Identification of Lewis A containing glycoproteins in plants and elucidation of the structure of plant Glycosylphosphatidylinositol anchors

Ph.D. Thesis

Beihammer Gernot

Vienna, March 2021

Supervisor: Assoc. Prof. Dr. Richard Strasser

Co-Supervisor: Univ. Prof. Dr. Friedrich Altmann

Acknowledgements

This research would not have been possible without the guidance of my supervisor Richard Strasser, who was so kind to provide me with the possibility to work on this project, supported me with scientific input and allowed me to follow the research at my own pace. Furthermore, I would like to thank all the members of Richard's research group who supported me during this project. Special thanks go to Christiane Veit, Ulrike Vavra, Yunji Shin, Kathrin Göritzer and Jennifer Schoberer, who all gave me help in scientific or technical questions and were responsible for the great atmosphere in the lab. My warm thanks also go out to the whole Department of Applied Genetics and Cell Biology and all the people who provided me with scientific input. In addition, I would like to give my warmest thanks to all the people in the BioTop PhD program, especially Margareta Furtmüller and Christa Jakopitsch.

Furthermore, I would like to thank Fritz Altmann and his research group, especially Daniel Maresch for his help with mass spectrometry and fruitful discussions. Many thanks also go to Rudolf Figl and Karin Polacsek for their support and effort.

I would also like to thank Els Van Damme, who provided me with the opportunity to work in her lab at Gent University for 6 months during this PhD thesis. The experience in the lab in Belgium has broadened my scientific horizon. Particular thanks go to Andrea Romero Pérez for her help and the experiments we conducted together. Furthermore, I would like to thank Isabel Verbeke for her support and patience.

Besides the people who have helped me on a professional level, I would like to thank my parents for their continuous help and support. I am especially grateful to my wife Julia, who shares my interest in plant glycobiology and has not only been a great help due to scientific input and discussions, but also by challenging me every day anew and due to her warm-hearted emotional support.

My thanks also go out to the FWF and the Belgian FWO for funding this project (project number I 2417).

Abstract

Terminal sugar residues on complex N-glycans affect the properties of the glycoproteins to which they are attached. In plants, the only known outer chain elongation of complex N-glycans are Lewis A structures. This glycan structure is conserved among plants, yet its abundance varies greatly between species and even tissues within the same organism. Lewis A bearing glycoproteins were isolated from *Arabidopsis thaliana*, *Nicotiana benthamiana* and *Oryza sativa*. As the majority of proteins identified in *A. thaliana* are involved in secondary cell wall biosynthesis, a potential phenotype of Lewis A deficient plants in terms of secondary cell wall composition was investigated. Plants lacking Lewis A glycans developed normally, however, and showed no alterations in cell wall constituents. Furthermore, a search of the 1001 genome database was conducted for natural accessions of *A. thaliana* containing mutations in the genes encoding the glycosyltransferases required for Lewis A biosynthesis. Lewis A levels were investigated in selected accessions, but only one line was found to be devoid of Lewis A structures while most tested plants contained comparable amounts of Lewis A.

Besides the role of Lewis A glycans, the structure of the glycan moiety of the plant Glycosylphosphatidylinositol (GPI) anchor was investigated. GPI anchored proteins are retained at the plasma membrane via their lipid moiety and play pivotal roles for plant fertility, stress response and signal transduction. Despite their significance, the structure of the GPI anchors has so far been shown for only one plant species, *Pyrus communis*. Here, the structure of the GPI glycan moiety was investigated in *A. thaliana* and *N. benthamiana*. GPI anchors in both plants were modified with a galactose β1,4-linked to the first mannose. The study presented here helps to close gaps in current knowledge of two important types of plant glycosylation, namely N-glycosylation and GPI anchors.

Kurzfassung

Die endständigen Zuckermoleküle auf komplexen N-Glykanen beeinflussen die Eigenschaften der Glykoproteine, an die sie gebunden sind. Die einzige in Pflanzen bekannte Verlängerung komplexer N-Glykane sind Lewis A Strukturen. Diese Glykanstruktur ist bei Pflanzen konserviert, wobei die Menge stark zwischen verschiedenen Spezies, aber auch zwischen verschiedenen Gewebetypen desselben Organismus schwankt. Lewis A tragende Glykoproteine aus *Arabidopsis thaliana*, *Nicotiana benthamiana* und *Oryza sativa* wurden isoliert. Da die große Mehrheit der in *A. thaliana* identifizierten Proteine eine Funktion in der Biosynthese sekundärer Pflanzenzellwände spielt, wurde ein entsprechenden Phenotyp in Pflanzen ohne Lewis A untersucht. Jedoch entwickelten sich Pflanzen ohne Lewis A normal und zeigten keine Besonderheiten bezüglich der Bestandteile der sekundären Zellwand. Des Weiteren wurde die 1001 genome database nach natürlichen Varianten von *A. thaliana* durchsucht, die Mutationen in den für die Lewis A Biosynthese benötigten Glykosyltransferasen enthalten. Nur eine Linie zeigte keine nachweisbaren Lewis A Strukturen, während andere Varianten vergleichbare Mengen an Lewis A enthielten.

Außerdem wurde die Struktur von pflanzlichen Glykosylphosphatidylinositol (GPI) Ankern untersucht. Mittels GPI verankerte Proteine werden an der Außenseite der Plasmamembran zurückgehalten und spielen in Pflanzen eine essenzielle Rolle für Fruchtbarkeit, Stressabwehr und Signalübertragung. Trotz deren Signifikanz wurde bisher nur die Struktur eines einzigen pflanzlichen GPI Ankers, aus *Pyrus communis,* erforscht. Hier wird die Untersuchung der Glykanuntereinheit der GPI Anker in *A. thaliana* und *N. benthamiana* beschrieben. In beiden Pflanzen waren die GPI Anker mit einer Galaktose, β1,4-gebunden an die erste Mannose, modifiziert. Die hier präsentierte Studie hilft Wissenslücken in zwei wichtigen Typen der Glykosylierung in Pflanzen zu schließen, nämlich in der N-Glykosylierung und bei GPI Ankern.

Table of Contents

Introduction

N-glycosylation

N-glycosylation is one of the most abundant modifications to proteins, whereby an oligosaccharide is attached to nascent polypeptide chains in the endoplasmic reticulum (ER). N-glycosylation can alter the properties of glycoproteins by affecting, among others, folding, receptor-ligand interactions or protein stability in a biological environment and is present in all domains of life. The oligosaccharide transferred to the protein is conserved among most eukaryotes but differs greatly in archaea and bacteria (Jarrell et al., 2014; Nothaft et al., 2013). As the glycoprotein is transported along the secretory pathway, the attached oligosaccharide is modified in a species-specific manner, yielding a wide variety of N-glycans in different organisms (Wang et al., 2017).

Precursor assembly and processing of plant N-glycans

In the course of N-glycosylation, a preassembled oligosaccharide precursor consisting of two N-Acetylglucosamines (GlcNAc), nine mannoses (Man) and three glucoses (Glc), $GlcNAc_2Man_9Glc_3$, is transferred to certain asparagine residues within the sequon N-X-S/T on polypeptide chains, where X can be any amino acid but proline. Biosynthesis of this precursor is catalyzed by the Asparagine linked glycosylation (ALG) proteins. It starts at the cytosolic side of the ER-membrane by the action of ALG7 and a heterodimeric complex consisting of ALG13 + ALG14, which transfer the two core GlcNAcs to Dolicholpyrophospate in β1,4-linkage. The activated sugar nucleotide UDP-GlcNAc serves as substrate for these reactions. Subsequently, the precursor is elongated by five mannoses catalyzed by ALG1, 2 and 11, which transfer mannose residues from GDP-mannose to the growing precursor. A so far unknown protein then flips the precursor oligosaccharide to the luminal side of the ER membrane, where four more mannose residues are attached by the action of ALG3, 9 and 12. Three glucoses are then transferred from GDP-glucose to the A-strand of the precursor catalyzed by ALG6, 8 and 10. A schematic representation of the assembled precursor is shown in **Figure 1**.

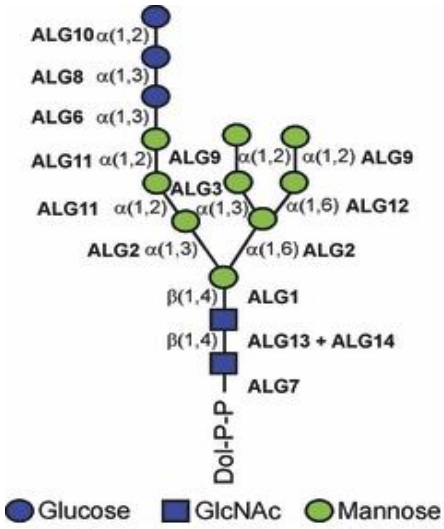


Figure 1. Structure of the oligosaccharide precursor transferred to proteins in eukaryotes. The linkages between the individual sugars as well as the glycosyltransferases catalyzing the precursor build up are shown. Taken from Farid et al., 2011.

The subsequent transfer of the preassembled sugar is catalyzed by the Oligosaccharyltransferase complex (OST). This occurs either co-translationally directly upon entry of the nascent polypeptide chain into the ER or post-translationally. In mammals, one of two different catalytic subunits can be present in the OST-complex: STT3A, which catalyzes the co-translational transfer, or STT3B, which catalyzes post-translational N-glycosylation of sites skipped by STT3A. Skipping of sites by STT3A can be due to multiple reasons, for example if the N-glycosylation site is located at the extreme C-terminal end of the protein or in close proximity of cysteine residues (Cherepanova et al., 2019). While the STT3A-containing OST is directly associated with the translocon, STT3B-OST is more distal from the translocon (Shrimal et al., 2017). Two homologues of the active site of OST were also detected in *A. thaliana*, AtSTT3A and AtSTT3B, but it is unclear whether they serve the same functions as their mammalian homologs (Niu et al., 2020).

Attachment of N-glycans is crucial for proper folding of glycoproteins. Upon transfer of the precursor to the peptide chain, the oligosaccharide is further processed. In a first step, the outer-most glucose is cleaved by Glucosidase I, followed by the removal of a second glucose-residue by Glucosidase II. The mono-glucosylated N-glycan is then bound by two chaperones, calnexin (CNX) and calreticulin (CRT), which are part of the quality control system to assure correct folding of glycoproteins. Cleavage of the last glucose residue by GCSII requires reorientation of the N-glycan in the active sight of the enzyme, allowing the glycoprotein to be subjected to glycan dependent folding in the meantime (Strasser et al., 2018). Upon cleavage of the last glucose residue, the glycoprotein dissociates from CNX and CRT and interacts with UDP-glucose:glycoprotein glucosyltransferase (UGGT), which conducts a quality check of correct protein folding. In case the glycoprotein is misfolded, UGGT transfers another glucose residue to the N-glycan, which allows in turn repeated binding to CNX and CRT. The misfolded glycoprotein can then undergo refolding. Terminally misfolded glycoproteins are degraded by the ER associated degradation (ERAD) system, while correctly folded glycoproteins are processed further. The last N-glycan processing reaction in the ER is the removal of one mannose residue catalyzed by α-mannosidase 3 (MNS3), resulting in $GlcNAc_2Man8$. Further processing steps take place in the Golgi Apparatus (**Figure 2**), where removal of three more mannoses occurs in the *cis*-Golgi. Afterwards, the N-Acetylglucosaminlytransferase GnTI attaches a GlcNAc residue in β1,2-linkage to the α1,3-linked mannose, followed by removal of two more mannoses and attachment of another GlcNAc residue to the α1,6-linked mannose by GnTII. Furthermore, core β1,2-xylose and α1,3-fucose are attached to the N-glycan, modifications that are specific for plants. The resulting structure, GnGnXF, can then be further elongated in the *trans*-Golgi by the sequential action of GALT1, a β1,3-galactosyltransferase, and FUT13, an α1,4-fucosyltransferase. The resulting terminal trisaccharide Fucα1,4(Galβ1,3)GlcNAc-R is also referred to as Lewis A, based on the nomenclature employed in the Lewis Blood group system.

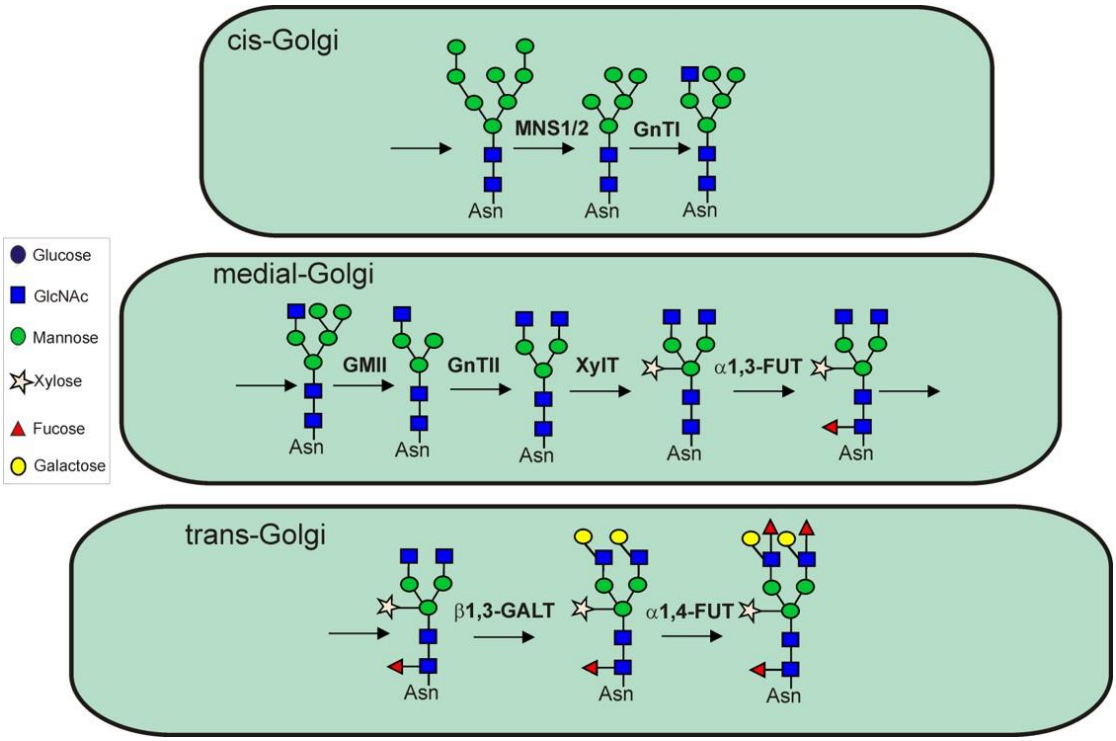


Figure 2. Processing of N-glycans on glycoproteins in the Golgi apparatus of plants. Modified from Schoberer and Strasser, 2011.

Lewis A structures in plants

The attachment of galactose and fucose, resulting in Lewis A bearing N-glycans, constitute the only known outer chain elongations of complex N-glycans in plants. Lewis A structures on complex N-glycans have been identified in all plant species investigated so far, including mosses (Koprivova et al., 2003), various food stuff (Wilson et al., 2001) and water plants (Maeda et al., 2016). In the dicot model plant *A. thaliana*, presence of Lewis A structures is tissue-dependent: while no Lewis A can be found in seedling leaves or rosette leaves, it is present in stems and siliques (Strasser et al., 2007). The transfer of the galactose residue catalyzed by GALT1 is the rate limiting step of Lewis A biosynthesis in *A. thaliana*. GALT1 has been reported to interact with Golgi α-mannosidase II (Schoberer et al., 2013), but not with FUT13. Interestingly, localization of FUT13 appears to be not solely limited to the *trans*-Golgi but has also been reported to be localized in the nuclear envelope (Rips et al., 2017). The reason for this dual targeting is, however, unclear.

The knowledge of which glycoproteins are decorated with Lewis A structures is scarce. Only four proteins have so far been identified in *A thaliana* as Lewis A bearing glycoproteins (Zhang et al., 2011; Zeng et al., 2018): Laccase4 (At2g38080), an uncharacterized GPI-anchored protein (At3g06035), DUF2921 (At1g52780) and Homologous to blue copper binding protein (At4g12880).

A Lewis A-deficient mutant line of *A. thaliana* displayed no obvious phenotype (Strasser et al., 2007). Due this lack of an obvious phenotype associated with lack of Lewis A structures in *A. thaliana* and the fact that only a very limited number of Lewis A bearing glycoproteins has been described, the role of Lewis A structures within plants has thus so far remained elusive.

Glycosylphosphatidylinositol anchor

Glycosylphosphatidylinositol (GPI) anchors are evolutionarily conserved post-translational modifications at the C-terminus of proteins that have been described in various eukaryotes and archaea (Kobayashi et al., 1997). Modification with a GPI anchor allows retention of proteins at the extracellular side of the plasma membrane. The backbone of the GPI anchor consists of an inositol phospholipid, a glycan moiety, and the bridging ethanolamine phosphate (EtNP) connecting the GPI-anchor with the C-terminus of the protein (**Figure 3**). The glycan moiety of the GPI-anchor has a conserved core structure consisting of 3 mannoses and a non-acetylated glucosamine (GlcN) residue. While this core structure appears to be conserved, a wide variety of side chain substitutions are found in different genera, but also in different tissues within the same organism (Tsai et al., 2012). Besides the side chain modifications, the type of lipid moiety of the GPI anchor varies between species (Kinoshita and Fujita, 2016).

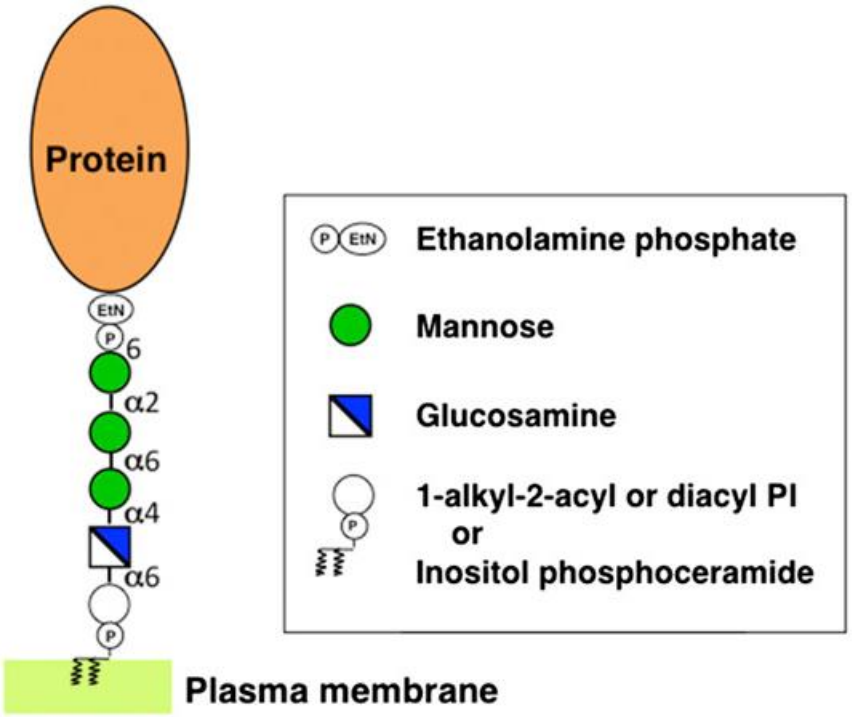


Figure 3. Core structure of the GPI-anchor consisting of an inositol phospholipid, a glycan moiety made up of a non-acetylated glucosamine and three mannose residues as well as the bridging ethanolamine phosphate. Modified from Kinoshita and Fujita, 2016.

GPI anchor biosynthesis in mammals

As in N-glycosylation, the preassembled GPI anchor is attached en-bloc to the protein in the ER and is further processed as the protein migrates through the secretory pathway. Biosynthesis of the GPI-precursor is shown in **Figure 4**. It begins at the cytosolic side of the ER membrane by attachment of GlcNAc to phosphatidyl inositol (PI). The activated sugar nucleotide UDP-GlcNAc serves as the substrate for this reaction, which is catalyzed by GPI-GlcNAc transferase, a heptameric protein complex in mammals and yeast (Murakami et al., 2005). In a subsequent step, GlcNAc is de-acetylated by PIG-L (Watanabe et al., 1999). PI-GlcN is then flipped to the luminal side of the ER membrane by a so far unknown protein. In the next step, the C2 of the inositol-ring is acylated by PIG-W, resulting in GlcN-(acyl)PI. The diacyl moiety of the precursor is then modified, resulting in mostly 1-alkyl-2-acyl PI and a minor fraction of diacyl PI (Kanzawa et al., 2009).

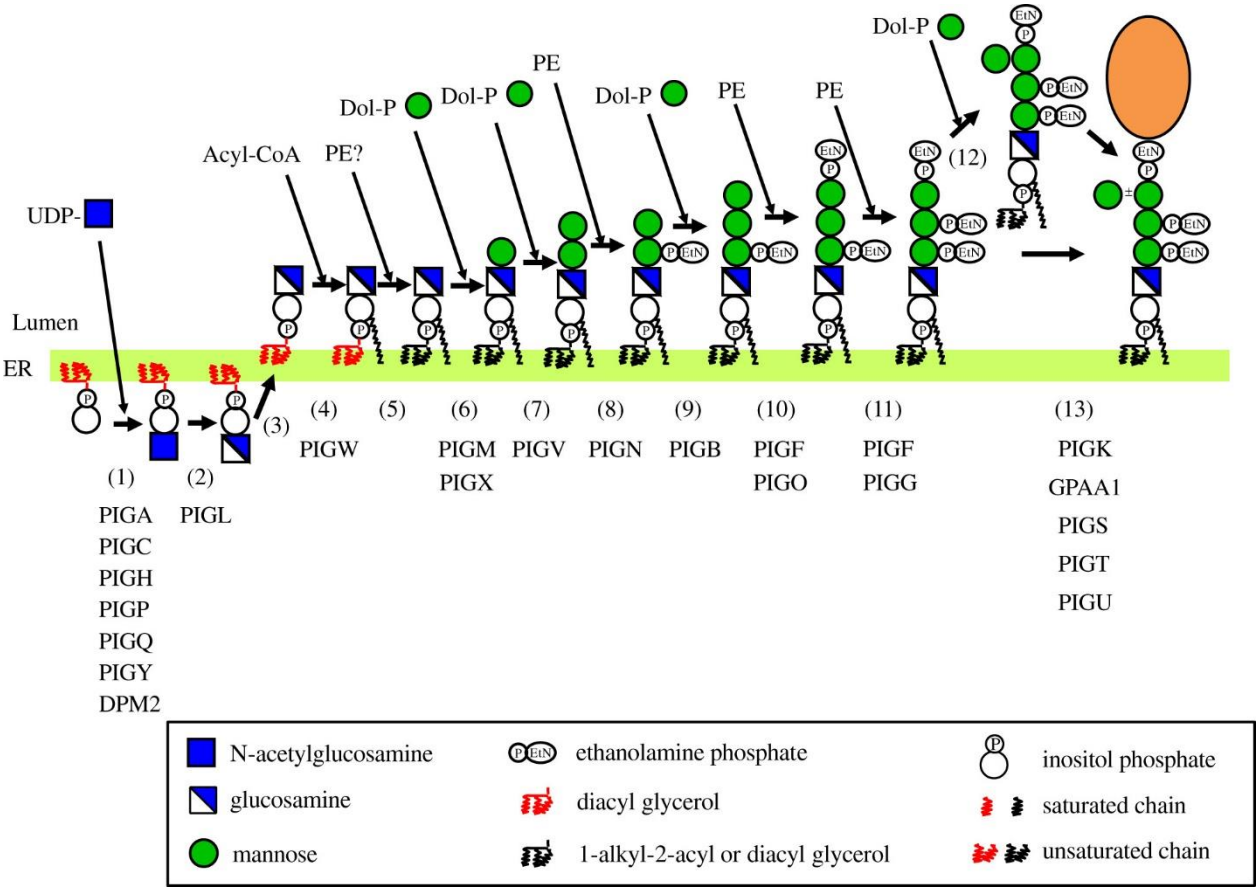


Figure 4. Biosynthesis of the GPI anchor and involved proteins in mammalian organisms. Taken from Kinoshita, 2020.

A dimeric protein complex, GPI mannosyltransferase 1, consisting of PIG-X and PIG-M then catalyzes the transfer of the first mannose from Dolichol-phosphate-mannose (Dol-P-Man) to the growing GPI-precursor in α1,4-linkage. The second mannose is attached in α1,6-linkage by GPI mannosyltransferase 2 (PIG-V). Subsequently, the first mannose is modified with EtNP by GPI EtNP transferase 1 (PIG-N). GPI mannosyltransferase 3, PIG-B, then catalyzes the transfer of the third mannose in α1,2-linkage from Dol-P-Man, followed by the attachment of the so called bridging EtNP by GPI EtNP transferase 2 (PIG-O). The final step in GPI-anchor biosynthesis is the attachment of another EtNP to the C6 of the second mannose, catalyzed by the dimeric GPI EtNP transferase 3, containing PIG-G and PIG-F (Shishioh et al., 2005).
The resulting EtNP-Man-(EtNP)Man-(EtNP)Man-GlcN-(acyl)PI represents the GPI-precursor that is attached to the C-terminus of proteins.
Attachment of the GPI-anchor is catalyzed by GPI-transamidase, a protein complex consisting of five subunits in mammals and yeast (Hong et al., 2003). After translocation of the proprotein into the ER, it is cleaved by PIG-K, a subunit of the GPI transamidase complex, between the ω-site and ω+1, where ω represents the amino acid to which the GPI-anchor is attached (Benghezal et al., 1996). Subsequently, the GPI-anchor is bound to the protein via the bridging EtNP. Proteins that are recognized by GPI transamidase do not contain a conserved peptide sequence, but rather share several common features around the ω site: Amino acids ω-11 to ω-1 (upstream of the ω site) tend to be unstructured and the sequence from ω-1 to ω+2 consists of small amino acids. Amino acids ω+3 to ω+9 are of moderate polarity, followed by a C-terminal stretch of hydrophobic amino acids (Yeats et al., 2018).

Processing of GPI anchors

After transfer of the GPI-anchor to the protein, it undergoes further modifications. In a first step, the previously attached acyl chain is again removed from the inositol via the action of the inositol deacetylase PGAP1. The transient inositol modification of GPI-anchors with acyl is required for action of EtNP transferase 2 (Murakami et al., 2005). Next, the EtNP linked to the second mannose is cleaved by EtNP phosphodiesterase PGAP5, localized at ER exit sites. The removal of both acyl and EtNP is required for interaction with p24 cargo receptor and COPII mediated transport to the Golgi Apparatus (Fujita et al., 2009). Upon arrival in the Golgi Apparatus, the unsaturated fatty acid at the sn2 position is first removed by PGAP3 and subsequently replaced by a saturated fatty acid, a reaction catalyzed by PGAP2 (Kinoshita and Fujita, 2016). Mammalian cells contain mainly stearic acid at the sn2 position. The yeast *Saccharomyces cerevisiae*, on the other hand, undergoes lipid remodeling already in the ER. It contains two different types of lipid moieties, either a diacylglycerol with a C26:0 chain at the sn2 position or a ceramide consisting of either phytosphingosine and a C26:0 fatty acid or phytosphingosine with monohydroxylated C26:0 (Fujita and Jigami, 2008). The lipid moiety of the only plant GPI anchor analyzed thus far from *Pyrus communis* consists of a ceramide containing phytosphingosine and tetracosanoic acid (Oxley and Bacic, 1999). Fatty acid remodeling is required for association of GPI-anchored proteins in lipid rafts (Maeda et al., 2007), likely due to decreased interaction of unsaturated fatty acids with cholesterol and glycosphingolipid, which are enriched in lipid rafts. The spatially close organization of certain proteins in lipid rafts has been suggested to facilitate processes such as signaling via increased binding of ligands to their receptors (Codini et al., 2021). Besides remodeling of the lipid moiety, GPI anchors can be further modified by addition of side chain sugars. Schematic representations of GPI anchors present on proteins from different genera are depicted in **Figure 5**.

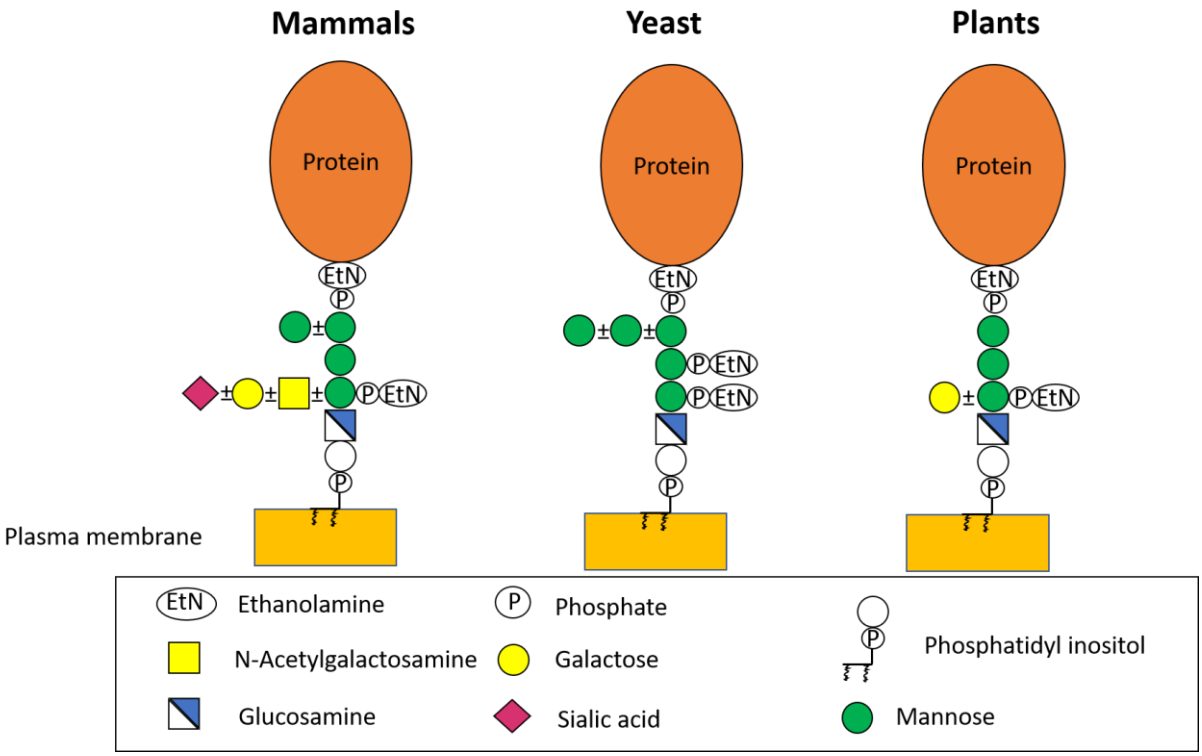


Figure 5. Schematic representation of the structure of mature GPI anchors found on proteins from mammals, yeast and plants.

In mammals, the side chain modifications are tissue dependent. The first mannose can be modified by the addition of up to three sugars, β1,4-GalNAc, β1,3-Gal and sialic acid. Furthermore, another mannose can be attached in α1,2-linkage to the mannose at position 3 (Taron et al., 2004), a modification also detected in yeast (*S. cerevisiae*). Yeasts furthermore attach a fifth mannose to 20-30 % of all GPI anchored proteins (Pittet and Conzelmann, 2007). The GPI anchor from *P. communis* contains a galactose β1,4-linked to the first mannose. Based on the different structures of GPI anchors in other organisms, it has been suggested that plants might also contain various, so far uncharacterized, different side chain modifications (Yeats et al., 2018).

Biological role of GPI anchored proteins

The abundance of GPI anchors on the cell surface as well as the functions fulfilled by GPI anchored proteins are diverse and vary between species. On protozoan parasites such as *Trypanosoma brucei*, GPI anchored proteins constitute a major fraction of all plasma membrane proteins and are essential for proliferation within the host (Nagamune et al., 2000). In *S. cerevisiae*, GPI anchored proteins are not only found linked to the plasma membrane, but also in the glucan network of the cell wall. For association with the cell wall, the lipid moiety is first cleaved, and the protein then covalently linked to glucan residues. Interestingly, this covalent linkage to glucans occurs in many cases only if the protein was first modified with a GPI anchor in the ER, variants lacking the GPI anchor signal are secreted (Lu et al., 1995). The various GPI anchored proteins in yeast are predominantly involved in stretch resistance or reshaping of the cell wall during bud emergence, mating, or cell separation (Pittet and Conzelmann, 2007). GPI anchored proteins in mammals are found in various tissues and are involved in a wide variety of processes including intercellular signaling, fertilization and cell adhesion (Kinoshita, 2020). In humans, deficiency to produce GPI anchored proteins, either due to impaired GPI precursor biosynthesis in the ER or processing of the GPI in the Golgi Apparatus, has been associated with developmental delay, intellectual deficiency, and other neurological abnormalities (Bellai-Dussault et al., 2018). A well-known example of a GPI anchored protein involved in disease is the scrapie isoform of the prion protein, Prp^{Sc}. This GPI anchored protein leads to misfolding of Prp^{C}, a GPI anchored protein mainly present on the surface of cells of the nervous system, and is the causative agent of Creutzfeldt-Jacob disease in humans or bovine spongiform encephalopathy in cattle (Wille and Requena, 2018).

In plants, it is estimated that approximately 1% of all proteins are modified with a GPI-anchor (Yeats et al., 2018). Various reports have been published about predicted GPI anchored proteins in *A. thaliana* (Borner et al., 2003; Elortza et al., 2003; Takahashi et al., 2016; Zhou, 2019), suggesting involvement of GPI anchored proteins in a plethora of functions. Besides playing vital roles in cell wall biosynthesis and anisotropic growth, GPI anchored proteins are furthermore involved in stress response to both biotic and abiotic stress, intercellular signal transduction and sexual reproduction.

Due to the lack of a cytoplasmic domain of GPI anchored proteins, they frequently interact with receptor like kinases (RLKs) for transduction of extracellular signals, such as plant hormones or pathogenesis related factors (Zhou, 2019). Besides their interaction with receptor like kinases for ligand recognition, GPI anchored proteins can also serve as chaperones for transport of interacting proteins to the desired subcellular localization, as described for LORELEI-LIKE-GPI-ANCHORED PROTEIN 1(LLG1), that escorts the receptor kinase FERONIA to the plasma membrane (Li et al., 2015).

Phenotypes of Arabidopsis mutants with impaired GPI anchor biosynthesis have been described: plants containing T-DNA insertions in the genes coding for SETH1 and SETH2, homologues of mammalian PIG-C and PIG-A, displayed reduced pollen germination and pollen tube elongation, combined with increased callose deposition in the pollen (Lalanne et al., 2004). Another line, *pnt1*, containing a mutated homologue of mammalian PIG-M, was reported to be embryo lethal and displayed altered abundance of cell wall constituents in undifferentiated callus (Gillmor et al., 2005).

Reports from yeast and mammals indicate that the attachment of a GPI anchor is a transient modification of certain proteins and that the lipid moiety can be cleaved to liberate them from the plasma membrane during protein maturation. This has also been described for two plant proteins, SKU5 (Sedbrook et al., 2002) and FLA4 (Xue et al., 2017), which were detected both in a soluble as well as in a membrane bound form. The enzyme(s) responsible for this GPI cleavage have not been identified yet in *A. thaliana*, but three putative GPI-specific phospholipase C genes have been proposed (Desnoyer et al., 2020).

Aims of this study

Despite the evolutionary conservation of Lewis A structures throughout the plant kingdom, the role of this N-glycan modification has not been elucidated so far. This is in part due to the low number of identified Lewis A bearing glycoproteins and the lack of an obvious phenotype associated with absence of Lewis A bearing N-glycans in *A. thaliana*. To close this gap in knowledge, one of the aims of this thesis was to identify glycoproteins decorated with Lewis A glycans in different plant species: *A. thaliana* and *O. sativa*, two model organisms of dicot and monocot plants, respectively, as well as *N. benthamiana*, a platform for production of recombinant proteins. To this end, a procedure for specific enrichment of Lewis A bearing glycoproteins had to be developed and optimized. Subsequently, the presence of Lewis A bearing N-glycans on the identified glycoproteins had to be confirmed. Besides identifying proteins modified with Lewis A structures, an Arabidopsis line containing T-DNA insertions in the genes encoding GALT1 and FUT13, the glycosyltransferases required for Lewis A biosynthesis, was characterized to obtain more phenotypic information. In addition, one of the topics addressed in this thesis is whether Lewis A structures are conserved among natural accessions of *A. thaliana*. The large part of what is known nowadays about Lewis A glycans stems from studies of Col-0 plants. This focus on only one line, however, does not reflect the great genomic variety of natural accessions. Therefore, accessions from the 1001 genome database containing mutations in GALT1, the enzyme catalyzing the rate limiting reaction of Lewis A biosynthesis, were investigated for presence of Lewis A.

Besides the role of Lewis A bearing N-glycans, another topic investigated during this PhD thesis was the structure of plant GPI anchors. Thus far, studies on GPI anchors have mainly focused on mammals, yeast and protists. To date, the structure of only one plant GPI-anchor has been elucidated on a protein from *P. communis* cell culture (Oxley and Bacic, 1999). Based on the structural diversity of GPI anchors attached to proteins in mammals, it has been hypothesized that also plants bear different GPI anchors (Yeats et al., 2018). Therefore, a construct bearing monomeric red fluorescent protein (mRFP) N-terminally linked to the GPI anchor signal from Arabidopsis COBRA protein was employed to study the structure of the GPI anchor in both seedlings of *A. thaliana* and leaves of *N. benthamiana*. The structure of the glycan moiety of the PIPC-cleavable GPI anchors was elucidated via mass spectrometry in combination with glycosidase digests.

Publications

Lewis A Glycans Are Present on Proteins Involved in Cell Wall Biosynthesis and Appear Evolutionarily Conserved Among Natural *Arabidopsis thaliana* Accessions

Beihammer G, Maresch D, Altmann F, Van Damme EJM, Strasser R

Original Research Article

Front. Plant Sci. 2021 Mar 11; 12:630891. doi: 10.3389/fpls.2021.630891

frontiers in Plant Science

ORIGINAL RESEARCH
published: 11 March 2021
doi: 10.3389/fpls.2021.630891

Lewis A Glycans Are Present on Proteins Involved in Cell Wall Biosynthesis and Appear Evolutionarily Conserved Among Natural *Arabidopsis thaliana* Accessions

*Gernot Beihammer[1], Daniel Maresch[2], Friedrich Altmann[2], Els J. M. Van Damme[3] and Richard Strasser[1]**

[1] Department of Applied Genetics and Cell Biology, Institute of Plant Biotechnology and Cell Biology, University of Natural Resources and Life Sciences, Vienna, Austria, [2] Division of Biochemistry, Department of Chemistry, University of Natural Resources and Life Sciences, Vienna, Austria, [3] Laboratory of Biochemistry and Glycobiology, Department of Biotechnology, Ghent University, Ghent, Belgium

OPEN ACCESS

Edited by:
Ning Li,
Hong Kong University of Science and Technology, Hong Kong

Reviewed by:
Kazuhito Fujiyama,
Osaka University, Japan
Alexander D. Frey,
Aalto University, Finland

***Correspondence:**
Richard Strasser
richard.strasser@boku.ac.at

Specialty section:
This article was submitted to Plant Proteomics and Protein Structural Biology, a section of the journal Frontiers in Plant Science

Received: 18 November 2020
Accepted: 18 January 2021
Published: 11 March 2021

Citation:
Beihammer G, Maresch D, Altmann F, Van Damme EJM and Strasser R (2021) Lewis A Glycans Are Present on Proteins Involved in Cell Wall Biosynthesis and Appear Evolutionarily Conserved Among Natural Arabidopsis thaliana Accessions. Front. Plant Sci. 12:630891. doi: 10.3389/fpls.2021.630891

N-glycosylation is a highly abundant protein modification present in all domains of life. Terminal sugar residues on complex-type N-glycans mediate various crucial biological processes in mammals such as cell-cell recognition or protein-ligand interactions. In plants, the Lewis A trisaccharide constitutes the only known outer-chain elongation of complex N-glycans. Lewis A containing complex N-glycans appear evolutionary conserved, having been identified in all plant species analyzed so far. Despite their ubiquitous occurrence, the biological function of this complex N-glycan modification is currently unknown. Here, we report the identification of Lewis A bearing glycoproteins from three different plant species: *Arabidopsis thaliana*, *Nicotiana benthamiana*, and *Oryza sativa*. Affinity purification via the JIM84 antibody, directed against Lewis A structures on complex plant N-glycans, was used to enrich Lewis A bearing glycoproteins, which were subsequently identified via nano-LC-MS. Selected identified proteins were recombinantly expressed and the presence of Lewis A confirmed via immunoblotting and site-specific N-glycan analysis. While the proteins identified in *O. sativa* are associated with diverse functions, proteins from *A. thaliana* and *N. benthamiana* are mainly involved in cell wall biosynthesis. However, a Lewis A-deficient mutant line of *A. thaliana* showed no change in abundance of cell wall constituents such as cellulose or lignin. Furthermore, we investigated the presence of Lewis A structures in selected accessions from the 1001 genome database containing amino acid variations in the enzymes required for Lewis A biosynthesis. Besides one relict line showing no detectable levels of Lewis A, the modification was present in all other tested accessions. The data provided here comprises the so far first attempt at identifying Lewis A bearing glycoproteins across different species and will help to shed more light on the role of Lewis A structures in plants.

Keywords: glycosylation, N-glycans, carbohydrate epitope, glycoproteomics, posttranslational modification

INTRODUCTION

N-glycosylation is an abundant protein modification present in all domains of life. N-glycosylation is initiated by the transfer of a preassembled oligosaccharide to certain asparagine residues of nascent polypeptide chains either co-translationally as they enter the lumen of the endoplasmic reticulum (ER), or post-translationally. The attached oligosaccharide is subsequently processed in the ER and Golgi, resulting in a wide variety of structures which can differ greatly between genera (Wang et al., 2017). The initial steps of N-glycan processing are evolutionarily conserved between plants and animals, i.e., cleavage of three glucoses and one mannose in the ER, and further removal of mannoses and attachment of N-acetylglucosamine (GlcNAc) in the *cis*-Golgi (Strasser, 2016). Differentiation of N-glycans between plants and animals starts in the medial-Golgi. Common processing steps include the cleavage of two more mannoses and the attachment of a second GlcNAc residue. Plants additionally add a xylose in β1,2-linkage to the N-glycan and a fucose residue is attached in α1,3-linkage to the core GlcNAc. The overall variety of N-glycan structures in plants is limited compared to animals. While animals produce biantennary or further branched N-glycans with additional sugars such as sialic acid, the only known outer chain elongation of complex N-glycans are Lewis A structures (Fitchette-Lainé et al., 1997). This trisaccharide consists of terminal GlcNAc, to which β1,3-galactose and α1,4-fucose are attached [Fucα1,4(Galβ1,3)GlcNAc-R]. The biosynthesis of Lewis A takes place in the *trans*-Golgi and is a sequential process, in which first the β1,3-galactosyltransferase (GALT1) attaches a galactose residue (Strasser et al., 2007) and afterwards the α1,4-fucosyltransferase (FUT13) attaches a fucose (Wilson, 2001, Léonard et al., 2002). In the model plant *A. thaliana*, Lewis A bearing complex N-glycans show organ-specific expression, being absent in leaves but present in stems and siliques. However, also in these plant organs they constitute only a minor fraction of the total N-glycan content (Strasser et al., 2007) indicating that only few proteins are modified with this trisaccharide. Overall, this modification appears to be evolutionarily conserved, having been identified in mosses such as *Physcomitrella patens* (Koprivova et al., 2003), a wide variety of foodstuff (Wilson et al., 2001, Okada et al., 2017), and on diverse water plants (Maeda et al., 2016). Despite its widespread occurrence, the biological role of Lewis A structures has not been identified. Lewis A-deficient plants of *A. thaliana* lack an obvious phenotype (Strasser et al., 2007; Strasser, 2016) and, so far, the knowledge about proteins bearing Lewis A is rather scarce. Two publications investigating the glycoproteome of Arabidopsis inflorescence stems (Zhang et al., 2011; Zeng et al., 2018) have identified in total four proteins decorated with Lewis A structures. Yet a study dedicated to the isolation of Lewis A bearing glycoproteins has not been published. Here, we describe the identification of glycoproteins modified with Lewis A structures from three different plant species and show that Lewis A structures represent an evolutionarily conserved modification in natural accessions of *A. thaliana*.

MATERIALS AND METHODS

Plant Material and Growth Conditions

Seeds of ecotypes and mutants of *A. thaliana* were sterilized in 4% (v/v) NaOCl solution prior to being sown to half-strength Murashige-Skoog (MS) plates containing 1% (v/w) sucrose and 0.8% (w/v) agar. Plants were grown under long-day conditions (16 h light/ 8 h dark) at 22°C and transferred to soil 7 days after sowing. The *galt1 fut13* double knockout mutant was obtained by crossing previously described lines *galt1* (N871760) and *fut13* (N567444) (Strasser et al., 2007). Except Col-0, all described ecotypes were ordered from the Nottingham Arabidopsis Stock Centre (NASC). Wild-type and ΔXT/FT *N. benthamiana* plants (Strasser et al., 2008) were grown at 24°C at 16 h light/8 h dark for 5 weeks. Shoots of *Oryza sativa japonica* cv Nipponbare were grown for 10 days as described (Lambin et al., 2020).

Isolation of Lewis A Bearing Glycoproteins

For isolation of Lewis A bearing glycoproteins, leaves and stems of 5-week-old *N. benthamiana*, stems and siliques of 6- and 8-week-old *A. thaliana* Col-0 plants, respectively, and shoots of 10-day-old *O. sativa* plants were used. Plant material was snap-frozen, grinded, and total proteins extracted using Radio-Immunoprecipitation assay (RIPA) buffer [50 mM Tris pH 8, 150 mM NaCl, 1% (v/v) Nonidet P-40, 0.5% (w/v) sodium deoxycholate, 0.1% (v/v) sodium dodecyl sulfate (SDS)]. To isolate Lewis A bearing glycoproteins, we used the Lewis A-specific JIM84 antibody (Horsley et al., 1993; Fitchette et al., 1999). To capture the IgM antibody, biotinylated anti-rat IgM antibody (Sigma-Aldrich) was incubated with streptavidin-agarose beads (Sigma-Aldrich) in the presence of RIPA dilution buffer (50 mM Tris pH 7.5, 150 mM NaCl, 5 mM EDTA). After binding of anti-rat IgM antibody to the resin, it was washed and incubated with JIM84 antibody. The resin with the two bound antibodies was then mixed with the total protein extracts. After incubation, the resin was washed three times with RIPA dilution buffer and transferred to a Micro Bio-Spin column (Bio-Rad). Bound proteins were subjected to an on-bead digest (Chrestensen et al., 2004) with slight modifications. Briefly, 10 mM dithiothreitol (DTT) in 100 mM NH_4HCO_3 buffer was pipetted to the agarose beads and incubated at 60°C for 15 min. 18 mM iodoacetamide in 100 mM NH_4HCO_3-buffer was added for alkylation and incubated for 30 min in the dark. Afterwards, trypsin was added (Sequencing Grade modified trypsin, Promega) to a final concentration of 3.5 ng/μL, incubated overnight at 37°C and peptide fragments were isolated on the next day via centrifugation. The supernatant was subjected to a final clean-up step using a C18 Hypersep cartridge (Thermo Fisher Scientific). The cartridge was washed three times with 65% (v/v) acetonitrile in 80 mM ammonium formate buffer pH 3, equilibrated in the same buffer and loaded with the eluted peptides in 100 mM NH_4HCO_3 buffer pH 7.8. After washing three times, peptides were eluted with 65% acetonitrile in pH 3 buffer and vacuum dried.

Identification of Lewis A Bearing Glycoproteins by Mass Spectrometry

Analysis of isolated Lewis A bearing glycoproteins was conducted using nano-LC-ESI-MS as previously described (Zámocký et al., 2020). In short, the peptide fragments were analyzed using a Dionex Ultimate 3000 system interfaced to a maXis 4G ETD QTOF system (Bruker). A Thermo Acclaim PepMap precolumn was coupled to a Thermo Acclaim PepMap 300 RSLC C18 separation column (2 μm particle size, 150 x 0.075 mm) for separation of peptides, running a gradient of 95% solvent A to 32% solvent B over the course of 60 min (solvent A: 0.1% formic acid in HQ-water, solvent B: 0.1% formic acid in acetonitrile), followed by a gradient from 32 to 70% solvent B for 10 min at 0.3 μL/min flow rate. CaptiveSpray nano Booster was used as ion source in positive ion mode. MS spectra were recorded in a range of 150–2200 m/z, and the six highest peaks selected for fragmentation in data dependent acquisition mode. For identification of proteins, the analysis files were converted to XML using Data Analysis 4.0 (Bruker) to allow MS/MS ion searches using MASCOT (embedded in Protein Scape 3.0, Bruker). For protein searches from *N. benthamiana*, a recently published database was used (Schiavinato et al., 2019). For Arabidopsis, the UniProt non-redundant *A. thaliana* database and for rice the non-redundant UniProt *O. sativa japonica* database were used. Proteins identified by at least two peptides and a protein score higher than 80 were accepted. For comparing the MS/MS data to the target sequence via X! Tandem (https://thegpm.org/TANDEM/), the following settings were used: reversed sequences no; check parent ions for charges 1, 2, and 3 yes; models found with peptide log e lower −1 and proteins log e lower −1; residue modifications: oxidation M, W, and deamidation N, Q; isotope error was considered; fragment type was set to monoisotopic; refinement was used with standard parameters; fragment mass error of 0.1 Da and ±7 ppm parent mass error; fragment types b and y ions; maximum parent ion charge of three; missed cleavage sites allowed was set to two; semi-cleavage yes.

Vector Construction

For expression of identified proteins in *A. thaliana* as well as in leaves of *N. benthamiana*, a modified version of the binary vector pPT2 (Strasser et al., 2005) was used. For expression of KORRIGAN in leaves of *N. benthamiana*, p29-Fc-KOR1 was used (Liebminger et al., 2013). For expression of COBL4, the coding sequence without endogenous signal peptide (bases 1-60) was amplified from cDNA of *A. thaliana* using primers At5g15630_1F (5′-TATATCTAGATATGATCCATT AGATCCTAGTGGTA-3′) and At5g15630_2R (5′-TATAA GATCTTCACCATATTGAGATGAATAGGAGA-3′), digested with XbaI/BglII and ligated into XbaI/BamHI-digested p117 (Shin et al., 2018). g6145 was amplified without the sequence encoding the signal peptide (bases 1–72) from cDNA of *N. benthamiana* using primers g6145_F1 (5′-TATATCTAGAGG AGATCCATTTAAGTTTTTTAACTT-3′) and g6145_R2(5′-TATAGGATCCCTAATAGAACACAGAAAAGATTGCA-3′), digested with XbaI/BamHI and ligated into XbaI/BamHI-digested p117. For expression of CEBIP, a codon-optimized version for *N. benthamiana* lacking the endogenous signal peptide coding sequence (bases 1–84) was ordered from GeneArt (Thermo Fisher Scientific). The DNA sequence was amplified using primers STRINGS_9F (5′-CTTCCGGCTCG TTTGTCTAGA-3′) and STRINGS_2R (5′-AAAAACCCTGGC GGGATCC-3′), digested with XbaI/BamHI and ligated into XbaI/BamHI-digested p117.

Recombinant Expression and Purification of Glycoproteins

For recombinant expression of identified proteins in *N. benthamiana*, syringe-mediated leaf infiltration of 5-week-old plants was used. *Agrobacterium tumefaciens* strain UIA143 (Strasser et al., 2005) carrying the respective plasmid for expression was grown over night at 29°C and adjusted to an OD_{600} of 0.15 on the next day in infiltration buffer (28 mM glucose, 50 mM 2-(4-morpholino)-ethanesulphonic acid (MES), 2 mM $Na_3PO_4.12H_2O$, 0.1 mM acetosyringone). p29-Fc-KOR1 was additionally co-infiltrated with silencing suppressor p19 (Garabagi et al., 2012) at an OD_{600} of 0.1. Leaf material was collected 2 days post infiltration (dpi) and snap-frozen. For stable expression in *A. thaliana,* plants were transformed using the floral dip method (Clough and Bent, 1998). Stems of homozygous plants were collected ∼5 weeks after sowing. Plant material was disrupted mechanically, and total proteins extracted using RIPA buffer. Recombinantly produced glycoproteins were purified via their fused tags, either GFP for KORRIGAN (GFP-Trap agarose, Chromotek) or mRFP for the other constructs (RFP-Trap agarose, Chromotek). After capture, proteins were eluted in 1.5x Laemmli buffer.

Deglycosylation and Immunoblotting

Purified proteins from *N. benthamiana* and Arabidopsis Col-0 plants were subjected to deglycosylation with Endoglycosidase H (Endo H, New England Biolabs) according to the manufacturer's instructions. For deglycosylation experiments with Peptide-N-glycosidase F (PNGaseF, New England Biolabs), proteins purified from ΔXT/FT-*N. benthamiana* (Strasser et al., 2008) and *A. thaliana fut11 fut12* (Strasser et al., 2004) were used. For immunoblotting, purified glycoproteins were separated on a 10% (v/v) polyacrylamide gel and transferred to a nitrocellulose membrane. Primary antibodies against GFP and RFP (both purchased from Chromotek) were diluted (1:2000) in PBS + 0.1% (v/v) Tween (PBST) + 1,5% (w/v) BSA, as secondary antibody anti-mouse IgG (Sigma-Aldrich) diluted (1:10.000) in PBST + 1% BSA was used. JIM84 was diluted (1:400) in PBST and anti-rat IgM antibody (Sigma-Aldrich) was used as secondary antibody. Detection via ECL substrate (Super Signal West PICO Plus Chemiluminescent substrate, Thermo Fisher Scientific) was monitored on a Fusion instrument (Vilber).

Site Specific N-glycan Analysis via Mass Spectrometry

For site-specific N-glycan analysis, purified proteins from wild-type *N. benthamiana* and *A. thaliana* Col-0 were

separated on a 10% polyacrylamide gel via SDS-PAGE, stained with Coomassie Brilliant Blue and excised from the gel. An in-gel digest was conducted as previously described (Kolarich and Altmann, 2000) with the exception, that a double-digest of trypsin and chymotrypsin was used for proteolysis.

Analysis of Cell Wall Constituents

Cell wall constituents were quantified as previously described (Corneillie et al., 2019) with slight modifications. Stems from fully grown Arabidopsis plants were collected after drying, the bottom-most 2 cm removed, and the next 15 cm used for analysis. Alcohol soluble residues were removed by sequential extraction in H_2O (30 min, 98°C), ethanol (30 min, 76°C), chloroform (30 min, 59°C), and acetone (30 min, 54°C). Remaining cell wall residues were weighed, and percentage of full stem weight calculated. For quantification of cellulose, half of the cell wall residue fraction was incubated in trifluoroacetic acid (TFA) for 3 h with shaking in-between to remove hemicellulose. After washing with acetone, the remaining pellet was dried and weighed. Updegraff-reagent (Updegraff, 1969) was added to the pellet and heated for 30 min at 100°C, washed once with H_2O and three times with acetone. After air drying overnight, the remaining crystalline cellulose was dissolved in 72% (v/v) sulphuric acid, incubated, H_2O added and centrifuged. The supernatant was diluted 1:10, mixed with anthrone reagent in sulphuric acid in a 96 well plate, heated to 80°C for 30 min and subsequently absorption was measured at 595 nm using a plate reader (Tecan, Infinite 200 Pro). Cellulose content in the samples was calculated based on a glucose–standard curve and percentage of cellulose per cell wall residues calculated. For determination of lignin, the second half of the cell wall residue pellet was mixed with 25% (v/v) acetyl-bromide solution in acetic acid, heated for 3 h at 50°C and centrifuged. The supernatant was diluted 1:20 in a solution containing 0.4 M NaOH and 18.75 mM hydroxylamine hydrochloride in acetic acid and absorption measured at 280 nm (Nanodrop, Thermo Fisher Scientific). Lignin content was then calculated as percentage of cell wall residues.

Root Growth Assays

Measurement of root growth dynamics and root growth under stress conditions was conducted as described (Dubiel et al., 2020). In short, seeds of *A. thaliana* Col-0 and mutant plants were sterilized, stratified for 3 days, and grown on half-strength MS medium. For analysis of root growth dynamics, root length was measured after 48, 96, 144, 192, and 240 h. For analysis of root growth under stress conditions, seedlings were transferred to either new plates (mock control), plates containing 150 mM NaCl (salt stress) or plates containing 12% (w/v) PEG6000 corresponding to a water potential of −0.5 MPa for simulation of drought stress via the PEG infusion method (Van Der Weele et al., 2000). After 10 days, plates were scanned, number of lateral root hairs counted, and root length measured using Image J (Schindelin et al., 2012).

RESULTS

Identification of Lewis A Bearing Glycoproteins

In this study, we set out to identify glycoproteins harboring the Lewis A trisaccharide on complex N-glycans in various organs of selected plants. To this end, we analyzed total protein extracts from siliques and stems of the dicot model plant *A. thaliana*, shoots of the monocot *O. sativa*, and leaves and stems from *N. benthamiana*, which is frequently used as a production platform for recombinant glycoproteins (Stoger et al., 2014). We isolated Lewis A bearing glycoproteins via the JIM84 antibody (Fitchette et al., 1999), conducted a tryptic digest and subsequently analyzed the resulting peptide fragments via peptide mapping using nano-LC-ESI-MS. To discriminate false positives from actual Lewis A bearing glycoproteins we included a negative control, for which we repeated the same isolation procedure in the absence of the JIM84 antibody. **Table 1** lists glycoproteins from *A. thaliana* and *O. sativa* that were identified at least twice in three replicates and were not detected in the negative control. We identified 10 proteins from *A. thaliana*, four of which we detected in both stems and siliques, namely Laccase 4, COBRA-like protein 4 (COBL4), Endoglucanase 25 (also known as KORRIGAN) and Protein trichome-birefringence like 3. All these proteins have numerous potential N-glycosylation sites and a predicted localization in the secretory pathway. Interestingly, many of the identified proteins have a proposed function in secondary cell wall biosynthesis. In addition, we detected Laccase 17, Probable inactive purple-acid-phosphatase and Protein trichome-birefringence-like 33 in siliques as well as O-fucosyltransferase family protein SUB1 in stems of *A. thaliana*. The list obtained from *O. sativa* comprises proteins of various functions. We detected two defense related proteins, Chitin elicitor-binding protein (CEBIP) and Germin-like protein 2–4. Furthermore, Dirigent protein, which is involved in biosynthesis of secondary metabolites, Os02g0615800, a receptor kinase, Beta-Ig-H3 domain-containing protein, a cell surface adhesion protein, and Os08g0503200, a glycerophosphodiester phosphodiesterase. In *N. benthamiana* (**Supplementary Table 1**) we detected homologs of Endoglucanase 25 as well as a homolog of a COBRA-like extracellular GPI-anchored protein. In addition, we detected homologs of glucosyl hydrolases as well as proteins involved in directional growth and cell expansion such as a mono-copper oxidase-like SKU5 protein (g6145).

The Identified Proteins Are Modified With Lewis A Structures When Recombinantly Expressed

We were not able to detect Lewis A bearing glycopeptides on the isolated peptide fractions from the three plant species. Thus, to confirm the presence of Lewis A bearing N-glycans on the identified proteins, we cloned one selected protein from each species into a binary expression vector and expressed it recombinantly as a fusion with a fluorescent protein in leaves of *N. benthamiana*, where, in contrast to *A. thaliana*, Lewis A structures are present (**Supplementary Figure 1**). After

TABLE 1 | Lewis A bearing glycoproteins from stems and siliques of *A. thaliana* Col-0 and shoots of *O. sativa japonica*.

UniProt Accession	Name	Potential N-glycosylation sites	Predicted Localization	Proposed Function
***A. thaliana* stems**				
O80434	Laccase 4 (IRX12)	14	Secreted	Lignin biosynthesis
Q9LFW3	COBRA-like protein 4 (COBL4)	8	Secreted (GPI-anchored)	Cellulose biosynthesis
Q38890	Endoglucanase 25 (KORRIGAN)	8	Plasma membrane	Cellulose biosynthesis
Q8LED3	Protein trichome birefringence-like 3 (TBL3)	6	Golgi apparatus	Deposition of cellulose
Q9LE45	O-fucosyltransferase family protein (SUB1)	11	Golgi apparatus	Transfer of glycosyl-residues
***A. thaliana* siliques**				
Q38890	Endoglucanase 25 (KORRIGAN)	8	Plasma membrane	Cellulose biosynthesis
O80434	Laccase 4 (IRX12)	14	Secreted	Lignin biosynthesis
Q9FJD5	Laccase 17 (LAC17)	15	Secreted	Lignin biosynthesis
Q9LFW3	COBRA-like protein 4 (COBL4)	8	Secreted (GPI-anchored)	Cellulose biosynthesis
Q9LMG7	Probable inactive purple acid phosphatase 2 (PAP2)	8	Secreted	Protein of unknown function
Q8LED3	Protein trichome birefringence-like 3 (TBL3)	6	Golgi apparatus	Secondary cell wall biosynthesis
F4IH21	Protein trichome birefringence-like 33 (TBL33)	6	Golgi apparatus	Secondary cell wall biosynthesis
***O. sativa* shoots**				
Q306J3	Dirigent protein (JAC1)	5	Secreted	Biosynthesis of lignans, flavonolignans, and alkaloids
Q8H8C7	Chitin elicitor-binding protein (CEBIP)	11	Plasma membrane	Chitin binding
Q6K7X0	Os02g0615800 protein	21	Plasma membrane	Receptor kinase
Q6F3A5	Beta-Ig-H3 domain-containing protein	2	Secreted	Cell surface adhesion protein
Q6ZFH9	Glycerophosphodiester Phospodiesterase	12	Secreted	Lipid metabolism
Q6ESF0	Germin-like protein 2-4	2	Secreted	Plant defense

Proteins were affinity-purified using the JIM84 antibody and identified via mass spectrometry using peptide mapping. The number of potential N-glycosylation sites is based on predictions using the NetNGlyc 1.0 Server (http://www.cbs.dtu.dk/services/NetNGlyc/). The localization and proposed function are based on information from the UniProt-database (https://www.uniprot.org/) and predictions using TargetP (http://www.cbs.dtu.dk/services/TargetP/).

purification, we confirmed the presence of Lewis A both via immunoblotting using the JIM84 antibody and via mass spectrometry using site-specific N-glycan analysis. **Figure 1** shows the results of immunoblotting for KORRIGAN from *A. thaliana*, g6145 from *N. benthamiana* and CEBIP from *O. sativa*. A shift is observed upon PNGase F treatment but not upon Endo H treatment, indicating that the majority of N-glycans on these recombinantly expressed glycoproteins are of the complex-type. Furthermore, all proteins show a signal on the JIM84-blot, which is not visible anymore in the PNGase F treated sample indicating that the Lewis A structure is present on N-glycans. Using a combination of trypsin and chymotrypsin for proteolysis, we could detect three peptides with N-glycosylation sites carrying Lewis A on both KORRIGAN and g6145 as well as one site on CEBIP (**Figure 2**). On all these N-glycosylation sites, Lewis A bearing N-glycans only constitute a minor fraction compared to other structures. We were, however, not able to detect all potential glycopeptides **(Supplementary Figure 2)**, and there might be additional N-glycosylation sites which also carry Lewis A structures.

Next, we compared the N-glycan microheterogeneity of an identified protein expressed either transiently in leaves of *N. benthamiana* or stably in stems of *A. thaliana*. The overall N-glycosylation pattern of COBL4 is very similar between the two plant species. For instance, sites carrying oligomannosidic N-glycans in *N. benthamiana* also carried the same type of N-glycans in *A. thaliana* (**Supplementary Figures 3**, **4**). Lewis A structures were more abundant on glycopeptides carrying N-glycosylation sites N154 and N162 in *N. benthamiana* compared to *A. thaliana*, where Lewis A glycans were hardly detectable. On the other hand, the glycopeptide with site N306 carried a complex N-glycan with Lewis A structures when expressed in *A. thaliana*, which were not found on COBL4 glycopeptides derived from *N. benthamiana*.

Lewis A-deficient *A. thaliana* Show Normal Stem Morphology and Root Development

Most proteins identified in both stems and siliques of *A. thaliana* are implicated in biosynthesis of the secondary plant cell wall. We thus hypothesized that plants lacking Lewis A structures on their

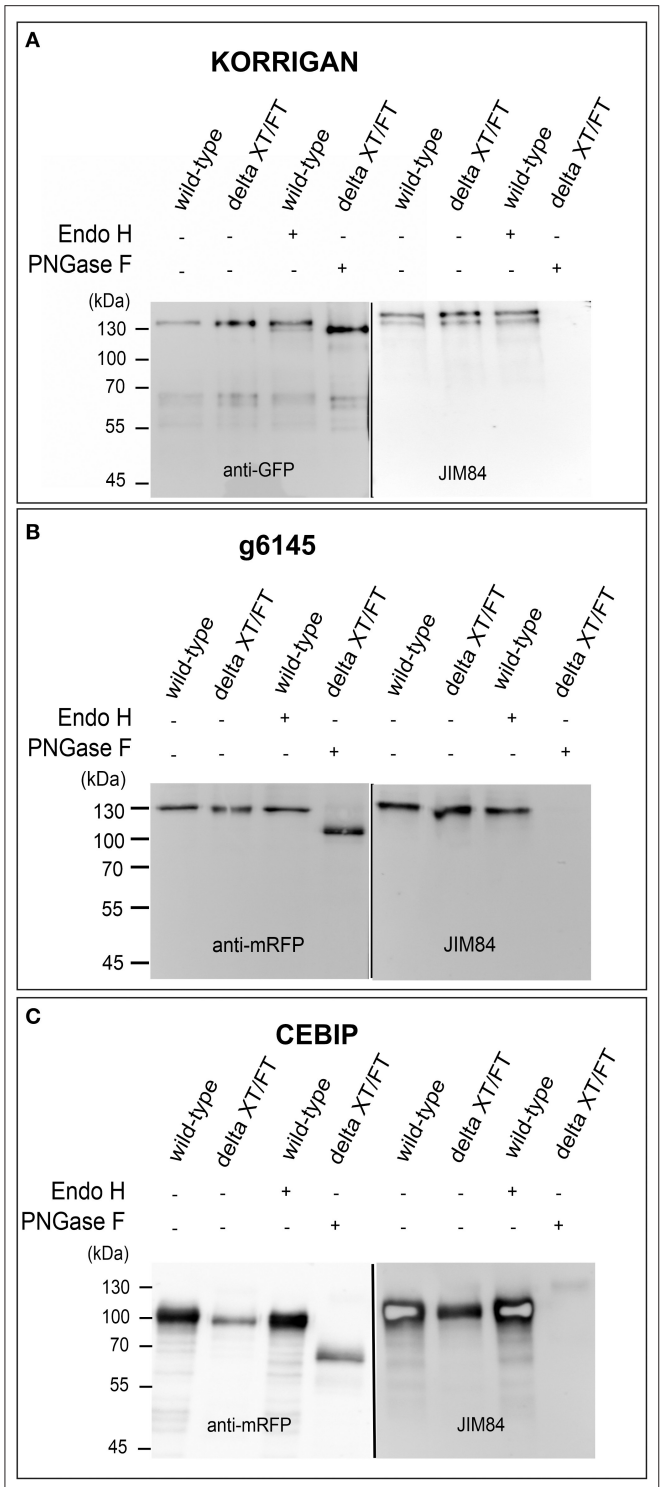


FIGURE 1 | Lewis A bearing complex N-glycans are present on recombinantly produced plant proteins. KORRIGAN **(A)**, g6145 **(B)**, and CEBIP **(C)** were transiently expressed in leaves of wild-type and ΔXT/FT plants of *N. benthamiana* and purified via their fused tag. Expression of proteins was confirmed via immunoblotting using an antibody directed against the fused tag (GFP for KORRIGAN and mRFP for g6145 and CEBIP). Presence of Lewis A was shown using the JIM84 antibody.

complex N-glycans might show alterations in stem development or morphology. To test this hypothesis, we made use of a line lacking Lewis A glycans due to T-DNA insertions in genes coding for GALT1 and FUT13 (**Supplementary Figure 5**). Stems of the *galt1 fut13* double mutant developed normally, exhibiting no differences in flowering time, stem length or mass of dried stems when compared to Col-0 wild-type plants (**Figure 3**). As a control, we included *rsw2-1* containing a missense mutation in the gene coding for KORRIGAN (Lane et al., 2001), resulting in a dwarfed phenotype and reduced cellulose content. Additionally, we analyzed the content and composition of fully grown dried stems. We analyzed cell wall residues (CWR) gravimetrically after sequential extraction of alcohol soluble residues. Lignin- and cellulose-levels within the fraction of alcohol insoluble residues were measured photometrically using the acetyl bromide method and the anthrone-assay, respectively. We could not find significant differences between *galt1 fut13* and Col-0, indicating that plants develop normal stem morphology in the absence of Lewis A structures.

A frequently observed phenotype in *A. thaliana* lines with mutations along the N-glycosylation pathway are defects in root growth or enhanced susceptibility toward abiotic challenge such as salt stress (Kang et al., 2008; Liebminger et al., 2009). To see whether *galt1 fut13* plants showed alterations in root growth we determined root growth dynamics by measuring root length at 2, 4, 6, 8, and 10 days after sowing (**Supplementary Figure 6**). The *galt1 fut13* plants showed no difference compared to Col-0. In addition, we analyzed the plants susceptibility to salt and drought stress (**Supplementary Figure 6**). Again, we could not detect any difference in the *galt1 fut13* line compared to wild-type, indicating normal root development of Lewis A-deficient Arabidopsis plants under abiotic stress conditions.

Functional *GALT1* and *FUT13* Genes Are Present in Most Natural Accessions of *A. thaliana*

Despite the lack of an obvious phenotype in Lewis A-deficient *A. thaliana* plants, taking into account the strong evolutionary conservation of the carbohydrate epitope we reasoned that the ability to synthesize Lewis A structures might still be an advantage for plants and play a yet unknown role for plant fitness. To examine whether there are Arabidopsis accessions containing SNPs, deletions or insertions in the genes coding for FUT13 and GALT1 we searched the 1001 genomes database using the POLYMORPH 1001 tool (https://tools.1001genomes.org/polymorph/). We found only one SNP with a predicted high impact, an alteration in the nucleotide sequence leading to the formation of a premature stop codon in the *FUT13* gene. This SNP is present in 55 accessions and results in the expression of a truncated FUT13 protein where the two C-terminal amino acids ("GV") are missing. Since there is some amino acid sequence variation at the C-terminus of FUT13 from different species we consider it unlikely that the truncation completely abolishes FUT13 function. In addition, polymorphisms were detected in the *FUT13* gene which might have an impact on FUT13

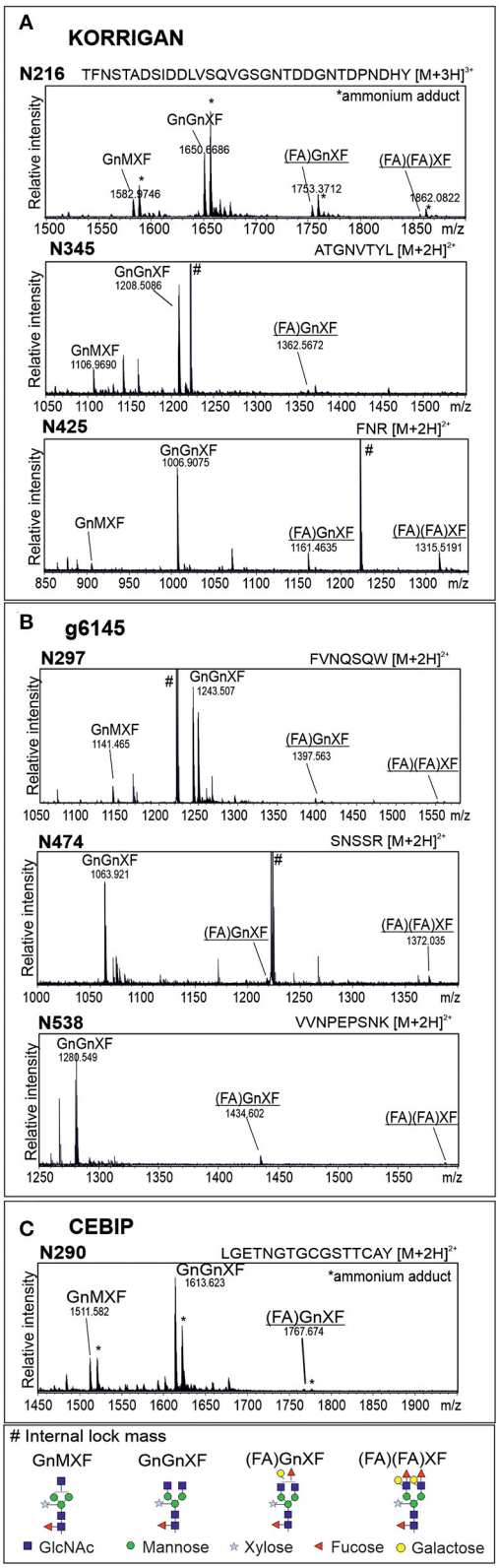


FIGURE 2 | Glycopeptides from recombinantly produced plant proteins display complex N-glycans with Lewis A structures. MS spectra of

(Continued)

FIGURE 2 | glycopeptides from KORRIGAN **(A)**, g6145 **(B)**, and CEBIP **(C)**. Proteins were recombinantly expressed in leaves of wild-type *N. benthamiana*, purified via the fused tag and subjected to proteolytic digest using trypsin and chymotrypsin. Nomenclature of N-glycans is according to the ProGlycAn system (http://www.proglycan.com/). Only one N-glycan isoform is indicated. Lewis A bearing glycans are underlined.

activity. As GALT1 catalyzes the rate limiting reaction of Lewis A biosynthesis, we decided to investigate polymorphisms in the *GALT1* gene in more detail. We found only one accession, IP-Orb-10, containing a premature stop codon instead of the codon encoding the last amino acid W643. **Figure 4** shows the amino acid sequence of GALT1 and highlights amino acids for which we found polymorphisms in the database. Polymorphisms leading to amino acid changes are found in the putative lectin domain and in the GALT1 catalytic domain. We selected different accessions containing amino acid changes in various regions of GALT1 (**Supplementary Table 2**) and analyzed the presence of Lewis A in total protein extracts from stems and siliques by immunoblotting (**Figures 4B,C; Supplementary Figure 7**). While most lines including IP-Orb-10 showed similar Lewis A levels as Col-0, some accessions such as Kulturen-1, MNF-Che-2, and MNF-Pin-39 displayed a slightly stronger signal with JIM84, especially in siliques. One line, PYL-6, showed only very low levels of Lewis A in both stems and siliques, and IP-Vim-0 appeared devoid of detectable Lewis A structures in siliques and displayed only a very faint signal in stems. To confirm the absence of Lewis A in IP-Vim-0 we repeated the JIM84 blot and compared the signal to extracts from Lewis A-deficient plants *galt1-fut13* and *galt1* (**Supplementary Figure 8**). We observed no JIM84 signal in IP-Vim-0 stems, indicating the absence of Lewis A structures in this line. In PYL-6 and IP-Vim-0 we found that GALT1 and FUT13 transcripts are expressed and we confirmed the polymorphisms in the region coding for the GALT1 catalytic domain in the IP-Vim-0 cDNA by sequencing (**Supplementary Figure 9**). In contrast to the result for the JIM84 antibody, both accessions show a signal with an antibody directed against N-glycans with β1,2-xylose and core α1,3-fucose (**Figures 4B,C**) indicating normal processing of oligomannosidic to complex N-glycans. Furthermore, we tested various ecotypes of *A. thaliana* and were able to detect Lewis A structures in all of them (**Supplementary Figure 10**). In summary, the data suggest that SNPs occurring in conserved regions of GALT1 are rare exceptions in natural accessions of *A. thaliana*.

DISCUSSION

The Lewis A trisaccharide on complex plant N-glycans is an evolutionarily conserved modification throughout the plant kingdom (Fitchette et al., 1999; Wilson et al., 2001; Léonard et al., 2004; Strasser et al., 2007). Here, we report the isolation of 10 proteins from *A. thaliana*, seven proteins from *O. sativa* and nine proteins from *N. benthamiana* modified with Lewis A. The relatively low number of identified Lewis A bearing glycoproteins can have two possible explanations: (1) either only a low number of specific glycoproteins are modified with

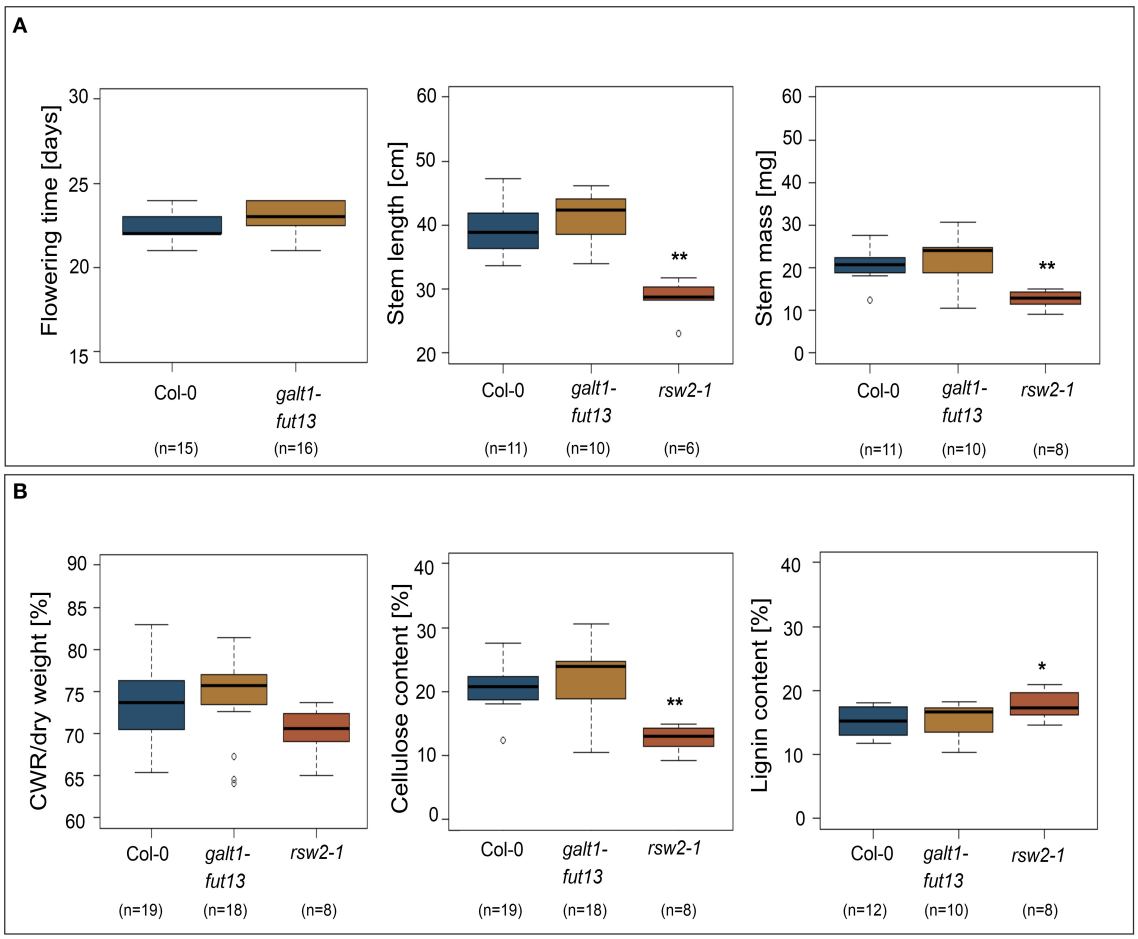


FIGURE 3 | Phenotypic comparison of wild-type and Lewis A-deficient (*galt1 fut13*) plants. *rsw2-1* plants used as controls. The bottom and top of the boxes represent 25 and 75 percent quartiles, respectively, while the horizontal line inside the box indicates the median. The whiskers indicate 1,5x interquartile range (IQR). Asterisks indicate statistically significant differences to Col-0 plants (*$p \leq 0.05$, **$p \leq 0.01$). **(A)** Comparison of flowering time, length of the plants primary stem, and mass of dried primary stems. **(B)** Content of the primary stems. Cell wall residues (CWR) were determined gravimetrically after sequential removal of alcohol soluble residues. Cellulose and lignin content were determined photometrically as percentage of CWR.

Lewis A structures or (2) Lewis A structures are present on various glycoproteins in the *trans*-Golgi, the plasma membrane and the apoplast, but occur in so small quantities that we could identify only the most abundant glycoproteins or the ones carrying complex N-glycans with high amounts of Lewis A, respectively. Based on these possibilities we cannot rule out that we only enriched a certain population of Lewis A bearing glycoproteins using the described purification approach. Overall, the information about Lewis A bearing glycoproteins published in literature is rather scarce. In a recent study investigating N-glycan microheterogeneity in, among others, inflorescence stems of *A. thaliana* (Zeng et al., 2018), the authors report the detection of biantennary complex N-glycans with Lewis A structures on three glycoproteins, among them Laccase 4 (At2g38080.1), which was also identified in our approach in both Arabidopsis stems and siliques. Furthermore, an uncharacterized potentially GPI-anchored protein (At3g06035) was detected. This protein was also found once in one of our experiments but was not included in our list which contains only proteins detected at least twice in three experiments. The third protein reported by the authors, DUF2921 (At1g52780), was not identified in our experiments. In another publication (Zhang et al., 2011), the authors were able to detect monoantennary Lewis A on one protein, homologous to blue copper binding protein (At4g12880). No data has so far been published about Lewis A-bearing glycans on endogenous proteins from *N. benthamiana*. However, Lewis A was reported to be present on recombinantly produced glycoproteins such as human erythropoietin (Castilho et al., 2013). In *O. sativa*, which carries considerable amounts of N-glycans with Lewis A modifications (Léonard et al., 2004), monoantennary Lewis A was detected on nucleotide pyrophosphatase/phosphodiesterases when overexpressed in rice cell culture (Kaneko et al., 2016). These proteins were not identified in our proteomics approach. While the proteins identified in rice show no clear pattern with regard to their proposed functions, the proteins identified in *A. thaliana*, and to a lesser extent also in *N. benthamiana*, show a clear bias toward cell wall biosynthesis. However, the Lewis A-deficient *galt1 fut13* line showed no differences in

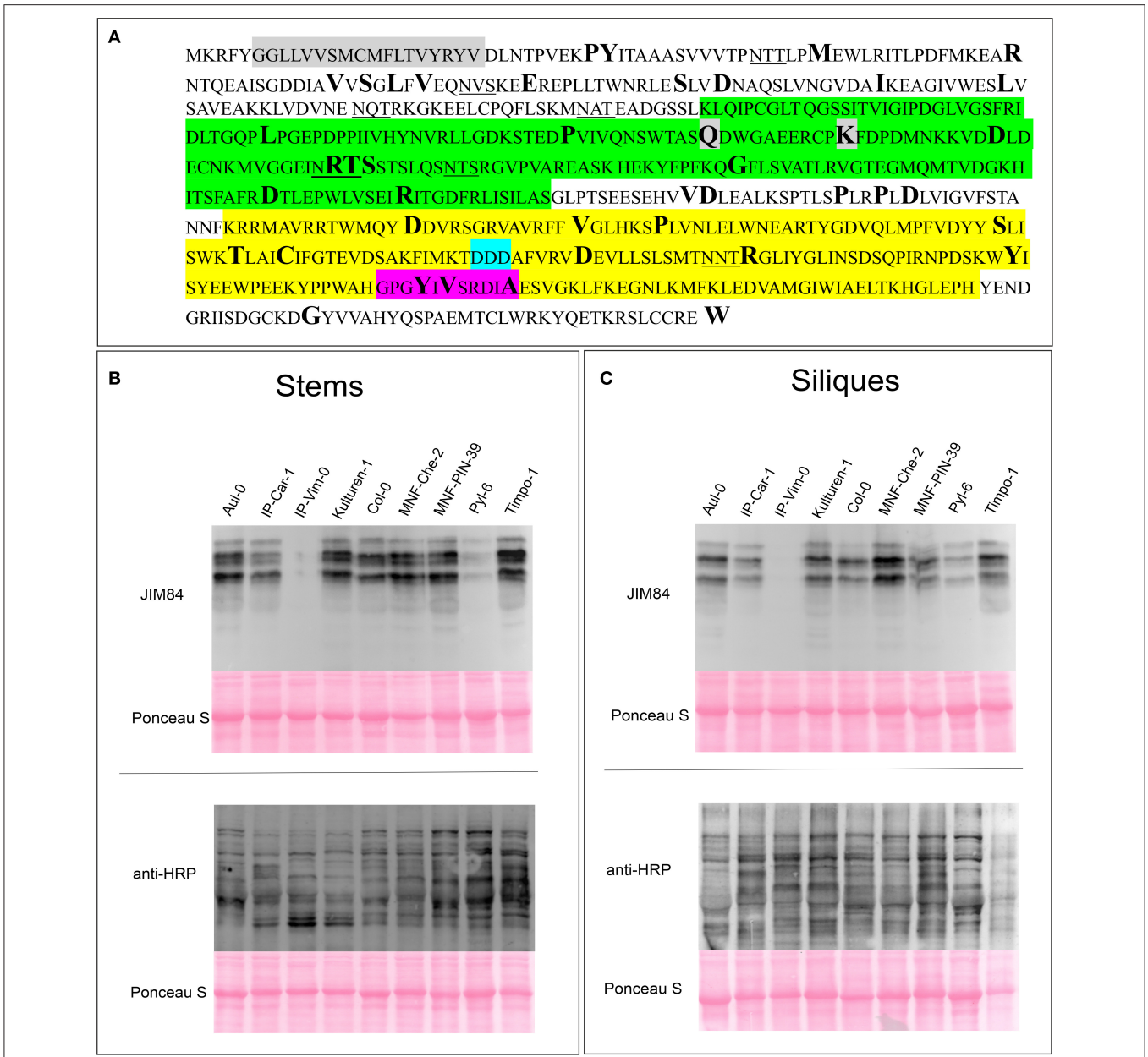


FIGURE 4 | Detection of Lewis A structures in Arabidopsis accessions carrying *GALT1* polymorphisms. **(A)** Protein sequence of GALT1 from Arabidopsis Col-0. The transmembrane domain is shown in gray, the putative galactoside-binding lectin domain in green, and the galactosyltransferase domain in yellow. Additionally, the DXD-motif in the catalytic site is highlighted in light blue and the conserved GxxYxxS as well as the adjacent DxA motif are highlighted in magenta. Underlined amino acids indicate N-glycosylation sequons. Sites depicted in bold show positions of amino acids, for which we could find variations in the 1001 genome database. **(B,C)** Immunoblot analysis of the indicated Arabidopsis accessions. Total protein extracts from stems and siliques, respectively, were analyzed with antibodies directed against Lewis A bearing (JIM84) and plant N-glycans carrying β1,2-xylose and core α1,3-fucose (anti-HRP).

respect to mass and length of the primary stem, flowering time or constituents of the secondary cell wall, such as cell wall residues, cellulose, and lignin. Furthermore, Lewis A-deficient plants showed normal root development, also under abiotic stress. This is in agreement with previous reports (Strasser et al., 2007; Rips et al., 2014; Basu et al., 2015) which described the lack of an observable phenotype. Furthermore, Basu and colleagues showed that GALT1 is the only member of the CAZy GT31-family not involved in hydroxyproline O-galactosylation of arabinogalactan-proteins (AGPs). While T-DNA insertion lines of other GALTs showed abnormalities of root hairs, seed coat mucilage, silique morphology or pollen tube development, no such phenotype was reported for GALT1-deficient plants (Basu et al., 2015). Interestingly though, FUT13 and GALT1

appear to be strongly conserved among the natural accessions found in the 1001 genome project. Out of 16 selected lines, only one line, IP-Vim-0, lacks detectable levels of Lewis A containing N-glycans in extracts from stems and siliques. This line has two polymorphisms in the *FUT13* gene that alter the FUT13 protein (T294K and 399 instead of 401 amino acids) and may also affect the FUT13 activity. While the truncated FUT13 sequence is found in 55 different accessions, the T294K polymorphism is only found in one other line IP-Vis-0. In the *GALT1* gene from IP-Vim-0 four polymorphisms are present that lead to changes of the amino acid sequence. Two of the changes (Q243L and K253N) are also found in Lewis A-positive lines such as MNF-Pin-39 and therefore do not inactivate GALT1. The other two polymorphisms are present in a conserved sequence motif that is present in the catalytic domain of CAZy GT31 family members (Qu et al., 2008; Petit et al., 2020). The conserved GxxYxxS motif (also called motif IV in the GT31 family members) constitutes the C-terminal part of a flexible G-loop. The loop is involved in donor substrate binding and is close to the active site of the enzyme. A DxA motif that is highly conserved in plant GT31 sequences (Qu et al., 2008) is directly following the GxxYxxS motif. While in Col-0 this sequence stretch is **G**PG**Y**IV**S**R**D**I**A,** it is **G**PG**N**IV**S**R**D**I**E** in IP-Vim-0. Both mutations very likely abolish the enzymatic activity of GALT1 and lead to the absence of Lewis A bearing N-glycans in IP-Vim-0. This line has been defined as a relict, meaning high genomic divergence from geographically proximal accessions (Alonso-Blanco et al., 2016) and may represent a rare exception of a plant line with an impaired pathway for the biosynthesis of Lewis A structures on complex N-glycans. Although we cannot rule out that other amino acid substitutions in GALT1 or FUT13 have a similar effect, it seems that the Lewis A trisaccharide structure is still highly conserved in Arabidopsis accessions occurring in different geographical regions. Together with the presence of the modification in all so far analyzed plant species these data hint at the existence of an evolutionary selection pressure to maintain the Lewis A glycan epitope. Despite many efforts, the role of this complex N-glycan modification remains a mystery in plants.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

GB, EV, and RS designed the experiments. GB and DM conducted the experiments. GB, DM, EV, FA, and RS analyzed the results. GB and RS wrote the paper. All authors have made a substantial and intellectual contribution to the work and approved it for publication.

FUNDING

This work was supported by the Austrian Science Fund (FWF) [Doctoral Program BioToP–Biomolecular Technology of Proteins (W1224)] and by the FWF Project I2417-B22.

ACKNOWLEDGMENTS

We thank Lukas Mach (Department of Applied Genetics and Cell Biology, University of Natural Resources and Life Sciences Vienna) for supplying the JIM84 antibody and helpful discussions as well as Malgorzata Dubiel (Department of Molecular Biotechnology, Ghent University) for help with the root length assays.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.630891/full#supplementary-material

REFERENCES

Alonso-Blanco, C., Andrade, J., Becker, C., Bemm, F., Bergelson, J., Borgwardt, K. M., et al. (2016). 1,135 genomes reveal the global pattern of polymorphism in *Arabidopsis thaliana*. *Cell* 166, 481–491. doi: 10.1016/j.cell.2016.05.063

Basu, D., Tian, L., Wang, W., Bobbs, S., Herock, H., Travers, A., et al. (2015). A small multigene hydroxyproline-O-galactosyltransferase family functions in arabinogalactan-protein glycosylation, growth and development in arabidopsis. *BMC Plant Biol.* 15:1–23. doi: 10.1186/s12870-015-0670-7

Castilho, A., Neumann, L., Gattinger, P., Strasser, R., Vorauer-Uhl, K., Sterovsky, K., et al. (2013). Generation of biologically active multi-sialylated recombinant human EPOFc in plants. *PLoS ONE* 8:e54836. doi: 10.1371/journal.pone.0054836

Chrestensen, C. A., Schroeder, J. M., Shabanowitz, J., Hunt, D. F., Pelo, J. W., Worthington, M. T., et al. (2004). MAPKAP kinase 2 phosphorylates tristetraprolin on *in vivo* sites including Ser178, a site required for 14-3-3 binding. *J. Biol. Chem.* 279, 10176–10184. doi: 10.1074/jbc.M310486200

Clough, S. J., and Bent, A.F. (1998). Floral dip: a simplified method for agrobacterium-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16, 735–743. doi: 10.1046/j.1365-313x.1998.00343.x

Corneillie, S., De Storme, N., Van Acker, R., Fangel, J. U., De Bruyne, M., De Rycke, R., et al. (2019). Polyploidy affects plant growth and alters cell wall composition. *Plant Physiol.* 179, 74–87. doi: 10.1104/pp.18.00967

Dubiel, M., Beeckman, T., Smagghe, G., and Van Damme, E. J. M. (2020). Arabidopsis lectin EULS3 is involved in ABA signaling in roots. *Front. Plant Sci.* 11:437. doi: 10.3389/fpls.2020.00437

Fitchette, A. C., Cabanes-Macheteau, M., Marvin, L., Martin, B., Satiat-Jeunemaitre, B., Gomord, V., et al. (1999). Biosynthesis and immunolocalization of lewis A-containing N-glycans in the plant cell. *Plant Physiol.* 121, 333–343. doi: 10.1104/pp.121.2.333

Fitchette-Lainé, A. C., Gomord, V., Cabanes, M., Michalski, J. C., Saint Macary, M., Foucher, B., et al. (1997). N-glycans harboring the lewis A epitope are expressed at the surface of plant cells. *Plant J.* 12, 1411–1417. doi: 10.1046/j.1365-313x.1997.12061411.x

Garabagi, F., Gilbert, E., Loos, A., Mclean, M. D., and Hall, J. C. (2012). Utility of the P19 suppressor of gene-silencing protein for production of therapeutic

antibodies in nicotiana expression hosts. *Plant Biotechnol. J.* 10, 1118–1128. doi: 10.1111/j.1467-7652.2012.00742.x

Horsley, D., Coleman, J., Evans, D., Crooks, K., Peart, J., Satiat-Jeunemaitre, B., et al. (1993). A monoclonal antibody, JIM 84, recognizes the golgi apparatus and plasma membrane in plant cells. *J. Exp. Bot.* 44, 223–229.

Kaneko, K., Takeshi, T., Inomata, T., Oikawa, K., Itoh, K., Hirose, K., et al. (2016). N-glycomic and microscopic subcellular localization analyses of NPP1, 2, and 6 strongly indicate that trans-golgi compartments participate in the golgi to plastid traffic of nucleotide pyrophosphatase/phosphodiesterases in rice. *Plant Cell Physiol.* 57, 1610–1628. doi: 10.1093/pcp/pcw089

Kang, J. S., Frank, J., Kang, C. H., Kajiura, H., Vikram, M., Ueda, A., et al. (2008). Salt tolerance of *Arabidopsis thaliana* requires maturation of N-glycosylated proteins in the golgi apparatus. *Proc. Natl. Acad. Sci. U.S.A.* 105:7893. doi: 10.1073/pnas.0800237105

Kolarich, D., and Altmann, F. (2000). N-glycan analysis by matrix-assisted laser desorption/ionization mass spectrometry of electrophoretically separated nonmammalian proteins: application to peanut allergen Ara h 1 and olive pollen allergen Ole e 1. *Anal. Biochem.* 285, 64–75. doi: 10.1006/abio.2000.4737

Koprivova, A., Altmann, F., Gorr, G., Kopriva, S., Reski, R., and Decker, E. L. (2003). N-glycosylation in the moss *Physcomitrella* patens is organized similarly to that in higher plants. *Plant Biol.* 5, 582–591. doi: 10.1055/s-2003-44721

Lambin, J., Demirel Asci, S., Dubiel, M., Tsaneva, M., Verbeke, I., Wytynck, P., et al. (2020). OsEUL lectin gene expression in rice: stress regulation, subcellular localization, and tissue specificity. *Front. Plant Sci.* 11:185. doi: 10.3389/fpls.2020.00185

Lane, D. R., Wiedemeier, A., Peng, L., Höfte, H., Vernhettes, S., Desprez, T., et al. (2001). Temperature-sensitive alleles of Rsw2 link the KORRIGAN endo-1,4-β-glucanase to cellulose synthesis and cytokinesis in *Arabidopsis*. *Plant Physiol.* 126, 278–288. doi: 10.1104/pp.126.1.278

Léonard, R., Costa, G., Darrambide, E., Lhernould, S., Fleurat-Lessard, P., Carlué, M., et al. (2002). The presence of lewis A epitopes in *Arabidopsis thaliana* glycoconjugates depends on an active A4-fucosyltransferase gene. *Glycobiology* 12, 299–306. doi: 10.1093/glycob/12.5.299

Léonard, R., Kolarich, D., Paschinger, K., Altmann, F., and Wilson, I. B. (2004). A genetic and structural analysis of the N-glycosylation capabilities. *Plant Mol Biol.* 55, 631–644. doi: 10.1007/s11103-004-1558-3

Liebminger, E., Grass, J., Altmann, F., Mach, L., and Strasser, R. (2013). Characterizing the link between glycosylation state and enzymatic activity of the endo-β1,4-glucanase KORRIGAN1 from *Arabidopsis thaliana*. *J. Biol. Chem.* 288, 22270–22280. doi: 10.1074/jbc.M113.475558

Liebminger, E., Hüttner, S., Vavra, U., Fischl, R., Schoberer, J., Grass, J., et al. (2009). Class I α-mannosidases are required for N-glycan processing and root development in *Arabidopsis thaliana*. *Plant Cell* 21, 3850–3867. doi: 10.1105/tpc.109.072363

Maeda, M., Misato, T., Takeo, Y., Vavricka, C. J., and Kimura, Y. (2016). Structural features of N-glycans linked to glycoproteins expressed in three kinds of water plants: predominant occurrence of the plant complex type N-glycans bearing lewis A epitope. *Carbohydr. Res.* 435, 50–57. doi: 10.1016/j.carres.2016.09.008

Okada, T., Ihara, H., Ito, R., and Ikeda, Y. (2017). Molecular cloning and functional expression of lewis type A1,3/A1,4-fucosyltransferase CDNAs from *Mangifera indica* L. *Phytochemistry* 144, 98–105. doi: 10.1016/j.phytochem.2017.08.021

Petit, D., Teppa, R. E., and Harduin-Lepers, A. (2020). A phylogenetic view and functional annotation of the animal B1,3-glycosyltransferases of the GT31 CAZy family. *Glycobiology* cwaa086. doi: 10.1093/glycob/cwaa086

Qu, Y., Egelund, J., Gilson, P. R., Houghton, F., Gleeson, P. A., Schultz, C. J., et al. (2008). Identification of a novel group of putative *Arabidopsis thaliana* β-(1,3)-galactosyltransferases. *Plant Mol. Biol.* 68, 43–59. doi: 10.1007/s11103-008-9351-3

Rips, S., Bentley, N., Jeong, I. S., Welch, J. L., von Schaewen, A., and Koiwa, H. (2014). Multiple N-glycans cooperate in the subcellular targeting and functioning of *Arabidopsis* KORRIGAN1. *Plant Cell* 26, 3792–3808. doi: 10.1105/tpc.114.129718

Schiavinato, M., Strasser, R., Mach, L., Dohm, J. C., and Himmelbauer, H. (2019). Genome and transcriptome characterization of the glycoengineered nicotiana benthamiana line ΔXT/FT. *BMC Genomics* 20:594. doi: 10.1186/s12864-019-5960-2

Schindelin, J., Arganda-Carrera, I., Frise, E., Verena, K., Mark, L., Tobias, P., et al. (2012). Fiji–an open platform for biological image analysis. *Nat. Methods* 9:676. doi: 10.1038/nmeth.2019

Shin, Y. J., Vavra, U., Veit, C., and Strasser, R. (2018). The glycan-dependent ERAD machinery degrades topologically diverse misfolded proteins. *Plant J.* 94, 246–259. doi: 10.1111/tpj.13851

Stoger, E., Fischer, R., Moloney, M., and Ma, J. K. C. (2014). Plant molecular pharming for the treatment of chronic and infectious diseases. *Annu. Rev. Plant Biol.* 65, 743–768. doi: 10.1146/annurev-arplant-050213-035850

Strasser, R. (2016). Plant protein glycosylation. *Glycobiology* 26, 926–939. doi: 10.1093/glycob/cww023

Strasser, R., Altmann, F., Mach, L., Glössl, J., and Steinkellner, H. (2004). Generation of *Arabidopsis thaliana* plants with complex N-glycans lacking B1,2-linked xylose and core B1,3-linked fucose. *FEBS Lett.* 561, 132–136. doi: 10.1016/S0014-5793(04)00150-4

Strasser, R., Bondili, J. S., Vavra, U., Schoberer, J., Svoboda, B., Glössl, J., et al. (2007). A unique B1,3-galactosyltransferase is indispensable for the biosynthesis of N-glycans containing lewis A structures in *Arabidopsis thaliana*. *Plant Cell* 19, 2278–2292. doi: 10.1105/tpc.107.052985

Strasser, R., Stadlmann, J., Sch?hs, M., Stiegler, G., Quendler, H., Mach, L., et al. (2008). Generation of glyco-engineered Nicotiana benthamiana for the production of monoclonalantibodies with a homogeneous human-like N-glycan structure. *Plant Biotechnol. J.* 6, 392–402. doi: 10.1111/j.1467-7652.2008.00330.x

Strasser, R., Stadlmann, J., Svoboda, B., Altmann, F., Glössl, J., and Mach, L. (2005). Molecular basis of N-acetylglucosaminyltransferase I deficiency in *Arabidopsis thaliana* plants lacking complex N-glycans. *Biochem. J.* 387, 385–391. doi: 10.1042/BJ20041686

Updegraff, D. M. (1969). Semimicro determination of cellulose in biological materials. *Anal. Biochem.* 32, 420–424. doi: 10.1016/S0003-2697(69)80009-6

Van Der Weele, C. M., Spollen, W. G., Sharp, R. E., and Baskin, T. I. (2000). Growth of *Arabidopsis thaliana* seedlings under water deficit studied by control of water potential in nutrient-agar media. *J. Exp. Bot.* 51, 1555–1562. doi: 10.1093/jexbot/51.350.1555

Wang, T., Hu, X. C., Cai, Z. P., Voglmeir, J., and Liu, L. (2017). Qualitative and quantitative analysis of carbohydrate modification on glycoproteins from seeds of *Ginkgo biloba*. *J. Agric. Food Chem.* 65, 7669–7679. doi: 10.1021/acs.jafc.7b01690

Wilson, I. B. H. (2001). Identification of a CDNA encoding a plant lewis-type A1,4-fucosyltransferase. *Glycoconj. J.* 18, 439–447. doi: 10.1023/A:1016030000527

Wilson, I. B. H., Zeleny, R., Kolarich, D., Staudacher, E., Stroop, C. J. M., Kamerling, J. P., et al. (2001). Analysis of Asn-linked glycans from vegetable foodstuffs: widespread occurrence of lewis A, core A1,3-linked fucose, and xylose substitutions. *Glycobiology* 11, 261–274. doi: 10.1093/glycob/11.4.261

Zámocký, M., Kamlárová, A., Maresch, D., Chovanová, K., Harichová, J., and Furtmüller, P. G. (2020). Hybrid heme peroxidases from rice blast fungus magnaporthe oryzae involved in defence against oxidative stress. *Antioxidants* 9, 1–20. doi: 10.3390/antiox9080655

Zeng, W., Ford, K. L., Bacic, A., and Heazlewood, J. L. (2018). N-linked glycan micro-heterogeneity in glycoproteins of *Arabidopsis*. *Mol. Cell. Proteomics* 17, 413–421. doi: 10.1074/mcp.RA117.000165

Zhang, Y., Giboulot, A., Zivy, M., Valot, B., Jamet, E., and Albenne, C. (2011). Combining various strategies to increase the coverage of the plant cell wall glycoproteome. *Phytochemistry* 72, 1109–1123. doi: 10.1016/j.phytochem.2010.10.019

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers

Supplementary Material

Lewis A glycans are present on proteins involved in cell wall biosynthesis and appear evolutionarily conserved among natural *Arabidopsis thaliana* accessions

Gernot Beihammer, Daniel Maresch, Friedrich Altmann, Els J.M. Van Damme, Richard Strasser

Contents

Table S1. List of proteins identified in leaves and stems of *N. benthamiana*. Proteins were affinity-purified using the JIM84 antibody and identified via mass spectrometry using peptide mapping. The number of potential N-glycosylation sites is based on predictions using the NetNGlyc 1.0 Server (http://www.cbs.dtu.dk/services/NetNGlyc/). The localization and proposed function are based on information from the UniProt-database (https://www.uniprot.org/) and predictions using TargetP (http://www.cbs.dtu.dk/services/TargetP/).

N. benthamiana stems				
Identifier	**Potential N-glycosylation sites**	**Predicted localization**	**Proposed function**	**Homologue of (protein from *A. thaliana*)**
g17123.t1	9	Plasma membrane	Cellulose biosynthesis	Endoglucanase 25
g6145.t1	14	Secreted (GPI-anchored)	Involved in directional growth processes	Mono-copper oxidase-like SKU5
g37143.t1	11	Golgi apparatus	Involved in coaction of cryptochrome and phytochrome	Calcium ion binding protein
g59335.t1	10	Plasma membrane	Cellulose biosynthesis	Endoglucanase 25
g99754.t1	9	Plasma membrane	Hydrolysis of glucosidic linkages	Glucan endo-1,3-beta-glucosidase 1
g54518.t1	10	Plasma membrane	Hydrolysis of glucosidic linkages	Glucan endo-1,3-beta-glucosidase 1
g23587.t1	9	Plasma membrane	Hydrolysis of glucosidic linkages	O-Glycosyl hydrolase family 17 protein
g8142.t1	9	Secreted (GPI-anchored)	Direction of cell expansion	COBRA-like extracellular GPI-anchored protein family
N. benthamiana leaves				
Identifier	**Potential N-glycosylation sites**	**Predicted localization**	**Proposed function**	**Homologue of (protein from *A. thaliana*)**
g6145.t1	14	Secreted (GPI-anchored)	Involved in directional growth processes	Mono-copper oxidase-like SKU5
g17123.t1	9	Plasma membrane	Cellulose biosynthesis	Endoglucanase 25
g23587.t1	9	Plasma membrane	Hydrolysis of glucosidic linkages	O-Glycosyl hydrolase family 17 protein
g54518.t1	10	Plasma membrane	Hydrolysis of glucosidic linkages	Glucan endo-1,3-beta-glucosidase 1
g21104.t1	7	Secreted	Hydrolysis of xylose from xyloglucan	Alpha-xylosidase 1

Table S2. Accessions of *A. thaliana* from the 1001 genome project for which Lewis A levels were monitored in stems and siliques. The *GALT1* (At1g26810) polymorphisms relative to the reference genome (https://tools.1001genomes.org/polymorph/) as well as the geographic location of their collection are listed. Only non-silent mutations are included in the list.

Accession ID	Name	Polymorphisms	Resulting amino acid variation	Country
265	PYL-6	833 G>A	R278Q	FRA
1925	MNF-Che-2	1256 A>C	D419A	USA
2016	MNF-Pin-39	728 A>T	Q243L	USA
		759 A>C	K253N	
		1492 G>A	D498N	
4857	UKSW06-257	728 A>T	Q243L	UK
		759 A>C	K253N	
		1421 G>A	C474Y	
6012	Eden-7	759 A>C	K253N	SWE
		1843 G>C	G612R	
7107	Durh-1	728 A>T	Q243L	UK
		759 A>C	K253N	
		1421 G>A	C474Y	
7520	Lp2-2	1052 G>A	R351Q	CZE
		759 A>C	K253N	
8240	Kulturen-1	619 C>T	L207P	SWE
		620 T>C		
		728 A>T	Q243L	
		759 A>C	K253N	
8249	Vimmerby	728 A>T	Q243L	SWE
		759 A>C	K253N	
		1390 A>T	S464C	
9530	IP-Car-1	355 A>T	I119F	ESP
		728 A>T	Q243L	
		759 A>C	K253N	
		1126 G>T	V376F	
9565	IP-Orb-10	1256 A>C	D419A	ESP
		1928 G>A	W643stop	
9598	IP-Vim-0	728 A>T	Q243L	ESP
		759 A>C	K253N	
		1657 T>A	Y553N	
		1679 C>A	A560E	
9647	Basen-1	728 A>T	Q243L	ITA
		759 A>C	K253N	
		1162 C>T	P388S	
9701	Ivano-1	759 A>C	K253N	BUL
		1171 C>T	P391S	
9822	Aul-0	1256 A>C	D419A	ESP
9968	Timpo-1	239 C>T	S80L	ITA

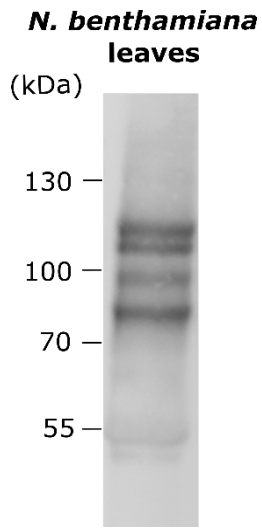


Figure S1. Immunoblot analysis of total protein extracts from leaves of *N. benthamiana*. Proteins were extracted using RIPA buffer, separated via SDS-PAGE and Lewis A bearing glycoproteins detected using the JIM84 antibody.

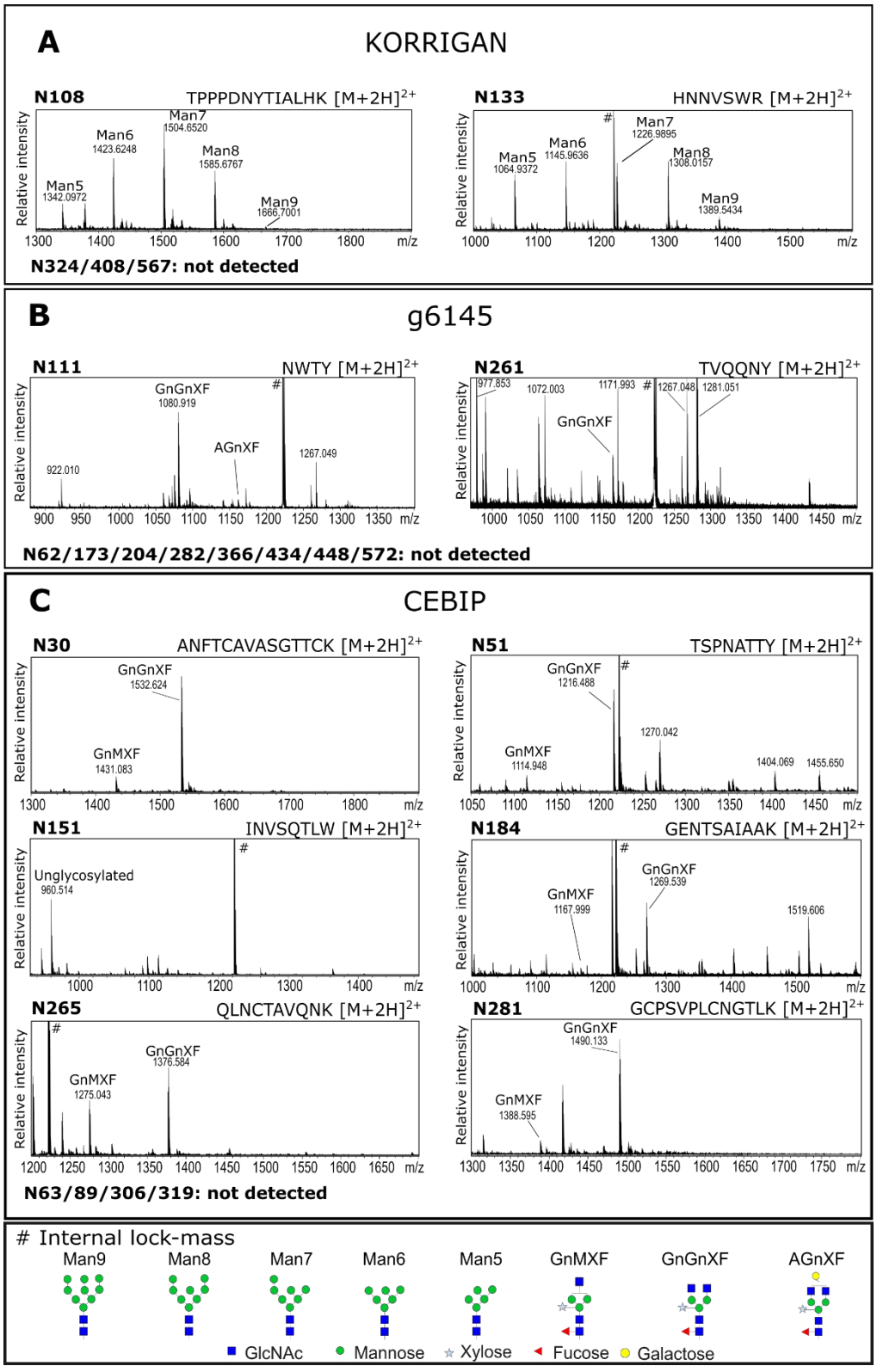


Figure S2. Site-specific N-glycan analysis of KORRIGAN, g6145 and CEBIP expressed in leaves of *N. benthamiana*. The spectra show sites where Lewis A structures were not detected. Proteins were

purified via their fused tag, separated via SDS-PAGE, excised from the gel and subjected to trypsin/chymotrypsin digestion and MS-analysis.

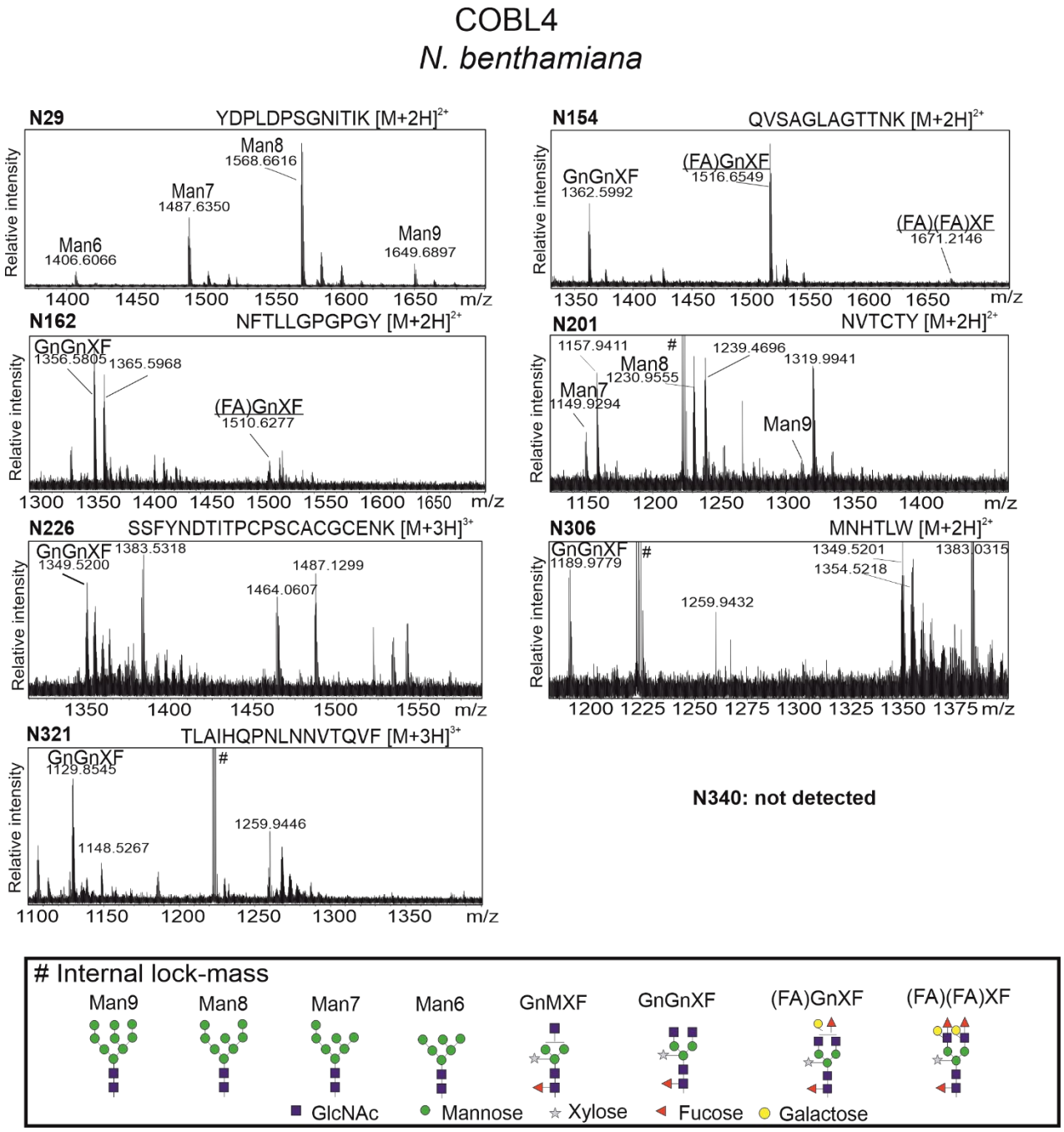


Figure S3. Site-specific N-glycan analysis of COBL4 purified from leaves of infiltrated *N. benthamiana*. Proteins were purified, separated via SDS-PAGE, excised from the gel and subjected to a trypsin/chymotrypsin digestion and MS-analysis. N-glycans decorated with Lewis A structures are underlined.

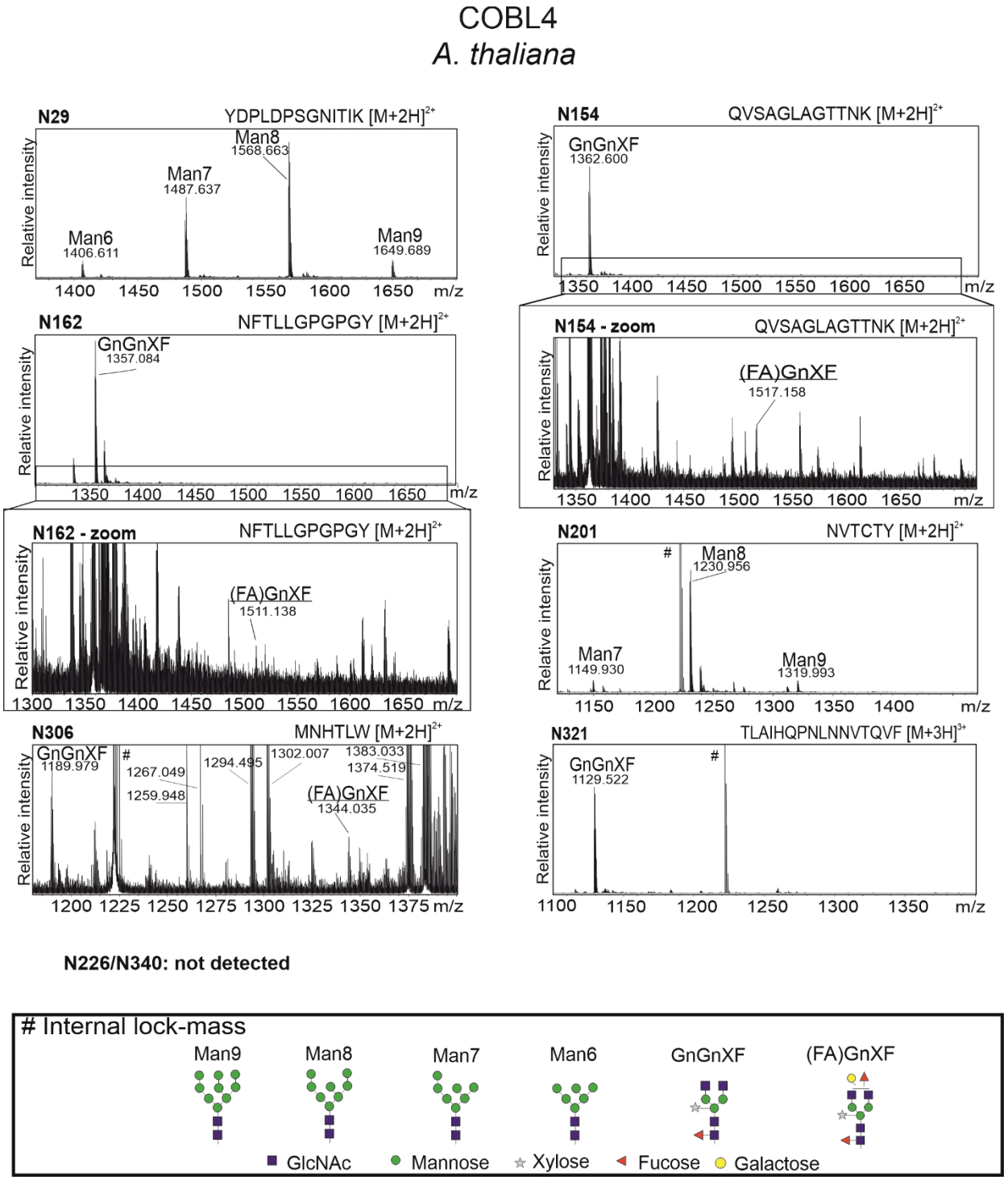


Figure S4. Site-specific N-glycan analysis of COBL4 purified from stems of stably expressing Arabidopsis Col-0 plants. Proteins were purified, separated via SDS-PAGE, excised from the gel and

subjected to a trypsin/chymotrypsin digestion and MS-analysis. N-glycans decorated with Lewis A structures are underlined.

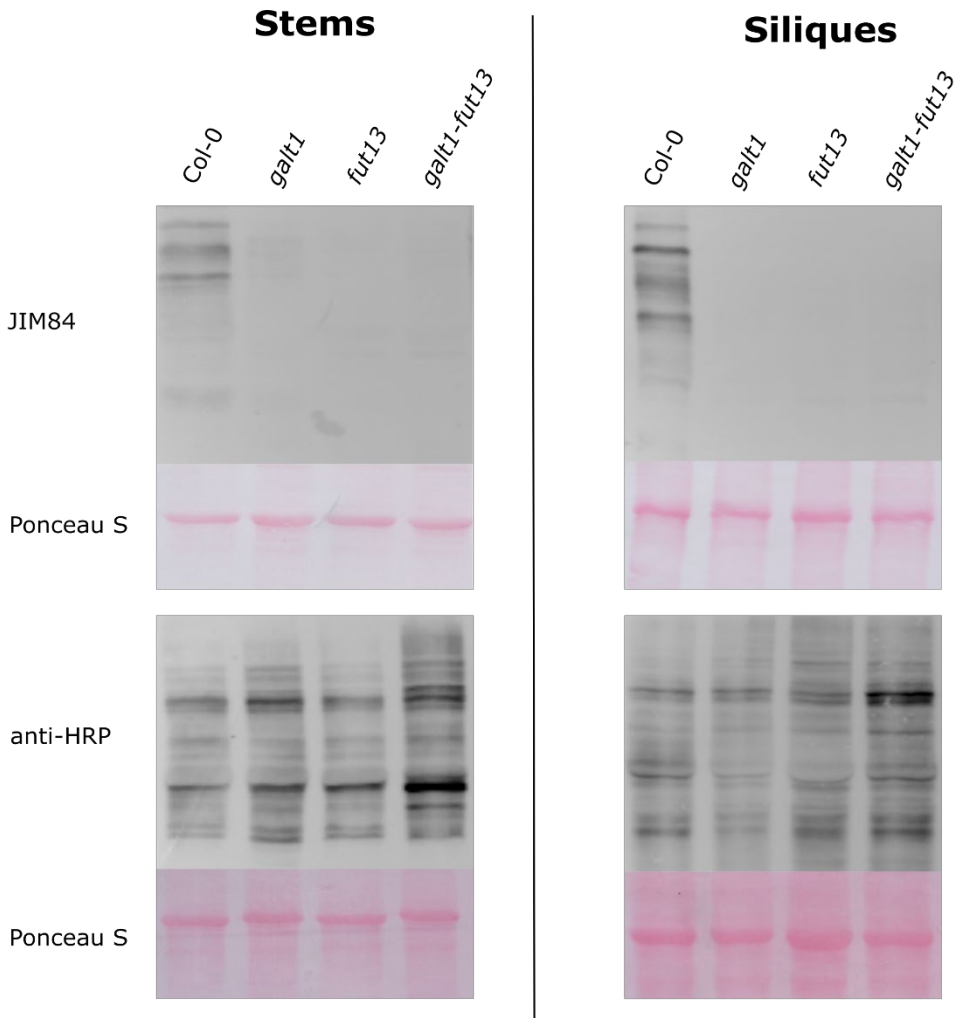


Figure S5. Immunoblot analysis of total protein extracts from stems and siliques of wild-type and Lewis A-deficient Arabidopsis plants, respectively. The JIM84 antibody was used to detect glycoproteins decorated with Lewis A structures, anti-HRP antibody for detection of plant N-glycans carrying β1,2-xylose and core α1,3-fucose.

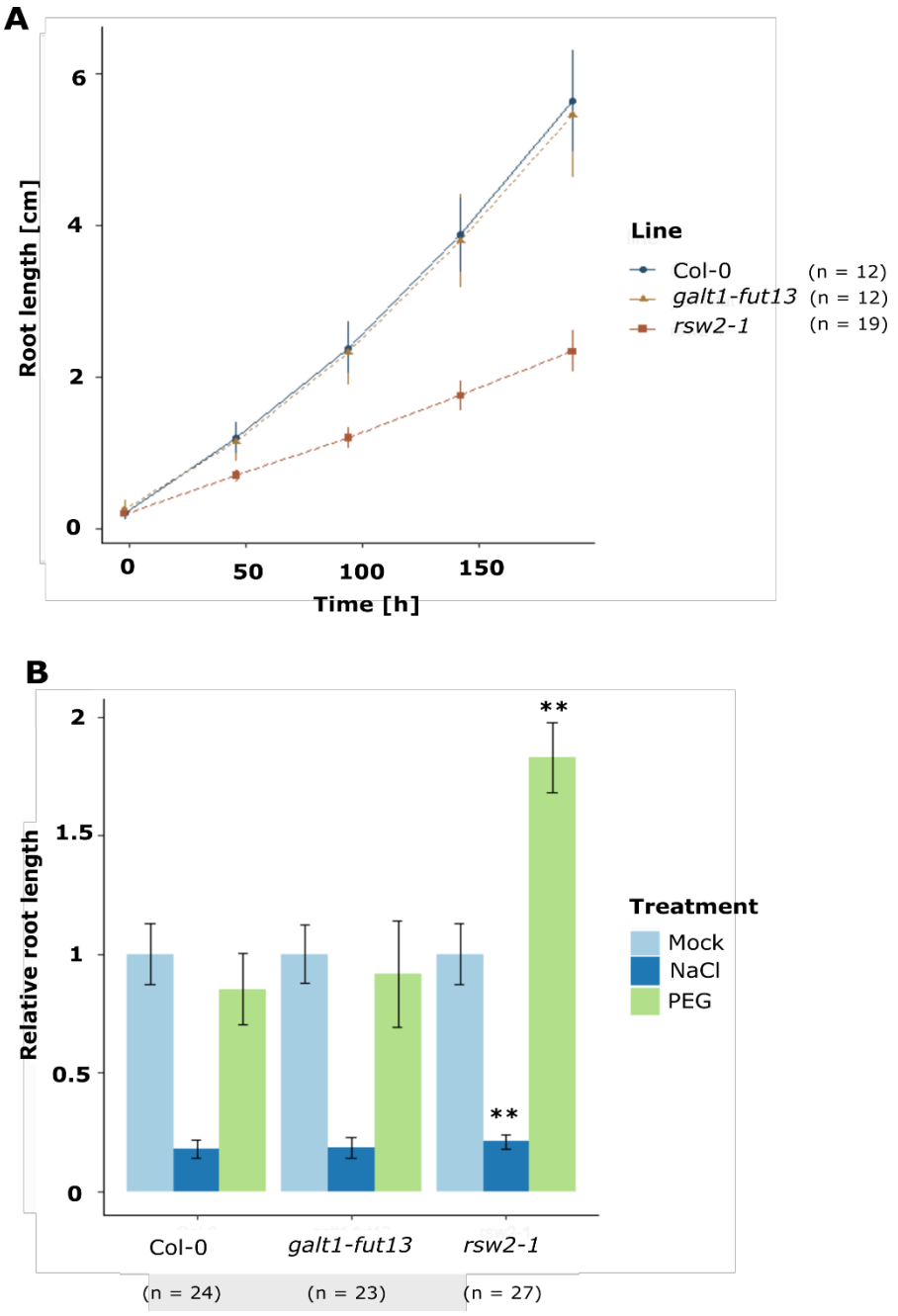


Figure S6. Root growth analysis of Col-0 and *galt1-fut13* plants. *rsw2-1* was included as a control. **(A)** Analysis of root growth dynamics of plants grown on half-strength MS-medium. Following stratification, root length was measured after 48, 96, 144, 192 and 240 hours. Error bars indicate standard deviation. **(B)** Abiotic stress treatment of seedlings using either 150 mM NaCl or PEG6000 at a water potential of -0.5 MPa to mimic drought stress. Error bars represent standard deviation, asterisks indicate statistically significant differences in the relative root lengths compared to Col-0 seedlings based on a Student's t-test (* $p \leq 0.05$, ** $p \leq 0.01$).

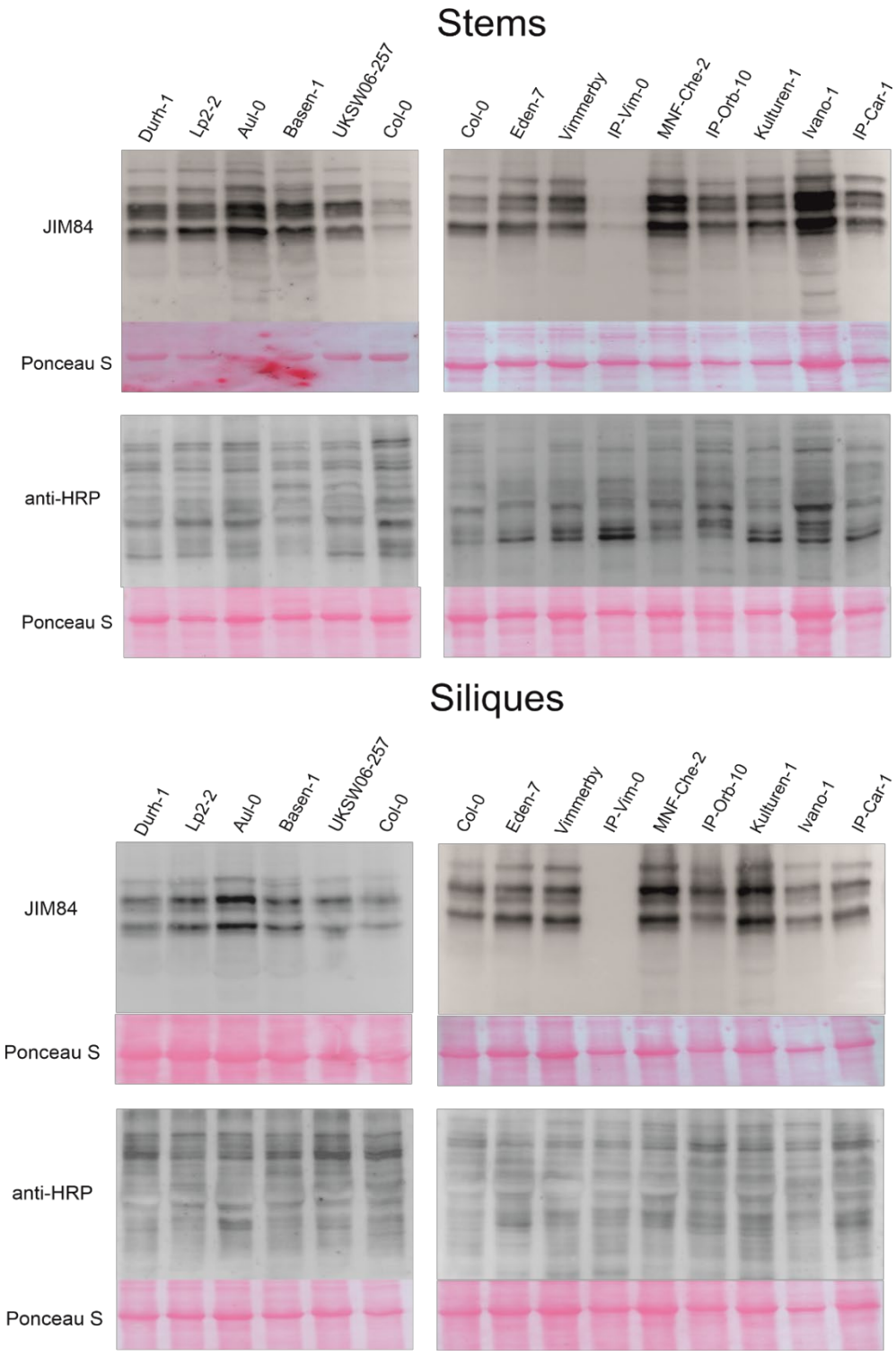


Figure S7. Immunoblot analysis of total protein extracts from stems and siliques of various accessions of *A. thaliana* containing amino acid variations in GALT1. Proteins were extracted using RIPA buffer, separated on an SDS-PAGE and either Lewis A bearing glycoproteins detected using the JIM84

antibody or plant N-glycans carrying β1,2-xylose and core α1,3-fucose detected using the anti-HRP antibody.

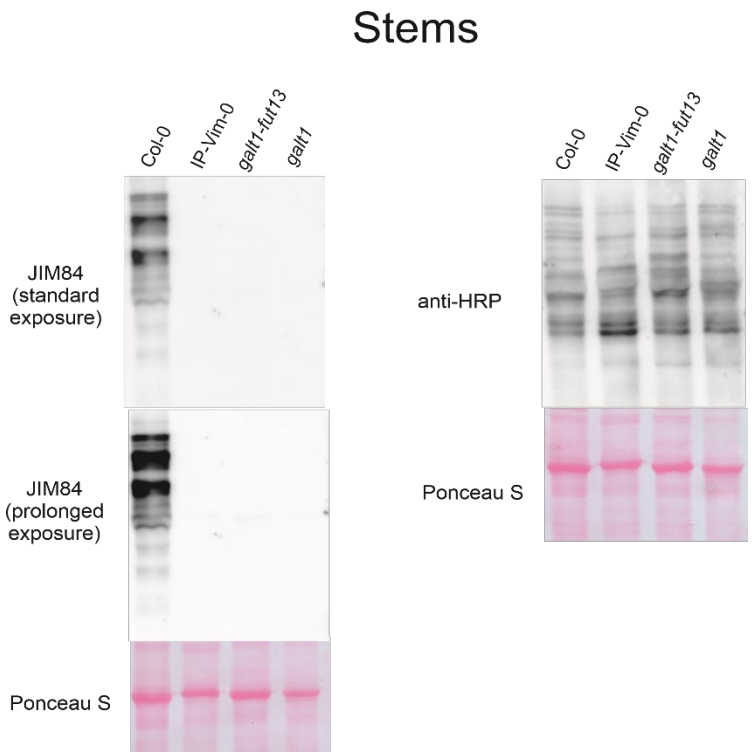
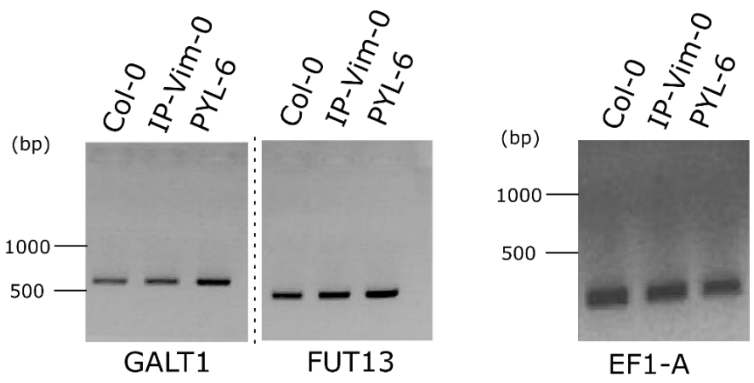


Figure S8. Immunoblot analysis of total protein extracts from stems of Col-0, IP-Vim-0 and Lewis A-deficient lines of *A. thaliana*. Proteins were extracted using RIPA buffer, separated on an SDS-PAGE and either Lewis A bearing glycoproteins detected using the JIM84 antibody or plant N-glycans carrying β1,2-xylose and core α1,3-fucose detected using the anti-HRP antibody.

A



B

Score	Expect	Identities	Gaps	Strand
1297 bits(702)	0.0	706/708(99%)	0/708(0%)	Plus/Plus

```
Query  5     AAGACGGATGGCTGTGAGGAGAACATGGATGCAGTATGATGATGTAAGATCTGGAAGAGT  64
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1218  AAGACGGATGGCTGTGAGGAGAACATGGATGCAGTATGATGATGTAAGATCTGGAAGAGT  1277

Query  65    TGCAGTACGCTTTTTTGTTGGCCTTCACAAAAGTCCTCTTGTTAACTTGGAACTCTGGAA  124
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1278  TGCAGTACGCTTTTTTGTTGGCCTTCACAAAAGTCCTCTTGTTAACTTGGAACTCTGGAA  1337

Query  125   CGAGGCTCGGACTTACGGTGATGTTCAGCTAATGCCCTTTGTTGATTATTACAGTCTCAT  184
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1338  CGAGGCTCGGACTTACGGTGATGTTCAGCTAATGCCCTTTGTTGATTATTACAGTCTCAT  1397

Query  185   CAGTTGGAAAACACTAGCCATCTGCATCTTCGGGACAGAGGTTGACTCAGCCAAGTTCAT  244
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1398  CAGTTGGAAAACACTAGCCATCTGCATCTTCGGGACAGAGGTTGACTCAGCCAAGTTCAT  1457

Query  245   CATGAAAACGGATGATGACGCCTTTGTTCGTGTAGATGAAGTGTTACTTTCTTTATCAAT  304
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1458  CATGAAAACGGATGATGACGCCTTTGTTCGTGTAGATGAAGTGTTACTTTCTTTATCAAT  1517

Query  305   GACCAACAACACTCGCGGGTTAATATACGGACTGATCAATTCCGACTCTCAACCTATTCG  364
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1518  GACCAACAACACTCGCGGGTTAATATACGGACTGATCAATTCCGACTCTCAACCTATTCG  1577

Query  365   AAACCCTGATAGCAAATGGTACATCAGTTATGAGGAATGGCCTGAAGAGAAATATCCACC  424
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1578  AAACCCTGATAGCAAATGGTACATCAGTTATGAGGAATGGCCTGAAGAGAAATATCCACC  1637

Query  425   ATGGGCGCATGGTCCAGGCAACATTGTATCTCGTGACATAGAAGAATCGGTTGGTAAGCT  484
             |||||||||||||||||| |||||||||||||||||||| |||||||||||||||||||
Sbjct  1638  ATGGGCGCATGGTCCAGGCTACATTGTATCTCGTGACATAGCAGAATCGGTTGGTAAGCT  1697

Query  485   TTTCAAAGAAGGAAACCTAAAGATGTTTAAGCTAGAAGATGTGGCAATGGGGATATGGAT  544
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1698  TTTCAAAGAAGGAAACCTAAAGATGTTTAAGCTAGAAGATGTGGCAATGGGGATATGGAT  1757

Query  545   AGCTGAGCTGACAAAACATGGACTCGAGCCTCATTACGAAAACGATGGAAGGATCATTAG  604
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1758  AGCTGAGCTGACAAAACATGGACTCGAGCCTCATTACGAAAACGATGGAAGGATCATTAG  1817

Query  605   TGATGGATGCAAGGATGGTTATGTGGTTGCTCATTACCAAAGCCCTGCCGAAATGACTTG  664
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1818  TGATGGATGCAAGGATGGTTATGTGGTTGCTCATTACCAAAGCCCTGCCGAAATGACTTG  1877

Query  665   CCTTTGGCGTAAATACCAAGAAACCAAACGCTCTCTTTGCTGCCGCGA  712
             ||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1878  CCTTTGGCGTAAATACCAAGAAACCAAACGCTCTCTTTGCTGCCGCGA  1925
```

Figure S9. Analysis of *GALT1* and *FUT13* transcripts in stems of Col-0, PYL-6 and IP-Vim-0. **(A)** PCR amplification of *GALT1* and *FUT13* cDNA obtained from RNA extracted from stems of the different ecotypes. *EF1-ALPHA* (EF1-A, At5g60390) was included as a reference gene. **(B)** Alignment of bases 1218-1925 of *GALT1* in Col-0 (lower sequence) and IP-Vim-0 (upper sequence) after sequencing. SNPs are underlined in black. The alignment was done using Blast (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome).

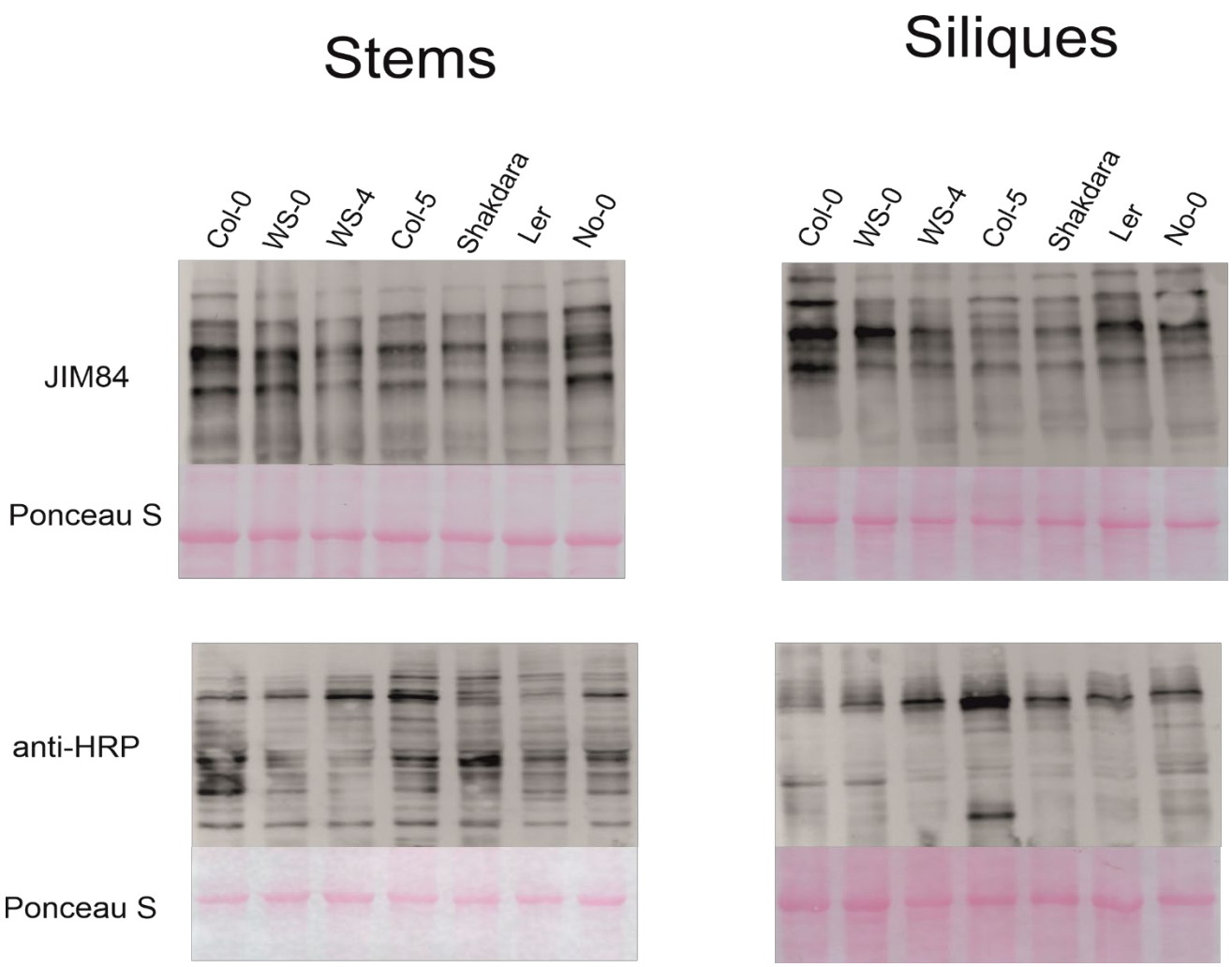


Figure S10. Immunoblot analysis of total protein extracts from stems and siliques of various ecotypes of *A. thaliana*. Proteins were extracted using RIPA buffer, separated on an SDS-PAGE and either Lewis A bearing glycoproteins detected using the JIM84 antibody or plant N-glycans carrying β1,2-xylose and core α1,3-fucose detected using the anti-HRP antibody.

Glycosylphosphatidylinositol-Anchor Synthesis in Plants: A Glycobiology Perspective

Beihammer G, Maresch D, Altmann F, Strasser R

Perspective Article

Front Plant Sci. 2020 Nov 17;11:611188. doi: 10.3389/fpls.2020.611188

frontiers in Plant Science

PERSPECTIVE
published: 17 November 2020
doi: 10.3389/fpls.2020.611188

Glycosylphosphatidylinositol-Anchor Synthesis in Plants: A Glycobiology Perspective

*Gernot Beihammer[1], Daniel Maresch[2], Friedrich Altmann[2] and Richard Strasser[1]**

[1]Department of Applied Genetics and Cell Biology, Institute of Plant Biotechnology and Cell Biology, University of Natural Resources and Life Sciences, Vienna, Austria, [2]Division of Biochemistry, Department of Chemistry, University of Natural Resources and Life Sciences, Vienna, Austria

More than 200 diverse secretory proteins from *Arabidopsis thaliana* carry a glycosylphosphatidylinositol (GPI) lipid anchor covalently attached to their carboxyl-terminus. The GPI-anchor contains a lipid-linked glycan backbone that is preassembled in the endoplasmic reticulum (ER) of plants and subsequently transferred to distinct proteins, which provides them with specific features. The GPI-anchored proteins exit the ER and are transported through the Golgi apparatus to the plasma membrane. In the Golgi, the glycan moiety can be further modified by the specific attachment of sugar residues. While these biosynthetic steps are already quite well understood in mammals and yeast, comparatively little is known in plants. In this perspective, we discuss the current knowledge about the biosynthesis of the GPI-anchor glycan moiety in the light of recent findings for mammalian GPI-anchor glycan modifications.

Keywords: posttranslational modification, mannose, glycosyltransferase, glycosylation, endoplasmic reticulum, glycosylphosphatidylinositol

OPEN ACCESS

Edited by:
Kris Vissenberg, University of Antwerp, Belgium

Reviewed by:
Taroh Kinoshita, Osaka University, Japan
Verena Kriechbaumer, Oxford Brookes University, United Kingdom

****Correspondence:***
Richard Strasser
richard.strasser@boku.ac.at

Specialty section:
This article was submitted to Plant Cell Biology, a section of the journal Frontiers in Plant Science

Received: *28 September 2020*
Accepted: *30 October 2020*
Published: *17 November 2020*

Citation:
Beihammer G, Maresch D, Altmann F and Strasser R (2020) Glycosylphosphatidylinositol-Anchor Synthesis in Plants: A Glycobiology Perspective. Front. Plant Sci. 11:611188. doi: 10.3389/fpls.2020.611188

INTRODUCTION

The attachment of glycosylphosphatidylinositol (GPI) is a common posttranslational modification for anchoring of proteins to the outer surface of the plasma membrane in eukaryotes. The conserved GPI moiety is composed of ethanolamine phosphate (EtNP), a conserved core glycan and phosphatidylinositol. The core glycan consists of three mannoses (Man) and one glucosamine (GlcN) residue that are linked to EtNP and phosphatidylinositol forming the GPI backbone structure EtNP-6Manα1-2Manα1-6Manα1-4GlcNα1-6*myo*-inositol-phospholipid (Kinoshita and Fujita, 2016; Liu and Fujita, 2020; **Figure 1**). Proteins destined to be GPI-anchored are translocated into the lumen of the endoplasmic reticulum (ER), the GPI attachment signal peptide at the C-terminus is cleaved off and the preassembled GPI is transferred en bloc to the last amino acid of the C-terminus called the ω site. The transfer is mediated by the GPI transamidase, a multi-subunit complex comprising five proteins. The attachment of GPI results in anchoring of the protein to the outer leaflet of the lipid bilayer. The modification with a GPI-anchor confers specific properties on proteins, such as efficient ER exit, sorting to the plasma membrane

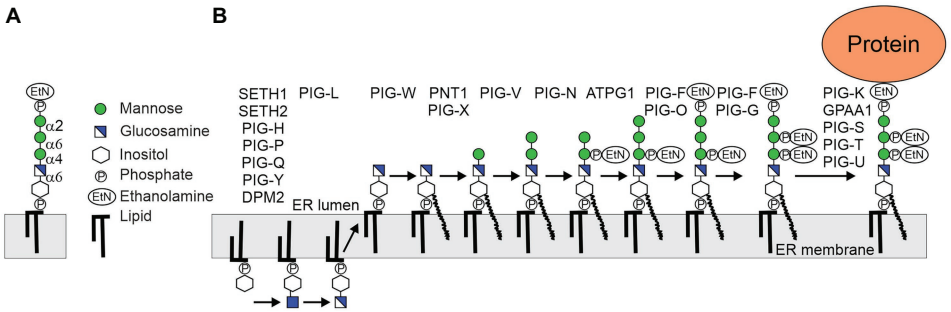


FIGURE 1 | (A) Structure of the conserved glycosylphosphatidylinositol (GPI)-anchor backbone. **(B)** Biosynthesis of plant GPI precursor in the endoplasmic reticulum (ER). The biosynthesis is a stepwise process initiated at the cytoplasmic side of the ER by the GPI-GlcNAc-transferase (GPI-GnT) multiprotein complex (SETH1, SETH2, PIG-H, PIG-P, PIG-Q, PIG-Y, and DPM2). The GlcNAc is de-N-acetylated by PIG-L and GlcN-PI flips to the luminal side of the ER. GPI mannosyltransferases PNT1 (homolog of mammalian PIG-M), PIG-V and ATPG1 (homolog of mammalian PIG-B) attach the three mannose (Man) residues to the backbone, which is further modified by attachment of ethanolamine phosphate (EtNP). The assembled precursor is transferred en bloc by the GPI transamidase complex (PIG-K, GPAA1, PIG-S, PIG-T, and PIG-U) to proteins. In addition to the bridging EtNP, one or two additional EtNP may be transferred to mannose residues by specific GPI-EtNP transferases that are present in *Arabidopsis thaliana* (Ellis et al., 2010; Luschnig and Seifert, 2011).

and association with specific membrane microdomains (Sikorska et al., 2016). In *Arabidopsis thaliana*, multiple protein families have been predicted by bioinformatic analysis to carry a GPI-anchor (Borner et al., 2002; Eisenhaber et al., 2003), and approximately 200 GPI-anchored proteins have been identified by different proteomics approaches (Borner et al., 2003; Elortza et al., 2006; Takahashi et al., 2016). Proteins that carry a GPI-anchor include the multi-copper oxidase-related protein SKU5 (Sedbrook et al., 2002), COBRA family proteins (Schindelman et al., 2001; Roudier et al., 2005; Brady et al., 2007), lipid-transfer proteins (LTPGs; Debono et al., 2009), and arabinogalactan proteins (AGPs; Oxley and Bacic, 1999; Shi et al., 2003; Xue et al., 2017). For a more comprehensive list of potential GPI-anchored proteins and phenotypes associated with mutants see recent reviews (Yeats et al., 2018; Zhou, 2019). Here, we focus on the biosynthesis of the GPI-anchor core glycan moiety and potential side chain modifications.

GPI-ANCHOR GLYCAN BIOSYNTHESIS IN MAMMALS AND YEAST

In mammalian cells, at least 150 proteins at the plasma membrane are attached to the cell surface by GPI-anchors and GPI-anchoring is essential for many biological processes including embryogenesis, fertilization, or the immune system (Kinoshita and Fujita, 2016). In *Saccharomyces cerevisiae*, more than 60 GPI-anchored proteins have been identified and GPI biosynthesis is required for the growth of yeast (Leidich et al., 1994). During biosynthesis and after attachment to proteins, the structures of lipid and glycan moieties from GPI-anchors are remodeled in the ER and in the Golgi apparatus.

Glycosylphosphatidylinositol biosynthesis is initiated at the cytosolic side of the ER by the transfer of GlcNAc from the nucleotide sugar UDP-GlcNAc to inositol to generate GlcNAc-PI (**Figure 1**). This step is catalyzed by the GPI-GlcNAc-transferase (GPI-GnT), a complex consisting of seven protein subunits in mammalian cells. GlcNAc-PI is de-N-acetylated to GlcN-PI by the deacetylase PIG-L, and GlcN-PI is flipped to the luminal side by an unknown process. In the ER lumen, GlcN-PI is acylated by the acyltransferase PIG-W and the lipid moiety is remodeled to generate GlcN-(acyl)PI. In the next step, the GPI α1,4-mannosyltransferase PIG-M and the GPI α1,6-mannosyltransferases PIG-V catalyze the sequential transfer of two Man residues to GlcN-(acyl)PI (Maeda et al., 2001; Kang et al., 2005). PIG-B, another GPI α1,2-mannosyltransferase, transfers the third mannose to generate Manα1-2Manα1-6Manα1-4GlcN-(acyl)PI (Takahashi et al., 1996). A GPI-EtNP-transferase transfers the so-called bridging EtNP that connects the protein and the glycan to the third mannose to generate EtNP-6Manα1-2Manα1-6Manα1-4GlcN-(acyl)PI (Kinoshita and Fujita, 2016). Two additional side-branch EtNPs are added to the first and second mannose residues. While the EtNP transfer to the first mannose takes place after the transfer of the second mannose, the modification of the second mannose with EtNP likely happens after the attachment of the bridging EtNP. The resulting structure is competent for transfer to proteins, but can be further modified by the attachment of a fourth mannose residue catalyzed by the GPI α1,2-mannosyltransferase PIG-Z. Like the ER-resident mannosyltransferases (ALG3, ALG9, and ALG11) involved in the assembly of the oligosaccharide precursor for N-glycosylation, all four GPI mannosyltransferases are multiple transmembrane proteins and use dolichol-phosphate-mannose (Dol-P-Man) as donor substrate. PIG-B and PIG-Z are like ALG9 and ALG12 members of CAZy family GT22. PIG-V which is distantly related to STT3, the catalytic subunit of the oligosaccharyltransferase complex, belongs to family GT76 (Kang et al., 2005). PIG-M is distantly related to ALG3 (GT58) and belongs to GT50 (Oriol et al., 2002). All these enzymes use dolichol-linked donor substrates and have 11–14 membrane spanning helices and conserved residues in luminal loops (Albuquerque-Wendt et al., 2019).

In yeast, an α1,2-linked mannose is attached to the third mannose and this modification is an essential biosynthetic step

in the assembly of yeast GPIs (Grimme et al., 2001). This mannosylation step takes place in the ER and precedes the attachment of the bridging EtNP. In mammals, some GPI-anchored proteins have also a fourth mannose residue in the same position that is transferred by PIG-Z (Taron et al., 2004; Hirata et al., 2018). In addition, the first mannose residue of mammalian GPI-anchors is often modified with N-acetylgalactosamine (GalNAc). In contrast to the previously mentioned glycosylation reactions, this step takes place in the Golgi apparatus and is catalyzed by the GPI GalNAc-transferase PGAP4 (Hirata et al., 2018). This enzyme uses UDP-GalNAc as donor substrate. The GalNAc side chain may be further elongated by incorporation of β1,3-linked galactose and α2,3-linked sialic acid. The Golgi-resident galactosyltransferase B3GALT4 catalyzes the transfer of a galactose from UDP-galactose to the side chain GalNAc residue (Wang et al., 2020). The GPI sialyltransferase was not identified yet, but recently it was found that in prion protein the sialic acid N-acetylneuraminic acid (Neu5Ac) is present in α2,3-linkage (Kobayashi et al., 2020).

GPI-ANCHOR CORE GLYCAN BIOSYNTHESIS IN PLANTS

In contrast to mammals and yeast, our knowledge about the different biosynthetic steps involved in the assembly of the GPI core glycan and possible side chain formations is limited (Yeats et al., 2018). Based on sequence comparison, *Arabidopsis* SETH1 and SETH2 have been identified as homologs of subunits PIG-C and PIG-A of the GPI-GnT complex (Lalanne et al., 2004). Disruption of SETH1 or SETH2 affects pollen germination and tube growth suggesting a role of GPI-anchored proteins in pollen function. PEANUT1 (PNT1) is the *Arabidopsis* homolog of PIG-M involved in the first mannosylation step (Gillmor et al., 2005; **Figure 1**). The *pnt1* mutant is embryo lethal, displays defects in cell wall biosynthesis and GPI-anchored proteins like SKU5 or COBRA are absent or strongly reduced in *pnt1* embryos or callus. Mammalian PIG-M forms a complex with PIG-X, which stabilizes the catalytic subunit PIG-M (Ashida et al., 2005). The *Arabidopsis* PIG-X homolog (At5g46850) has not been characterized but likely has a similar function. An *Arabidopsis* PIG-X knockout could display a less severe phenotype than *pnt1* because PIG-M expression is not completely abolished in mammalian cells lacking PIG-X (Ashida et al., 2005). *Arabidopsis* lines lacking PIG-V, the GPI mannosyltransferase catalyzing the transfer of the second mannose, have not been described and PIG-V has not been biochemically characterized. Like *pnt1*, a complete PIG-V knockout will block the biosynthesis of the GPI backbone and is thus likely embryo lethal. *Arabidopsis* ABNORMAL POLLEN TUBE GUIDANCE1 (APTG1) can functionally replace the yeast PIG-B homolog that transfers the third mannose to the GPI precursor (Dai et al., 2014). In line with the essential function in plants, the *aptg1* mutant showed embryo lethality. APTG1 is an integral membrane protein located in the ER and plants with disrupted APTG1 expression display mislocalization of GPI-anchored proteins. Together these studies provide a clear and consistent insight into the enzymes involved in the core glycan biosynthesis and their biological function in plants.

SIDE CHAIN MODIFICATIONS IN PLANTS

The complete chemical structure of a mammalian GPI-anchor was published in 1988 (Homans et al., 1988). In plants, only one study reported the structure of a GPI-anchored protein (Oxley and Bacic, 1999). The GPI-anchor of an AGP isolated from *Pyrus communis* suspension cells has a glycan core that is identical to the one from mammals and yeast. Instead of a GalNAc at the same position, the first mannose of the core glycan was partially modified with a β-linked galactose.

To see if the presence of the β-linked galactose is common in plants, we transiently expressed RFP fused to the C-terminal GPI attachment signal peptide from *Arabidopsis* COBRA transiently in *Nicotiana benthamiana* leaves or in transgenic *Arabidopsis* (Strasser et al., 2005), purified the GPI-anchored RFP-COB1 reporter protein and subjected the PI-PLC released C-terminal tryptic peptide to LC-ESI-MS/MS analysis (Kolarich and Altmann, 2000) (**Figure 2**). RFP-COB1 was present in the plasma membrane and MS-spectra of the terminal peptide from the two different plant species were obtained. The MS-analysis revealed masses in both species corresponding to the presence of the GPI core backbone without any additional side chain EtNP modifications. The presence of the respective EtNP transferases in the *Arabidopsis* genome (Ellis et al., 2010; Luschnig and Seifert, 2011) suggests that this modification is removed during the remodeling of the GPI-anchor rather than being absent in plants. In mammalian cells, the side chain EtNP is removed by the phosphoesterase PGAP5 (Fujita et al., 2009). While yeast contain two PGAP5 homologs (CDC1 and TED1), there is one (At1g53710) so far uncharacterized PGAP5 homolog in *Arabidopsis*. Structures composed of four hexoses were detected in RFP-COB1 from *A. thaliana* (**Figure 2**) and *N. benthamiana* (**Supplementary Figure S1**), which is consistent with the three sequential mannoses in the core glycan. The fourth hexose is a side chain modification that is likely the previously described β-linked galactose. Alternatively, the hexose could be a mannose residue α-linked to the third mannose that has been described in yeast and mammals. When, we digested the peptide from RFP-COB1 with a β-galactosidase, a single hexose (mass Δ162) was quantitatively removed from the *Arabidopsis* derived peptide (**Figure 2**). For the peptide from *N. benthamiana*, we performed an additional α-mannosidase digestion because the β-galactosidase treatment did not fully remove the hexose (**Supplementary Figure S2**). The α-mannosidase treatment did not alter the glycan composition of the GPI-anchor suggesting that the fourth hexose is a galactose. It seems likely that the galactose is bound in β1,4-linkage because the galactosidase used for the digestion exhibits high specificity for this type of linkage (Zeleny et al., 1997). Closer investigation of the MS/MS-spectra of RFP-COB1 subjected to β-galactosidase treatment revealed that the mass corresponding to GlcN-PI+2xHex was absent when compared to the mock incubated control, indicating that the galactose is bound to the GlcN-linked mannose

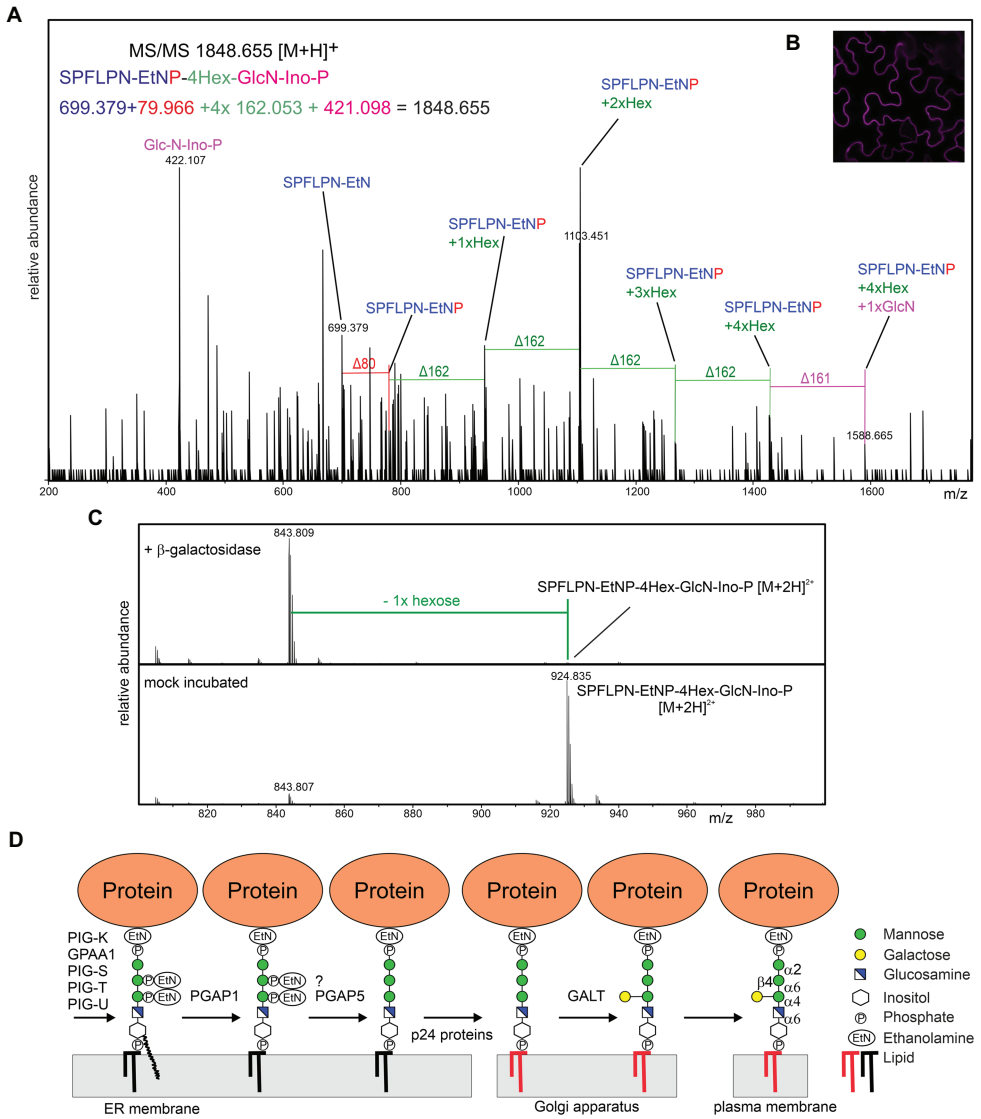


FIGURE 2 | (A) LC-ESI-MS/MS analysis of the C-terminal peptide from RFP-COB1. RFP-COB1 was expressed in *A. thaliana* seedlings, captured by binding to RFP-Trap resin and subjected to PI-PLC and trypsin digestion. **(B)** RFP-COB1 is found at the plasma membrane. RFP-COB1 was transiently expressed in *Nicotiana benthamiana* leaf epidermal cells and analyzed by confocal microscopy. **(C)** LC-ESI-MS analysis of β-galactosidase digested GPI-anchor derived from *A. thaliana* expressed RFP-COB1. **(D)** The transferred GPI backbone is further modified by different enzymes, including inositol deacylation, the remodeling of the lipid portion (depicted by the change in color from black to red in the illustration) and the EtNP removal, which is likely required for recognition by p24 cargo receptor proteins and efficient ER exit. In the proposed model, the plant-specific side chain modification is transferred in the Golgi by an unknown β-galactosyltransferase (GALT).

(**Supplementary Figure S3**). These findings are consistent with the structure from AGP isolated from pear cells (Oxley and Bacic, 1999) indicating that attachment of a single galactose in β-linkage is a common side chain formation of the GPI core glycan in plants.

DISCUSSION

The detection of a single hexose as GPI side chain modification raises several questions: which glycosyltransferase catalyzes this step, which donor substrate is used in the reaction and in which subcellular compartment is the side chain modified. The glycosyltransferase family CAZy GT22 (α1,2-mannosyltransferases) contains only three *Arabidopsis* proteins. APTG1, the only homolog of PIG-B and the two enzymes (ALG12 and ALG9) involved in the assembly of the lipid-linked oligosaccharide precursor for N-glycosylation. The lack of a PIG-Z homolog in plants (Ellis et al., 2010; Luschnig and Seifert, 2011) is consistent with the absence of a fourth mannose residue attached to the GPI glycan core in the ER unless APTG1 transfers a second mannose residue as it has

been suggested for PIG-B (Wang et al., 2020). Many ER-resident glycosyltransferases are integral membrane proteins that use dolichol-phosphate-linked sugars as donor substrate. A Dol-P-galactose has not been described and ER-resident multiple transmembrane domain-containing galactosyltransferases are not known. Therefore, we suggest that the side chain modification takes place in the Golgi apparatus of plants and involves an unknown β-galactosyltransferase that uses UDP-galactose as donor substrate. Golgi glycosyltransferases are typically type II membrane proteins with a single transmembrane domain and a short N-terminal cytoplasmic region (Schoberer and Strasser, 2011). The recently identified GPI-GalNAc-transferase PGAP4 has a different structure with two additional tandem transmembrane domains (Hirata et al., 2018). This peculiar structure likely facilitates the interaction with the membrane-anchored substrate. While there is no PGAP4 homolog present in the *Arabidopsis* proteome, it is possible that the unknown plant GPI-galactosyltransferase has a similar membrane topology. On the other hand, B3GALT4 that transfers a galactose to the GalNAc has a common type II membrane topology. B3GALT4 is distantly related to *Arabidopsis* hydroxyproline O-galactosyltransferases that are members of the GT31 family (Basu et al., 2015). Plants have a large family of GT31 galactosyltransferases with still poorly characterized function (Showalter and Basu, 2016) that are involved in different pathways, including N-glycan processing (Strasser et al., 2007). One of those galactosyltransferases from GT31 could be responsible for the side chain formation of GPI core glycan.

What is the function of the GPI side chain modification? Like for many glycan modifications such as Lewis A-type structures, the function of the attached GPI side chain galactose is currently unknown (Strasser, 2016). In mammalian brain, galactosylated and sialylated GPI-anchors are more abundant than in other tissues. However, not all GPI-anchored proteins are modified to the same extent with some proteins having only GalNAc instead of additional galactose and sialic acid modifications (Kobayashi et al., 2020). Sialylation of the human prion protein side chain may contribute to the pathology of prion disease (Bate et al., 2016). The specific function of side chain glycosylation in mammals and the conserved nature of the identified core glycan modification in plants suggests that the β-linked galactose has a biological role that needs to be unraveled in future studies.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

GB and RS designed the experiments and wrote the paper. GB and DM conducted the experiments. GB, DM, FA, and RS analyzed the results. All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

This work was supported by the Austrian Science Fund [FWF; Doctoral Program BioToP–Biomolecular Technology of Proteins (W1224)] and by the FWF Projects I2417-B22 and P31920-B32.

ACKNOWLEDGMENTS

We thank Yun-Ji Shin and Ulrike Vavra (Department of Applied Genetics and Cell Biology) for help with cloning, protein purification, and confocal microscopy.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.611188/full#supplementary-material

REFERENCES

Albuquerque-Wendt, A., Hütte, H. J., Buettner, F. F. R., Routier, F. H., and Bakker, H. (2019). Membrane topological model of glycosyltransferases of the GT-C superfamily. *Int. J. Mol. Sci.* 20:4842. doi: 10.3390/ijms20194842

Ashida, H., Hong, Y., Murakami, Y., Shishioh, N., Sugimoto, N., Kim, Y. U., et al. (2005). Mammalian PIG-X and yeast Pbn1p are the essential components of glycosylphosphatidylinositol-mannosyltransferase I. *Mol. Biol. Cell* 16, 1439–1448. doi: 10.1091/mbc.e04-09-0802

Basu, D., Tian, L., Wang, W., Bobbs, S., Herock, H., Travers, A., et al. (2015). A small multigene hydroxyproline-O-galactosyltransferase family functions in arabinogalactan-protein glycosylation, growth and development in *Arabidopsis*. *BMC Plant Biol.* 15:295. doi: 10.1186/s12870-015-0670-7

Bate, C., Nolan, W., and Williams, A. (2016). Sialic acid on the glycosylphosphatidylinositol anchor regulates PrP-mediated cell signaling and prion formation. *J. Biol. Chem.* 291, 160–170. doi: 10.1074/jbc.M115.672394

Borner, G. H., Lilley, K. S., Stevens, T. J., and Dupree, P. (2003). Identification of glycosylphosphatidylinositol-anchored proteins in *Arabidopsis*. A proteomic and genomic analysis. *Plant Physiol.* 132, 568–577. doi: 10.1104/pp.103.021170

Borner, G. H., Sherrier, D. J., Stevens, T. J., Arkin, I. T., and Dupree, P. (2002). Prediction of glycosylphosphatidylinositol-anchored proteins in *Arabidopsis*. A genomic analysis. *Plant Physiol.* 129, 486–499. doi: 10.1104/pp.010884

Brady, S. M., Song, S., Dhugga, K. S., Rafalski, J. A., and Benfey, P. N. (2007). Combining expression and comparative evolutionary analysis. The COBRA gene family. *Plant Physiol.* 143, 172–187. doi: 10.1104/pp.106.087262

Dai, X. R., Gao, X. Q., Chen, G. H., Tang, L. L., Wang, H., and Zhang, X. S. (2014). Abnormal pollen tube guidance1, an endoplasmic reticulum-localized mannosyltransferase homolog of glycosylphosphatidylinositol10 in yeast and phosphatidylinositol glycan anchor biosynthesis B in human, is required for *Arabidopsis* pollen tube micropylar guidance and embryo development. *Plant Physiol.* 165, 1544–1556. doi: 10.1104/pp.114.236133

Debono, A., Yeats, T. H., Rose, J. K., Bird, D., Jetter, R., Kunst, L., et al. (2009). *Arabidopsis* LTPG is a glycosylphosphatidylinositol-anchored lipid transfer protein required for export of lipids to the plant surface. *Plant Cell* 21, 1230–1238. doi: 10.1105/tpc.108.064451

Eisenhaber, B., Wildpaner, M., Schultz, C. J., Borner, G. H., Dupree, P., and Eisenhaber, F. (2003). Glycosylphosphatidylinositol lipid anchoring of plant proteins. Sensitive prediction from sequence- and genome-wide studies for *Arabidopsis* and rice. *Plant Physiol.* 133, 1691–1701. doi: 10.1104/pp.103.023580

Ellis, M., Egelund, J., Schultz, C. J., and Bacic, A. (2010). Arabinogalactan-proteins: key regulators at the cell surface? *Plant Physiol.* 153, 403–419. doi: 10.1104/pp.110.156000

Elortza, F., Mohammed, S., Bunkenborg, J., Foster, L. J., Nühse, T. S., Brodbeck, U., et al. (2006). Modification-specific proteomics of plasma membrane proteins: identification and characterization of glycosylphosphatidylinositol-anchored proteins released upon phospholipase D treatment. *J. Proteome Res.* 5, 935–943. doi: 10.1021/pr050419u

Fujita, M., Maeda, Y., Ra, M., Yamaguchi, Y., Taguchi, R., and Kinoshita, T. (2009). GPI glycan remodeling by PGAP5 regulates transport of GPI-anchored proteins from the ER to the golgi. *Cell* 139, 352–365. doi: 10.1016/j.cell.2009.08.040

Gillmor, C. S., Lukowitz, W., Brininstool, G., Sedbrook, J. C., Hamann, T., Poindexter, P., et al. (2005). Glycosylphosphatidylinositol-anchored proteins are required for cell wall synthesis and morphogenesis in *Arabidopsis*. *Plant Cell* 17, 1128–1140. doi: 10.1105/tpc.105.031815

Grimme, S. J., Westfall, B. A., Wiedman, J. M., Taron, C. H., and Orlean, P. (2001). The essential Smp3 protein is required for addition of the side-branching fourth mannose during assembly of yeast glycosylphosphatidylinositols. *J. Biol. Chem.* 276, 27731–27739. doi: 10.1074/jbc.M101986200

Hirata, T., Mishra, S. K., Nakamura, S., Saito, K., Motooka, D., Takada, Y., et al. (2018). Identification of a golgi GPI-N-acetylgalactosamine transferase with tandem transmembrane regions in the catalytic domain. *Nat. Commun.* 9:405. doi: 10.1038/s41467-017-02799-0

Homans, S. W., Ferguson, M. A., Dwek, R. A., Rademacher, T. W., Anand, R., and Williams, A. F. (1988). Complete structure of the glycosyl phosphatidylinositol membrane anchor of rat brain Thy-1 glycoprotein. *Nature* 333, 269–272. doi: 10.1038/333269a0

Kang, J. Y., Hong, Y., Ashida, H., Shishioh, N., Murakami, Y., Morita, Y. S., et al. (2005). PIG-V involved in transferring the second mannose in glycosylphosphatidylinositol. *J. Biol. Chem.* 280, 9489–9497. doi: 10.1074/jbc.M413867200

Kinoshita, T., and Fujita, M. (2016). Biosynthesis of GPI-anchored proteins: special emphasis on GPI lipid remodelling. *J. Lipid Res.* 57, 6–24. doi: 10.1194/jlr.R063313

Kobayashi, A., Hirata, T., Nishikaze, T., Ninomiya, A., Maki, Y., Takada, Y., et al. (2020). α2,3 linkage of sialic acid to a GPI anchor and an unpredicted GPI attachment site in human prion protein. *J. Biol. Chem.* 295, 7789–7798. doi: 10.1074/jbc.RA120.013444

Kolarich, D., and Altmann, F. (2000). N-glycan analysis by matrix-assisted laser desorption/ionization mass spectrometry of electrophoretically separated nonmammalian proteins: application to peanut allergen Ara h 1 and olive pollen allergen ole e 1. *Anal. Biochem.* 285, 64–75. doi: 10.1006/abio.2000.4737

Lalanne, E., Honys, D., Johnson, A., Borner, G. H., Lilley, K. S., Dupree, P., et al. (2004). SETH1 and SETH2, two components of the glycosylphosphatidylinositol anchor biosynthetic pathway, are required for pollen germination and tube growth in *Arabidopsis*. *Plant Cell* 16, 229–240. doi: 10.1105/tpc.014407

Leidich, S. D., Drapp, D. A., and Orlean, P. (1994). A conditionally lethal yeast mutant blocked at the first step in glycosyl phosphatidylinositol anchor synthesis. *J. Biol. Chem.* 269, 10193–10196.

Liu, Y. S., and Fujita, M. (2020). Mammalian GPI-anchor modifications and the enzymes involved. *Biochem. Soc. Trans.* 48, 1129–1138. doi: 10.1042/BST20191142

Luschnig, C., and Seifert, G. (2011). "Posttranslational modifications of plasma membrane proteins and their implications for plant growth and development" in *The plant plasma membrane, plant cell monographs 19*. eds. A. S. Murphy, W. Peer and B. Schulz (Springer-Verlag Berlin Heidelberg), 109–128.

Maeda, Y., Watanabe, R., Harris, C. L., Hong, Y., Ohishi, K., Kinoshita, K., et al. (2001). PIG-M transfers the first mannose to glycosylphosphatidylinositol on the lumenal side of the ER. *EMBO J.* 20, 250–261. doi: 10.1093/emboj/20.1.250

Oriol, R., Martinez-Duncker, I., Chantret, I., Mollicone, R., and Codogno, P. (2002). Common origin and evolution of glycosyltransferases using Dol-P-monosaccharides as donor substrate. *Mol. Biol. Evol.* 19, 1451–1463. doi: 10.1093/oxfordjournals.molbev.a004208

Oxley, D., and Bacic, A. (1999). Structure of the glycosylphosphatidylinositol anchor of an arabinogalactan protein from *Pyrus communis* suspension-cultured cells. *Proc. Natl. Acad. Sci. U. S. A.* 96, 14246–14251. doi: 10.1073/pnas.96.25.14246

Roudier, F., Fernandez, A., Fujita, M., Himmelspach, R., Borner, G., Schindelman, G., et al. (2005). COBRA, an *Arabidopsis* extracellular glycosyl-phosphatidyl inositol-anchored protein, specifically controls highly anisotropic expansion through its involvement in cellulose microfibril orientation. *Plant Cell* 17, 1749–1763. doi: 10.1105/tpc.105.031732

Schindelman, G., Morikami, A., Jung, J., Baskin, T. I., Carpita, N. C., Derbyshire, P., et al. (2001). COBRA encodes a putative GPI-anchored protein, which is polarly localized and necessary for oriented cell expansion in *Arabidopsis*. *Genes Dev.* 15, 1115–1127. doi: 10.1101/gad.879101

Schoberer, J., and Strasser, R. (2011). Sub-compartmental organization of golgi-resident N-glycan processing enzymes in plants. *Mol. Plant* 4, 220–228. doi: 10.1093/mp/ssq082

Sedbrook, J. C., Carroll, K. L., Hung, K. F., Masson, P. H., and Somerville, C. R. (2002). The *Arabidopsis* SKU5 gene encodes an extracellular glycosyl phosphatidylinositol-anchored glycoprotein involved in directional root growth. *Plant Cell* 14, 1635–1648. doi: 10.1105/tpc.002360

Shi, H., Kim, Y., Guo, Y., Stevenson, B., and Zhu, J. (2003). The *Arabidopsis* SOS5 locus encodes a putative cell surface adhesion protein and is required for normal cell expansion. *Plant Cell* 15, 19–32. doi: 10.1105/tpc.007872

Showalter, A. M., and Basu, D. (2016). Extensin and arabinogalactan-protein biosynthesis: glycosyltransferases, research challenges, and biosensors. *Front. Plant Sci.* 7:814. doi: 10.3389/fpls.2016.00814

Sikorska, N., Lemus, L., Aguilera-Romero, A., Manzano-Lopez, J., Riezman, H., Muñiz, M., et al. (2016). Limited ER quality control for GPI-anchored proteins. *J. Cell Biol.* 213, 693–704. doi: 10.1083/jcb.201602010

Strasser, R. (2016). Plant protein glycosylation. *Glycobiology* 26, 926–939. doi: 10.1093/glycob/cww023

Strasser, R., Bondili, J., Vavra, U., Schoberer, J., Svoboda, B., Glössl, J., et al. (2007). A unique beta1,3-galactosyltransferase is indispensable for the biosynthesis of N-glycans containing Lewis a structures in *Arabidopsis thaliana*. *Plant Cell* 19, 2278–2292. doi: 10.1105/tpc.107.052985

Strasser, R., Stadlmann, J., Svoboda, B., Altmann, F., Glössl, J., and Mach, L. (2005). Molecular basis of N-acetylglucosaminyltransferase I deficiency in *Arabidopsis thaliana* plants lacking complex N-glycans. *Biochem. J.* 387, 385–391. doi: 10.1042/BJ20041686

Takahashi, M., Inoue, N., Ohishi, K., Maeda, Y., Nakamura, N., Endo, Y., et al. (1996). PIG-B, a membrane protein of the endoplasmic reticulum with a large lumenal domain, is involved in transferring the third mannose of the GPI anchor. *EMBO J.* 15, 4254–4261. doi: 10.1002/j.1460-2075.1996.tb00800.x

Takahashi, D., Kawamura, Y., and Uemura, M. (2016). Cold acclimation is accompanied by complex responses of glycosylphosphatidylinositol (GPI)-anchored proteins in *Arabidopsis*. *J. Exp. Bot.* 67, 5203–5215. doi: 10.1093/jxb/erw279

Taron, B. W., Colussi, P. A., Wiedman, J. M., Orlean, P., and Taron, C. H. (2004). Human Smp3p adds a fourth mannose to yeast and human glycosylphosphatidylinositol precursors in vivo. *J. Biol. Chem.* 279, 36083–36092. doi: 10.1074/jbc.M405081200

Wang, Y., Maeda, Y., Liu, Y. S., Takada, Y., Ninomiya, A., Hirata, T., et al. (2020). Cross-talks of glycosylphosphatidylinositol biosynthesis with glycosphingolipid biosynthesis and ER-associated degradation. *Nat. Commun.* 11:860. doi: 10.1038/s41467-020-14678-2

Xue, H., Veit, C., Abas, L., Tryfona, T., Maresch, D., Ricardi, M. M., et al. (2017). *Arabidopsis thaliana* FLA4 functions as a glycan-stabilized soluble factor via its carboxy-proximal fasciclin 1 domain. *Plant J.* 91, 613–630. doi: 10.1111/tpj.13591

Yeats, T. H., Bacic, A., and Johnson, K. L. (2018). Plant glycosylphosphatidylinositol anchored proteins at the plasma membrane-cell wall nexus. *J. Integr. Plant Biol.* 60, 649–669. doi: 10.1111/jipb.12659

Zeleny, R., Altmann, F., and Praznik, W. (1997). A capillary electrophoretic study on the specificity of β-galactosidases from *Aspergillus oryzae*, *Escherichia*

coli, *Streptococcus pneumoniae* and *Canavalia ensiformis* (Jack bean). *Anal. Biochem.* 246, 96–101. doi: 10.1006/abio.1996.9973

Zhou, K. (2019). Glycosylphosphatidylinositol-anchored proteins in *Arabidopsis* and one of their common roles in signaling transduction. *Front. Plant Sci.* 10:1022. doi: 10.3389/fpls.2019.01022

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer VK declared a past co-authorship with several of the authors FA and RS to the handling editor.

GPI-anchor synthesis in plants – a glycobiology perspective

Gernot Beihammer, Daniel Maresch, Friedrich Altmann, Richard Strasser

Supplementary Material

Contents

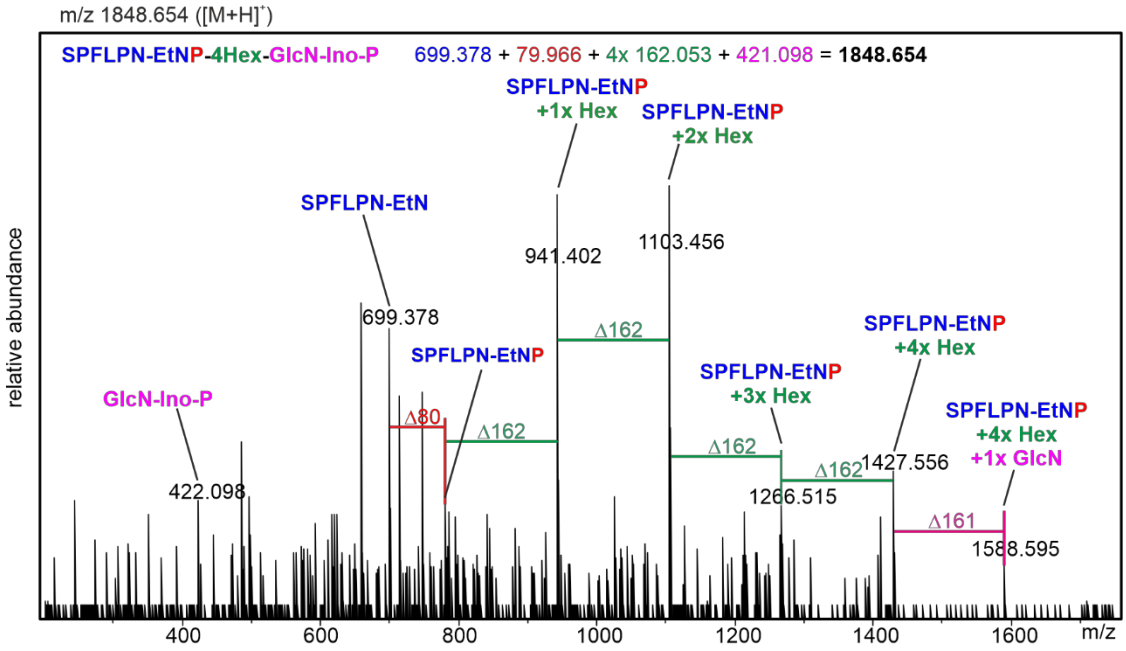


Figure S1. LC-ESI-MS/MS analysis of the C-terminal peptide from *N. benthamiana* RFP-COB1. RFP-COB1 was transiently expressed by agroinfiltration of leaves, captured by binding to RFP-Trap resin and subjected to PI-PLC and trypsin digestion. A representative analysis from two biological replicates is shown.

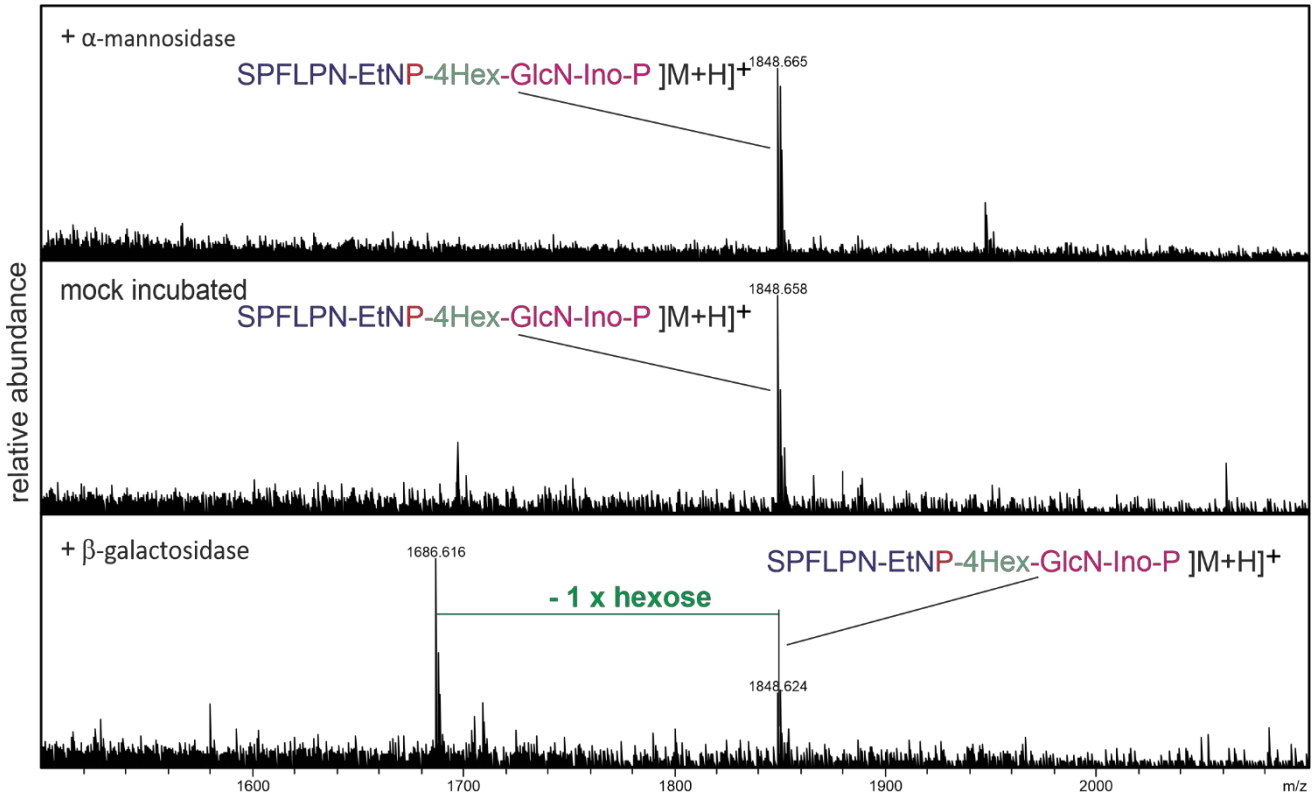


Figure S2. LC-ESI-MS analysis of α-mannosidase or β-galactosidase digested GPI-anchor derived from *N. benthamiana* expressed RFP-COB1. The β-galactosidase digestion was carried out two times with identical results. A representative analysis is shown.

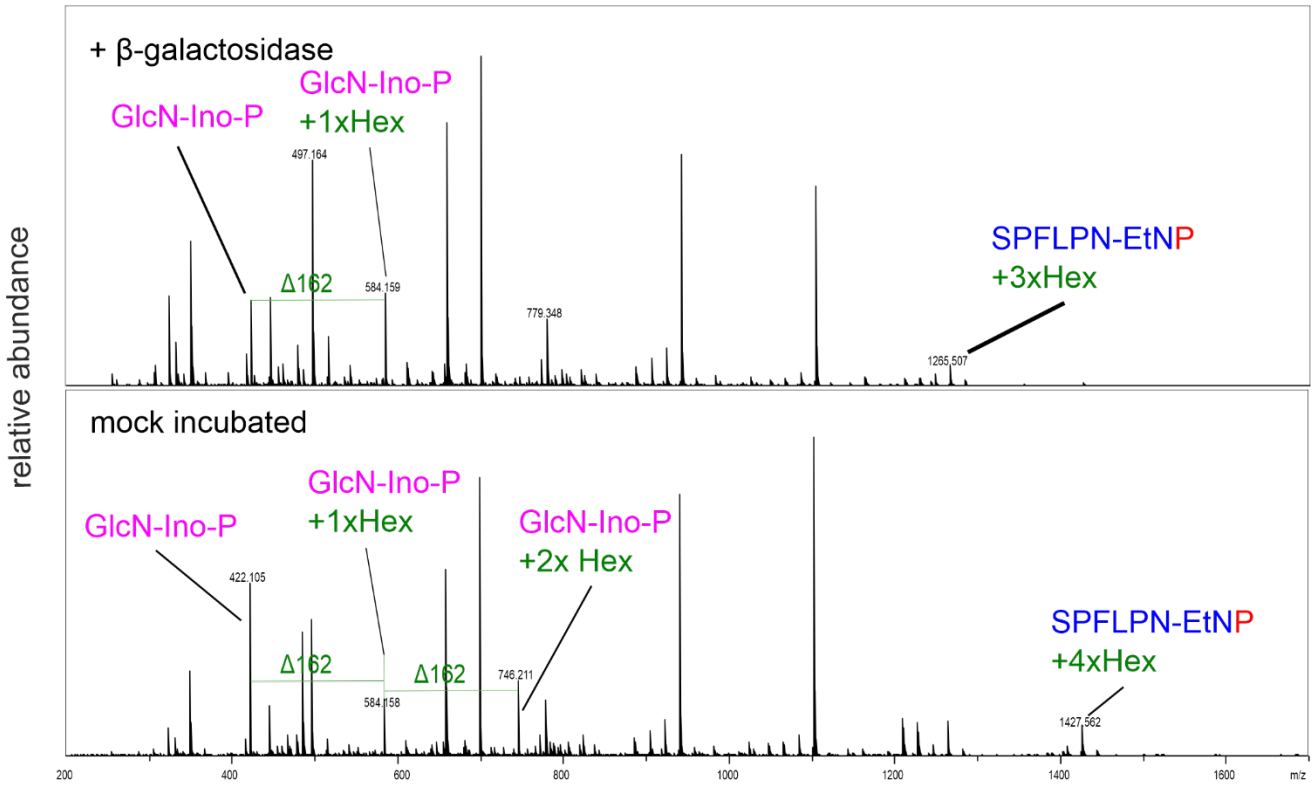


Figure S3. LC-ESI-MS/MS analysis of β-galactosidase digested and mock-incubated GPI-anchor derived from RFP-COB1 expressed in *A. thaliana* seedlings. A representative analysis from two technical replicates is shown.

Supplementary Methods

The C-terminal peptide of COBRA_ARATH was amplified from genomic DNA of *A. thaliana* using primers cob1_5F (5'-TATAGGATCCCCTTTTCTTCCCAACGGTGGTTCC-3')/cob1_6R (5'-TATACTCGAGTTAGGCAGAGAAGAAGAAAAAGAC-3') and cloned into a modified version of the pPT2-vector (Strasser *et al.* 2005) containing the Arabidopsis ubiquitin10-promotor and N-terminal mRFP. Reporter proteins were either purified from leaves of *N. benthamiana* following agroinfiltration or from 14-day old seedings of stably transformed *A. thaliana*. Plant samples were snap frozen, grinded and the membrane fraction collected by centrifugation in 100 mM Tris, 25 %(w/v) sucrose and 5 % (v/v) glycerol buffer (pH 7.4) at 65000 g for 1 h at 4 °C. After washing in 100 mM Tris, 150 mM NaCl and 0.5 mM EDTA buffer (pH 7.4), the membrane pellet was resuspended in 20 mM Tris, 10 mM NaCl, 0.1 mM $CaCl_2$, 1% (v/v) Phosphoinositide Phospholipase C (PIPLC, P6466, Thermo Fisher Scientific) (pH7.4) and incubated for 15 min. After cleavage of lipids from RFP-GPI, the protein was purified using RFP trap (Chromotek), separated by SDS-PAGE, excised from the gel, reduced, S-carbamidomethylated and subjected to tryptic digest as described previously (Kolarich and Altmann 2000). The resulting GPI-bearing peptide fragment was analyzed via LC-ESI-MS/MS using a Bruker Maxis 4G Q-TOF instrument. For enzymatic deglycosylation of the GPI anchor, either β-galactosidase from *Aspergillus oryzae* (Zeleny *et al.* 1997) or α-mannosidase from *Canavalia ensiformis* (M7257, Sigma-Aldrich) was used. RFP-GPI was purified once from *A. thaliana* and two times from *N. benthamiana*.

References

Kolarich, D. and Altmann, F. (2000). N-Glycan analysis by matrix-assisted laser desorption/ionization mass spectrometry of electrophoretically separated nonmammalian proteins: application to peanut allergen Ara h 1 and olive pollen allergen Ole e 1. Anal Biochem. 285, 64-75.

Strasser, R., Stadlmann, J., Svoboda, B., Altmann, F., Glössl, J. and Mach, L. (2005). Molecular basis of N-acetylglucosaminyltransferase I deficiency in *Arabidopsis thaliana* plants lacking complex N-glycans. Biochem J. 387, 385-91.

Zeleny, R., Altmann, F. and Praznik, W. (1997). A capillary electrophoretic study on the specificity of beta-galactosidases from *Aspergillus oryzae*, *Escherichia coli*, *Streptococcus pneumoniae*, and *Canavalia ensiformis* (jack bean). Anal Biochem. 246, 96-101.

Discussion

The aim of this thesis was to gain information about the role of Lewis A structures in plants and to elucidate the structure of the plant GPI anchor. In a first step, Lewis A bearing glycoproteins were identified from different tissues of *A. thaliana*, *N. benthamiana* and *O. sativa* after an affinity purification using the JIM84 antibody. This antibody was raised against a cell wall extract from carrot (*Daucus carota*) suspension culture (Horsley et al., 1993) and recognizes Lewis a bearing plant N-glycans (Fitchette et al., 1999). The exact binding epitope for this antibody is, however, unknown. In *A. thaliana*, the identified proteins primarily had a function in secondary cell wall biosynthesis. Similarly, identified Lewis A bearing proteins from *N. benthamiana* were involved in cell wall biosynthesis and organization as well as anisotropic growth. Proteins identified in *O. sativa* on the other hand were of diverse functions.

Arabidopsis knockout lines of some of the identified proteins have been characterized in previous studies. Mutant lines lacking functional COBRA-like protein 4 (COBL4), *irx6*, Protein trichome birefringence-like 3 (TBL3), *tbr*, or KORRIGAN, *rsw2-1*, contained decreased levels of cellulose (Brown et al., 2005; Bischoff et al., 2010; Lane et al., 2001), while a double knockout of LACCASE4 and LACCASE17 showed decreased levels of lignin when grown under continuous light (Berthet et al., 2011). Despite this apparent bias of Lewis A bearing glycoproteins towards secondary cell wall biosynthesis, Lewis A deficient plants developed normally and showed no alterations in constituents of the secondary cell wall. Based on these results, it appears that the presence of Lewis A on these proteins is not necessary for their function. This is in agreement with a previous study conducted on KORRIGAN (Liebminger et al., 2013), in which it was shown that the presence of complex N-glycans on recombinantly produced KORRIGAN was not required for its function.

Interestingly, a search of the 1001 genome project for accessions containing mutations in the gene encoding GALT1 suggested that Lewis A structures are evolutionarily conserved among natural accessions of *A. thaliana*. Only one line devoid of Lewis A bearing N-glycans could be identified among the investigated lines. Based on this, it can be assumed that presence of Lewis A provides the plants with an advantage in terms of fitness. The accessions investigated in this study had originally been picked in very different environments, thus allowing no conclusion as to what function Lewis A structures play within the plants.

Besides investigating Lewis A bearing N-glycans, an analysis of the glycan moiety of plant GPI anchors was conducted. Up to this point, the structure of only one plant GPI anchor, namely from *P. communis* (Oxley and Bacic, 1999), had been solved. A reporter construct consisting of mRFP N-terminally linked to the GPI attachment signal of COBRA protein from *A. thaliana* was cloned and could be expressed in seedlings of *A. thaliana* and leaves of *N. benthamiana*. Purification via the mRFP tag and subsequent digest with PIPLC and trypsin yielded a product consisting of the peptide fragment SPFLPN, EtNP, 4 hexoses, GlcN and inositol phosphate. The fourth hexose was found to be a galactose, most likely linked to the first mannose, based on results of a β-galactosidase digest. Interestingly, the shift upon β-galactosidase treatment was not complete for the *N. benthamiana* derived reporter protein. Thus, it cannot be excluded that a fraction of the GPI anchors has a different hexose attached, which is not mannose or galactose. Alternatively, the fraction of GPI anchors resistant to β-galactosidase treatment could bear a modification that inhibits cleavage of the galactose and was not detected by mass spectrometry due to technical limitations of the method. The structure of the GPI glycan moiety found in *A. thaliana* and *N. benthamiana* corresponds to the one elucidated for *P. communis*. This could hint towards a conserved structure among plants, but further analysis of different, evolutionary more distant, plant species is required. Another interesting topic to address in future studies is the composition of the lipid moiety, which was not investigated here.

References

Bellai-Dussault K, Nguyen TTM, Baratang N V., Jimenez-Cruz DA, Campeau PM. Clinical variability in inherited glycosylphosphatidylinositol deficiency disorders. *Clin Genet*. 2019;95(1):112-121. doi:10.1111/cge.13425

Benghezal M, Benachour A, Rusconi S, Aebi M, Conzelmann A. Yeast Gpi8p is essential for GPI anchor attachment onto proteins. *EMBO J*. 1996;15(23):6575-6583. doi:10.1002/j.1460-2075.1996.tb01048.x

Berthet S, Demont-Caulet N, Pollet B, et al. Disruption of LACCASE4 and 17 results in tissue-specific alterations to lignification of Arabidopsis thaliana stems. *Plant Cell*. 2011;23(3):1124-1137. doi:10.1105/tpc.110.082792

Bischoff V, Nita S, Neumetzler L, et al. TRICHOME BIREFRINGENCE and its homolog AT5G01360 encode plant-specific DUF231 proteins required for cellulose biosynthesis in arabidopsis. *Plant Physiol*. 2010;153(2):590-602. doi:10.1104/pp.110.153320

Borner GHH, Lilley KS, Stevens TJ, Dupree P. Identification of glycosylphosphatidylinositol-anchored proteins in Arabidopsis. A proteomic and genomic analysis. *Plant Physiol*. 2003;132(2):568-577. doi:10.1104/pp.103.021170

Brown DM, Zeef LAH, Ellis J, Goodacre R, Turner SR. Identification of novel genes in Arabidopsis involved in secondary cell wall formation using expression profiling and reverse genetics. *Plant Cell*. 2005;17(8):2281-2295. doi:10.1105/tpc.105.031542

Cherepanova NA, Venev S V., Leszyk JD, Shaffer SA, Gilmore R. Quantitative glycoproteomics reveals new classes of STT3A- And STT3B-dependent N-glycosylation sites. *J Cell Biol*. 2019;218(8):2782-2796. doi:10.1083/JCB.201904004

Codini M, Garcia-Gil M, Albi E. Cholesterol and sphingolipid enriched lipid rafts as therapeutic targets in cancer. *Int J Mol Sci*. 2021;22(2):1-17. doi:10.3390/ijms22020726

Desnoyer N, Palanivelu R. Bridging the GAPs in plant reproduction: a comparison of plant and animal GPI-anchored proteins. *Plant Reprod*. 2020;33(3-4):129-142. doi:10.1007/s00497-020-00395-9

Elortza F, Nühse TS, Foster LJ, Stensballe A, Peck SC, Jensen ON. Proteomic analysis of glycosylphosphatidylinositol-anchored membrane proteins. *Mol Cell Proteomics*. 2003;2(12):1261-1270. doi:10.1074/mcp.M300079-MCP200

Farid A, Pabst M, Schoberer J, Vienna LS, Altmann F, Vienna LS. Arabidopsis thaliana alpha1,2-glucosyltransferase (ALG10) is required for efficient N-glycosylation and leaf growth. 2011;(June):313-325. doi:10.1111/j.1365-313X.2011.04688.x

Fitchette AC, Cabanes-Macheteau M, Marvin L, et al. Biosynthesis and immunolocalization of Lewis a-containing N-glycans in the plant cell. *Plant Physiol*. 1999;121(2):333-343. doi:10.1104/pp.121.2.333

Fujita M, Jigami Y. Lipid remodeling of GPI-anchored proteins and its function. *Biochim Biophys Acta - Gen Subj*. 2008;1780(3):410-420. doi:10.1016/j.bbagen.2007.08.009

Fujita M, Maeda Y, Ra M, Yamaguchi Y, Taguchi R, Kinoshita T. GPI Glycan Remodeling by PGAP5 Regulates Transport of GPI-Anchored Proteins from the ER to the Golgi. *Cell*. 2009;139(2):352-365. doi:10.1016/j.cell.2009.08.040

Gillmor C, Lukowitz W, Brininstool G, Sedbrook JC, Hamann T, Poindexter P, et al. Glycosylphosphatidylinositol-anchored proteins are required for cell wall synthesis and morphogenesis in arabidopsis. Plant Cell. 2005;17(4):1128–40. doi:10.1105/tpc.105.031815

Hong Y, Ohishi K, Kang JY, Tanaka S, Inoue N, Nishimura J, Maeda Y, Kinoshita T. Human PIG-U and Yeast Cdc91p Are the Fifth subunit of GPI Transamidase That Attaches GPI-Anchors to Proteins. *Molecular Biology of the Cell* 2003;14:1780-1789. doi: 10.1091/mbc.e02-12-0794.

Horsley D, Coleman J, Evans D, Crooks K, Peart J, Satiat-Jeunmaitre B, Hawes C. A Monoclonal Antibody, JIM 84, Recognizes the Golgi Apparatus and Plasma Membrane in Plant Cells. *Journal of Experimental Botany* 1993;44: 223-229.

Jarrell KF, Ding Y, Meyer BH, Albers SV, Kaminski L, Eichler J. N-Linked Glycosylation in *Archaea*: a Structural, Functional, and Genetic Analysis. *Microbiology and Molecular Biology Reviews.* 2014;78:304-341. doi:10.1128/MMBR.00052-13

Kanzawa N, Maeda Y, Ogiso H, Murakami Y, Taguchi R, Kinoshita T. Peroxisome dependency of alkyl-containing GPI-anchor biosynthesis in the endoplasmic reticulum. *Proc Natl Acad Sci U S A*. 2009;106(42):17711-17716. doi:10.1073/pnas.0904762106

Kinoshita T. Biosynthesis and biology of mammalian GPI-anchored proteins. *Open Biol*. 2020;10(3). doi:10.1098/rsob.190290

Kinoshita T, Fujita M. Biosynthesis of GPI-anchored proteins : special emphasis on GPI lipid remodeling. 2016;57. doi:10.1194/jlr.R063313

Kobayashi T, Nishizaki R, Ikezawa H. The presence of CPI-linked protein(s) in an archaeobacterium, Sulfolobus acidocaldarius, closely related to eukaryotes. *Biochim Biophys Acta - Gen Subj*. 1997;1334(1):1-4. doi:10.1016/S0304-4165(96)00099-2

Koprivova A, Altmann F, Gorr G, Kopriva S, Reski R, Decker EL. N-Glycosylation in the Moss Physcomitrella patens is Organized Similarly to that in Higher Plants. *Plant Biol*. 2003;5(6):582-591. doi:10.1055/s-2003-44721

Lalanne E, Honys D, Johnson A, et al. SETH1 and SETH2, Two Components of the Glycosylphosphatidylinositol Anchor Biosynthetic Pathway, Are Required for Pollen Germination and Tube Growth in Arabidopsis. *Plant Cell*. 2004;16(1):229-240. doi:10.1105/tpc.014407

Lane DR, Wiedemeier A, Peng L, et al. Temperature-sensitive alleles of rsw2 link the KORRIGAN endo-1,4-β-glucanase to cellulose synthesis and cytokinesis in Arabidopsis. *Plant Physiol*. 2001;126(1):278-288. doi:10.1104/pp.126.1.278

Li C, Yeh F, Cheung AY, et al. Glycosylphosphatidylinositol-anchored proteins as chaperones and co-receptors for FERONIA receptor kinase signaling in Arabidopsis. 2015;1:1-21. doi:10.7554/eLife.06587

Liebminger E, Grass J, Altmann F, Mach L, Strasser R. Characterizing the link between glycosylation state and enzymatic activity of the endo-β1,4-glucanase KORRIGAN1 from Arabidopsis thaliana. *J Biol Chem*. 2013;288(31):22270-22280. doi:10.1074/jbc.M113.475558

Lu CF, Montijn RC, Brown JL, et al. Glycosyl phosphatidylinositol-dependent cross-linking of α-agglutinin and β1,6-glucan in the Saccharomyces cerevisiae cell wall. *J Cell Biol*. 1995;128(3):333-340. doi:10.1083/jcb.128.3.333

Maeda M, Tani M, Yoshiie T, Vavricka CJ, Kimura Y. Structural features of N -glycans linked to glycoproteins expressed in three kinds of water plants : Predominant occurrence of the plant complex type N -glycans bearing Lewis a epitope. *Carbohydr Res*. 2016;435:50-57. doi:10.1016/j.carres.2016.09.008

Maeda Y, Ra M, Yamaguchi Y, Taguchi R, Kinoshita T., Fujita M. GPI Glycan Remodeling by PGAP5 Regulates Transport of GPI-Anchored Proteins from the ER to the Golgi. *Cell* 2007;139(2):352–65. doi.10.1091/mbc.E06–10–0885

Murakami Y, Siripanyaphinyo U, Hong Y, Tashita Y, Maeda Y, Kinoshita H. The Initial Enzyme for Glycosylphosphatidylinositol Biosynthesis Requires PIG-Y, a Seventh Component. *Molecular Biology of the Cell* 2005 (16): 5236-5246. doi.10.1091/mbc.E05-08-0743.

Nagamune K, Nozaki T, Maeda Y, et al. Critical roles of glycosylphosphatidylinositol for Trypanosoma brucei. *Proc Natl Acad Sci U S A*. 2000;97(19):10336-10341. doi:10.1073/pnas.180230697

Niu G, Shao Z, Liu C, Chen T, Jiao Q, Hong Z. Comparative and evolutionary analyses of the divergence of plant oligosaccharyltransferase STT3 isoforms. *FEBS Open Bio*. 2020;10(3):468-483. doi:10.1002/2211-5463.12804

Nothaft H, Szymanski CM. Bacterial Protein N -Glycosylation : New Perspectives and Applications *. *The Journal of Biological Chemistry* 2013.288(10):6912-6920. doi:10.1074/jbc.R112.417857

Oxley D, Bacic A. Structure of the glycosylphosphatidylinositol anchor of an arabinogalactan protein from Pyrus communis suspension-cultured cells. *Proc Natl Acad Sci U S A*. 1999.96(25):14246-14251. doi:10.1073/pnas.96.25.14246

Pittet M, Conzelmann A. Biosynthesis and function of GPI proteins in the yeast Saccharomyces cerevisiae. *Biochim Biophys Acta - Mol Cell Biol Lipids*. 2007.1771(3):405-420. doi:10.1016/j.bbalip.2006.05.015

Rips S, Frank M, Elting A, Niklas J., von Schaewen A. Golgi α 1,4-fucosyltransferase of Arabidopsis thaliana partially localizes at the nuclear envelope. *Traffic* 2017.18:646-657. doi:10.1111/tra.12506

Schoberer J, Liebminger E, Botchway SW, Strasser R, Hawes C. Time-Resolved Fluorescence Imaging Reveals Differential Interactions of N -Glycan Processing

Enzymes across the Golgi Stack in Planta. *Plant Physiology* 2013;161:1737-1754. doi:10.1104/pp.112.210757

Schoberer J, Strasser R. Sub-Compartmental Organization of Golgi-Resident N-Glycan Processing Enzymes in Plants. *Molecular Plant* 2011;4(2). doi:10.1093/mp/ssq082

Sedbrook JC, Carroll KL, Hung KF, Masson PH, Somerville CR. The Arabidopsis SKU5 gene encodes an extracellular glycosyl phosphatidylinositol-anchored glycoprotein involved in directional root growth. *Plant Cell*. 2002.14(7):1635-1648. doi:10.1105/tpc.002360

Shishioh N, Hong Y, Ohishi K, Ashida H, Maeda Y, Kinoshita T. GPI7 is the second partner of PIG-F and involved in modification of glycosylphosphatidylinositol. *J Biol Chem*. 2005.280(10):9728-9734. doi:10.1074/jbc.M413755200

Shrimal S, Cherepanova NA, Gilmore R. DC2 and KCP2 mediate the interaction between the oligosaccharyltransferase and the ER translocon. *J Cell Biol*. 2017.216(11):3625-3638. doi:10.1083/jcb.201702159

Strasser R, Bondili JS, Vavra U, Schoberer J, Svoboda B, Glössl J, et al. A unique β1,3-galactosyltransferase is indispensable for the biosynthesis of N-glycans containing Lewis a structures in Arabidopsis thaliana. *Plant Cell* 2007.19(7):2278–92. doi:10.1105/tpc.107.052985

Strasser R. Protein Quality Control in the Endoplasmic Reticulum of Plants. *Annu Rev Plants Biol Plant Biol*. 2018.69:147-172. doi:10.1146/annurev-arplant-042817-040331

Takahashi D, Kawamura Y, Uemura M. Cold acclimation is accompanied by complex responses of glycosylphosphatidylinositol (GPI)-anchored proteins in Arabidopsis. *J Exp Bot*. 2016;67(17):5203-5215. doi:10.1093/jxb/erw279

Taron BW, Colussi PA, Wiedman JM, Orlean P, Taron CH. Human Smp3p adds a fourth mannose to yeast and human glycosylphosphatidylinositol precursors in vivo. *J Biol Chem*. 2004.279(34):36083-36092. doi:10.1074/jbc.M405081200

Tsai YH, Liu X, Seeberger PH. Chemical biology of glycosylphosphatidylinositol anchors. *Angew Chemie - Int Ed*. 2012.51(46):11438-11456. doi:10.1002/anie.201203912

Wang P, Wang H, Gai J, et al. Evolution of protein N -glycosylation process in Golgi apparatus which shapes diversity of protein N -glycan structures in plants, animals and fungi. *Scientific Reports* 2017.7, 40301. doi:10.1038/srep40301

Watanabe R, Ohishi K, Maeda Y, Nakamura N, Kinoshita T. Mammalian PIG-L and its yeast homologue Gpi12p are N-acetylglucosaminylphosphatidylinositol de-N-acetylases essential in glycosylphosphatidylinositol biosynthesis. *Biochem J*. 1999.339(1):185-192. doi:10.1042/0264-6021:3390185

Wille H, Requena JR. The structure of PrPsc prions. *Pathogens* 2018.7(1):1-11. doi:10.3390/pathogens7010020

Wilson IBH, Zeleny R, Kolarich D, Staudacher E, Stroop CJM. Analysis of Asn-linked glycans from vegetable foodstuffs: widespread occurrence of Lewis a , core α 1,3 -

linked fucose and xylose substitutions. *Glycobiology* 2001.11(4): 261-274. doi:10.1093/glycob/11.4.261

Xue H, Veit C, Abas L, et al. Arabidopsis thaliana FLA4 functions as a glycan-stabilized soluble factor via its carboxy-proximal Fasciclin 1 domain. *Plant J*. 2017.91(4):613-630. doi:10.1111/tpj.13591

Yeats TH, Bacic A, Johnson KL. Plant glycosylphosphatidylinositol anchored proteins at the plasma membrane-cell wall nexus. *J Integr Plant Biol*. 2018.60(8):649-669. doi:10.1111/jipb.12659

Zeng W, Ford KL, Bacic A, Heazlewood JL. N-linked glycan micro-heterogeneity in glycoproteins of arabidopsis. *Mol Cell Proteomics* 2018;17(3):413-421. doi:10.1074/mcp.RA117.000165

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* 2011.72(10):1109-1123. doi:10.1016/j.phytochem.2010.10.019

Zhou K. Glycosylphosphatidylinositol-Anchored Proteins in Arabidopsis and One of Their Common Roles in Signaling Transduction. *Front Plant Sci*. 2019.10:1-20. doi:10.3389/fpls.2019.01022

Appendix I

An oligosaccharyltransferase from Leishmania major increases the N-glycan occupancy on recombinant glycoproteins produced in Nicotiana benthamiana

Castilho A, **Beihammer G**, Pfeiffer C, Göritzer K, Montero-Morales L, Vavra U, Maresch D, Grünwald-Gruber C, Altmann F, Steinkellner H, Strasser R

Original Research article

Plant Biotechnol J. 2018 Oct;16(10):1700-1709. doi: 10.1111/pbi.12906

Plant Biotechnology Journal

Plant Biotechnology Journal (2018) **16**, pp. 1700–1709 doi: 10.1111/pbi.12906

An oligosaccharyltransferase from *Leishmania major* increases the N-glycan occupancy on recombinant glycoproteins produced in *Nicotiana benthamiana*

Alexandra Castilho[1], Gernot Beihammer[1], Christina Pfeiffer[1], Kathrin Göritzer[1], Laura Montero-Morales[1], Ulrike Vavra[1], Daniel Maresch[2], Clemens Grünwald-Gruber[2], Friedrich Altmann[2], Herta Steinkellner[1] and Richard Strasser[1,*]

[1]*Department of Applied Genetics and Cell Biology, University of Natural Resources and Life Sciences, Vienna, Austria*
[2]*Department of Chemistry, University of Natural Resources and Life Sciences, Vienna, Austria*

Received 31 August 2017;
revised 15 December 2017;
accepted 6 February 2018.
*Correspondence (Tel 0043-1-47654-94145; fax 0043-1-47654-94009; email richard.strasser@boku.ac.at)

Keywords: glyco-engineering, N-glycosylation, *Nicotiana benthamiana*, oligosaccharyltransferase, plant-made pharmaceuticals.

Summary

N-glycosylation is critical for recombinant glycoprotein production as it influences the heterogeneity of products and affects their biological function. In most eukaryotes, the oligosaccharyltransferase is the central-protein complex facilitating the N-glycosylation of proteins in the lumen of the endoplasmic reticulum (ER). Not all potential N-glycosylation sites are recognized *in vivo* and the site occupancy can vary in different expression systems, resulting in underglycosylation of recombinant glycoproteins. To overcome this limitation in plants, we expressed LmSTT3D, a single-subunit oligosaccharyltransferase from the protozoan *Leishmania major* transiently in *Nicotiana benthamiana,* a well-established production platform for recombinant proteins. A fluorescent protein-tagged LmSTT3D variant was predominately found in the ER and co-located with plant oligosaccharyltransferase subunits. Co-expression of LmSTT3D with immunoglobulins and other recombinant human glycoproteins resulted in a substantially increased N-glycosylation site occupancy on all N-glycosylation sites except those that were already more than 90% occupied. Our results show that the heterologous expression of LmSTT3D is a versatile tool to increase N-glycosylation efficiency in plants.

Introduction

Asparagine (N)-linked glycosylation is a major co- and post-translational modification of proteins entering the secretory pathway. Many recombinant biopharmaceuticals for therapeutic use in humans are N-glycosylated, and distinct N-glycan structures play crucial roles for their *in vivo* efficacy (Jefferis, 2009; Zacchi and Schulz, 2016). Yet, the extent of N-glycan attachment to a distinct glycosylation site may vary greatly giving rise to the formation of incompletely glycosylated proteins with potentially unwanted characteristics. N-glycans are important for protein folding and protein stability and specifically modulate protein–protein interactions. For erythropoietin (EPO), it has been shown that the *in vivo* biological activity correlates with the number of N-linked glycans (Elliott *et al.*, 2004) and nonglycosylated monoclonal antibodies display reduced or complete loss of immune receptor binding (Nose and Wigzell, 1983; Walker *et al.*, 1989).

In all eukaryotes, a hallmark of N-glycosylation is the *en bloc* transfer of a common preassembled oligosaccharide ($Glc_3Man_9GlcNAc_2$) from the lipid carrier dolichol pyrophosphate to selected asparagine residues in the sequence Asn-X-Ser/Thr (X any amino acid except proline) of nascent polypeptides (Aebi, 2013; Zielinska *et al.*, 2010). The transfer of the oligosaccharide takes place in the lumen of the ER and is catalysed by the oligosaccharyltransferase (OST) complex. In yeast and mammals, OST is a multimeric membrane-bound protein complex (Kelleher and Gilmore, 2006) consisting of one catalytically active subunit (STT3) and several different noncatalytic subunits that contribute to N-glycosylation by regulation of the substrate specificity, stability or assembly of the complex (Knauer and Lehle, 1999; Mohorko *et al.*, 2011; Yan and Lennarz, 2002). The organization of the OST complex is more complex in metazoans than in yeast, and different subunit compositions have been described (Mohorko *et al.*, 2011; Roboti and High, 2012; Shibatani *et al.*, 2005). Mammals harbour two different catalytic STT3 isoforms (STT3A and STT3B) that are present in distinct OST complexes (Ruiz-Canada *et al.*, 2009; Shrimal *et al.*, 2013, 2015). The STT3A/STT3B-containing complexes have overlapping and isoform specific functions and differ in their catalytic activity and acceptor substrate selectivity. While STT3A is predominately involved in co-translational glycosylation, STT3B displays a preference for post-translational glycosylation. By contrast, some unicellular parasites like *Leishmania major* or *Trypanosoma brucei* have several STT3 copies, but lack other noncatalytic subunits of the yeast or mammalian OST complex (Kelleher and Gilmore, 2006; Samuelson *et al.*, 2005). These single-subunit OST enzymes display distinct protein acceptor and oligosaccharide donor specificities (Izquierdo *et al.*, 2009; Nasab *et al.*, 2008).

N-glycosylation in plants requires a similar heteromeric OST complex, which is still poorly described (Strasser, 2016). *Arabidopsis thaliana* has two catalytic subunits, termed STT3A and STT3B (Koiwa *et al.*, 2003). STT3A-deficient plants are viable, but display a protein underglycosylation defect that disturbs the biogenesis of different proteins including the heavily glycosylated pattern recognition receptor EF-TU RECEPTOR (EFR), the endo-β1,4-glucanase KORRIGAN1 (KOR1/RSW2) (Kang *et al.*, 2008) or

the myrosinase TGG1 (Koiwa *et al.*, 2003; Nekrasov *et al.*, 2009; Saijo *et al.*, 2009). Moreover, the *A. thaliana stt3a stt3b* double knockout mutant is gametophytic lethal (Koiwa *et al.*, 2003) highlighting the importance of the catalytic OST subunits for protein N-glycosylation in plants.

Plants are increasingly used as production hosts for recombinant human glycoproteins intended for therapeutic use. The majority of the approved recombinant biopharmaceuticals like monoclonal antibodies are glycoproteins and N-glycosylation modulates, for example, the IgG function by affecting the binding affinity to receptors on immune cells. In recent years, enormous efforts have been made to engineer plant-based expression hosts for the production of glycoproteins with targeted glycan profiles (Dicker *et al.*, 2016; Hanania *et al.*, 2017; Kallolimath *et al.*, 2016; Li *et al.*, 2016; Limkul *et al.*, 2016; Loos *et al.*, 2014; Mercx *et al.*, 2017; Strasser *et al.*, 2014). These advances reduced the variation of N-glycan structures on a given site (microheterogeneity) thereby contributing to product homogeneity and consistency. Besides a recent patent application (WO2014195011A1) little/no attempts have been made that address differences in N-glycosylation site occupancy on recombinant proteins (macroheterogeneity). Here, we present a strategy to overcome underglycosylation at N-glycosylation sites on different recombinant glycoproteins when transiently expressed in *N. benthamiana*. We found that the expression of the single-subunit oligosaccharyltransferase STT3D from *L. major* (LmSTT3D) substantially improves the N-glycosylation efficiency on different transiently expressed recombinant glycoproteins.

Results

Recombinant IgG and an Fc-fusion protein display considerable underglycosylation

Previous studies have shown that the single N-glycosylation site at position Asn297 from the heavy chain of different recombinant IgG molecules is frequently underglycosylated when transiently expressed in *N. benthamiana* (Bendandi *et al.*, 2010; Loos *et al.*, 2015; Strasser *et al.*, 2008; Zeitlin *et al.*, 2016) (Figure 1a). We expressed a monoclonal IgG antibody transiently in wild-type as well as in the glyco-engineered ΔXT/FT *N. benthamiana* line which is widely used as expression host for recombinant glycoproteins (Strasser *et al.*, 2008, 2014). LC-ESI-MS analysis of the proteolytically digested heavy chain showed the presence of considerable amounts of the unglycosylated peptide in both expression hosts (Figure 1b). To investigate this variation in N-glycosylation site occupancy more in detail and to better visualize the difference between glycosylated and nonglycosylated variants, we generated an expression construct where the Fc-domain from the IgG heavy chain lacking a variable region is fused to a signal peptide for targeting to the secretory pathway (SP-Fc). Upon SDS-PAGE under reducing conditions and subsequent immunoblotting, the expressed SP-Fc protein migrates at approximately 35 kDa and a faster migrating band at approximately 33 kDa is clearly detectable (Figure 1c). When digested with PNGase F which cleaves off the single N-glycan, the deglycosylated band co-

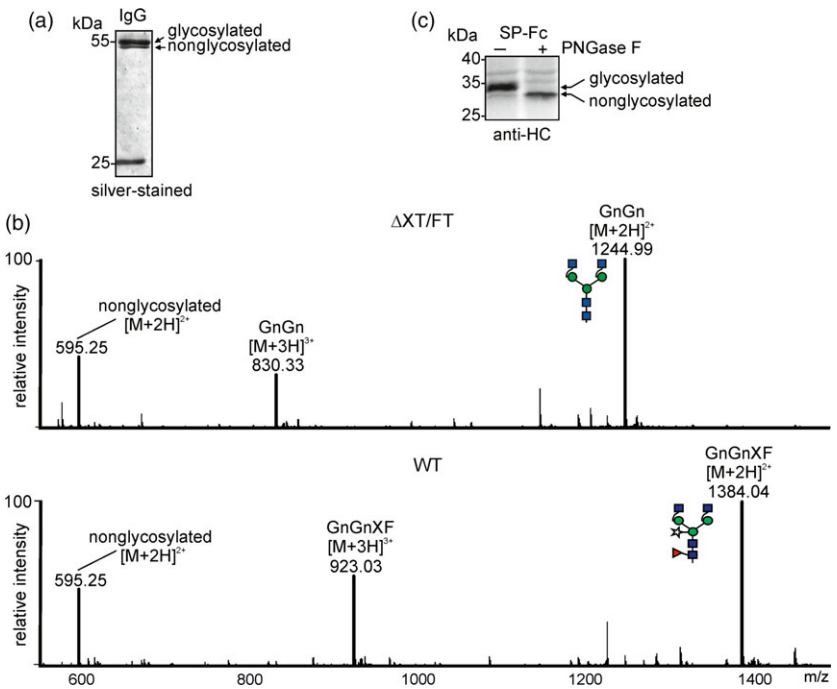


Figure 1 Underglycosylation is observed on transiently expressed IgG and on SP-Fc. (a) A monoclonal IgG antibody was transiently expressed in *N. benthamiana* wild-type plants. The IgG protein was purified 2 days after infiltration, separated by SDS-PAGE and silver-stained. The presence of glycosylated and nonglycosylated variants is indicated. The 25 kDa band represents the light chain. (b) A monoclonal antibody purified from *N. benthamiana* wild-type (WT) or ΔXT/FT was digested with trypsin and subjected to LC-ESI-MS analysis. The mass $[M + 2H]^{2+}$ of the nonglycosylated peptide EEQYNSTYR carrying the Fc-N-glycosylation site (Asn297) and the major glycosylated peaks $[M + 2H]^{2+}$ and $[M + 3H]^{3+}$ are depicted. Peak labels were made according to the ProGlycAn system (www.proglycan.com), and the glycan illustrations are drawn according to the nomenclature from the Consortium for Functional Glycomics. (c) SP-Fc was expressed in *N. benthamiana* ΔXT/FT, and protein was extracted 24 h after infiltration and subjected to PNGase F digestion. Immunoblot detection was performed with anti-IgG antibodies.

migrates with the faster migrating band of the undigested SP-Fc. Mass spectrometry-based analysis of the glycosylation occupancy corroborates these findings for incomplete glycosylation (data not shown).

LmSTT3D from the protist *Leishmania major* is retained in the ER of plants

Leishmania major harbours four paralogues (termed LmSTT3A-D) of the single-subunit OST. In previous studies, it has been shown that the LmSTT3D from the protist *L. major* can rescue the growth and N-glycosylation defects observed in *Saccharomyces cerevisiae* lacking a functional STT3 protein (Nasab *et al.*, 2008) and improves N-glycosylation efficiency of recombinant proteins expressed in *Pichia pastoris* (Choi *et al.*, 2012). Consequently, we hypothesized that LmSTT3D activity may overcome the observed inefficient N-glycosylation of IgG in our plant-based expression system. To test this assumption, we generated a binary vector for expression of a codon-optimized LmSTT3D variant fused to GFP (LmSTT3D-GFP, Figure 2a) and transiently expressed the protein in *N. benthamiana*. According to the proposed topology model for LmSTT3D, the catalytic region close to the C-terminus faces the lumen of the ER, similar to the predictions for *A. thaliana* STT3A (Figure 2b). On immunoblots, a single band of expected size is detectable for LmSTT3D-GFP (Figure 2c). In wild-type leaf epidermal cells, ER-labelling was visible under the confocal microscope 1 day after infiltration (Figure 2d). Two and 3 days after infiltration, ER and additional puncta were detectable which represent Golgi bodies as well as undefined vesicular structures. In contrast to that, *A. thaliana* STT3A-GFP was only observed in the ER. Co-localization with the ER-resident OST4B-mRFP, a subunit of the plant oligosaccharyltransferase complex (Farid *et al.*, 2013), or the *cis*/medial Golgi-marker GnTI-mRFP (Schoberer *et al.*, 2013) confirmed the subcellular localization of LmSTT3D-GFP (Figure 2d) suggesting that LmSTT3D-GFP is incompletely retained in the ER.

LmSTT3D enhances the N-glycosylation occupancy of recombinant SP-Fc and IgG

In the next experiments, we examined whether LmSTT3D-GFP can improve the N-glycosylation efficiency of SP-Fc and IgG when transiently co-expressed in *N. benthamiana*. In the presence of LmSTT3D-GFP, the faster migrating band of SP-Fc disappeared, indicating an enhanced occupancy of glycosylation site Asn297 (Figure 3a and b). MS-based analysis of peptides/glycopeptides derived from proteolytically digested SP-Fc expressed in ΔXT/FT demonstrated that the co-expression of LmSTT3D-GFP drastically reduced the amount of the nonglycosylated variant (Figure 3c, Table 1). The N-glycan composition of the recombinantly expressed proteins was not altered upon LmSTT3D-GFP co-expression. The major N-glycan peak corresponds to processed complex N-glycans (GnGn: $GlcNAc_2Man_3GlcNAc_2$) (Figure 3c) indicating that LmSTT3D-GFP co-expression does not interfere with complex N-glycan processing of SP-Fc in the Golgi. The same result was obtained for an IgG co-expressed with LmSTT3D-GFP (Figure S1 and Table 1). Intact MS analysis of the fully assembled IgG 2G12 revealed further that in the absence of LmSTT3D-GFP, nonglycosylated as well as hemi-glycosylated (only one of the two heavy chains carries an N-glycan) forms are present. Co-expression of LmSTT3D-GFP leads to an increase in fully assembled IgG with two N-glycans, one attached to each heavy chain (Figure 4). In summary, our data show that LmSTT3D-GFP co-expression increases the N-glycosylation site occupancy of SP-Fc and IgG.

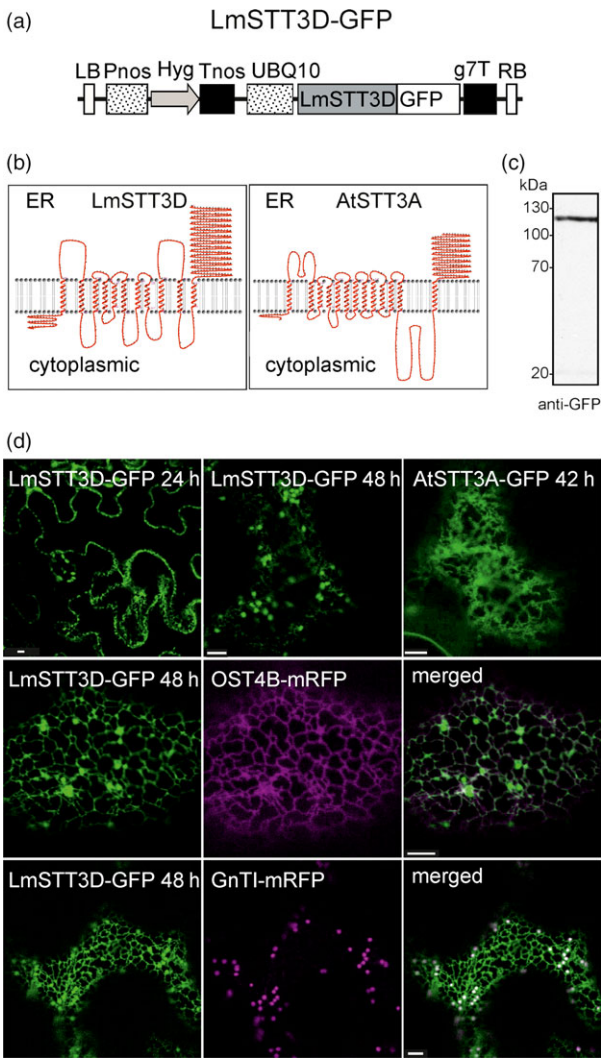


Figure 2 LmSTT3D-GFP accumulates in the ER and Golgi. (a) Schematic representation of the UBQ10:LmSTT3D-GFP expression vector. LB: left border; Pnos: nopaline synthase gene promoter; Hyg: hygromycin B phosphotransferase gene; Tnos: nopaline synthase gene terminator; UBQ10: *A. thaliana* ubiquitin-10 promoter; LmSTT3D: *L. major* catalytic OST subunit STT3D open reading frame; GFP: green fluorescent protein; g7T: agrobacterium gene 7 terminator; RB: right border. (b) Topology of LmSTT3D and *A. thaliana* STT3A (AtSTT3A). The transmembrane domain regions and their topology were obtained using the HMMTop prediction program (http://www.enzim.hu/hmmtop/). The illustration was generated using TMRPres2D visualization (http://bioinformatics.biol.uoa.gr/TMRPres2D/). (c) Immunoblot (with anti-GFP antibodies) of LmSTT3D-GFP transiently expressed in *N. benthamiana*. (d) LmSTT3D-GFP was either expressed alone or in combination with the ER-marker OST4B-mRFP and the Golgi-marker GnTI-mRFP in *N. benthamiana* leaf epidermal cells. Analysis of fluorescent proteins was carried out by confocal laser scanning microscopy at the indicated time after infiltration. Bars = 5 μm. Expression of AtSTT3A-GFP is shown for comparison.

LmSTT3D improves the N-glycosylation efficiency of different recombinant glycoproteins

We found that LmSTT3D-GFP co-expression is a suitable tool to increase the N-glycosylation efficiency of SP-Fc and IgG. To

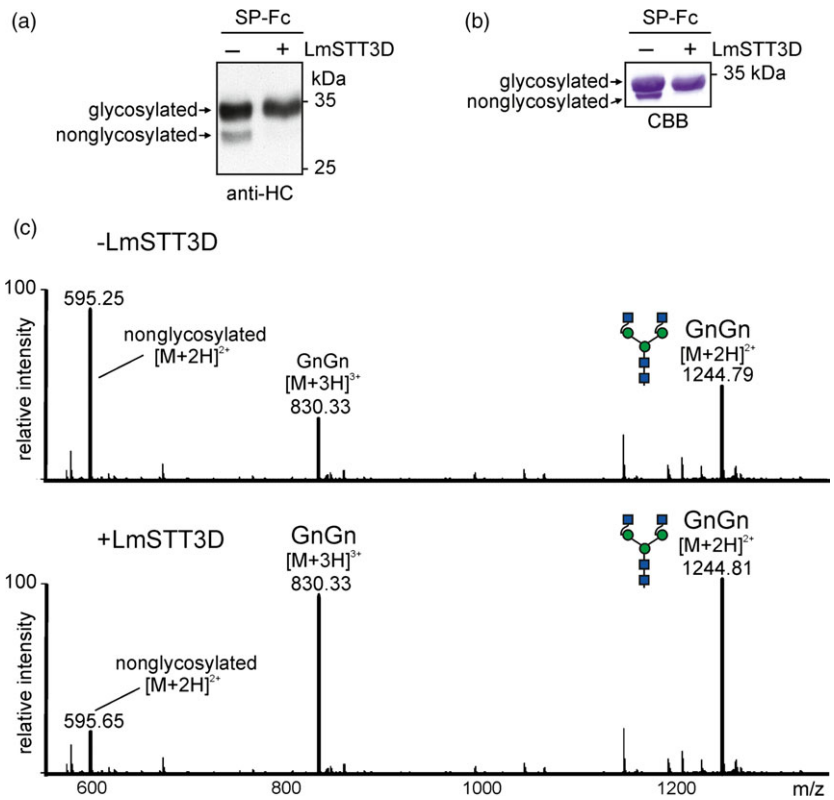


Figure 3 LmSTT3D-GFP co-expression increases the N-glycosylation site occupancy on SP-Fc. SP-Fc was transiently expressed in *N. benthamiana* leaves together with UBQ10:LmSTT3D-GFP. (a) Proteins were extracted from wild-type 2 days postinfiltration and subjected to SDS-PAGE and immunoblotting using anti-IgG heavy chain (HC) antibodies. (b) SDS-PAGE and Coomassie Brilliant Blue staining of SP-Fc purified from ΔXT/FT. (c) LC-ESI-MS analysis of tryptic glycopeptides from SP-Fc expressed in ΔXT/FT. In the shown spectra, the peak at 595.25 ($[M + 2H]^{2+}$) is assigned to the nonglycosylated peptide EEQYNSTYR, and the peaks at 830.33 ($[M + 3H]^{3+}$) and 1244.8 ($[M + 2H]^{2+}$) are assigned to the complex N-glycan GnGn (nomenclature according to the ProGlycAn system: www.proglycan.com).

Table 1 Comparison of the LmSTT3D-GFP effect on N-glycosylation of recombinant glycoproteins expressed in ΔXT/FT.

Protein	N-glycosylation site (GS)	% glycosylated	% glycosylated + LmSTT3D	% increase	Number of repetitions
Fc	GS1 NST	56 ± 3	93 ± 2	66	3
IgG	GS1 NST	87 ± 5	98 ± 4	13	3
IgE	GS3 NKT	20 ± 7	36 ± 3	80	2
	GS5 NLT	40 ± 11	91 ± 3	128	
	GS6 NHS	<2 ± 3	63 ± 4	>1000	
IgA1	GS1 NLT	96 ± 0	94 ± 4		2
	GS2 NVS	59 ± 1	95 ± 4	60	
EPO-Fc	GS1 NIT	60 ± 5	81 ± 3	35	2
	GS3 NSS	95 ± 2	93 ± 0		
	GS4 NST	90 ± 1	99 ± 2	10	
IFN-γ	GS1 + GS2	20 ± 7	67 ± 12	235	5

Mean values + standard deviation from independent experiments (biological replicates) are shown. The glycosylation site occupancy of IFN-γ was calculated by quantification of bands from immunoblots. All other values are derived from MS-based quantification of peptides from purified proteins. Please note, due to an incomplete proteolytic digestion, no reliable quantification of GS2 from EPO-Fc could be performed.

extend our findings, we tested the impact of LmSTT3D-GFP on the N-glycosylation site occupancy of other transiently expressed mammalian glycoproteins carrying multiple glycosylation sites (GS). First, we co-expressed recombinant IgE (7 GS) and IgA1 (2 GS) together with LmSTT3D-GFP. These immunoglobulins have been recently expressed in *N. benthamiana* and contain N-glycosylation sites that were partially occupied (Göritzer *et al.*, 2017; Montero-Morales *et al.*, 2017). GS1, GS2, GS4 and GS7 are fully occupied on recombinant IgE. GS3 and GS5, on the other hand, display partial glycosylation and GS6 is normally not occupied (Montero-Morales *et al.*, 2017; Plomp *et al.*, 2014). In the presence of LmSTT3D-GFP, we observed a slight shift in the migration position of the IgE heavy chain (Figure 5a). MS-based analysis of IgE glycosylation sites confirmed an increase in N-glycosylation site occupancy for the IgE glycosylation sites that were previously found to be incompletely glycosylated in the absence of LmSTT3D-GFP (80% increase for GS3 and 128% for GS5, Table 1). Interestingly, GS6 becomes N-glycosylated in the presence of LmSTT3D-GFP and more than half of the purified IgE is now glycosylated at this particular site (Table 1). The N-glycans found on IgE GS6 were mainly processed complex type N-glycans that are commonly found on plant-produced recombinant glycoproteins indicating that the LmSTT3D-mediated transfer does not lead to altered N-glycan processing (Figure S2). GS2 at the C-terminus of IgA1 is normally incompletely N-glycosylated when expressed in plants or human cells (Göritzer *et al.*, 2017). In the

presence of LmSTT3D-GFP, however, we detected a reduction in the incompletely glycosylated alpha heavy chain from IgA1 (Figure 5b) and MS-based quantification revealed almost complete glycosylation of the sequon in the C-terminal tailpiece (Table 1) without affecting N-glycan processing (Figure S3).

To see whether the positive effect of LmSTT3D can also be observed with recombinant glycoproteins that are not related to immunoglobulins, we transiently expressed human EPO-Fc (Castilho *et al.*, 2011) and the cytokine interferon-γ (IFN-γ). Differences in SDS-PAGE migration of EPO-Fc were observed when LmSTT3D-GFP was co-expressed indicating a reduction in underglycosylation (Figure 5c). Quantification of glycopeptides derived from EPO showed a clear increase at glycosylation site one (Table 1). As observed for Fc glycosylation, the N-glycan profiles were virtually identical in the absence or presence of LmSTT3D-GFP. Mainly the fully processed GnGn glycans were present on all three N-glycosylation sites of EPO as well as on the Fc site when expressed in ΔXT/FT (Figure S4 and data not shown), suggesting that LmSTT3D-GFP expression does not interfere with N-glycan processing.

The effect of LmSTT3D on the N-glycosylation site occupancy of IFN-γ which carries two N-glycosylation sites (Asn25 and Asn97, Figure S5) was examined by immunoblots of a variant carrying a C-terminal HA-tag. IFN-γ-HA shows three bands on immunoblots indicating that it is incompletely glycosylated (no N-glycan, a single N-glycan or fully glycosylated with two N-glycans) when transiently expressed in *N. benthamiana* wild-type and ΔXT/FT. Co-expression of LmSTT3D-GFP resulted in the appearance of a major protein band representing the fully glycosylated protein that could be converted to the nonglycosylated IFN-γ-HA by PNGase F digestion (Figure 5d and e, Table 1). These data show that LmSTT3D co-expression improves the N-glycosylation site occupancy of numerous glycoproteins.

LmSTT3D-GFP-HDEL is efficiently retained in the ER

While the co-expression of LmSTT3D-GFP resulted in a significant improvement of the N-glycosylation efficiency on various proteins, the overlapping occurrence in the ER and Golgi bodies suggests that part of the protein is not functional due to the mislocalization. Consequently, we examined whether a LmSTT3D variant with increased ER accumulation improves its functionality. Thus, a construct was generated which expressed LmSTT3D-GFP with a C-terminal HDEL tetrapeptide for ER retrieval (Figure 6a). Imaging by confocal microscopy showed that the attachment of the HDEL motif leads to an efficient steady-state distribution of LmSTT3D-GFP-HDEL in the ER in *N. benthamiana* leaf epidermal cells (Figure 6b). No signal was observed in Golgi bodies. To analyse the functionality of LmSTT3D-GFP-HDEL, we co-expressed it together with IFN-γ-HA and analysed the glycosylation status. Like LmSTT3D-GFP, the LmSTT3D-GFP-HDEL variant was functional and improved the N-glycosylation site occupancy of IFN-γ-HA. Quantification of bands from immunoblots showed no difference between LmSTT3D-GFP and LmSTT3D-GFP-HDEL (Figure 6c) suggesting that the expression of the fully ER-retained LmSTT3D-GFP-HDEL variant does not increase N-glycosylation efficiency compared to the incompletely retained variant.

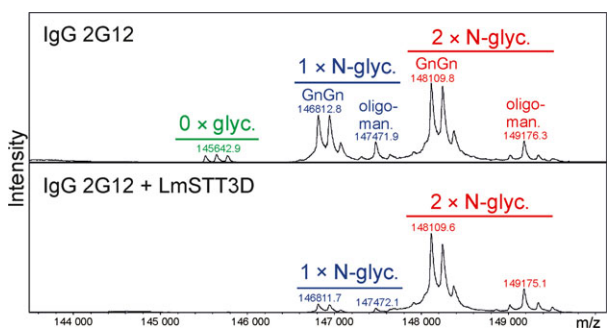


Figure 4 The N-glycan site occupancy of fully assembled IgG in the presence or absence of LmSTT3D-GFP was determined using LC-ESI-MS. The peaks corresponding to unglycosylated (green), hemi-glycosylated (blue, one N-glycan) and fully glycosylated (red, two N-glycans) IgG (HIV-neutralizing antibody 2G12) are highlighted. Multiple peaks represent different glycoforms (complex N-glycan GnGn, oligomannosidic glycans) and variations in the clipping of C-terminal lysine. Please note, the clipping of lysine is not found on variants carrying oligomannosidic N-glycans indicating that this processing reaction occurs in a post-ER compartment.

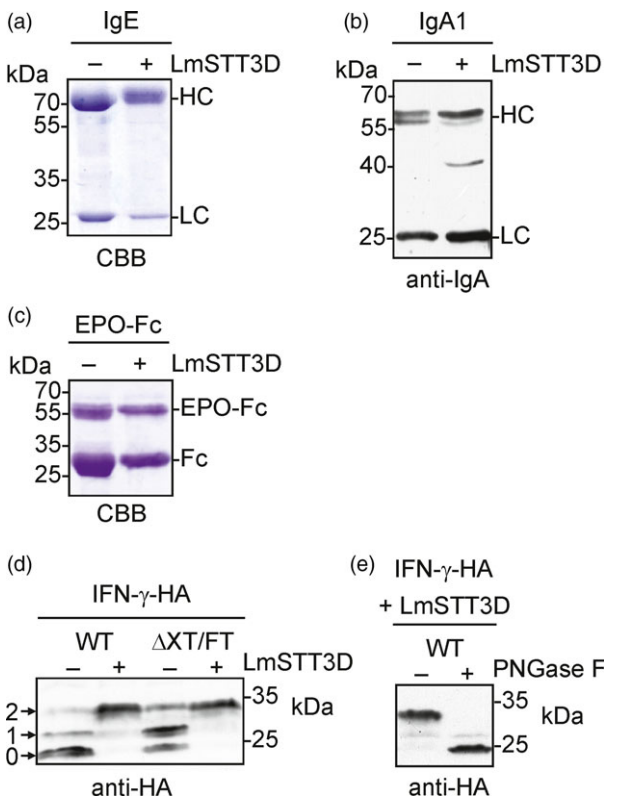


Figure 5 LmSTT3D-GFP enhances the N-glycosylation efficiency of different recombinant proteins when transiently expressed in *N. benthamiana*. (a) Human IgE was expressed in ΔXT/FT in the presence (+) or absence (–) of LmSTT3D-GFP, and the purified IgE was analysed by SDS-PAGE and Coomassie Brilliant Blue (CBB) staining. (b) Human IgA1 was expressed in ΔXT/FT, and total protein extracts were analysed by immunoblotting with antibodies against the alpha heavy chain and the kappa light chain (anti-IgA). (c) EPO-Fc was expressed in ΔXT/FT, purified and subjected to SDS-PAGE and CBB staining. (d) Expression of IFN-γ-HA in the presence (+) or absence (–) of LmSTT3D-GFP. Protein extracts were subjected to SDS-PAGE and immunoblotting using anti-HA antibodies. The migration position of the nonglycosylated (0), mono- (1) and di-glycosylated (2) IFN-γ-HA protein is indicated. (e) PNGase F digestion of IFN-γ-HA co-expressed with LmSTT3D-GFP.

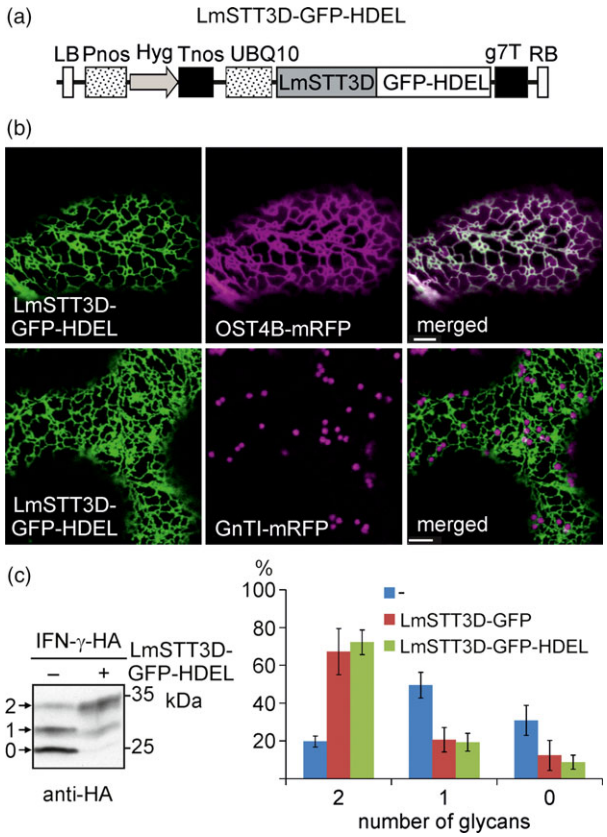


Figure 6 Attachment of the HDEL tetrapeptide improves ER localization of LmSTT3D-GFP and does not interfere with its functionality. (a) Schematic illustration of the UBQ10:LmSTT3D-GFP-HDEL expression vector. For abbreviations, see legend of Figure 2. (b) LmSTT3D-GFP-HDEL was either expressed alone or in combination with the ER-marker OST4B-mRFP in wild-type leaf epidermal cells. Images were acquired two days postinfiltration. Bars = 5 μm. (c) SDS-PAGE and immunoblotting of IFN-γ-HA expressed in ΔXT/FT in the presence (+) or absence (–) of LmSTT3D-GFP-HDEL. Quantification of IFN-γ-HA protein bands upon expression in ΔXT/FT. The diagram shows mean values plus standard deviation from at least five biological replicates.

Discussion

N-glycans have a strong impact on protein folding, stability and influence the interaction with other proteins. N-glycosylation of recombinant biopharmaceuticals is critical for product quality (Reusch and Tejada, 2015). For example, the single N-glycan from the IgG Fc-domain modulates immune effector functions and unglycosylated IgG variants display drastically reduced affinity for Fcγ-receptors (Ferrara *et al.*, 2011; Jefferis and Lund, 2002; Shields *et al.*, 2002). Moreover, the presence of additional N-glycans can improve the *in vivo* half-life and activity of recombinant biopharmaceuticals. This has been impressively demonstrated for a hyperglycosylated EPO variant (darbepoetin alfa) that is glycosylated at two additionally introduced N-glycosylation sites and has been approved for treatment of anaemia (Elliott *et al.*, 2003). EPO or IgG from human serum and recombinant variants thereof expressed in mammalian cells are typically very efficiently glycosylated (Table S1). By contrast, *N. benthamiana* (Table 1) and to a certain degree also other expression systems such as *P. pastoris* (Choi *et al.*, 2012) or insect cells (Sareneva *et al.*, 1995) display more variation in N-glycosylation efficiency. Despite the documented importance of proper N-glycosylation site occupancy, comparably few studies have so far approached the diversity caused by the absence of glycans at particular sites on recombinant glycoproteins. Previous studies have shown that LmSTT3D, the single catalytic subunit from the protozoan *L. major*, can replace the function of the endogenous STT3 subunit from *S. cerevisiae* and complements growth and N-glycosylation defects associated with OST deficiency (Nasab *et al.*, 2008). Biochemical and genetic evidence indicates that LmSTT3D is functionally independent and not integrated into the native OST complex when heterologously expressed (Hese *et al.*, 2009; Nasab *et al.*, 2008).

Recombinant IFN-γ has been approved for treatment of different human diseases and is a promising candidate for cancer immunotherapy (Razaghi *et al.*, 2016). Glycosylation of IFN-γ is important for its proteolytic stability, secretion and circulatory half-life (Bocci *et al.*, 1985; Sareneva *et al.*, 1993, 1995, 1996). Transient expression in *N. benthamiana* indicates that IFN-γ is inefficiently glycosylated in the absence of LmSTT3D. We currently do not know whether both sites are equally affected or whether Asn97, which is present in an α-helical region, is less occupied as has been suggested for IFN-γ from human cells (Sareneva *et al.*, 1996). In contrast to *N. benthamiana*, approximately two-thirds of native human IFN-γ is fully glycosylated (Rinderknecht *et al.*, 1984; Sareneva *et al.*, 1995) and recombinant IFN-γ expressed in CHO fed-batch cultures displays low amounts of nonglycosylated protein (Wong *et al.*, 2010). The reason for this discrepancy in N-glycosylation efficiency between mammalian cells and plants is currently unknown, but may reflect differences in the composition and function of the OST complex. Notably, in the presence of LmSTT3D, similar levels of fully glycosylated IFN-γ are obtained in plants and on the naturally occurring protein (Table 1 and Table S1).

We found that not all analysed N-glycosylation sites were equally well glycosylated upon LmSTT3D co-expression. Glycosylation of GS3 from EPO-Fc or GS1 from IgA1, which were already efficiently occupied in the absence of LmSTT3D, was not improved. On the other hand, GS2 from IgA1, which is only partially glycosylated when expressed in plants or human cells (Göritzer *et al.*, 2017), could be completely modified with N-glycans upon LmSTT3D expression. This site is likely post-translationally modified in mammalian cells and plays an important role in the assembly of dimeric IgA1 (Atkin *et al.*, 1996). Consequently, our data indicate that LmSTT3D preferentially glycosylates certain N-glycosylation sites which has also been recognized in a previous study (Nasab *et al.*, 2008). The precise sequence or conformational constraints influencing LmSTT3D-dependent glycosylation are unknown. Remarkably, the co-expression of LmSTT3D resulted in the glycosylation of IgE GS6. This is in contrast to native serum or recombinantly produced IgE from human cells (Montero-Morales *et al.*, 2017; Plomp *et al.*, 2014), indicating that LmSTT3D has a more relaxed substrate specificity and recognizes glycosylation sites that are normally not used by the mammalian OST complex. Due to the various biological roles of N-glycans, the functional relevance of an additional N-glycan is difficult to predict. LmSTT3D co-expression facilitates the production of non-natural glycoproteins that can be used to test the influence on physicochemical properties of proteins and known protein interactions in future studies.

A further increase in the glycosylation efficiency may be achieved by stable expression of LmSTT3D in *N. benthamiana* or

incorporation of LmSTT3D into multicassette expression vectors used for transient expression together with a glycoprotein of interest to ensure that all cells express LmSTT3D. The stable expression of the KDEL tagged variant, which is at least equally functional when expressed with IFN-γ or IgG (data not shown), will less likely interfere with N-glycan processing in the Golgi or overall Golgi organization and function.

Apart from the expression levels, interaction with ER-resident proteins, polypeptide substrate specificity or enzyme kinetics of the catalytic STT3 subunit, glycosylation efficiency may be controlled by supply of the preassembled lipid-linked oligosaccharide substrate. Deprivation of glucose from CHO cell cultures reduced the amounts of lipid-linked oligosaccharides resulting in the expression of nonglycosylated monoclonal antibodies (Liu *et al.*, 2014). In addition to optimized metabolic parameters, a limitation in donor substrate availability may be the result of inefficient lipid-linked oligosaccharide transfer into the ER. This shortcoming may be overcome by co-expression of an artificial flippase (Parsaie Nasab *et al.*, 2013). Moreover, it is well known that protein intrinsic structural constraints strongly influence the N-glycosylation efficiency at distinct sites. For example, the presence of a serine instead of a threonine in the consensus site N-X-S/T is less preferred by the OST complex and N-X-T sites are more frequently glycosylated in organisms from different eukaryotic domains of life (Zielinska *et al.*, 2012). Exchange of amino acids in the sequon or at adjacent sites of the polypeptide can drastically alter the glycosylation site occupancy (Murray *et al.*, 2015). For a recombinant elastase expressed in *P. pastoris,* a change of the sequon from N-X-S to N-X-T resulted in an increased glycosylation efficiency that was accompanied by higher production levels of the recombinant glycoprotein (Han *et al.*, 2015). By contrast, mutagenesis of flanking amino acids and generation of an optimized aromatic sequon with increased glycosylation efficiency negatively affected the secretion of IFN-γ expressed in human cells and caused variability in protein expression of another glycoprotein (Huang *et al.*, 2017). Likewise, antibody engineering by generation of an aromatic sequon (FANST instead of the canonical QYNST) improved the thermal stability of the antibody, but reduced the affinity to specific Fcγ-receptors (Chen *et al.*, 2016). These studies highlight impressively that protein engineering at glycosylation sites can have various consequences leading to reduced productivity or altered product quality. Consequently, strategies aiming at an improvement of N-glycosylation by engineering of the OST complex are very promising and relevant for different plant-based expression platforms (Hamorsky *et al.*, 2015; Rademacher *et al.*, 2008; Vamvaka *et al.*, 2016; Van Droogenbroeck *et al.*, 2007). Further advances require a better understanding of the OST complex composition and molecular function of the individual subunits. Taken together, our findings demonstrate that transient LmSTT3D expression is a robust extension of currently existing glyco-engineering approaches and should be integrated into production processes to reduce product heterogeneity and improve biological activities related to N-glycosylation of recombinant glycoproteins.

Experimental procedures

Cloning of expression vectors

The expression constructs for IgG 2G12 (Schähs *et al.*, 2007), EPO-Fc (Castilho *et al.*, 2011), IgE (Montero-Morales *et al.*, 2017) and IgA1 (Göritzer *et al.*, 2017) were described previously. To generate the SP-Fc expression vector, the DNA fragment coding for GCSI-CTS-Fc was amplified from GCSI-CTS-GFPglyc (Schoberer *et al.*, 2009) by PCR using primers GCSI-7F (TATATCTA GAATGACCGGAGCTAGCCGTCGGAGC) and Fc-6R (TATACTC GAGTTATTTACCCGGAGACAGGGAGAGG). The PCR product was digested with XbaI/XhoI and cloned into XbaI/SalI sites of p47 (Hüttner *et al.*, 2014) to generated p71-GCSI-Fc. Subsequently, the chitinase signal peptide was amplified from *N. benthamiana* cDNA by PCR using Nb-Chi-F1 (TATATCTA GAATGAGGCTTAGAGAATTCACAG) and Nb-Chi-R2 (TATAG GATCCTGCCGAGGCAGAGAGTAGGAGAGA), XbaI/BamHI digested and cloned into XbaI/BamHI digested p71-GCSI-Fc, resulting in p71-SP-Fc. For IFN-γ expression, a codon-optimized DNA fragment encoding human IFN-γ was synthetized by GeneArt Gene Synthesis (Thermo Fisher Scientific). The synthetic DNA fragment was XbaI/BamHI digested and cloned into the XbaI/BamHI sites of expression vector p43. The vector p43 is a derivative of expression vector p27 (Strasser *et al.*, 2007) whereby the CaMV 35S promoter was replaced by the *A. thaliana* UBQ10 promoter and a sequence encoding a 3x HA-tag for C-terminal fusion was inserted upstream of the terminator sequence (Figure S5). To generate the LmSTT3D-GFP expression vector a codon-optimized open reading frame coding for *L. major* STT3D (Nasab *et al.*, 2008) was synthetized by GeneArt Gene Synthesis. The LmSTT3D open reading frame was excised by XbaI/BamHI digestion and cloned into XbaI/BamHI sites of p47 or p56. Vector p56 is derived from p47 by replacement of the GFP coding sequence with the one for GFP-HDEL. For generation of the STT3A-GFP expression vector p20-STT3A, the *A. thaliana* STT3A coding region was amplified by PCR as described previously (Farid *et al.*, 2013) and cloned into XbaI/BamHI digested plasmid p20F (Schoberer *et al.*, 2009).

Transient expression and immunoblot analysis

All plant expression vectors were transformed into *Agrobacterium tumefaciens* (strain UIA143) (Farrand *et al.*, 1989). Syringe-mediated agroinfiltration was used for transient expression in leaves of 4- to 5-week-old *N. benthamiana* grown on soil under long-day conditions (16 h light/8 h dark) at 25°C. At the indicated time points, leaf pieces were harvested from infiltrated plants, and total protein extracts were prepared and subjected to SDS-PAGE followed by silver staining (Strasser *et al.*, 2004) or immunoblotting as described in detail previously (Shin *et al.*, 2017). IgG and Fc-containing fragments were monitored with anti-human IgG (H+L)-horseradish peroxidase antibody (Promega, Mannheim, Germany), IgA with anti-alpha chain/anti-kappa-chain antibodies and IFN-γ-HA with anti-HA antibodies. For deglycosylation, protein extracts were incubated with peptide-N-glycosidase F (PNGase F) (New England Biolabs, Frankfurt am Main, Germany) according to the manufacturer's procedure. Quantification of gel bands on immunoblots was performed with a ChemiDoc imager (Bio-Rad, Vienna, Austria) and Quantity One 1D analysis software (Bio-Rad).

For detection of LmSTT3D-GFP on immunoblots, leaf material was harvested 48 h after infiltration of *N. benthamiana* leaves. Proteins were extracted with 1 × Laemmli sample buffer supplemented with 6M urea and incubated at 37°C for 5 min. The fusion protein was detected with anti-GFP horseradish peroxidase (MACS Miltenyi Biotec, Bergisch Gladbach, Germany) antibodies.

Confocal imaging of fluorescent protein fusions

Leaves of 4- to 5-week-old *N. benthamiana* were infiltrated with agrobacterium suspensions carrying binary plant expression vectors for expression of GFP- or mRFP-tagged proteins with the following optical densities (OD_{600}): 0.1 for p47-LmSTT3D (LmSTT3D-GFP), p56-LmSTT3D (LmSTT3D-GFP-HDEL), p20-STT3A (AtSTT3A-GFP). Agrobacteria carrying the expression constructs p31-OST4B (OST4B-mRFP, ER-marker) (Farid *et al.*, 2013) and p31-GnTI (GnTI-mRFP, Golgi-marker) (Schoberer *et al.*, 2013) were infiltrated with $OD_{600} = 0.05$. Confocal images were acquired 1 and 2 days postinfiltration on a Leica SP5 II confocal microscope using the Leica LAS AF software system (http://www.leica.com). Dual-colour image acquisition of cells expressing both GFP and mRFP was performed simultaneously. Postacquisition image processing was performed in Adobe PHOTOSHOP CS5.

LC-ESI-MS analysis

The full-length heavy chain from IgGs and SP-Fc was purified from the protein extract by binding to rProtein A Sepharose™ Fast Flow (GE Healthcare Europe, Vienna, Austria). Purified protein was subjected to SDS-PAGE under reducing conditions and Coomassie Brilliant Blue staining. The corresponding protein band was excised from the gel, destained, carbamidomethylated, in-gel trypsin digested and analysed by liquid chromatography electrospray ionization mass spectrometry (LC-ESI-MS), as described in detail previously (Stadlmann *et al.*, 2008). A detailed explanation of N-glycan abbreviations can be found at http://www.proglycan.com. Protein purification and MS-based analysis of (glyco) peptides from EPO-Fc, IgE and IgA1 were described in detail recently (Castilho *et al.*, 2011; Göritzer *et al.*, 2017; Montero-Morales *et al.*, 2017). Site occupancy was calculated from the peak area of nonglycosylated versus the sum of the peak areas of all glycoforms including relevant adduct ions and observed charged states. The principal suitability of this 'peak sum' approach under the conditions applied was verified in two stages. First, selected samples were subjected to deglycosylation with peptide-N-glycosidase A (ProGlycAn, Vienna, Austria) and the ratio of the Asn vs. the Asp containing glycopeptides, which separate in RP-HPLC, was measured. Peptides differing in charged amino acids may have differing mass spectrometric responses, and hence, in a second stage, the Fc tandem peptide EEQYNSTYREEQYDSTYR (JP peptides, Berlin, Germany) was digested with trypsin to obtain an equimolar mixture of the Asn and the Asp form of the Fc (glyco-)peptide. These measurements showed that the 'peak sum' approach gave reliable values with a possible overestimation of nonglycosylation of a very few percentages especially in the case of low underglycosylation. We assume the situation for other glycoproteins to be comparable to that with IgG Fc.

Mass spectrometric analysis of fully assembled IgGs

The purified IgGs were directly injected to a LC-ESI-MS system (LC: Dionex Ultimate 3000 LC). A gradient from 20% to 80% acetonitrile in 0.05% trifluoroacetic acid (using a Thermo ProSwift™ RP-4H column (0.2 × 250 mm)) at a flow rate of 8 μL/min was applied (30-minute gradient time). Detection was performed with a Q-TOF instrument (Bruker maXis 4G) equipped with the standard ESI source in positive ion, MS mode (range: 750–5000 Da). Instrument calibration was performed using ESI calibration mixture (Agilent). Data were processed using Data Analysis 4.0 (Bruker), and the spectrum was deconvoluted by MaxEnt.

Acknowledgements

We thank Michaela Bogner, Martina Dicker, Thomas Hackl, Jennifer Schoberer, Yun-Ji Shin and Christiane Veit for technical support and the BOKU-VIBT Imaging Centre for access and expertise. We thank Victor Klimyuk from Icon Genetics GmbH for the kind gift of expression vectors used for IgE expression. This work was supported by grants from the Austrian Science Fund (FWF): I2417-B22 (to RS), P28673-B28 (to HS) and the Doctoral Program BioToP (Grant W 1224).

Conflict of interest

The authors declared that they have no conflict of interests.

References

Aebi, M. (2013) N-linked protein glycosylation in the ER. *Biochim. Biophys. Acta*, **1833**, 2430–2437.

Atkin, J.D., Pleass, R.J., Owens, R.J. and Woof, J.M. (1996) Mutagenesis of the human IgA1 heavy chain tailpiece that prevents dimer assembly. *J. Immunol.* **157**, 156–159.

Bendandi, M., Marillonnet, S., Kandzia, R., Thieme, F., Nickstadt, A., Herz, S., Fröde, R. *et al.* (2010) Rapid, high-yield production in plants of individualized idiotype vaccines for non-Hodgkin's lymphoma. *Ann. Oncol.* **21**, 2420–2427.

Bocci, V., Pacini, A., Pessina, G.P., Paulesu, L., Muscettola, M. and Lunghetti, G. (1985) Catabolic sites of human interferon-gamma. *J. Gen. Virol.* **66**, 887–891.

Castilho, A., Gattinger, P., Grass, J., Jez, J., Pabst, M., Altmann, F., Gorfer, M. *et al.* (2011) N-glycosylation engineering of plants for the biosynthesis of glycoproteins with bisected and branched complex N-glycans. *Glycobiology*, **21**, 813–823.

Chen, W., Kong, L., Connelly, S., Dendle, J.M., Liu, Y., Wilson, I.A., Powers, E.T. *et al.* (2016) Stabilizing the CH2 domain of an antibody by engineering in an enhanced aromatic sequon. *ACS Chem. Biol.* **11**, 1852–1861.

Choi, B.K., Warburton, S., Lin, H., Patel, R., Boldogh, I., Meehl, M., d'Anjou, M. *et al.* (2012) Improvement of N-glycan site occupancy of therapeutic glycoproteins produced in *Pichia pastoris*. *Appl. Microbiol. Biotechnol.* **95**, 671–682.

Dicker, M., Tschofen, M., Maresch, D., Konig, J., Juarez, P., Orzaez, D., Altmann, F. *et al.* (2016) Transient Glyco-engineering to produce recombinant IgA1 with defined N- and O-Glycans in plants. *Front Plant Sci.* **7**, 18.

Elliott, S., Lorenzini, T., Asher, S., Aoki, K., Brankow, D., Buck, L., Busse, L. *et al.* (2003) Enhancement of therapeutic protein in vivo activities through glycoengineering. *Nat. Biotechnol.* **21**, 414–421.

Elliott, S., Egrie, J., Browne, J., Lorenzini, T., Busse, L., Rogers, N. and Ponting, I. (2004) Control of rHuEPO biological activity: the role of carbohydrate. *Exp. Hematol.* **32**, 1146–1155.

Farid, A., Malinovsky, F.G., Veit, C., Schoberer, J., Zipfel, C. and Strasser, R. (2013) Specialized roles of the conserved subunit OST3/6 of the oligosaccharyltransferase complex in innate immunity and tolerance to abiotic stresses. *Plant Physiol.* **162**, 24–38.

Farrand, S.K., O'Morchoe, S.P. and McCutchan, J. (1989) Construction of an Agrobacterium tumefaciens C58 recA mutant. *J. Bacteriol.* **171**, 5314–5321.

Ferrara, C., Grau, S., Jäger, C., Sondermann, P., Brünker, P., Waldhauer, I., Hennig, M. *et al.* (2011) Unique carbohydrate-carbohydrate interactions are required for high affinity binding between FcgammaRIII and antibodies lacking core fucose. *Proc. Natl. Acad. Sci. USA*, **108**, 12669–12674.

Göritzer, K., Maresch, D., Altmann, F., Obinger, C. and Strasser, R. (2017) Exploring site-specific N-Glycosylation of HEK293 and plant-produced human IgA isotypes. *J. Proteome Res.* **16**, 2560–2570.

Hamorsky, K.T., Kouokam, J.C., Jurkiewicz, J.M., Nelson, B., Moore, L.J., Husk, A.S., Kajiura, H. *et al.* (2015) N-glycosylation of cholera toxin B subunit in *Nicotiana benthamiana*: impacts on host stress response, production yield and vaccine potential. *Sci. Rep.* **5**, 8003.

Han, M., Wang, W., Wang, X., Liu, X., Cao, H., Tao, Y. and Yu, X. (2015) Enhanced expression of recombinant elastase in *Pichia pastoris* through the substitution of Thr for Ser in Asn-Xaa-Ser sequons. *Appl. Biochem. Biotechnol.* **175**, 428–435.

Hanania, U., Ariel, T., Tekoah, Y., Fux, L., Sheva, M., Gubbay, Y., Weiss, M. *et al.* (2017) Establishment of a tobacco BY2 cell line devoid of plant-specific xylose and fucose as a platform for the production of biotherapeutic proteins. *Plant Biotechnol. J.* **15**, 1120–1129.

Hese, K., Otto, C., Routier, F.H. and Lehle, L. (2009) The yeast oligosaccharyltransferase complex can be replaced by STT3 from *Leishmania major*. *Glycobiology*, **19**, 160–171.

Huang, Y.W., Yang, H.I., Wu, Y.T., Hsu, T.L., Lin, T.W., Kelly, J.W. and Wong, C.H. (2017) Residues comprising the enhanced aromatic sequon influence protein N-glycosylation efficiency. *J. Am. Chem. Soc.* **139**, 12947–12955.

Hüttner, S., Veit, C., Vavra, U., Schoberer, J., Dicker, M., Maresch, D., Altmann, F. *et al.* (2014) A context-independent N-glycan signal targets the misfolded extracellular domain of Arabidopsis STRUBBELIG to endoplasmic-reticulum-associated degradation. *Biochem. J.* **464**, 401–411.

Izquierdo, L., Atrih, A., Rodrigues, J., Jones, D. and Ferguson, M. (2009) *Trypanosoma brucei* UDP-glucose:glycoprotein glucosyltransferase has unusual substrate specificity and protects the parasite from stress. *Eukaryot. Cell*, **8**, 230–240.

Jefferis, R. (2009) Glycosylation as a strategy to improve antibody-based therapeutics. *Nat. Rev. Drug. Discov.* **8**, 226–234.

Jefferis, R. and Lund, J. (2002) Interaction sites on human IgG-Fc for FcgammaR: current models. *Immunol. Lett.* **82**, 57–65.

Kallolimath, S., Castilho, A., Strasser, R., Grunwald-Gruber, C., Altmann, F., Strubl, S., Galuska, C.E. *et al.* (2016) Engineering of complex protein sialylation in plants. *Proc. Natl. Acad. Sci. USA*, **113**, 9498–9503.

Kang, J., Frank, J., Kang, C., Kajiura, H., Vikram, M., Ueda, A., Kim, S. *et al.* (2008) Salt tolerance of Arabidopsis thaliana requires maturation of N-glycosylated proteins in the Golgi apparatus. *Proc. Natl. Acad. Sci. USA*, **105**, 5933–5938.

Kelleher, D. and Gilmore, R. (2006) An evolving view of the eukaryotic oligosaccharyltransferase. *Glycobiology*, **16**, 47R–62R.

Knauer, R. and Lehle, L. (1999) The oligosaccharyltransferase complex from yeast. *Biochim. Biophys. Acta*, **1426**, 259–273.

Koiwa, H., Li, F., McCully, M., Mendoza, I., Koizumi, N., Manabe, Y., Nakagawa, Y. *et al.* (2003) The STT3a subunit isoform of the Arabidopsis oligosaccharyltransferase controls adaptive responses to salt/osmotic stress. *Plant Cell*, **15**, 2273–2284.

Li, J., Stoddard, T.J., Demorest, Z.L., Lavoie, P.O., Luo, S., Clasen, B.M., Cedrone, F. *et al.* (2016) Multiplexed, targeted gene editing in *Nicotiana benthamiana* for glyco-engineering and monoclonal antibody production. *Plant Biotechnol. J.* **14**, 533–542.

Limkul, J., Iizuka, S., Sato, Y., Misaki, R., Ohashi, T. and Fujiyama, K. (2016) The production of human glucocerebrosidase in glyco-engineered *Nicotiana benthamiana* plants. *Plant Biotechnol. J.* **14**, 1682–1694.

Liu, B., Spearman, M., Doering, J., Lattová, E., Perreault, H. and Butler, M. (2014) The availability of glucose to CHO cells affects the intracellular lipid-linked oligosaccharide distribution, site occupancy and the N-glycosylation profile of a monoclonal antibody. *J. Biotechnol.* **170**, 17–27.

Loos, A., Gruber, C., Altmann, F., Mehofer, U., Hensel, F., Grandits, M., Oostenbrink, C. *et al.* (2014) Expression and glycoengineering of functionally active heteromultimeric IgM in plants. *Proc. Natl. Acad. Sci. USA*, **111**, 6263–6268.

Loos, A., Gach, J.S., Hackl, T., Maresch, D., Henkel, T., Porodko, A., Bui-Minh, D. *et al.* (2015) Glycan modulation and sulfoengineering of anti-HIV-1 monoclonal antibody PG9 in plants. *Proc. Natl. Acad. Sci. USA*, **112**, 12675–12680.

Mercx, S., Smargiasso, N., Chaumont, F., De Pauw, E., Boutry, M. and Navarre, C. (2017) Inactivation of the β(1,2)-xylosyltransferase and the α(1,3)-fucosyltransferase genes in *Nicotiana tabacum* BY-2 Cells by a Multiplex CRISPR/Cas9 Strategy Results in Glycoproteins without Plant-Specific Glycans. *Front Plant Sci.* **8**, 403.

Mohorko, E., Glockshuber, R. and Aebi, M. (2011) Oligosaccharyltransferase: the central enzyme of N-linked protein glycosylation. *J. Inherit. Metab. Dis.* **34**, 869–878.

Montero-Morales, L., Maresch, D., Castilho, A., Turupcu, A., Ilieva, K.M., Crescioli, S., Karagiannis, S.N. *et al.* (2017) Recombinant plant-derived human IgE glycoproteomics. *J. Proteomics*, **161**, 81–87.

Murray, A.N., Chen, W., Antonopoulos, A., Hanson, S.R., Wiseman, R.L., Dell, A., Haslam, S.M. *et al.* (2015) Enhanced aromatic sequons increase oligosaccharyltransferase glycosylation efficiency and glycan homogeneity. *Chem. Biol.* **22**, 1052–1062.

Nasab, F.P., Schulz, B.L., Gamarro, F., Parodi, A.J. and Aebi, M. (2008) All in one: *Leishmania major* STT3 proteins substitute for the whole oligosaccharyltransferase complex in Saccharomyces cerevisiae. *Mol. Biol. Cell*, **19**, 3758–3768.

Nekrasov, V., Li, J., Batoux, M., Roux, M., Chu, Z., Lacombe, S., Rougon, A. *et al.* (2009) Control of the pattern-recognition receptor EFR by an ER protein complex in plant immunity. *EMBO J.* **28**, 3428–3438.

Nose, M. and Wigzell, H. (1983) Biological significance of carbohydrate chains on monoclonal antibodies. *Proc. Natl. Acad. Sci. USA*, **80**, 6632–6636.

Parsaie Nasab, F., Aebi, M., Bernhard, G. and Frey, A.D. (2013) A combined system for engineering glycosylation efficiency and glycan structure in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **79**, 997–1007.

Plomp, R., Hensbergen, P.J., Rombouts, Y., Zauner, G., Dragan, I., Koeleman, C.A., Deelder, A.M. *et al.* (2014) Site-specific N-glycosylation analysis of human immunoglobulin E. *J. Proteome Res.* **13**, 536–546.

Rademacher, T., Sack, M., Arcalis, E., Stadlmann, J., Balzer, S., Altmann, F., Quendler, H. *et al.* (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.

Razaghi, A., Owens, L. and Heimann, K. (2016) Review of the recombinant human interferon gamma as an immunotherapeutic: Impacts of production platforms and glycosylation. *J. Biotechnol.* **240**, 48–60.

Reusch, D. and Tejada, M.L. (2015) Fc glycans of therapeutic antibodies as critical quality attributes. *Glycobiology*, **25**, 1325–1334.

Rinderknecht, E., O'Connor, B.H. and Rodriguez, H. (1984) Natural human interferon-gamma. Complete amino acid sequence and determination of sites of glycosylation. *J. Biol. Chem.* **259**, 6790–6797.

Roboti, P. and High, S. (2012) The oligosaccharyltransferase subunits OST48, DAD1 and KCP2 function as ubiquitous and selective modulators of mammalian N-glycosylation. *J. Cell Sci.* **125**, 3474–3484.

Ruiz-Canada, C., Kelleher, D.J. and Gilmore, R. (2009) Cotranslational and posttranslational N-glycosylation of polypeptides by distinct mammalian OST isoforms. *Cell*, **136**, 272–283.

Saijo, Y., Tintor, N., Lu, X., Rauf, P., Pajerowska-Mukhtar, K., Häweker, H., Dong, X. *et al.* (2009) Receptor quality control in the endoplasmic reticulum for plant innate immunity. *EMBO J.* **28**, 3439–3449.

Samuelson, J., Banerjee, S., Magnelli, P., Cui, J., Kelleher, D.J., Gilmore, R. and Robbins, P.W. (2005) The diversity of dolichol-linked precursors to Asn-linked glycans likely results from secondary loss of sets of glycosyltransferases. *Proc. Natl. Acad. Sci. USA*, **102**, 1548–1553.

Sareneva, T., Cantell, K., Pyhälä, L., Pirhonen, J. and Julkunen, I. (1993) Effect of carbohydrates on the pharmacokinetics of human interferon-gamma. *J. Interferon Res.* **13**, 267–269.

Sareneva, T., Pirhonen, J., Cantell, K. and Julkunen, I. (1995) N-glycosylation of human interferon-gamma: glycans at Asn-25 are critical for protease resistance. *Biochem. J.* **308**, 9–14.

Sareneva, T., Mørtz, E., Tölö, H., Roepstorff, P. and Julkunen, I. (1996) Biosynthesis and N-glycosylation of human interferon-gamma. Asn25 and Asn97 differ markedly in how efficiently they are glycosylated and in their oligosaccharide composition. *Eur. J. Biochem.* **242**, 191–200.

Schähs, M., Strasser, R., Stadlmann, J., Kunert, R., Rademacher, T. and Steinkellner, H. (2007) Production of a monoclonal antibody in plants with a humanized N-glycosylation pattern. *Plant Biotechnol. J.* **5**, 657–663.

Schoberer, J., Vavra, U., Stadlmann, J., Hawes, C., Mach, L., Steinkellner, H. and Strasser, R. (2009) Arginine/lysine residues in the cytoplasmic tail promote ER export of plant glycosylation enzymes. *Traffic*, **10**, 101–115.

Schoberer, J., Liebminger, E., Botchway, S.W., Strasser, R. and Hawes, C. (2013) Time-resolved fluorescence imaging reveals differential interactions of N-glycan processing enzymes across the Golgi stack in planta. *Plant Physiol.* **161**, 1737–1754.

Shibatani, T., David, L.L., McCormack, A.L., Frueh, K. and Skach, W.R. (2005) Proteomic analysis of mammalian oligosaccharyltransferase reveals multiple subcomplexes that contain Sec61, TRAP, and two potential new subunits. *Biochemistry*, **44**, 5982–5992.

Shields, R.L., Lai, J., Keck, R., O'Connell, L.Y., Hong, K., Meng, Y.G., Weikert, S.H. *et al.* (2002) Lack of fucose on human IgG1 N-linked oligosaccharide improves binding to human Fcgamma RIII and antibody-dependent cellular toxicity. *J. Biol. Chem.* **277**, 26733–26740.

Shin, Y.J., Castilho, A., Dicker, M., Sádio, F., Vavra, U., Grünwald-Gruber, C., Kwon, T.H. *et al.* (2017) Reduced paucimannosidic N-glycan formation by suppression of a specific β-hexosaminidase from *Nicotiana benthamiana*. *Plant Biotechnol. J.* **15**, 197–206.

Shrimal, S., Trueman, S.F. and Gilmore, R. (2013) Extreme C-terminal sites are posttranslocationally glycosylated by the STT3B isoform of the OST. *J. Cell Biol.* **201**, 81–95.

Shrimal, S., Cherepanova, N.A. and Gilmore, R. (2015) Cotranslational and posttranslocational N-glycosylation of proteins in the endoplasmic reticulum. *Semin. Cell Dev. Biol.* **41**, 71–78.

Stadlmann, J., Pabst, M., Kolarich, D., Kunert, R. and Altmann, F. (2008) Analysis of immunoglobulin glycosylation by LC-ESI-MS of glycopeptides and oligosaccharides. *Proteomics*, **8**, 2858–2871.

Strasser, R. (2016) Plant protein glycosylation. *Glycobiology*, **26**, 926–939.

Strasser, R., Altmann, F., Mach, L., Glössl, J. and Steinkellner, H. (2004) Generation of *Arabidopsis thaliana* plants with complex N-glycans lacking beta1,2-linked xylose and core alpha1,3-linked fucose. *FEBS Lett.* **561**, 132–136.

Strasser, R., Bondili, J., Schoberer, J., Svoboda, B., Liebminger, E., Glössl, J., Altmann, F. *et al.* (2007) Enzymatic properties and subcellular localization of Arabidopsis beta-N-acetylhexosaminidases. *Plant Physiol.* **145**, 5–16.

Strasser, R., Stadlmann, J., Schähs, M., Stiegler, G., Quendler, H., Mach, L., Glössl, J. *et al.* (2008) Generation of glyco-engineered *Nicotiana benthamiana* for the production of monoclonal antibodies with a homogeneous human-like N-glycan structure. *Plant Biotechnol. J.* **6**, 392–402.

Strasser, R., Altmann, F. and Steinkellner, H. (2014) Controlled glycosylation of plant-produced recombinant proteins. *Curr. Opin. Biotechnol.* **30C**, 95–100.

Vamvaka, E., Twyman, R.M., Murad, A.M., Melnik, S., Teh, A.Y., Arcalis, E., Altmann, F. *et al.* (2016) Rice endosperm produces an underglycosylated and potent form of the HIV-neutralizing monoclonal antibody 2G12. *Plant Biotechnol. J.* **14**, 97–108.

Van Droogenbroeck, B., Cao, J., Stadlmann, J., Altmann, F., Colanesi, S., Hillmer, S., Robinson, D.G. *et al.* (2007) Aberrant localization and underglycosylation of highly accumulating single-chain Fv-Fc antibodies in transgenic Arabidopsis seeds. *Proc. Natl. Acad. Sci. USA*, **104**, 1430–1435.

Walker, M.R., Lund, J., Thompson, K.M. and Jefferis, R. (1989) Aglycosylation of human IgG1 and IgG3 monoclonal antibodies can eliminate recognition by human cells expressing Fc gamma RI and/or Fc gamma RII receptors. *Biochem. J.* **259**, 347–353.

Wong, D.C., Wong, N.S., Goh, J.S., May, L.M. and Yap, M.G. (2010) Profiling of N-glycosylation gene expression in CHO cell fed-batch cultures. *Biotechnol. Bioeng.* **107**, 516–528.

Yan, Q. and Lennarz, W.J. (2002) Studies on the function of oligosaccharyl transferase subunits. Stt3p is directly involved in the glycosylation process. *J. Biol. Chem.* **277**, 47692–47700.

Zacchi, L.F. and Schulz, B.L. (2016) N-glycoprotein macroheterogeneity: biological implications and proteomic characterization. *Glycoconj. J.* **33**, 359–376.

Zeitlin, L., Geisbert, J.B., Deer, D.J., Fenton, K.A., Bohorov, O., Bohorova, N., Goodman, C. *et al.* (2016) Monoclonal antibody therapy for Junin virus infection. *Proc. Natl. Acad. Sci. USA*, **13**, 4458–4463.

Zielinska, D.F., Gnad, F., Wiśniewski, J.R. and Mann, M. (2010) Precision mapping of an in vivo N-glycoproteome reveals rigid topological and sequence constraints. *Cell*, **141**, 897–907.

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.

Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:

Figure S1 Mass spectra of the IgG glycopeptide in the presence or absence of LmSTT3D-GFP.

Figure S2 Mass spectra of IgE glycopeptides in the presence or absence of LmSTT3D-GFP.

Figure S3 Mass spectra of IgA1 glycopeptides in the presence or absence of LmSTT3D-GFP.

Figure S4 Mass spectra of the EPO-Fc glycopeptides harbouring glycosylation site 1 in the presence or absence of LmSTT3D-GFP.

Figure S5 Schematic illustration of the IFN-γ-HA expression construct and the corresponding amino acid sequence.

Table S1 Comparison of the N-glycosylation site occupancy of native and recombinant glycoproteins.

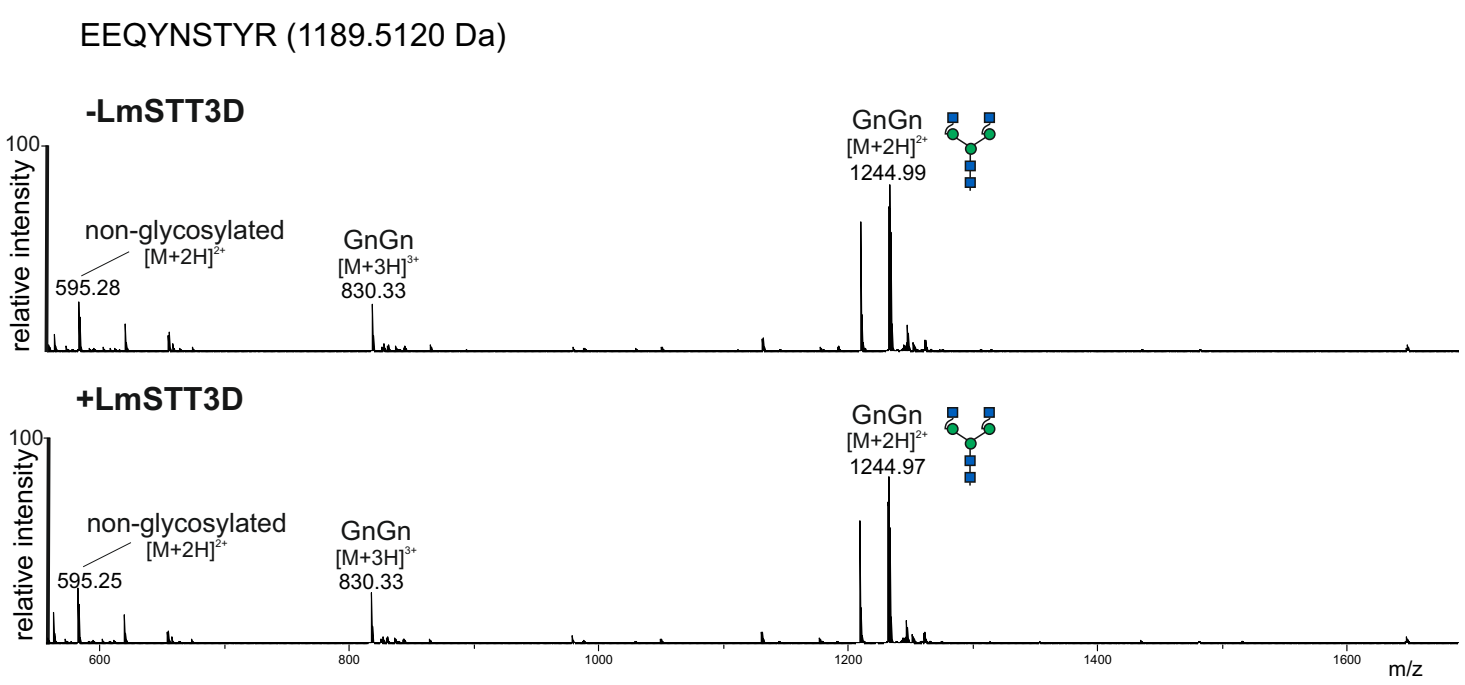


Figure S1 Mass spectra of (glyco)peptides from IgG. IgG was transiently expressed in *N. benthamiana* ΔXT/FT in the presence or absence of LmSTT3D-GFP. Purified IgG was trypsin digested and analysed by LC-ESI-MS.

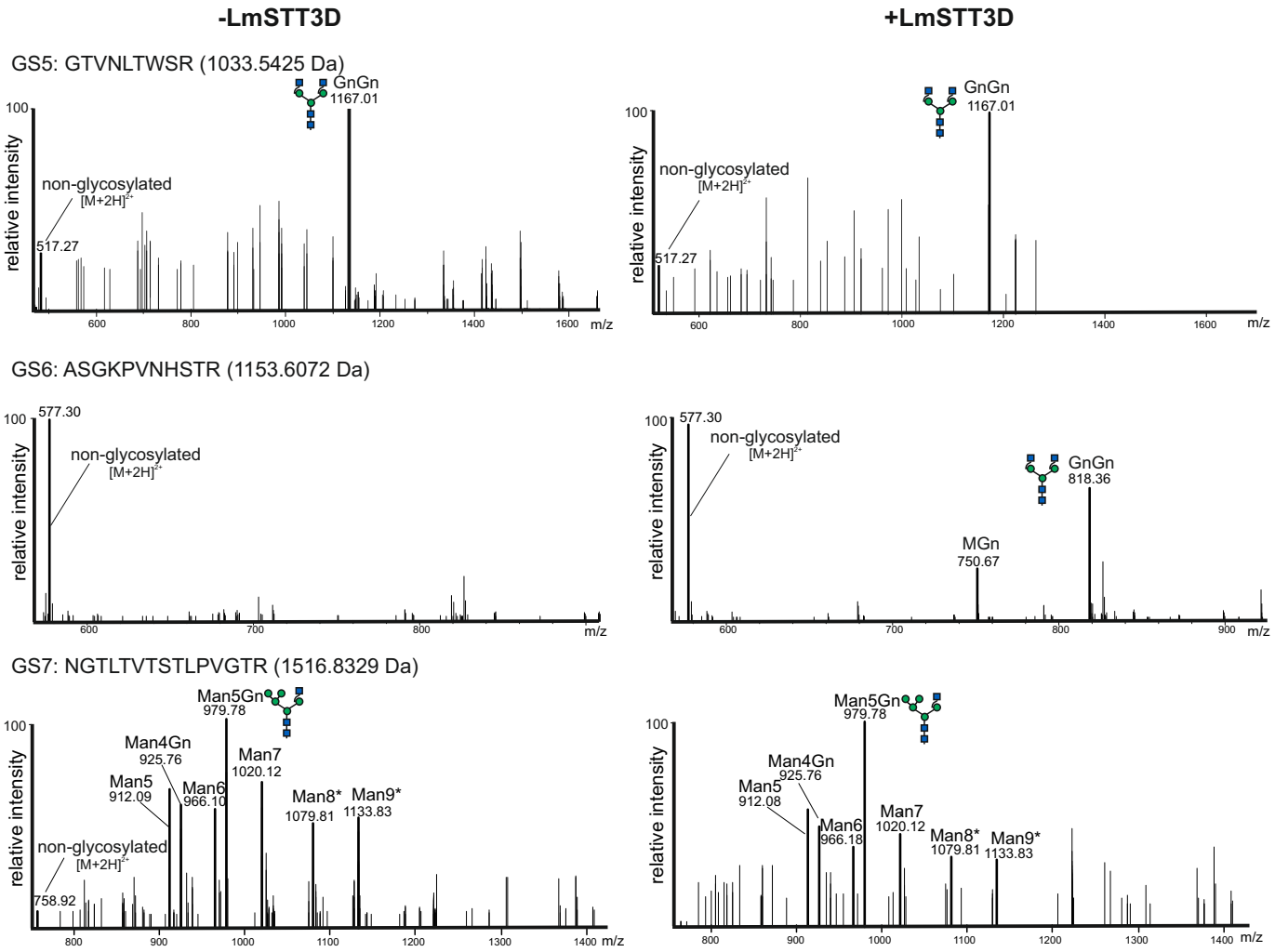


Figure S2 Mass-spectra of IgE (glyco)peptides harbouring glycosylation sites (GS) 5, 6 and 7. IgE was transiently expressed in *N. benthamiana* ΔXT/FT in the presence or absence of LmSTT3D-GFP. Purified IgE was trypsin digested and analysed by LC-ESI-MS. The mass range displays the major glycoforms. Except for the glycopeptides harbouring GS5, the peaks corresponding to triple charged glycopeptides ($[M+3H]^{3+}$) are indicated. The major glycoform on GS5 as well as all non-glycosylated peptides are shown in their double charged ($[M+2H]^{2+}$) state. Ammonia adducts are denoted by an asterisk.

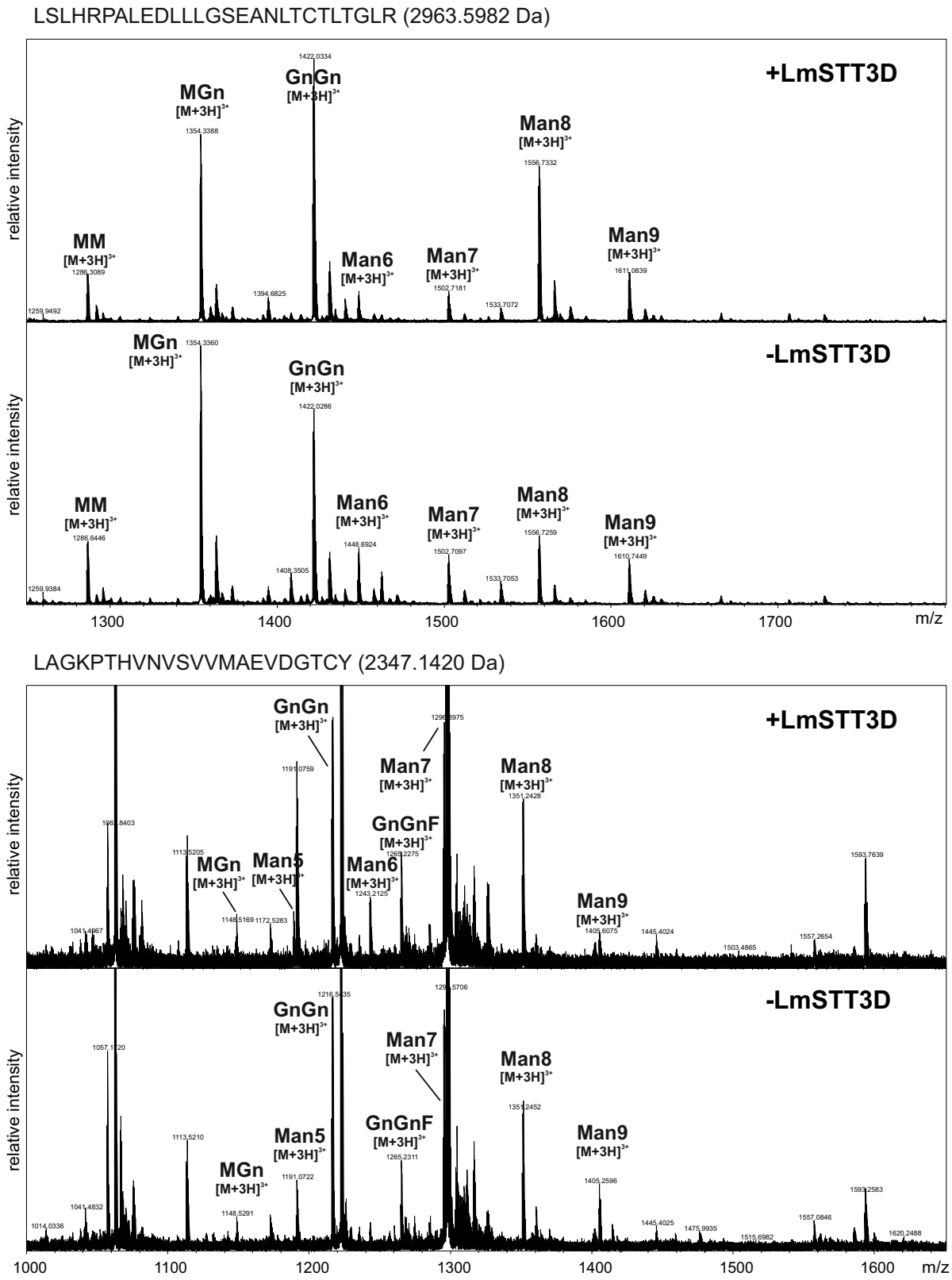


Figure S3 Mass-spectra of IgA1 (glyco)peptides. IgA1 was transiently expressed in *N. benthamiana* ΔXT/FT in the presence or absence of LmSTT3D-GFP. The purified IgA1 heavy chain was trypsin digested and analysed by LC-ESI-MS.

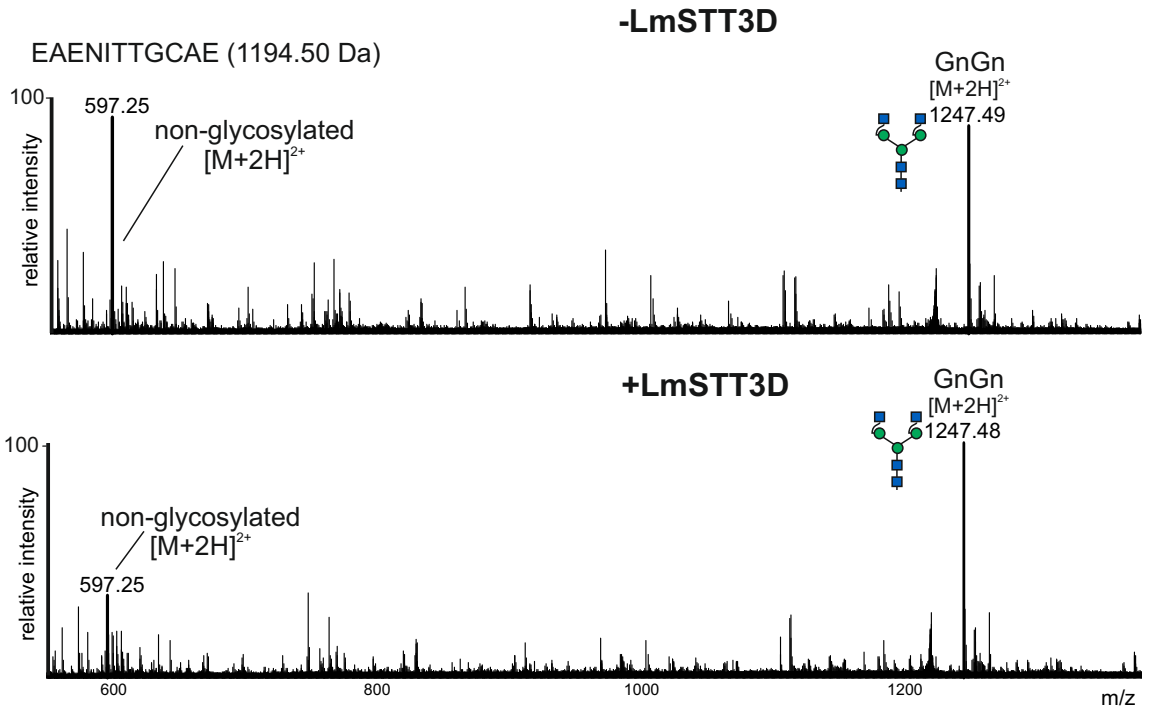
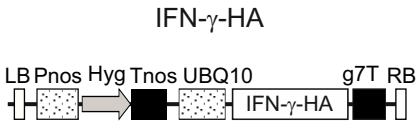


Figure S4 Mass-spectra of EPO-Fc (glyco)peptides. EPO-Fc was transiently expressed in *N. benthamiana* ΔXT/FT in the presence or absence of LmSTT3D-GFP. Purified EPO-Fc was trypsin+GluC digested and analysed by LC-ESI-MS. The mass range displaying the major glycoform from the peptide corresponding to glycosylation site 1 (due to the iodoacetamide treatment the cysteine is present as carbamidomethyl-cysteine) is shown.

(a)



(b)

MANKHLSLSLFLVLLGLSASLAQDPYVKEAENLKKYFNAGHSDVADNGTLFLGILKNWKEESDRKIMQ
SQIVSFYFKLFKNFKDDQSIQKSVETIKEDMNVKFFNSNKKKRDDFEKLTNYSVTDLNVQRKAIHELI
QVMAELSPAAKTGKRKRSQMLFRGRRASQGSYPYDVPDYASLYPYDVPDYASLYPYDVPDYASL

Figure S5 Schematic illustration of the IFN-γ–HA expression construct (a) and the IFN-γ–HA amino acid sequence (b). (a) LB: left border; Pnos: nopaline synthase gene promoter; Hyg: hygromycin B phosphotransferase gene; Tnos: nopaline synthase gene terminator; UBQ10: *A. thaliana* ubiquitin-10 promoter; IFN-γ-HA: coding sequence of human interferon γ fused to a 3x hemagglutinin (HA) tag; g7T: agrobacterium gene 7 terminator; RB: right border. (b) Amino acid sequence of IFN-γ–HA. The signal peptide from barely alpha-amylase is marked in green. The two N-glycosylation sites (NGT and NYS) are marked in blue. A short dipeptide linker is marked in red and the 3x HA tag is highlighted in grey.

Table S1 Comparison of the N-glycosylation site occupancy of native and recombinant glycoproteins.

Protein	N-glycosylation site (GS)	source	% glycosylated	Reference
Fc	GS1 NST	HEK293	> 90	Jez et al. 2012 and unplublished
IgG	GS1 NST	CHO	99	Rustandi et al. 2008
	GS1 NST	human serum	~100	Karnoup et al. 2007
IgE	GS3 NKT	HEK293	80	Montero-Morales et al. 2017
	GS5 NLT	HEK293	91	
	GS6 NHS	HEK293	0	
	GS3 NKT	human serum	75	Montero-Morales et al. 2017
	GS5 NLT	human serum	85	
	GS6 NHS	human serum	0	
	GS3 NKT	human serum	80	Plomp et al. 2014
	GS5 NLT	human serum	98	
IgA1	GS1 NLT	HEK293	99	Göritzer et al. 2017
	GS2 NVS	HEK293	64	
	GS1 NLT	human colostrum	< 50	Huang et al. 2015
	GS2 NVS	human colostrum	< 10	
	GS1 NLT	human serum	85	Hülsmeier et al. 2016
EPO-Fc	all sites	CHO	100	Taschwer et al. 2012
EPO	all sites	CHO	100	Gong et al. 2013
	all sites	*P. pastoris*	100	
	all sites	human serum	~100	Skibeli et al 2001
IFN-γ	GS1 + GS2	CHO	65	Wong et al. 2010
	GS1 + GS2	human blood cells	~67	Rinderknecht et al. 1984

Please note for some glycoproteins no precise data for site-specific N-glycosylation site occupancy could be obtained from literature and for mammalian cells glycosylation efficiency is dependent on culture conditions.

References

Gong, B., Burnina, I., Stadheim, T.A. and Li, H. (2013) Glycosylation characterization of recombinant human erythropoietin produced in glycoengineered Pichia pastoris by mass spectrometry. J Mass Spectrom **48**, 1308-1317.

Göritzer, K., Maresch, D., Altmann, F., Obinger, C. and Strasser, R. (2017) Exploring Site-Specific N-Glycosylation of HEK293 and Plant-Produced Human IgA Isotypes. *J Proteome Res* **16**, 2560-2570.

Huang, J., Guerrero, A., Parker, E., Strum, J.S., Smilowitz, J.T., German, J.B. and Lebrilla, C.B. (2015) Site-specific glycosylation of secretory immunoglobulin A from human colostrum. J Proteome Res **14**, 1335-1349.

Hülsmeier, A.J., Tobler, M., Burda, P. and Hennet, T. (2016) Glycosylation site occupancy in health, congenital disorder of glycosylation and fatty liver disease. Sci Rep **6**, 33927.

Jez, J., Antes, B., Castilho, A., Kainer, M., Wiederkum, S., Grass, J., Rüker, F., Woisetschläger, M. and Steinkellner, H. (2012) Significant impact of single N-glycan residues on the biological activity of Fc-based antibody-like fragments. J Biol Chem **287**, 24313-24319.

Karnoup, A.S., Kuppannan, K. and Young, S.A. (2007) A novel HPLC-UV-MS method for quantitative analysis of protein glycosylation. J Chromatogr B Analyt Technol Biomed Life Sci **859**, 178-191.

Montero-Morales, L., Maresch, D., Castilho, A., Turupcu, A., Ilieva, K.M., Crescioli, S., Karagiannis, S.N., Lupinek, C., Oostenbrink, C., Altmann, F. and Steinkellner, H. (2017) Recombinant plant-derived human IgE glycoproteomics. *J Proteomics* **161**, 81-87.

Plomp, R., Hensbergen, P.J., Rombouts, Y., Zauner, G., Dragan, I., Koeleman, C.A., Deelder, A.M. and Wuhrer, M. (2014) Site-specific N-glycosylation analysis of human immunoglobulin E. J Proteome Res 13, 536-546.

Rinderknecht, E., O'Connor, B.H. and Rodriguez, H. (1984) Natural human interferon-gamma. Complete amino acid sequence and determination of sites of glycosylation. J Biol Chem **259**, 6790-6797.

Rustandi, R.R., Washabaugh, M.W. and Wang, Y. (2008) Applications of CE SDS gel in development of biopharmaceutical antibody-based products. Electrophoresis **29**, 3612-3620.

Skibeli, V., Nissen-Lie, G. and Torjesen, P. (2001) Sugar profiling proves that human serum erythropoietin differs from recombinant human erythropoietin. Blood **98**, 3626-3634.

Taschwer, M., Hackl, M., Hernández Bort, J.A., Leitner, C., Kumar, N., Puc, U., Grass, J., Papst, M., Kunert, R., Altmann, F. and Borth, N. (2012) Growth, productivity and protein glycosylation in a CHO EpoFc producer cell line adapted to glutamine-free growth. J Biotechnol **157**, 295-303.

Wong, D.C., Wong, N.S., Goh, J.S., May, L.M. and Yap, M.G. (2010) Profiling of N-glycosylation gene expression in CHO cell fed-batch cultures. *Biotechnol Bioeng* **107**, 516-528.

Appendix II

Curriculum vitae

Personal Information

Name: Gernot Beihammer

Date of birth: 07.06.1988

Place of birth: Sankt Johann in Tirol

Address: Hannovergasse 35/6, 1200 Wien

Email: Gernot.beihammer@boku.ac.at

Mobile: 00436504452982

Education

since 07/16	**Member of the PhD program BioTop** (Biomolecular Technology of Proteins) University of Natural Resources and Life Sciences Vienna, Austria PhD Thesis: "Identification of Lewis A containing glycoproteins in plants and elucidation of the structure of plant Glycosylphosphatidylinositol anchors."
10/2012-07/2016	**Master Program Biotechnology** University of Natural Resources and Life Sciences Vienna, Austria Master Thesis: "Characterization of the N-glycosylation efficiency of recombinant glycoproteins produced in plants."
10/08-07/12	**Bachelor's Program Food Science and Biotechnology** University of Natural Resources and Life Sciences Vienna, Austria Bachelor Thesis: "Assessment of phosphorus mobilization by soil bacteria via ICP-MS analysis and LC-ICP-MS speciation"
09/1998-06/16	**Bundesgymnasium Sankt Johann in Tirol (High School)**

Publications

SCI Publications

Castilho A, **Beihammer G**, Pfeiffer C, Göritzer K**,** Montero-Morales L, Vavra U, Maresch D, Grünwald-Gruber C, Altmann F**,** Steinkellner H**,** Strasser R. An oligosaccharyltransferase from Leishmania major increases the N-glycan occupancy on recombinant glycoproteins produced in *Nicotiana benthamiana*. *Plant Biotechnol. J.* 2018. 16:1700-1709. doi: 10.1111/pbi.12906

Beihammer G, Maresch D, Altmann F, Strasser R. Glycosylphosphatidylinositol-Anchor Synthesis in Plants: A Glycobiology Perspective. *Front. Plant Sci*. 2020. 11:611188. doi: 10.3389/fpls.2020.611188

Beihammer G, Maresch D, Altmann F, Van Damme EJM, Strasser R. Lewis A Glycans Are Present on Proteins Involved in Cell Wall Biosynthesis and Appear Evolutionarily Conserved Among Natural *Arabidopsis thaliana* Accessions. *Front. Plant Sci.* 2021. 12:630891. doi: 10.3389/fpls.2021.630891

Poster Presentations at conferences

Beihammer G, Romero Pérez A, Van Damme E and Strasser R (2019) Investigating a putative role for Lewis A - structures in plant pathogen-defence (Poster) [25th International Symposium on Glycoconjugates, Milan, Italy, 25-31 August 2019]

Beihammer G, Maresch D, Altmann F and Strasser R **(**2018) Investigating the role of Lewis A-structures for plant pathogen defence (Poster) [29th Joint Glycobiology meeting 2018, Ghent, Belgium, 21-23 October 2018]

Beihammer G, Maresch D, Altmann F and Strasser R (2018) Investigating the role of Lewis A-structures in plants (Poster) [Plant Physiology and Biochemistry, Vienna, Austria, 9-10 July 2018]

Appendix III

Eidesstattliche Erklärung

Ich erkläre ehrenwörtlich, dass ich die vorliegende Arbeit selbständig und ohne fremde Hilfe verfasst habe, andere als die angegebenen Quellen nicht verwendet habe und die den benutzten Quellen wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Wien, 1.4.21

Ort, Datum

Unterschrift