco, 600 16th Street, Genentech Hall Room N474A, San Francisco, CA 94143-2240. Tel.: 415-476-5189, Fax: 415 502 1655; E-mail: chalkley@cgl.ucsf.edu.

REFERENCES

1. Apweiler, R., Hermjakob, H., and Sharon, N. (1999) On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim. Biophys. Acta* **1473**, 4–8
2. Gomord, V., Fitchette, A. C., Menu-Bouaouiche, L., Saint-Jore-Dupas, C., Plasson, C., Michaud, D., and Faye, L. (2010) Plant-specific glycosylation patterns in the context of therapeutic protein production. *Plant Biotechnol. J.* **8**, 564–587
3. Etzler, M. E., and Mohnen, D. (2009) Viridiplantae. In *Essentials of Glycobiology*, 2nd ed. Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W., and Etzler, M. E., (eds.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. 653 p.
4. Jacob, G. S., Welby, J. K., Scudder, P. R., Kirmaier, C., Abbas, S. Z., Howard, S. C., Keene, J. L., Schmuke, J. J., Broschat, K., and Steining, C. (1995) Studies on selectin-carbohydrate interactions. *Adv. Exp. Med. Biol.* **376**, 283–290
5. Baum, L. G., Garner, O. B., Schaefer, K., and Lee, B. (2014) Microbe-host interactions are positively and negatively regulated by galectin–glycan interactions. *Front. Immunol.* **5**, 284
6. Woronowicz, A., Amith, S. R., De Vusser, K., Laroy, W., Contreras, R., Basta, S., and Szewczuk, M. R. (2007) Dependence of neurotrophic factor activation of Trk tyrosine kinase receptors on cellular sialidase. *Glycobiology* **17**, 10–24
7. Amith, S. R., Jayanth, P., Franchuk, S., Siddiqui, S., Seyrantepe, V., Gee, K., Basta, S., Beyaert, R., Pshezhetsky, A. V., and Szewczuk, M. R. (2009) Dependence of pathogen molecule-induced toll-like receptor activation and cell function on Neu1 sialidase. *Glycoconj. J.* **26**, 1197–1212
8. Li, H., and d'Anjou, M. (2009) Pharmacological significance of glycosylation in therapeutic proteins. *Curr. Opin. Biotechnol.* **20**, 678–684
9. Koiwa, H., Li, F., McCully, M. G., Mendoza, I., Koizumi, N., Manabe, Y., Nakagawa, Y., Zhu, J., Rus, A., Pardo, J. M., Bressan, R. A., and Hasegawa, P. M. (2003) The STT3a subunit isoform of the *Arabidopsis* oligosaccharyltransferase controls adaptive responses to salt/osmotic stress. *Plant Cell* **15**, 2273–2284
10. Lerouxel, O., Mouille, G., Andeme-Onzighi, C., Bruyant, M. P., Seveno, M., Loutelier-Bourhis, C., Driouich, A., Hofte, H., and Lerouge, P. (2005) Mutants in defective glycosylation, an *Arabidopsis* homolog of an oligosaccharyltransferase complex subunit, show protein underglycosylation and defects in cell differentiation and growth. *Plant J.* **42**, 455–468
11. Kang, J. S., Frank, J., Kang, C. H., Kajiura, H., Vikram, M., Ueda, A., Kim, S., Bahk, J. D., Triplett, B., Fujiyama, K., Lee, S. Y., von Schaewen, A., and Koiwa, H. (2008) Salt tolerance of *Arabidopsis thaliana* requires maturation of N-glycosylated proteins in the Golgi apparatus. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 5933–5938
12. von Schaewen, A., Sturm, A., O'Neill, J., and Chrispeels, M. J. (1993) Isolation of a mutant *Arabidopsis* plant that lacks N-acetyl glucosaminyl transferase I and is unable to synthesize Golgi-modified complex N-linked glycans. *Plant Physiol.* **102**, 1109–1118
13. Fanata, W. I., Lee, K. H., Son, B. H., Yoo, J. Y., Harmoko, R., Ko, K. S., Ramasamy, N. K., Kim, K. H., Oh, D. B., Jung, H. S., Kim, J. Y., Lee, S. Y., and Lee, K. O. (2013) N-glycan maturation is crucial for cytokinin-mediated development and cellulose synthesis in *Oryza sativa*. *Plant J.* **73**, 966–979
14. van Ree, R., Cabanes-Macheteau, M., Akkerdaas, J., Milazzo, J. P., Loutelier-Bourhis, C., Rayon, C., Villalba, M., Koppelman, S., Aalberse, R., Rodriguez, R., Faye, L., and Lerouge, P. (2000) Beta(1,2)-xylose and alpha(1,3)-fucose residues have a strong contribution in IgE binding to plant glycoallergens. *J. Biol. Chem.* **275**, 11451–11458
15. Hart, G. W., Housley, M. P., and Slawson, C. (2007) Cycling of O-linked

- beta-N-acetylglucosamine on nucleocytoplasmic proteins. *Nature* **446**, 1017–1022
16. Hardivillé, S., and Hart, G. W. (2014) Nutrient regulation of signaling, transcription, and cell physiology by O-GlcNAcylation. *Cell Metab.* **20**, 208–213
 17. Xu, S. L., Chalkley, R. J., Wang, Z. Y., and Burlingame, A. L. (2012) Identification of O-linked beta-D-N-acetylglucosamine-modified proteins from *Arabidopsis*. *Methods Mol. Biol.* **876**, 33–45
 18. Peter-Katalinić, J. (2005) Methods in enzymology: O-glycosylation of proteins. *Methods Enzymol.* **405**, 139–171
 19. Song, W., Mentink, R. A., Henquet, M. G., Cordewener, J. H., van Dijk, A. D., Bosch, D., America, A. H., and van der Krol, A. R. (2013) N-glycan occupancy of *Arabidopsis* N-glycoproteins. *J. Proteomics* **93**, 343–355
 20. Zielinska, D. F., Gnad, F., Schropp, K., Wiśniewski, J. R., and Mann, M. (2012) Mapping N-glycosylation sites across seven evolutionarily distant species reveals a divergent substrate proteome despite a common core machinery. *Mol. Cell* **46**, 542–548
 21. Vosseller, K., Trinidad, J. C., Chalkley, R. J., Specht, C. G., Thalhammer, A., Lynn, A. J., Snedecor, J. O., Guan, S., Medzihradzky, K. F., Maltby, D. A., Schoepfer, R., and Burlingame, A. L. (2006) O-linked N-acetylglucosamine proteomics of postsynaptic density preparations using lectin weak affinity chromatography and mass spectrometry. *Mol. Cell. Proteomics* **5**, 923–934
 22. Trinidad, J. C., Schoepfer, R., Burlingame, A. L., and Medzihradzky, K. F. (2013) N- and O-glycosylation in the murine synaptosome. *Mol. Cell. Proteomics* **12**, 3474–3488
 23. Trinidad, J. C., Barkan, D. T., Gullede, B. F., Thalhammer, A., Sali, A., Schoepfer, R., and Burlingame, A. L. (2012) Global identification and characterization of both O-GlcNAcylation and phosphorylation at the murine synapse. *Mol. Cell. Proteomics* **11**, 215–229
 24. Vizcaino, J. A., Deutsch, E. W., Wang, R., Csordas, A., Reisinger, F., Ríos, D., Dianes, J. A., Sun, Z., Farrah, T., Bandeira, N., Binz, P. A., Xenarios, I., Eisenacher, M., Mayer, G., Gatto, L., Campos, A., Chalkley, R. J., Kraus, H. J., Albar, J. P., Martinez-Bartolomé, S., Apweiler, R., Omenn, G. S., Martens, L., Jones, A. R., and Hermjakob, H. (2014) ProteomeX-change provides globally coordinated proteomics data submission and dissemination. *Nat. Biotechnol.* **32**, 223–226
 25. Medzihradzky, K. F., Kaasik, K., and Chalkley, R. J. (2015) Tissue-specific glycosylation at the glycopeptide level. *Mol. Cell. Proteomics* **14**, 2103–2110
 26. Baker, P. R., and Chalkley, R. J. (2014) MS-Viewer: A web-based spectral viewer for proteomics results. *Mol. Cell. Proteomics* **13**, 1392–1396
 27. Medzihradzky, K. F., Kaasik, K., and Chalkley, R. J. (2015) Characterizing sialic acid variants at the glycopeptide level. *Anal. Chem.* **87**, 3064–3071
 28. Kajijura, H., Okamoto, T., Misaki, R., Matsuura, Y., and Fujiyama, K. (2012) *Arabidopsis* beta1,2-xylosyltransferase: substrate specificity and participation in the plant-specific N-glycosylation pathway. *J. Biosci. Bioeng.* **113**, 48–54
 29. Zhang, H., Huang, R. Y., Jalili, P. R., Irungu, J. W., Nicol, G. R., Ray, K. B., Rohrs, H. W., and Gross, M. L. (2010) Improved mass spectrometric characterization of protein glycosylation reveals unusual glycosylation of maize-derived bovine trypsin. *Anal. Chem.* **82**, 10095–10101
 30. Chandler, K. B., Brnakova, Z., Sanda, M., Wang, S., Stalnak, S. H., Bridger, R., Zhao, P., Wells, L., Edwards, N. J., and Goldman, R. (2014) Site-specific glycan microheterogeneity of inter-alpha-trypsin inhibitor heavy chain H4. *J. Proteome Res.* **13**, 3314–3329
 31. Kim, Y. C., Jahren, N., Stone, M. D., Udeshi, N. D., Markowski, T. W., Witthuhn, B. A., Shabanowitz, J., Hunt, D. F., and Olszewski, N. E. (2013) Identification and origin of N-linked beta-D-N-acetylglucosamine monosaccharide modifications on *Arabidopsis* proteins. *Plant Physiol.* **161**, 455–464
 32. Floss, D. M., Sack, M., Stadlmann, J., Rademacher, T., Scheller, J., Stöger, E., Fischer, R., and Conrad, U. (2008) Biochemical and functional characterization of anti-HIV antibody-ELP fusion proteins from transgenic plants. *Plant Biotechnol. J.* **6**, 379–391
 33. Rademacher, T., Sack, M., Arcalis, E., Stadlmann, J., Balzer, S., Altmann, F., Quendler, H., Stiegler, G., Kunert, R., Fischer, R., and Stoger, E. (2008) Recombinant antibody 2G12 produced in maize endosperm efficiently neutralizes HIV-1 and contains predominantly single-GlcNAc N-glycans. *Plant Biotechnol. J.* **6**, 189–201
 34. Liebminger, E., Grass, J., Jez, J., Neumann, L., Altmann, F., and Strasser, R. (2012) Myrosinases TGG1 and TGG2 from *Arabidopsis thaliana* contain exclusively oligomannosidic N-glycans. *Phytochemistry* **84**, 24–30