

available at www.sciencedirect.comwww.elsevier.com/locate/jprot

Review

N-glycoproteomics in plants: Perspectives and challenges

Wei Song^{a,b}, Maurice G.L. Henquet^a, Remco A. Mentink^b, Aalt Jan van Dijk^a,
Jan H.G. Cordewener^a, Dirk Bosch^a, Antoine H.P. America^a, Alexander R. van der Krol^{b,*}

^aPlant Research International, Wageningen University and Research Centre, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands

^bLaboratory of Plant Physiology, Wageningen University and Research Centre, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands

ARTICLE INFO

Article history:

Received 6 January 2011

Accepted 2 May 2011

Available online 12 May 2011

Keywords:

N-glycosylation

Glycoproteomics

N-glycans

Glycopeptides

Cell specific proteomics

Mass spectrometry

ABSTRACT

In eukaryotes, proteins that are secreted into the ER are mostly modified by N-glycans on consensus NXS/T sites. The N-linked glycan subsequently undergoes varying degrees of processing by enzymes which are spatially distributed over the ER and the Golgi apparatus. The post-ER N-glycan processing to complex glycans differs between animals and plants, with consequences for N-glycan and glycopeptide isolation and characterization of plant glycoproteins. Here we describe some recent developments in plant glycoproteomics and illustrate how general and plant specific technologies may be used to address different important biological questions.

© 2011 Elsevier B.V. All rights reserved.

Contents

1. N-glycosylation of proteins imported into ER	1464
2. Functions of N-glycans in the ER	1465
3. Functions of N-glycans beyond the ER	1465
4. Strategies in N-glycoprotein analysis: N-glycomics	1465
5. Strategies in N-glycoprotein analysis: N-glycoproteomics	1467
5.1. Gel based glycoprotein analysis.	1467
5.2. Complexity reduction by selective glyco-proteomics	1467
5.3. Lectin based selection of glycoproteins and glycopeptides	1468
5.4. Hydrazide based selection of glycoproteins and glycopeptides	1469

Abbreviations: ER, endoplasmatic reticulum; cgl, complex glycan less; GlcNAc, N-acetylglucosamine; OST, oligosaccharyltransferase; ERAD, ER-associated degradation; PNGase, peptide-N-glycosidase; UPR, unfolded protein response; MALDI-TOF MS, matrix assisted laser desorption/ionization time-of-flight mass spectrometry; PAGE, polyacrylamide gel electrophoresis; ConA, concanavalin A; WGA, wheat germ agglutinin; AIL, *Artocarpus integrifolia* lectin; PNA, peanut agglutinin; GalNAc, N-acetylgalactosamine; HRP, horseradish peroxidase; LC-MS/MS, liquid chromatography tandem mass spectrometry; CE, capillary electrophoresis; LTP, lipid transfer protein; RBCS, ribulose biphosphate carboxylase; GO, gene ontology.

* Corresponding author. Tel.: +31 317 482448.

E-mail address: sander.vanderkrol@wur.nl (A.R. van der Krol).

1874-3919/\$ – see front matter © 2011 Elsevier B.V. All rights reserved.

doi:10.1016/j.jprot.2011.05.007

6.	Applications of N-glycoproteome analysis	1470
6.1.	N-glycan type mapping	1470
6.2.	N-glycan site occupancy mapping for transmembrane topology verification	1470
6.3.	N-glycan site occupancy mapping for prediction tools	1470
7.	Unique opportunities for N-glycoproteomics in plants	1470
7.1.	Full N-glycoproteome mapping	1470
7.2.	Cell specific tagging of the complex N-glycoproteome	1471
8.	Bottlenecks in glycoproteomics	1471
8.1.	Multiple glycoprotein products	1471
8.2.	Peptide identification	1471
8.3.	Non-robust PNGase A	1471
9.	Summary and perspectives	1471
	Acknowledgment	1472
	Appendix A. Supplementary data	1472
	References	1472

1. N-glycosylation of proteins imported into ER

The process of N-glycosylation in plants was recently reviewed in the context of therapeutic glycoprotein production in plants by Gomord et al. (2010) [1]. The process starts with dolichol lipid-linked glycan biosynthesis on the cytosolic side of the ER membrane. After flipping to the lumen side of the ER glycan synthesis continues to the full Glc₃Man₉GlcNAc₂ structure [2]. This fully assembled glycan is linked to the protein backbone via an amide bond (N-linked: N-glycosylation) of an asparagine residue in an Asn-X-Ser/Thr (NxS/T) motif; where X can be any amino acid, except proline [3]. Transfer from the lipid to the protein occurs during translation

and transfer of secreted proteins into the ER at the Sec61 transmembrane channel complex by the oligosaccharyltransferase (OST) complex [4,5]. In *Arabidopsis*, only three from the nine different components of the OST complex in yeast have been identified [6]. Approximately 70% of all NxS/T sites on secreted proteins are indeed glycosylated [7,8], suggesting additional requirements for N-glycosylation such as flanking peptide sequence, rate of the protein folding inside the ER lumen and capacity of the OST glycosylation machinery [9–11].

In the ER the terminal glucose and some mannoses are trimmed from the N-glycan (see next section). The lipid linked glycan biosynthesis and the first modification steps in ER and the Golgi apparatus are conserved between higher eukaryotic

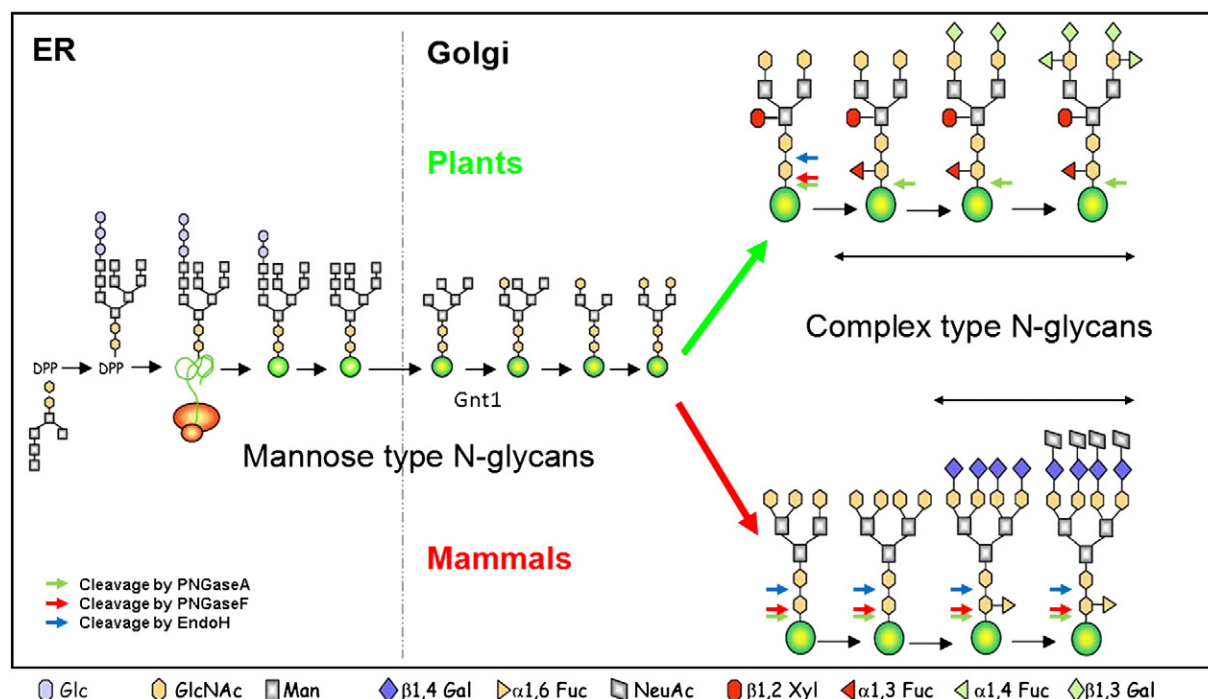


Fig. 1 – N-glycosylation in plants and animals. Differences in glycosylation pattern between plants and mammals appear in the Golgi apparatus. The most notable plant-specific N-glycan modifications are the α -1,3 fucose attached to the first GlcNAc residue and the β -1,2 xylose attached to the central mannose residue. Mammalian N-glycans may be modified by α -1,6 fucoses as opposed to the α -1,3 of plants, as well as β -1,4 galactose residues, attached to terminal GlcNAc moieties. To these galactose residues sialic acid groups may be attached. (Adapted from Ref. [12]).

species, but further modifications in the Golgi apparatus may differ between species [12]. In the Golgi apparatus the first modification is the removal of one to four α -1,2 mannose residues by α -mannosidase I [13]. Up until this point, the N-linked glycans are termed high-mannose N-glycans structures. The first obligatory step in the synthesis of complex glycans is the addition of an N-acetylglucosamine moiety to the α -1,3 mannose branch by N-acetylglucosaminetransferase I (GnTI) (Fig. 1, [14]) which in plants subsequently allows for two independent N-glycan core modifications: an α -1,3 linked fucose residue can be connected to the innermost GlcNAc residue and/or a β -1,2 xylose can be attached to the central mannose residue [15]. In mammals, the complex N-glycans have an α -1,6 fucose instead of an α -1,3 fucose linked to the innermost GlcNAc of the glycan core, while no β -1,2 xylose core modifications occur in mammals (Fig. 1) and glycans may be further modified by a β -1,4 galactose to which additional sialic acid residues may be attached [16–18]. Since many of these reactions do not go to completion, a very heterogeneous end-glycan profile may exist on a specific glycoprotein in a tissue. The difference in core glycan modifications between plants and animals is reviewed in detail by Gomord [1] and of importance in glycoproteomics because the enzyme that is frequently used for deglycosylation of mammalian glycoproteins (PNGase F) does not recognize the plant glycan structure modified by an α -1,3 linked fucose.

2. Functions of N-glycans in the ER

In the ER, the mannose type N-glycan on the imported protein supposedly aids the protein folding process. Moreover, the attached glycan serves as a tag, which marks the quality of the protein folding state [19]. The misfolded protein is retained in the ER or, when transported to the Golgi apparatus may be recycled back to the ER, till correctly folded [20]. However, a slow N-glycan trimming activity in the ER serves as a timer mechanism; when the N-glycan is trimmed too far, this indicates a problem with folding and the trimmed N-glycan with exposed hydrophobic domains, is recognized by ER associated degradation machinery (ERAD) [21]. The ERAD protein complex binds to the misfolded glycoprotein and facilitates export to the cytosol where the N-glycan is removed by PNGases and the protein is targeted for degradation by proteasomes [20,22,23]. Recently a set of N-glycan modifying mannosidase encoding genes (Mns 1–3) were identified in *Arabidopsis* [24]. The triple mutant of the three *mns* genes forms only short roots, with radially swollen cortical cells and displays alterations in development of cell walls and aerial plant parts, suggesting that several glycoproteins involved in cell wall formation are affected by the mutations. The effect of the mutations could be linked to the role of mannose trimming in glycoprotein folding quality control in the ER and ER-associated degradation of misfolded glycoproteins [24]. Our knowledge of protein folding quality control and ERAD is mostly from yeast and mammalian systems, but recently the role of ERAD related processes in plants was reviewed by Liu and Howell [25]. Because of the prominent role of N-glycans in the unfolded protein responses (UPR) and associated ERAD, glycoproteomics can play a key role in further elucidation of these processes in plants.

3. Functions of N-glycans beyond the ER

In mammalian cells the glycans on the glycoproteins have been proven to be involved in a wide range of biological functions such as receptor binding, cell signaling, protein folding, subcellular distribution and localization, protein stability, endocytosis, immune recognition, inflammation and pathogenicity [5,26–30]. In contrast, only very little is known about the role of N-glycan processing beyond the ER in plants. Only for isolated plant proteins the function of the processed N-glycan has been investigated in detail. N-glycans on plant glycoproteins have been shown to affect catalytic activity, thermostability and folding [31] or subcellular localization and secretion [32]. Recently it was found that N-glycans may also play a role in plant pathogen interactions [33] and in functional pattern recognition receptors [34]. Some plant glycoproteins have a special N-glycan structure known as a “Lewis a” antigen [35,36]. In mammals these glycan structures are located on glycoproteins at the cell surface and function in cell–cell recognition and adhesion processes [15]. In plants these structures are also at the cell surface, but the function of this specific glycan structure in plants remains unknown.

Mutations in mammalian cells that affect N-glycan processing are linked to very severe phenotypes (Type II congenital disorders [37,38]). In contrast, eliminating overall modifications to complex type N-glycans on plant glycoproteins has no or little effect on plant growth. For instance, the *Arabidopsis cgl* mutant is unable to modify N-glycans on glycoproteins in the Golgi apparatus to complex glycans due to a non-functional GnTI gene [39,40]. In the *cgl* mutant all glycoproteins which normally have complex N-glycans now have mannose type N-glycans. However, plants grow normal and only display a higher sensitivity to salt stress [41]. Also, when complex glycosylation was artificially suppressed by RNAi in tobacco and duckweed, the plants did not show an obvious growth phenotype [42–44]. There is also an altered N-glycan profile in the *Arabidopsis* glycosylation mutant *alg3-2* [45,46]. In the *alg3* mutant the N-glycans of proteins in the ER are replaced by aberrant mannose type N-glycans. Although it could be expected that this affects recognition of misfolded proteins in the ER, the *alg3* mutant shows no severe effect on plant growth [46]. Similar disorders in mammals result in severe phenotypes. The difference of the impact of mutated complex glycosylation in mammals and plants could be related to the fact that in mammalian cells this part of the glycoproteome has essential functions in protein–protein interactions in the circulatory system, while in plants there is only very limited long range transport of proteins.

4. Strategies in N-glycoprotein analysis: N-glycomics

There are two parts in the analysis of glycoproteins: one is related to the structure profiling of the N-glycan on the protein and the other is related to identification of the proteins which are modified by N-glycans. The strategies that can be used in glycoprotein analysis are summarized in Fig. 2. In this figure,

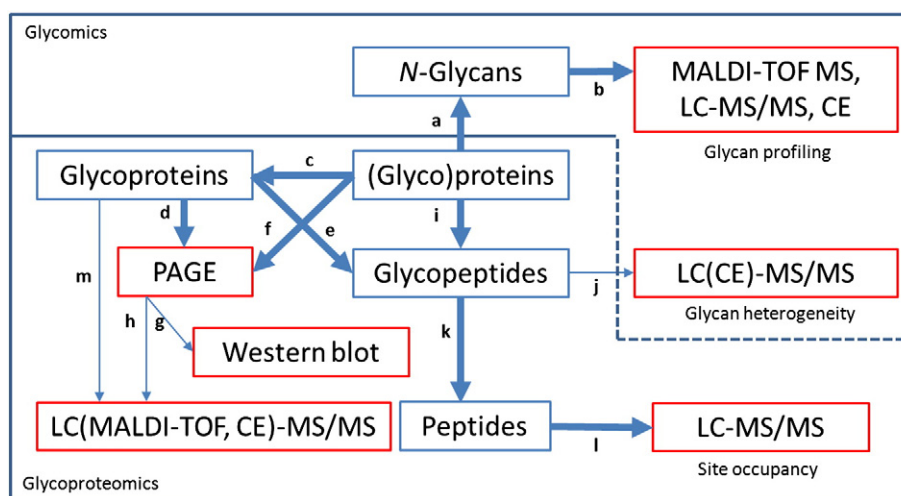


Fig. 2 – Strategies in N-glycomics and N-glycoproteomics. The thick arrows indicate strategies which potentially allow for bulk analysis of N-glycoproteins/glycans, while the thin arrows indicate strategies which are usually limited to individually selected, purified N-glycoproteins. Path a: release of N-glycans by PNGase treatment; b: N-glycans analysis by MALDI-TOF MS, LC-MS/MS or capillary electrophoresis (CE); c: isolation of N-glycoproteins subfraction from whole (glyco)proteins pool; d: PAGE size separation of N-glycoprotein subfraction; e: tryptic digestion of N-glycoprotein subfraction; f: PAGE size separation of whole (glyco)proteins pool; g: specific N-glycoprotein detection by Western blot; h: analysis of tryptic digest of selected N-glycoprotein by LC-MS/MS, MALDI-TOF MS or CE-MS/MS after PAGE separation; i: N-glycopeptides isolation by selective affinity capture; j: N-glycan analysis of purified glycopeptides by LC-MS/MS or CE-MS/MS; k: release of peptides from N-glycopeptides by PNGase treatment; l: mapping the N-glyco-site occupancy of peptides obtained from path k by LC-MS/MS; m: analysis of intact N-glycoprotein by LC-MS/MS, MALDI-TOF MS or CE-MS/MS.

path a–b and i–j describe the different options for obtaining information on the N-glycan structure based on MALDI-TOF MS, MALDI-TOF/TOF MS/MS, or LC-MS/MS analysis of either the cleaved N-glycans or the glycopeptides (glycomics).

Alternatively, capillary electrophoresis (CE) can be used for oligosaccharides profiling analysis. CE has the advantage of efficient separation and short run times allowing for HTP analysis of multiple samples [47–51]. For path i–j information

Table 1 – N-glycans profile in different plants, plant tissues and of isolated glycoprotein (numbers are % of the detected total N-glycan pool); data were compiled from Refs. [46,56,87]. Secr Ab: protein from transgene expressing antibody with secretion signal; ER Ab: protein from transgene expressing antibody with ER retention signal; n.d.: not detectable.

	Tobacco leaf			A. thaliana leaf		A. thaliana seed	A. thaliana seed
	Young	Old	Secrated antibody	Young	Old	Total	ER-retained antibody
"Mannose-type" N-glycans							
Man3GlcNAc2						n.d.	
Man4GlcNAc2			2.0			n.d.	
Man5GlcNAc2	3.0	4.0	4.0	26.7	31.7	21.3	
Man6GlcNAc2	5.0	5.0		4.7	7.5	9.3	
Man7GlcNAc2	3.0	2.0	4.0	4.5	5.7	4.7	26.6
Man8GlcNAc2		2.0	4.0	5.9	5.2	8.9	73.4
Man9GlcNAc2			1.0	2.2	1.8	n.d.	
Man3XylGlcNAc2	6.0	10.0	4.0	8.3	5.0	4.8	
Man4XylGlcNAc2						2.1	
GlcNAcMan3GlcNAc2			4.0				
GlcNAcMan3XylGlcNAc2	3.0	3.0	3.0	8.5	7.7	n.d.	
GlcNAc2Man3GlcNAc2			10.0				
GlcNAc2Man3XylGlcNAc2			4.0				
"Complex-type" N-glycans							
Man3XylFucGlcNAc2	29.0	39.0	14.0	15.3	11.3	41.2	
Man3FucGlcNAc2		4.0	1.0				
GlcNAcMan3FucGlcNAc2			2.0				
GlcNAcMan3XylFucGlcNAc2	42.0	31.0	13.0	8.7	9.1	3.5	
GlcNAc2Man3FucGlcNAc2			5.0				
GlcNAc2Man3XylFucGlcNAc2	8.0		23.0	15.1	15.0	4.3	
Total	100	100	100	100	100	100	100

about the mass (amino acid sequence) of the peptide to which the glycan is attached is necessary for establishing the glycan structure. The other branches relate to different strategies in identifying glycoproteins or glycopeptides (glycoproteomics). In Fig. 2 the thick arrows indicate strategies which potentially allow for bulk analysis of N-glycoproteins/N-glycans; the thin arrows indicate strategies which are usually limited to individually selected glycoproteins and thus are less compatible with an “omics” type of approach.

The total set of N-glycans present in the total glycoprotein pool of a plant tissue may be analyzed by MALDI-TOF MS, LC-MS/MS or CE after release of the N-glycans from the protein(s) by PNGase A treatment (Fig. 2 path a–b). For glycan identification, there is a free accessible tool called “GlycoWorkbench” available online, which can give glycan interpretation and annotation based on the glycan mass spectrometry data [52]. An example of such a MALDI-TOF MS analysis of the N-glycan pool from two different plant tissues (leaf and seed) is shown in Table 1. The relative abundance of the different glycan structures clearly differs between leaves and seeds, and also changes upon aging in leaves. Alternatively, a specific glycoprotein may be isolated from a plant tissue extract, as shown in Table 1. Structures can be released with PNGase A for analysis with different approaches (Fig. 2, path a–b), or the (purified) protein can be digested with trypsin and analyzed by LC-MS/MS, MALDI-TOF MS or CE-MS/MS to determine the mass of the glycopeptide (Fig. 2, path h). A bioinformatics tool called “ProTerNyc” can be used to calculate all putative N-glycopeptides and identify them within the mass data obtained from mass spectrometry [53]. In many cases, a single glycoprotein shows a range of different glycan structures [54,55], even when there is only a single N-glycosylation site present [56]. This most likely is the result of different subcellular resident times in ER and/or cis-, medial-, or trans-Golgi, resulting in different N-glycan processing intermediates. The N-glycan composition of glycoproteins isolated from different tissues (leaf, seedling, and seed) also shows distinct differences, which may be a reflection of different glycoprotein composition in these tissues (data not shown). Remarkably, there are large differences in the ratio of “mannose” and complex type glycans between different plant species. It could be that this is related to differences in ER:Golgi/plasma membrane ratio in large (tobacco) and small (*Arabidopsis*) plant cells. Experiments with the inhibitor of N-glycosylation tunicamycin shows that glycoproteins have a turnover varying from hours to days [34]. However, whether the changes in the N-glycan composition in leaves upon aging are a reflection of different glycoproteins being expressed or reflect a maturation of N-glycans on the existing glycoproteins is presently not known.

5. Strategies in N-glycoprotein analysis: N-glycoproteomics

5.1. Gel based glycoprotein analysis

Most information of individual plant glycoproteins is coming from PAGE based glycoprotein analysis (Fig. 2, path d, g, h). To test whether a protein is glycosylated it can be treated with

PNGase or Endo H, which (under denaturing conditions) can remove the glycan(s) from the glycoprotein. Treated and untreated protein fractions can be analyzed on Western blot after size separation on PAGE and difference in mobility of a specific glycoprotein can be detected if specific antibodies against this protein are available. This approach only tells whether a protein is glycosylated. It provides no information on the attached N-glycan structure beyond the level of mannose type or complex type N-glycan (by comparing sensitivity to PNGase A, PNGase F and Endo H: Fig. 1), nor information on glycosylation site occupancy. Moreover, it is limited to single protein analysis, so not applicable for “omics” approaches in glycoprotein analysis.

N-glycan modifications on proteins within a protein extract may be mapped by PAGE and Western blotting, using glycan specific lectins, antibodies or fluorescent staining techniques [57]. Lectins are a group of proteins capable of binding various forms of carbohydrates with different specificities. The lectin ConA is often used to detect glycoproteins. It binds to mannose type N-glycans, but not to plant complex-type N-glycans [57]. Some lectins may also bind to part of the O-glycoproteome. For instance, wheat germ agglutinin (WGA) binds to some GlcNAc- or sialic acid-containing N-glycans, but also to clusters of α -GalNAc O-glycans in mucin-like glycoprotein(s) of *Caenorhabditis elegans* [58]. The selective staining of part of the glycoproteome may make it easier to detect subtle changes in a total protein extract when glycoproteins are specifically affected. The anti-HRP polyclonal antibody directed against the horseradish peroxidase (HRP) recognizes the fucose and xylose residues on complex type N-glycans. Anti-HRP is therefore often used to detect proteins with complex N-glycans on Western blots. Fig. 3 shows an example of detection with anti-HRP of the complex glycoproteome in leaves of WT *Arabidopsis* plants (Fig. 3, lanes 1, 3 and 5) and leaves of the glycosylation mutant *cgl* [14,59], which is blocked in the formation of complex N-glycans (Fig. 3, lanes 2, 4 and 8). Glycoproteins may also be detected on gel by covalently linking a fluorescent molecule to periodate-oxidized carbohydrate groups (e.g. Pro-Q Emerald, Molecular Probes™). This staining method enables detection of even low abundant glycoproteins on gel, but does not distinguish between different types of N-glycans or between N- and O-glycans [60].

5.2. Complexity reduction by selective glyco-proteomics

Identification of a complete proteome of a tissue by LC-MS/MS is mostly hampered by the high complexity of the peptide mixture that is obtained after digestion with trypsin. This can be solved in part by fractionation of the peptide mixture using multi-dimensional LC-MS/MS. However, this makes comparison of numerous protein samples extra difficult. Proteomics by direct LC-MS/MS is therefore more feasible with selected proteome fractions with reduced peptide complexity, such as from isolated membranes or isolated glycoproteome fractions using different types of affinity capturing. For instance, only 3% of the tryptic fragments of the human proteome contain N-glycosylation sites; yet these peptides with an N-glycan represent the majority of extracellular proteins (over 70%) [61]. The putative N-glycoproteome of a plant species can be predicted from its genome sequence information. Of the 32,825 proteins that have

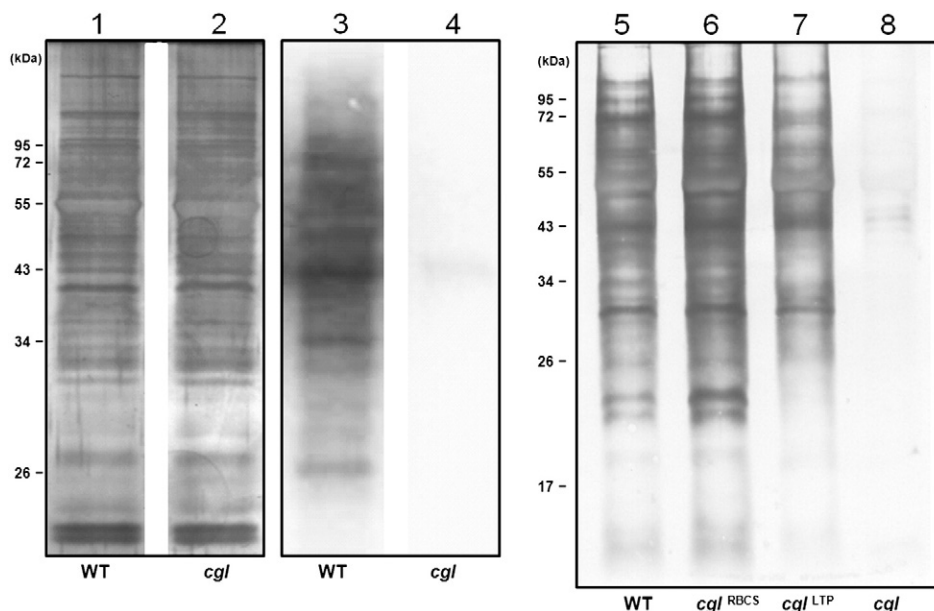


Fig. 3 – Complex N-glycan tagging in epidermal and mesophyll cells of *Arabidopsis* leaves. Lanes 1–2: SDS–PAGE of total protein extracts from leaves of *Arabidopsis* wild type (lane 1) and *cgl* mutant (lane 2) as detected by silver staining. Lanes 3–4: Western blot of identical WT (lane 3) and *cgl* (lane 4) fractions and staining of the complex glycoproteome using anti-HRP; Lanes 5–8: Western blot analysis of the complex glycoproteome (probe anti-HRP) of a leaf tissue of wild type *Arabidopsis* (lane 5), leaf tissue of *cgl* mutant complemented with RBCS–GnTI (lane 6) or LTP1–GnTI (lane 7), leaf tissue of *cgl* mutant (lane 8). For technical details see Supplement.

been annotated in the *Arabidopsis* genome (TAIR9, www.arabidopsis.org), 5879 proteins have a secretion signal peptide (SP) at the N-terminus predicted by TargetP, indicating that these are putative secreted proteins and potential candidates for N-glycosylation. Of this putative secreted proteome, 3597 proteins actually contain one or more NxS/T consensus glycosylation sites. Because not every NxS/T consensus site on secreted proteins will be occupied by an N-glycan, the predicted 3597 proteins constitute the maximum estimate of the N-glycoproteome. Fig. 5 compares the annotated biological functions of the entire *Arabidopsis* proteome with that of the predicted secreted N-glycoproteome (see also Table 2). The pie-charts show that

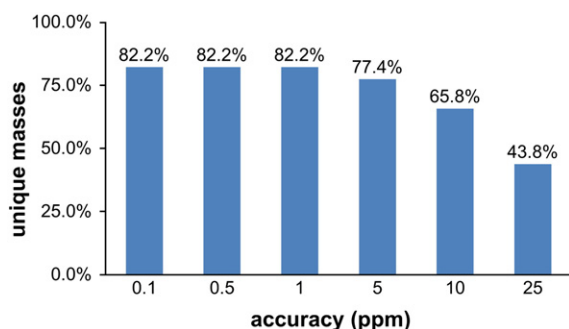


Fig. 4 – Percentage of unique putative glycopeptide masses per mass accuracy (ppm). Comparison of measured mass of putative glycopeptides and a database of predicted glycopeptide masses, based on the *Arabidopsis* sequence data from TAIR9. The percentage indicates how many of the masses can be uniquely assigned to a glycoprotein at the given accuracy of mass measurement.

the putative secreted N-glycoproteome is enriched for hydrolase, kinase and receptor signaling proteins. Especially the receptor kinases with leucine rich repeats (which have been implemented in plant pathogen defense reactions [62–65]), have multiple (up to 32) consensus N-glycosylation sites, suggesting that isolation and analysis of the N-glycoproteome may be of assistance in studying plant pathogen interactions.

After in silico trypsin digestion, the proteins of the predicted N-glycoproteome of *Arabidopsis* render 12,399 peptide fragments which contain one or more NxS/T sites. A tryptic digest of the full *Arabidopsis* proteome results in 1,665,744 (predicted) fragments and a selection of only the glycopeptides would give approximately a 100-fold reduction in peptide complexity, not regarding the additional complexity formed by different glycan structures. When glycopeptides from *Arabidopsis* can be selectively isolated and subsequently deglycosylated, the measured mass of the peptides (without the PTM) can be matched to a database of predicted (glyco)peptide masses (based on the tryptic peptide masses from all secreted proteins with a consensus NxS/T site). When the mass alone would be measured with an accuracy of ~1 ppm, a maximum of 82% of the peptides can be matched unequivocally to a single glycoprotein (see Fig. 4). For the remainder 18% of predicted (glyco)peptides additional information, like MS/MS sequence information is required for further identification, or protein isoforms could not be discriminated at (glyco)peptide sequence level.

5.3. Lectin based selection of glycoproteins and glycopeptides

Antibodies or lectins can not only be used for detection of glycoproteins, but may also be used for isolation of glycoprotein

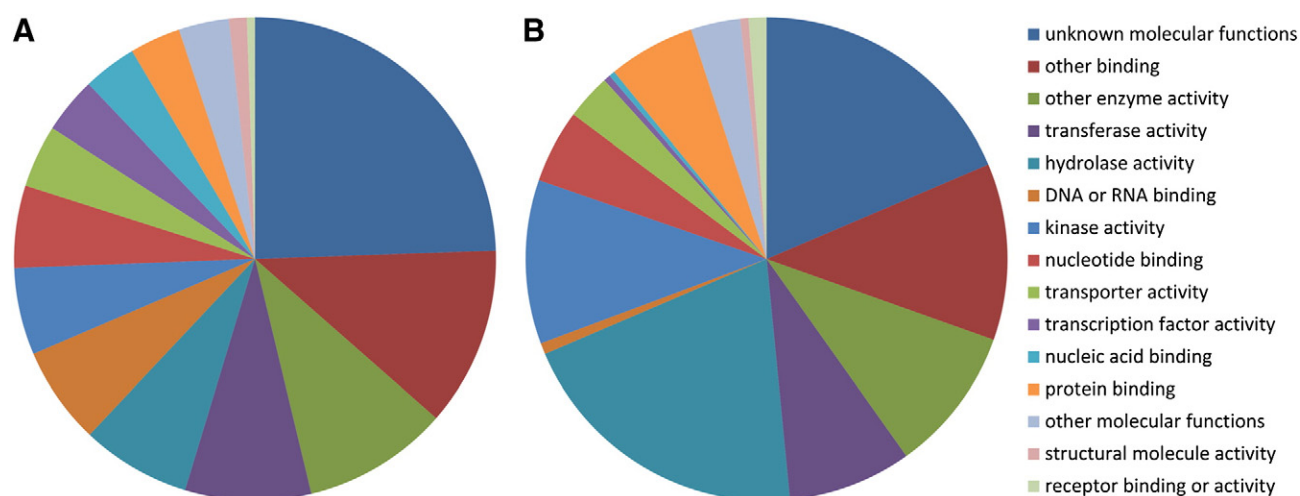


Fig. 5 – Ontology of *Arabidopsis* proteome and putative secreted N-glycoproteome. (A) Gene ontology for molecular function of the full *Arabidopsis* proteome from TAIR9; (B) Gene ontology for molecular function of the putative secreted N-glycoproteome of *Arabidopsis*[88] based on signal peptides (SignalP: www.cbs.dtu.dk/services/SignalP/) and NxS/T consensus site.

pools. The isolated glycoprotein subfraction may then be size fractionated on PAGE, 2D-PAGE or size-exclusion chromatography (SEC), and protein bands of interest may be isolated for analysis by LC-MS/MS, MALDI-TOF MS or CE-MS/MS (Fig. 2, path c–d–h). Alternatively, the glycoproteins can be identified directly by those different approaches without PAGE separation (Fig. 2, path m). Lectins are useful in selecting for groups of proteins carrying a specific type of glycan, but do not tolerate strong denaturing conditions for the interaction with glycans. Therefore, affinity capturing of intact glycoproteins may miss those proteins for which the glycan moiety is masked by the protein structure. The selectivity of lectins for certain glycans that affects detection on Western blots (see above) also affects the selection of the glycoprotein pool from a total protein extract. The lectin ConA is often used to capture proteins with mannose type glycans as ConA does not bind to complex type N-glycans [57]. WGA can be used to isolate glycoproteins with a terminal GlcNAc on N-glycans, but glycoproteins modified by O-glyco-

sylation may also bind to WGA [66]. The lectins AIL and PNA can be used to capture Gal-containing N- and O-glycoproteins. AIL is assumed to recognize mostly α -Gal, while PNA recognizes α -Gal and β -Gal [67,68]. Recently, different lectin affinity purification methods were combined for fractionation and identification of cell wall associated glycoproteins [69]. After tryptic digestion of the isolated protein fractions, peptides were identified by either LC-MS/MS or MALDI-TOF MS, resulting in the identification of 127 glycoproteins [69]. Of these 127 glycoproteins, 101 proteins were previously identified in a cell wall proteome analysis of 11-day-old etiolated hypocotyls of *Arabidopsis thaliana*[70].

5.4. Hydrazide based selection of glycoproteins and glycopeptides

Glycopeptides may also be captured by linking them through the glycan moiety to hydrazide resin (Fig. 2, path i–k–l). For this, the glycans of a tryptic digest of the isolated protein sample are first

Table 2 – Gene ontology annotations for the whole proteome (all #) and putative N-glycoproteome (glyco #) of *Arabidopsis*. Enrichment shows the increase (fold change) of proteins with specific molecular function in putative N-glycoproteome compared to whole proteome.

GO molecular function	All #	%	Glyco #	%	Enrichment
Unknown molecular functions	10,893	24.45%	744	18.60%	
Other binding	5345	12.00%	473	11.83%	
Other enzyme activity	4359	9.79%	389	9.73%	
Transferase activity	3764	8.45%	332	8.30%	
Hydrolase activity	3273	7.35%	805	20.13%	2.72
DNA or RNA binding	2910	6.53%	30	0.75%	
Kinase activity	2597	5.83%	439	10.98%	1.87
Nucleotide binding	2456	5.51%	196	4.90%	
Transporter activity	1864	4.18%	122	3.05%	
Transcription factor activity	1679	3.77%	18	0.45%	
Nucleic acid binding	1610	3.61%	16	0.40%	
Protein binding	1518	3.41%	233	5.83%	1.71
Other molecular functions	1502	3.37%	133	3.33%	
Structural molecule activity	536	1.20%	22	0.55%	
Receptor binding or activity	239	0.54%	48	1.20%	2.22
Total	44,545	100%	4000	100%	

oxidized to create functional aldehyde groups and subsequently covalently linked to hydrazide resin [71]. After washing away the non-linked peptides, the glycopeptides that are linked to the resin with N-glycans can be released by PNGase treatment, which results in a cleavage between the N-glycan and the peptide backbone. Although both PNGase A and PNGase F work for glycoproteins from mammals (Fig. 1), PNGase F is used mostly because it is (commercially) available at much higher concentration than PNGase A and can therefore be used in a much more robust way. PNGase F only cleaves “mannose type” N-glycans, whereas peptides linked by complex N-glycans (with a glycan core modified by α -1,3 fucose) will remain linked to the resin. To release the complex N-glycan containing peptides from the resin, the less efficient PNGase A has to be used. For quantitative analysis of N-linked glycopeptides and comparison between samples we use hydrazide based isolation of glycopeptides, followed by label free quantitative LC-MS/MS. Alternatively, differential stable isotope labeling of samples can be used [71]. In our hands, even after extensive washing, also non-glycopeptides are released from the hydrazide resin upon PNGase F treatment. The presence of an NxS/T glycosylation consensus site, in combination with a 1 Da mass shift (N $\rightarrow$ D conversion) of the peptide, can be used to distinguish between real glycopeptides and contaminating non-glycopeptides.

6. Applications of N-glycoproteome analysis

6.1. N-glycan type mapping

N-glycan type mapping can give information on the subcellular location or transport of a glycoprotein. Because the different N-glycan modifying enzymes are spatially distributed over the ER and cis-, medial- or trans-Golgi, the N-glycan structure itself contains information about the subcellular location or transport pathway of a glycoprotein. For instance, when a glycoprotein only contains mannose type N-glycans it is indicative of ER localization or of direct transport to subcellular compartments that by-pass the Golgi apparatus. When a protein contains one or more complex N-glycans, it is an indication that the protein is resident in or has passed through the Golgi apparatus (e.g. many of the plasma membrane proteins are modified by complex N-glycans). Evidence of glycosylation of plant proteins is mostly obtained through gel-shift based assays in which the mobility of the target protein before and after PNGase treatment is compared. When treatment of the protein with either PNGase F or endoglucanase H (both do not cleave complex N-glycans of which the core is modified by fucose) or PNGase A (cleaves all N-glycans) results in a gel mobility shift, this can be taken as proof for the presence of glycan and a rough indication of the glycan type may be obtained. As mentioned above, such information is indicative for subcellular location or protein transport pathway. However, on the intact protein not all attached N-glycans may be accessible for deglycosylation. The use of some reducing and/or denaturing agent may be required to make buried glycans accessible for enzymatic action. Furthermore, gel-shift based assays do not provide information on the site of N-glycan attachment and only limited information on the number and type of N-glycans that are attached to the protein.

6.2. N-glycan site occupancy mapping for transmembrane topology verification

For secreted proteins with multiple membrane spanning domains, the position of an N-glycan can give information on the membrane spanning domain topology, because N-glycan attachment is at the lumen site of the ER [2,23,72]. Indeed, the topology of membrane proteins is often verified by introducing a consensus N-glycosylation sites in specific loops, located between the transmembrane spanning domains, and monitor the occurrence of N-glycosylation when expressed in the cell [73,74]. N-glycoproteomics aimed at mapping N-glycan site occupancy of glycoproteins may therefore help to validate the correct topology of membrane proteins with one or more transmembrane spanning domains. Knowing the correct topology of membrane proteins is important for assigning the correct location of enzymatic activity (ER-lumen, cytosol, apoplast etc.) [75–79].

6.3. N-glycan site occupancy mapping for prediction tools

Accurate prediction of N-glycan site occupancy on secreted proteins is of use for development of plants as production platform of highly value pharmaceutical glycoproteins like antibodies or vaccines with uniform glycosylation patterns. The efficiency with which the glycan is transferred onto the consensus sequence is determined by (1) the serine or threonine in the consensus sequence NxS/T, (2) the amino acid x in the consensus sequence, (3) the location of the NxS/T motif within the polypeptide chain and its proximity to the C terminus, (4) the amino acids flanking the NxS/T sequence, and (5) the polypeptide structure surrounding the consensus sequence [80]. A prediction tool was developed for variable site-occupancy classification of N-linked glycosylation using artificial neural networks, based on only 48 NxS/T sites from individual proteins of which site occupancy had been retrieved from literature [11]. For such prediction tools, true glycoproteomics would greatly increase the input that is used for entraining the neural network and could thus vastly improve the accuracy of such prediction tools.

7. Unique opportunities for N-glycoproteomics in plants

7.1. Full N-glycoproteome mapping

The difference between high mannose versus complex-type glycans can be used to discriminate proteins that have passed the Golgi apparatus or not (yet). For the study of changes in the glycoproteome of the ER and ER derived compartments that bypass protein transport through the Golgi apparatus, the selectiveness of PNGase F is actually an advantage when glycopeptides are captured through their glycan and can only be released by PNGase F treatment. In protein extracts from wild type plants, the expensive PNGase A enzyme (which can cleave all N-glycans) is inefficient in releasing N-glycopeptides (our observation). This limits an integral analysis of all glycopeptides captured through their N-glycans. Because PNGase F can only cleave mannose-type N-glycans it will

mostly release glycopeptides from ER compartments in wild type plant protein extracts. For an inventory of the full N-glycoproteome, mutant plants in which N-glycan modification to complex glycan has been blocked can be used [14]. In this genetic background glycoproteins only contain mannose type N-glycans and therefore, all glycopeptides can be released by PNGase F from their N-glycans. Alternatively, for plant species for which such mutants are not available, different knockout strategies can be used to block complex glycosylation [1,42–44]. The isolated glycopeptides can be analyzed by label free quantitative LC–MS/MS comparing mass peak intensities between (groups of) samples. Alternatively, glycopeptides may be labeled during the PNGase treatment. When the deglycosylation reaction is performed in ^{18}O -water, ^{18}O is incorporated at the N-glycosylation site and samples treated with PNGase F in normal water and ^{18}O -water can be easily compared quantitatively [81].

7.2. Cell specific tagging of the complex N-glycoproteome

As a tool to compare the glycoprotein pools of different cell types in plants, we used the *Arabidopsis* complex-glycan-less (*cgl*) mutant [14,59]. By restoring N-glycan processing in this mutant by expressing the wild type GnTI gene under control of a cell specific promoters ([89] LTP1 for epidermal expression and RBCS for mesophyll specific expression), the biosynthesis of complex N-glycans is only restored in epidermal cells or mesophyll cells, respectively. Only in the cells with wild-type GnTI activity, glycoproteins carry complex type N-glycans with the xylose and fucose epitopes, which allows the study, characterization and isolation of the “complex-type” glycoproteome of the specified cells. Fig. 3 shows as example the difference in complex glycoproteome detected in *Arabidopsis* leaves. Lanes 1 and 2 show the equal (silver) staining of WT and *cgl* protein that was loaded, while lanes 3 and 4 show the Western blot detection of proteins with complex N-glycans using anti-HRP. Lane 4 shows that in the *cgl* mutant background virtually no proteins are detected by anti-HRP. Lanes 5–8 show an example of the cell type specific tagging with complex N-glycans in epidermal and mesophyll cells of *Arabidopsis* leaf tissue. The detected glycoprotein pattern in lanes 6 and 7 shows there is quite some overlap of the complex glycoproteins in the two cell types. Even though Western blots only allow for the detection of abundant glycoproteins, already for each cell type uniquely tagged proteins are detected. We are currently optimizing the protocols for isolating the peptides that are tagged with complex N-glycans from *Arabidopsis* leaf tissue for further protein identification.

8. Bottlenecks in glycoproteomics

8.1. Multiple glycoprotein products

The complex chemical structure, the different chemical properties of the peptide and glycan moieties and relative low abundance of most glycoproteins in plants, make analysis of glycoproteins or glycopeptides a difficult task. Integral approaches to glycoprotein analysis are further complicated by the dynamics of the glycosylation process itself, resulting in

variable glycan structures on the same glycoprotein. The already relatively low signal of a single glycopeptide with different attached glycan moieties may therefore be distributed over multiple peaks at different retention times using liquid chromatography, impeding detection or efficient identification in complex mixtures.

8.2. Peptide identification

The 1 Da (in water) or 2 Da (in ^{18}O -water) mass shift from the PNGase treatment, being diagnostic for the glycosylation site, could alternatively originate from a deamidation of asparagine. This needs to be checked by control samples (e.g. detecting the “normal” level of deamidation). Furthermore, because many glycoproteins contain only one or few glycosylation sites, linking the isolated glycopeptide mass to the presence and identity of a specific protein is no longer supported by multiple peptide masses. While common guidelines for publication require the identification of proteins by multiple peptides [82,83], this is not feasible in many cases for affinity selected single peptides. We argue that if good quality MS/MS information is provided per peptide, these “one-hit” identifications are correct and valuable.

8.3. Non-robust PNGase A

General N-glycoproteomics research for plant glycoproteins would profit a lot with isolation of the PNGase A gene and biotechnological production of this more versatile PNGase. It would enable more efficient release of glycopeptides independent of glycosylation mutant background. We mentioned the unique possibility of performing cell specific glycan tagging in plants, by cell specific complementation of the *cgl* mutant. Analysis of the proteins of the targeted cell type that are tagged by complex glycans, using the hydrazide capturing method, requires removal of “mannose” type N-glycopeptides with PNGase F and subsequent release with PNGase A. This procedure is still not effective enough and we are currently testing whether isolation of the complex glycopeptides may be enriched by prior removal of all mannose type N-glycans from the peptide mix before binding to the hydrazide resin.

9. Summary and perspectives

With the new developments at the level of software and state of the art analytical equipment, it becomes feasible to dive deeply inside of the functional importance of N-glycosylation in plants. Recent newly developed MS technologies like electron capture dissociation (ECD) and electron transfer dissociation (ETD) which results in peptide backbone fragmentation while keeping the post-translational modification intact may further improve characterization of glycopeptides [84–86]. The glycoproteome, and especially the part of the glycoproteome in the plasma membrane, which is (mostly) modified by complex N-glycans, is highly enriched in components that interact with extracellular signals during plant development, but also extracellular signals emanating from invading pathogens. Glycoproteins may only be one small part of these biological processes, but the tools that are being developed to study this part of the proteome may be of

great assistance to enhance our insight into these important biological processes.

Acknowledgment

This project was carried out within the research program of the Centre of BioSystems Genomics (CBSG) which is part of the Netherlands Genomics Initiative/Netherlands Organization for Scientific Research. We thank the (anonymous) reviewers for their constructive comments.

Appendix A. Supplementary data

Supplementary data to this article can be found online at [doi:10.1016/j.jprot.2011.05.007](https://doi.org/10.1016/j.jprot.2011.05.007).

REFERENCES

- [1] Gomord V, Fitchette AC, Menu Bouaouiche L, et al. Plant specific glycosylation patterns in the context of therapeutic protein production. *Plant Biotechnol J* 2010;8: 564–87.
- [2] Trombetta ES. The contribution of N-glycans and their processing in the endoplasmic reticulum to glycoprotein biosynthesis. *Glycobiology* 2003;13:77R–91R.
- [3] Bause E. Structural requirements of N-glycosylation of proteins. Studies with proline peptides as conformational probes. *Biochem J* 1983;209:331–6.
- [4] Knauer R, Lehle L. The oligosaccharyltransferase complex from *Saccharomyces cerevisiae*. Isolation of the OST6 gene, its synthetic interaction with OST3, and analysis of the native complex. *J Biol Chem* 1999;274:17249–56.
- [5] Lehle L, Strahl S, Tanner W. Protein glycosylation, conserved from yeast to man: a model organism helps elucidate congenital human diseases. *Angew Chem Int Ed Engl* 2006;45: 6802–18.
- [6] Lerouxel O, Mouille G, Andeme-Onzighi C, et al. Mutants in defective glycosylation, an *Arabidopsis* homolog of an oligosaccharyltransferase complex subunit, show protein underglycosylation and defects in cell differentiation and growth. *Plant J* 2005;42:455–68.
- [7] Apweiler R, Hermjakob H, Sharon N. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim Biophys Acta* 1999;1473:4–8.
- [8] Petrescu A, Milac A, Petrescu S, Dwek R, Wormald M. Statistical analysis of the protein environment of N-glycosylation sites: implications for occupancy, structure, and folding. *Glycobiology* 2004;14:103.
- [9] Kasturi L, Chen H, Shakin-Eshleman SH. Regulation of N-linked core glycosylation: use of a site-directed mutagenesis approach to identify Asn–Xaa–Ser/Thr sequons that are poor oligosaccharide acceptors. *Biochem J* 1997;323(Pt 2):415–9.
- [10] Mellquist JL, Kasturi L, Spitalnik SL, Shakin-Eshleman SH. The amino acid following an asn–X–Ser/Thr sequon is an important determinant of N-linked core glycosylation efficiency. *Biochemistry* 1998;37:6833–7.
- [11] Senger RS, Karim MN. Variable site-occupancy classification of N-linked glycosylation using artificial neural networks. *Biotechnol Prog* 2005;21:1653–62.
- [12] Faye L, Boulaflois A, Benchabane M, Gomord V, Michaud D. Protein modifications in the plant secretory pathway: current status and practical implications in molecular pharming. *Vaccine* 2005;23:1770–8.
- [13] Szumilo T, Kaushal GP, Elbein AD. Demonstration of GlcNAc transferase I in plants. *Biochem Biophys Res Commun* 1986;134:1395–403.
- [14] von Schaewen A, Sturm A, O'Neill J, Chrispeels M. 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* 1993;102:1109.
- [15] Lerouge P, Cabanes-Macheteau M, Rayon C, Fischette-Laine AC, Gomord V, Faye L. N-glycoprotein biosynthesis in plants: recent developments and future trends. *Plant Mol Biol* 1998;38:31–48.
- [16] van Ree R, Cabanes-Macheteau M, Akkerdaas J, et al. Beta(1,2)-xylose and alpha(1,3)-fucose residues have a strong contribution in IgE binding to plant glycoallergens. *J Biol Chem* 2000;275:11451–8.
- [17] Fotisch K, Vieths S. N- and O-linked oligosaccharides of allergenic glycoproteins. *Glycoconj J* 2001;18:373–90.
- [18] Bencurova M, Hemmer W, Focke-Tejkl M, Wilson IB, Altmann F. Specificity of IgG and IgE antibodies against plant and insect glycoprotein glycans determined with artificial glycoforms of human transferrin. *Glycobiology* 2004;14: 457–66.
- [19] Helenius A, Aebi M. Roles of N-linked glycans in the endoplasmic reticulum. *Annu Rev Biochem* 2004;73:1019–49.
- [20] Spiro R. Role of N-linked polymannose oligosaccharides in targeting glycoproteins for endoplasmic reticulum-associated degradation. *Cell Mol Life Sci* 2004;61: 1025–41.
- [21] Hampton RY. ER-associated degradation in protein quality control and cellular regulation. *Curr Opin Cell Biol* 2002;14: 476–82.
- [22] Yoshida Y. A novel role for N-glycans in the ERAD system. *J Biochem* 2003;134:183–90.
- [23] Helenius A, Aebi M. Roles of N-linked glycans in the endoplasmic reticulum. *Annu Rev Biochem* 2004;73:1019–49.
- [24] Liebminger E, Veit C, Mach L, Strasser R. Mannose trimming reactions in the early stages of the N-glycan processing pathway. *Plant Signal Behav* 2010;5:476–8.
- [25] Liu JX, Howell SH. Endoplasmic reticulum protein quality control and its relationship to environmental stress responses in plants. *Plant Cell* 2010;22:2930–42.
- [26] Klenk H, Garten W. Host cell proteases controlling virus pathogenicity. *Trends Microbiol* 1994;2:39–43.
- [27] Ihara Y, Sakamoto Y, Mihara M, Shimizu K, Taniguchi N. Overexpression of N-acetylglucosaminyltransferase III disrupts the tyrosine phosphorylation of Trk with resultant signaling dysfunction in PC12 cells treated with nerve growth factor. *J Biol Chem* 1997;272:9629.
- [28] Takahashi M, Tsuda T, Ikeda Y, Honke K, Taniguchi N. Role of N-glycans in growth factor signaling. *Glycoconj J* 2003;20: 207–12.
- [29] Green R, Stone E, Tenno M, Lehtonen E, Farquhar M, Marth J. Mammalian N-glycan branching protects against innate immune self-recognition and inflammation in autoimmune disease pathogenesis. *Immunity* 2007;27:308–20.
- [30] Chatterjee S, Kolmakova A, Rajesh M. Regulation of lactosylceramide synthase (glucosylceramide 14 galactosyltransferase); implication as a drug target. *Curr Drug Targets* 2008;9:272–81.
- [31] Lige B, Ma S, van Huystee RB. The effects of the site-directed removal of N-glycosylation from cationic peanut peroxidase on its function. *Arch Biochem Biophys* 2001;386:17–24.
- [32] Ceriotti A, Duranti M, Bollini R. Effects of N-glycosylation on the folding and structure of plant proteins. *J Exp Bot* 1998;49: 1091–103.

- [33] Pattison RJ, Amtmann A. N-glycan production in the endoplasmic reticulum of plants. *Trends Plant Sci* 2009;14: 92–9.
- [34] Haweker H, Rips S, Koiwa H, et al. Pattern recognition receptors require N-glycosylation to mediate plant immunity. *J Biol Chem* 2010;285:4629–36.
- [35] Fitchette-Laine AC, Gomord V, Cabanes M, Michalski JC, Saint Macary M, Foucher B, et al. N-glycans harboring the Lewis a epitope are expressed at the surface of plant cells. *Plant J* 1997;12:1411–7.
- [36] Melo NS, Nimtz M, Conradt HS, Fevereiro PS, Costa J. Identification of the human Lewis(a) carbohydrate motif in a secretory peroxidase from a plant cell suspension culture (*Vaccinium myrtillus* L.). *FEBS Lett* 1997;415:186–91.
- [37] Schollen E, Grunewald S, Keldermans L, Albrecht B, Korner C, Matthijs G. CDG-Id caused by homozygosity for an ALG3 mutation due to segmental maternal isodisomy UPD3(q21.3-qter). *Eur J Med Genet* 2005;48:153–8.
- [38] Grunewald S, Matthijs G, Jaeken J. Congenital disorders of glycosylation: a review. *Pediatr Res* 2002;52:618–24.
- [39] Vonschaewen A, Sturm A, Oneill J, Chrispeels MJ. 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* 1993;102:1109–18.
- [40] Strasser R, Stadlmann J, Svoboda B, Altmann F, Glossl J, Mach L. Molecular basis of N-acetylglucosaminyltransferase I deficiency in *Arabidopsis thaliana* plants lacking complex N-glycans. *Biochem J* 2005;387:385–91.
- [41] Kang JS, Frank J, Kang CH, et al. Salt tolerance of *Arabidopsis thaliana* requires maturation of N-glycosylated proteins in the Golgi apparatus. *Proc Natl Acad Sci USA* 2008;105:5933–8.
- [42] Cox KM, Sterling JD, Regan JT, et al. Glycan optimization of a human monoclonal antibody in the aquatic plant *Lemna minor*. *Nat Biotechnol* 2006;24:1591–7.
- [43] Strasser R, Stadlmann J, Schahs M, et al. Generation of glyco-engineered *Nicotiana benthamiana* for the production of monoclonal antibodies with a homogeneous human-like N-glycan structure. *Plant Biotechnol J* 2008;6: 392–402.
- [44] Daskalova SM, Radder JE, Cichacz ZA, et al. Engineering of *N. benthamiana* L. plants for production of N-acetylgalactosamine-glycosylated proteins—towards development of a plant-based platform for production of protein therapeutics with mucin type O-glycosylation. *BMC Biotechnol* 2010;10:62.
- [45] von Schaewen A, Sturm A, O'Neill J, Chrispeels MJ. 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* 1993;102:1109–18.
- [46] Henquet M, Lehle L, Schreuder M, et al. Identification of the gene encoding the alpha 1, 3-mannosyltransferase (ALG3) in *Arabidopsis* and characterization of downstream N-glycan processing. *Plant Cell Online* 2008;20:1652.
- [47] Taverna M, Baillet A, Biou D, Schlüter M, Werner R, Ferrier D. Analysis of carbohydrate mediated heterogeneity and characterization of N linked oligosaccharides of glycoproteins by high performance capillary electrophoresis. *Electrophoresis* 1992;13:359–66.
- [48] Chen FTA, Evangelista RA. Profiling glycoprotein N linked oligosaccharide by capillary electrophoresis. *Electrophoresis* 1998;19:2639–44.
- [49] Karamanos NK, Hjerpe A. Strategies for analysis and structure characterization of glycans/proteoglycans by capillary electrophoresis. Their diagnostic and biopharmaceutical importance. *Biomed Chromatogr* 1999;13:507–12.
- [50] Kamoda S, Nomura C, Kinoshita M, et al. Profiling analysis of oligosaccharides in antibody pharmaceuticals by capillary electrophoresis. *J Chromatogr A* 2004;1050:211–6.
- [51] Balaguer E, Neusüss C. Glycoprotein characterization combining intact protein and glycan analysis by capillary electrophoresis–electrospray ionization–mass spectrometry. *Anal Chem* 2006;78:5384–93.
- [52] Ceroni A, Maass K, Geyer H, Geyer R, Dell A, Haslam SM. GlycoWorkbench: a tool for the computer-assisted annotation of mass spectra of glycans. *J Proteome Res* 2008;7: 1650–9.
- [53] Albenne C, Canut H, Boudart G, Zhang Y, San Clemente H, Pont-Lezica R, et al. Plant cell wall proteomics: mass spectrometry data, a trove for research on protein structure/function relationships. *Mol Plant* 2009;2:977–89.
- [54] Triguero A, Cabrera G, Cremata JA, Yuen CT, Wheeler J, Ramirez NI. Plant-derived mouse IgG monoclonal antibody fused to KDEL endoplasmic reticulum-retention signal is N-glycosylated homogeneously throughout the plant with mostly high-mannose-type N-glycans. *Plant Biotechnol J* 2005;3:449–57.
- [55] Lim JM, Aoki K, Angel P, et al. Mapping glycans onto specific N-Linked glycosylation sites of *Pyrus communis* PGIP redefines the interface for EPG–PGIP interactions. *J Proteome Res* 2008;8: 673–80.
- [56] Henquet M, Eigenhuijsen J, Hesselink T, et al. Characterization of the single-chain Fv–Fc antibody MBP10 produced in *Arabidopsis* alg3 mutant seeds. *Transgenic Res* 2010;1–10.
- [57] Varki A. *Essentials of glycobiology*. Cold Spring Harbor Laboratory Press; 1999.
- [58] Natsuka S, Kawaguchi M, Wada Y, Ichikawa A, Ikura K, Hase S. Characterization of wheat germ agglutinin ligand on soluble glycoproteins in *Caenorhabditis elegans*. *J Biochem* 2005;138: 209.
- [59] Strasser R, Stadlmann J, Svoboda B, Altmann F, Glossl J, Mach L. Molecular basis of N-acetylglucosaminyltransferase I deficiency in *Arabidopsis thaliana* plants lacking complex N-glycans. *Biochem J* 2005;387:385.
- [60] Zhang Y, Huang L, Wang Z. Techniques for staining glycoprotein on gel electrophoresis or electroblotting. *Progress Chemistry* 2008;20:1158–64.
- [61] Zhang H, Loriaux P, Eng J, et al. UniPep—a database for human N-linked glycosites: a resource for biomarker discovery. *Genome Biol* 2006;7:R73.
- [62] dzierski K, Montgomery J, Curtis J, Handman E. Leucine-rich repeats in host-pathogen interactions. *Arch Immunol Ther Exp* 2004;52:104–12.
- [63] Shanmugam V. Role of extracytoplasmic leucine rich repeat proteins in plant defence mechanisms. *Microbiol Res* 2005;160:83–94.
- [64] Chinchilla D, Zipfel C, Robatzek S, et al. A flagellin-induced complex of the receptor FLS2 and BAK1 initiates plant defence. *Nature* 2007;448:497–500.
- [65] Dodds PN, Rathjen JP. Plant immunity: towards an integrated view of plant–pathogen interactions. *Nat Rev Genet* 2010;11: 539–48.
- [66] Jackson S, Tjian R. O-glycosylation of eukaryotic transcription factors: implications for mechanisms of transcriptional regulation. *Cell* 1988;55:125–33.
- [67] Banerjee R, Mande S, Ganesh V, et al. Crystal structure of peanut lectin, a protein with an unusual quaternary structure. *Proceedings National Academy Sciences USA* 1994;91:227.
- [68] Arockia Jeyaprakash A, Jayashree G, Mahanta S, et al. Structural basis for the energetics of jacalin–sugar interactions: promiscuity versus specificity. *J Mol Biol* 2005;347:181–8.
- [69] Zhang Y, Giboulot A, Zivy M, Valot B, Jamet E, Albenne C. Combining various strategies to increase the coverage of the plant cell wall glycoproteome. *Phytochemistry* 2010.

- [70] Irshad M, Canut H, Borderies G, Pont-Lezica R, Jamet E. A new picture of cell wall protein dynamics in elongating cells of *Arabidopsis thaliana*: confirmed actors and newcomers. *BMC Plant Biol* 2008;8:94.
- [71] Tian Y, Zhou Y, Elliott S, Aebersold R, Zhang H. Solid-phase extraction of N-linked glycopeptides. *Nat Protoc* 2007;2:334–9.
- [72] Moremen KW, Molinari M. N-linked glycan recognition and processing: the molecular basis of endoplasmic reticulum quality control. *Curr Opin Struct Biol* 2006;16:592–599.
- [73] Wo ZG, Oswald RE. Transmembrane topology of two kainate receptor subunits revealed by N-glycosylation. *Proceedings National Academy Sciences USA* 1994;91:7154.
- [74] Kato Y, Sakaguchi M, Mori Y, et al. Evidence in support of a four transmembrane-pore-transmembrane topology model for the *Arabidopsis thaliana* Na⁺/K⁺ translocating AtHKT1 protein, a member of the superfamily of K⁺ transporters. *Proceedings National Academy Sciences USA* 2001;98:6488.
- [75] Kimura Y, Hess D, Sturm A. The N-glycans of jack bean alpha-mannosidase. Structure, topology and function. *Eur J Biochem* 1999;264:168–75.
- [76] Mühlenhoff M, Manegold A, Windfuhr M, Gotza B, Gerardy-Schahn R. The impact of N-glycosylation on the functions of polysialyltransferases. *J Biol Chem* 2001;276:34066.
- [77] Yusa A, Kitajima K, Habuchi O. N-linked oligosaccharides are required to produce and stabilize the active form of chondroitin 4-sulphotransferase-1. *Biochem J* 2005;388:115.
- [78] Yusa A, Kitajima K, Habuchi O. N-linked oligosaccharides on chondroitin 6-sulfotransferase-1 are required for production of the active enzyme, Golgi localization, and sulfotransferase activity toward keratan sulfate. *J Biol Chem* 2006;281:20393.
- [79] Uemura S, Kurose T, Suzuki T, et al. Substitution of the N-glycan function in glycosyltransferases by specific amino acids: ST3Gal-V as a model enzyme. *Glycobiology* 2006;16:258.
- [80] Jones J, Krag SS, Betenbaugh MJ. Controlling N-linked glycan site occupancy. *Biochimica Biophysica Acta BBA General Subjects* 2005;1726:121–37.
- [81] Shakey Q, Bates B, Wu J. An approach to quantifying N-linked glycoproteins by enzyme-catalyzed 18O3-labeling of solid-phase enriched glycopeptides. *Anal Chem* 2010;82:7722–8.
- [82] Nesvizhskii AI. A survey of computational methods and error rate estimation procedures for peptide and protein identification in shotgun proteomics. *J Proteomics* 2010;73:2092–123.
- [83] Alves G, Ogurtsov AY, Yu YK. Assigning statistical significance to proteotypic peptides via database searches. *J Proteomics* 2010;74:199–211.
- [84] Syka JEP, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proceedings National Academy Sciences USA* 2004;101:9528.
- [85] Ahn NG, Shabb JB, Old WM, Resing KA. Achieving in-depth proteomics profiling by mass spectrometry. *ACS Chem Biol* 2007;2:39–52.
- [86] Jackson SN, Dutta S, Woods AS. The use of ECD/ETD to identify the site of electrostatic interaction in noncovalent complexes. *J Am Soc Mass Spectrom* 2009;20:176–9.
- [87] Elbers IJ, Stoopen GM, Bakker H, et al. Influence of growth conditions and developmental stage on N-glycan heterogeneity of transgenic immunoglobulin G and endogenous proteins in tobacco leaves. *Plant Physiol* 2001;126:1314–22.
- [88] Berardini TZ, Mundodi S, Reiser L, et al. Functional annotation of the *Arabidopsis* genome using controlled vocabularies. *Plant Physiol* 2004;135:745.
- [89] Thoma S, Hecht U, Kippers A, Botella J, De Vries S, Somerville C. Tissue-specific expression of a gene encoding a cell wall-localized lipid transfer protein from *Arabidopsis*. *Plant Physiol* 1994;105:35–45.