Therapeutic properties of recombinant proteins are very often affected by the composition and heterogeneity of their glycans. Conventional expression systems for recombinant pharmaceutical proteins typically do not address this problem and produce a mixture of glycoforms that are neither identical to human glycans nor optimized for enhanced efficacy. In terms of glycosylation, plants offer certain advantages over mammalian cells as the N-glycosylation pathway of plants is comparably simple and a typical mammalian O-glycosylation pathway is not present at all. During the last ten years we have developed a plant-based expression platform for the generation of recombinant glycoproteins with defined N-glycans. Now we have extended our tool-box for glycoengineering in the tobacco related species *Nicotiana benthamiana* toward the production of tailored mucin-type O-glycans on recombinant proteins.

Introduction

In recent years plants have emerged as interesting new hosts for the production of pharmaceutical proteins as they provide several advantages compared with conventional mammalian cell-based expression systems. Cultivation of plants is comparably simple and inexpensive. Plants offer virtually unrestricted scalability and plants are considered safe as they lack human pathogens. In addition, major advances in expression technologies, downstream processing and the implementation of GMP manufacturing make plants attractive alternatives for the production of recombinant biopharmaceuticals. This development is also reflected by the recent marketing approval of a plant-produced recombinant α-glucocerebrosidase for enzyme-replacement therapy in humans. Within the different plant-based expression systems transient production using viral-based vectors is very promising as it is very flexible and allows high-yield protein production within a few days. This approach is, for example, optimal suited for the production of recombinant therapeutics to treat infectious diseases. A monoclonal antibody against Ebola virus has recently been produced by transient expression in *N. benthamiana*.

Another advantage of plants is their ability to assemble proteins and perform posttranslational modifications similar to mammalian cells. Glycosylation is the most important posttranslational modification of recombinant biopharmaceuticals and many different classes of recombinant drugs like monoclonal antibodies, growth factors or hormones are glycosylated. In mammalian cells two major types of protein glycosylation exist: N-glycosylation, which is initiated by transfer of a preassembled oligosaccharide to an asparagine residue present in unfolded or partially folded polypeptides and O-glycosylation, which involves the transfer of a single sugar residue to a serine or threonine of a fully folded protein. Both types of glycosylation are subjected to extensive processing reactions resulting in the generation of a large variety of different glycans on proteins. It has been clearly demonstrated that certain glycoforms drastically affect the efficacy of recombinant proteins. Yet,
the conventional expression hosts like CHO cells produce only sub-optimal O-glycans and frequently heterogeneity of N- and O-glycans is observed on recombinant therapeutics. For example, the presence of a core α-1,6-fucose residue on the N-glycan of the IgG heavy chain affects the binding to the Fc receptors and subsequently alters the effector functions of IgG. In addition, there are raising concerns regarding the presence of non-human glycosylation like the incorporation of N-glycolylneuraminic (Neu5Gc) into glycosylated biopharmaceuticals produced in some mammalian cell lines.1

In contrast to N-glycosylation, the biological function of distinct O-glycan structures on recombinant glycoproteins is still poorly understood. Given the fact that certain biopharmaceutical proteins, like for example IgAs or erythropoietin (EPO), are O-glycosylated, a production platform for tailored O-glycans might not only improve product quality by reducing glycan heterogeneity but could also enhance the biological activities of recombinant therapeutics. Similar to mammalian N-glycosylation, O-glycan initiation and processing is rather a complex process. In humans, for example, initiation of mucin-type O-glycosylation can be performed by up to 20 different UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferases (GalNAc-Ts) which catalyze the transfer of a single N-acetylgalactosamine (GalNAc) residue to serine/threonine amino acids (GalNAcβ1→Ser/Thr). These GalNAc-Ts have only partially overlapping substrate specificities and their regulation as well as their precise biological role in mammals is still not entirely clear. Further elongation of GalNAc residues is performed by Golgi-located glycosyltransferases and results in different O-glycan structures including core 1 to core 4 structures. The uncontrolled expression of these endogenous glycosyltransferases in mammalian cell-based production hosts prevents the generation of defined structures leading to the formation of rather heterogeneous O-glycans on recombinant proteins. Consequently, there is a demand for novel production systems or engineered conventional expression hosts that result in the production of tailored O-glycans on recombinant proteins.

N-Glycan Engineering in Plants

Recombinant glycoproteins typically contain an N-terminal signal peptide for targeting to the secretory pathway. Upon emergence of the polypeptide in the ER-lumen the oligosaccharide transferase complex scans the amino acid sequence for the presence of N-glycosylation sites (Asn-X-Thr/Ser) and transfers the precursor assembled GlcMan9GlcNAc2 oligosaccharide to an asparagine residue within the consensus site. In the ER the terminal glucose residues are removed and the trimming of distinct mannose residues is initiated by class 1 α-mannosidases (MNS proteins). These N-glycan processing events are linked to ER-quality control processes like the calnexin/calreticulin cycle and the ER-associated degradation (ERAD) pathway. Additional mannose residues are then removed by Golgi-located MNS1/MNS2 and the Man/GlcNAc α3,4N-glycanase, structure (Fig. 1) is processed into complex N-glycans by the action of N-acetylglcosaminyltransferase I (GnTI), Golgi-α2-mannosidase II (GMI) and N-acetylglcosaminyltransferase II (GnTII). Further modification of plant complex N-glycans is typically performed by β1,2-xyllosyltransferase and core α1,3-fucosyltransferase, which add β1,2-xyllose and core α1,3-fucose, respectively. These two glycosyltransferases are not present in mammals and the administration of glycoprotein therapeutics with xylosylated and fucosylated N-glycans might therefore lead to adverse immunogenic reactions in mammals. As a consequence the N-glycosylation pathway of several plant species has been engineered in a way to eliminate the expression of the corresponding non-mammalian glycosyltransferases. The glyco-engineered N. benthamiana plants known as ΔXF plants produce predominantly GlcNAcβ1→Manα1,3GlcNAc2 structures on recombinant proteins. In subsequent studies this complex N-glycan has been processed by the incorporation of β1,4-linked galactose, terminal α2,6-linked sialic acid, core α1,6-linked fucose and the formation of tri- and tetra-antennary glycans. All these common mammalian complex N-glycan modifications are not found in plants and have to be engineered to make plants optimal hosts for the production of recombinant glycoprotein therapeutics.

O-glycan Engineering in Plants

While the early steps of the N-glycan processing pathway are highly conserved in mammals and plants the O-glycosylation capacity is completely different. O-glycosylation in plants is frequently found on hydroxyl groups of hydroxyproline (Hyp) residues and only to a lesser extent on other amino acids like serines. The hydroxyl groups of the amino acids can be modified by addition of a large arabino-oligogalactan polysaccharide, a short arabinoxylan chain, or a single galactose residue, but GalNAc is typically not present. Accordingly, mucin-type O-glycosylation has not been found on plant-produced recombinant proteins. Plants are therefore very well suited for the development of mammalian O-glycosylation pathway.

In the first attempt to initiate O-GalNAc formation in plants aTerritoria entertwistula UDP-GlcNAc 4-epimerase and aCantharidina elegans UDP-GlcNAc/UDP-GalNAc transporter were co-expressed with human GalNAc-T2 in N. benthamiana. Mass spectrometry revealed the attachment of a single GalNAc residue to an endogenous plant protein and lectin blotting indicated the incorporation of GalNAc residues into a co-expressed recombinant protein containing a mucin-type sequence motif with several serine/threonine residues. In a similar study, it was shown by structural analysis that transient expression of Pseudomonas aeruginosa GlcNAc C4-epimerase and human GalNAc-T2 in N. benthamiana are sufficient for O-GalNAc formation on a polypeptide containing a mucin-type tandem repeat. This strategy was also applied to generate transgenic Arabidopsis thaliana plants and tobacco BY2 cells with the capacity for the production of O-linked GalNAc on recombinant proteins. Very recently, we demonstrated that GalNAc can be attached to the single O-glycosylation site present in human EPO. In our effort,
A common terminal modification of mucin-type O-linked glycans is the attachment of sialic acid (N-acetylneuraminic acid: Neu5Ac). As plants do not contain this sialic acid the whole biosynthetic pathway for conversion of the precursor UDP-GlcNAc to CMP-sialic acid and its transport into the Golgi has to be transferred into plants (Fig. 1).23,36 Golgi-resident sialyltransferases transfer sialic acid from CMP-sialic acid to protein-linked glycans. On urinary and recombinant EPO the most frequent core 1 extension is the incorporation of two sialic acid residues resulting in the formation of disialyl-core 1 structures. The generation of this O-linked glycan on plant-derived EPO-Fc was achieved by co-expression of the two mammalian sialyltransferases ST3Gal-I and ST6GalNAc-III/IV the formation of the Gal\(\beta_1\)-3GalNAcα-Ser/Thr core 1 structure (also known as T antigen) (Fig. 1). In humans, there is a single C1GALT1 enzyme that generates this commonly found mucin-type O-glycan. Importantly, proper formation of the T antigen in mammals requires the presence of the chaperone COSMC, which prevents aggregation and subsequent degradation of C1GALT1 in the ER.35 In N. benthamiana co-expression of COSMC with human C1GALT1 did not result in the incorporation of significant amounts of galactose suggesting that human COSMC is either not functional in plants or not correctly localized. By contrast, transient expression of the Drosophila C1GALT1 along with human GalNAcT2 in N. benthamiana resulted in the formation of core 1 structures on EPO-Fc.31

Figure 1. Schematic presentation of the engineered O- and N-glycosylation pathways in N. benthamiana ΔXF plants. Only the Golgi processing steps are shown. Heterologous expressed proteins for the generation of disialylated O-glycans are shown in red (GNE: UDP-N-acetylglucosamine 2-epimerase/N-acetylmannosamine kinase; NANS: N-acetylneuraminic acid phosphate synthetase; CMAS: CMP-N-acetylneuraminic acid synthetase; CST: CMP-NeuAc transporter). Additional expression of two mammalian glycosyltransferases (GaIT: \(\beta_1,4\)-galactosyltransferase; ST: \(\alpha_2,6\)-sialyltransferase) is required for complex N-glycan modification and expression of mammalian N-acetylglucosaminyltransferase IV (GnTIV) and N-acetylglucosaminyltransferase V (GnTV) is required for the generation of multi-antennary complex N-glycans (these mammalian enzymes are depicted in blue). Endogenous plant proteins involved in early N-glycan processing reactions (MNS1/2: Golgi-\(\alpha\)-mannosidase I; GnTI: N-acetylglucosaminyltransferase I; GMII: Golgi-\(\alpha\)-mannosidase II; GnTII: N-acetylglucosaminyltransferase II) or conversion and transport of nucleotide sugars for O- and N-glycan formation are shown in black (NANP: Neu5Ac-9-phosphate phosphatase).
with human GalNAc-T2 and Drosophila C.6GalNAct. In total, the implementation of this O-glycan modification required the coordinated expression of eight non-plant proteins in N. benthamiana. All of them were transiently expressed by infiltration of leaves with Agrobacterium tumefaciens containing the individual expression constructs.

In contrast to natural glycosyltransferases like EPO or IgAs contain both types of glycosylation. It is therefore essential to develop a plant-based production platform for generation of tailored O- and N-glycans. For EPO the terminal sialic acid residues present on N-linked glycans are of utmost importance for its in vivo function. In N. benthamiana, the additional expression of human β-1,4-galactosyltransferase (GaT1) and rat α2,6-sialyltransferase (ST) resulted in successful manipulation of both glycosylation pathways and the formation of sialylated N- and O-glycans. Importantly, our experimental data showed that the different mammalian glycosyltransferases for attachment of sugars to N- or O-glycans do not cross-react indicating that they retain their specificity when heterologously expressed in plants.

**Future of O-glycan Engineering in Plants**

Due to the absence of naturally occurring mucin-type O-glycosylation plants are very well suited for the generation of highly homogenous O-glycan structures. However, despite all the progress there are still some limitations. One obstacle that has to be overcome in plants is the formation of Hyp residues in proximity to exposed O-glycosylation sites on recombinant mammalian proteins. These hydroxylated proline residues have been detected on stable and transiently expressed mucin-derived peptides and close to Ser-126 on EPO-Fc. The Hyp residues serve as starting points for the attachment of plant-specific O-glycosylation and subsequently increase the risk of unwanted side effects like immunogenic or allergenic reactions when present on recombinant biopharmaceuticals. Different strategies have been proposed to prevent the formation of plant-specific Hyp glycosylation. In particular, knockout or knockdown of the corresponding prolyl-4 hydroxylase(s) holds great promise to achieve a robust system for expression of different types of O-glycosylated proteins without plant specific Hyp-glycosylation.

The plant-produced O-glycan structures are still not homogenous and complete conversion of all transferred GalNAc-residues to distalylated core 1 structure has not been achieved yet. Apart from appropriate expression and localization of the heterologous proteins efficient glycosylation depends on the availability of sufficient amounts of nucleotide sugars. One strategy for improvement is therefore a better control of nucleotide sugar levels. In particular, an increase in CMP-sialic acid should be beneficial for optimization of the sialylation capacity of engineered plants. CMP-sialic acid originates from endogenous UDP-GlcNAc and in mammalian cells its formation is tightly controlled and subjected to feedback regulation. Feedback inhibition can be prevented by expression of a mutated version of UDP-N-acetylglucosamine 2-epimerase/N-acetylmannotransferase kinase (GNE), the key enzyme for CMP-sialic acid formation. Apart from metabolic engineering of the nucleotide sugar level it would also be highly desirable to develop tools for controlling the Golgi transport and steady-state levels of the nucleotide sugars in the Golgi.

For the generation of sialylated N- and O-glycans we have simultaneously expressed ten individual constructs and the recombinant glycoprotein by infiltration of a single Agrobacterium mixture into N. benthamiana leaves. While a further increase of constructs in the infiltration mixture is still possible, e.g., required for generation of branched sialylated N-glycans, it would be very convenient to combine several of the required proteins and express them from multi-cassette expression vectors. For example, the three enzymes for for CMP-sialic acid biosynthesis and the corresponding Golgi transporter could all be produced from a single construct leading to a more consistent supply with the donor substrate and as a consequence to increased amounts of fully sialylated glycoproteins.

Future developments will also concentrate on O-glycan engineering steps that have not been performed so far, like the formation of core 2, 3 and 4 structures. Apart from incorporation of galactose, GalNAc and sialic acid these O-glycans can contain GlcNAc and fucose in different linkages. The glycosyltransferases that transfer these sugars are typically not present in plants and have to be expressed similar to the glycosylation mixture for core 1 structure formation and capping with sialic acid. Since the biosynthesis of mucin-type O-glycans occurs in a step-wise manner it is crucial to maintain the appropriate order of modification steps when performed in plants. The sub-Golgi localization of glycosyltransferases in the cis, medial and trans-Golgi plays a central role in the control of glycosylation. Golgi-targeting signals of mammalian glycosyltransferases are functional in plants. However, subtle differences in sub-Golgi targeting have been observed for some mammalian glycosyltransferases resulting in altered glycan profiles on plant-produced recombinant proteins. For the generation of sialylated N-glycans in plants the future will show whether these developments can be translated into industrial applications for the production of next-generation biopharmaceuticals.

**Disclosure of Potential Conflicts of Interest**

No potential conflicts of interest were disclosed.

**Acknowledgements**

Related work in my group is supported by a grant from the Federal Ministry of Transport, Innovation and Technology (bmvit) and Austrian Science Fund (FWF): TRP 242-B20.
 References

1. Fischer R, Schildberg F, Shabalin S, Tzevat RM, Dessaud E. BMPR iv recruit circularized phosphoenolpyruvate carboxylase
2. Luscombe NM, Liao J, Palatnik J, et al. N-glycan trimming by GRP is essential for functional glycosylation.
3. Li X, De la Torre-Colindres M, Leroux C, et al. N-Acetylation of a polypeptide is required for efficient folding of a soluble ectodomain.
4. Lin V, Hidalgo-Leonard R, Supattapone S, et al. Generation of N-glycan trimming variants with truncated complex N-glycans.
5. Malmberg K, Michel V, Glauert J, et al. N-linked oligosaccharide trimming improves binding to human FcγR IIa by B cell lines. J Biol Chem 2008; 283:13283-13291.
6. Steinbeker J, Wek R, Brüning K, et al. Regulation of mammalian embryonic stem cell pluripotency by treating tobacco plants with N-glycan trimming agents. J Plant Physiol 2012; 169:2024-2032.
7. Schiel Y, Luy N, Borm L, et al. Amyloidogenic O-glycosylation: a classification of the polypeptide GalNAc-transferase gene family. Biochim Biophys Acta 2009; 1794:1613-1619.
8. Matsui K, Nishimura K, Yamashita T, et al. Generation of glycoengineered Nicotiana benthamiana for the production of non-Hodgkin's lymphoma. Ann Oncol 2012; 23:2420-2427.
9. Li X, Hidalgo-Leonard R, Supattapone S, et al. N-Acetylation of a polypeptide is required for efficient folding of a soluble ectodomain. J Biol Chem 2008; 283:13283-13291.
10. Schiel Y, Luy N, Borm L, et al. Amyloidogenic O-glycosylation: a classification of the polypeptide GalNAc-transferase gene family. Biochim Biophys Acta 2009; 1794:1613-1619.
11. Matsui K, Nishimura K, Yamashita T, et al. Generation of glycoengineered Nicotiana benthamiana for the production of non-Hodgkin's lymphoma. Ann Oncol 2012; 23:2420-2427.
12. Li X, Hidalgo-Leonard R, Supattapone S, et al. N-Acetylation of a polypeptide is required for efficient folding of a soluble ectodomain. J Biol Chem 2008; 283:13283-13291.
13. Schiel Y, Luy N, Borm L, et al. Amyloidogenic O-glycosylation: a classification of the polypeptide GalNAc-transferase gene family. Biochim Biophys Acta 2009; 1794:1613-1619.
14. Matsui K, Nishimura K, Yamashita T, et al. Generation of glycoengineered Nicotiana benthamiana for the production of non-Hodgkin's lymphoma. Ann Oncol 2012; 23:2420-2427.
15. Li X, Hidalgo-Leonard R, Supattapone S, et al. Generation of glycoengineered Nicotiana benthamiana for the production of non-Hodgkin's lymphoma. Ann Oncol 2012; 23:2420-2427.
42. Bork K, Kastner W, Winklmayr W, Heinzloren R. Enhanced sialylation of EPO by overexpression of UDP-GlcNAc 2-epimerase/ManAc kinase containing a sialuria mutation in CHO cells. FEBS Lett 2007; 591:497-8; PMID:17586195; http://dx.doi.org/10.1016/S0014-5793(07)01836-0.

43. Nagah B, Van Damme EL, Calloïcquet N, Zehou L, Taweret J, Dinhue F, et al. Biopolysaccharide activity, new COPD-relevant EPO/N, from a newly transformed Nicotiana benthamiana plant presenting transaminase N-glycan variant. J Biotechnol 2012; 168:242-50; PMID:22630611; http://dx.doi.org/10.1016/j.jbiotec.2012.03.005.

44. 44. Beilc C, Gosselin R, Kandla R, Marfikour S. Golden gene shuffling: a one-pot DNA shuffling method based on type II restriction enzymes. Plant Sci 2009; 177:53; PMID:19406741; http://dx.doi.org/10.1016/j.plantsci.2009.05.053.

45. Sarret-Polignani A, Zakari E, Zardales M, Janet P, Fernández-del-Castillo A, Gómez A, et al. GoldenHind: an iterative cloning system for standardized assembly of recombinant genetic modules. Plant Sci 2011; 179:78; http://dx.doi.org/10.1016/j.plantsci.2011.02.012.