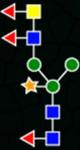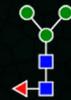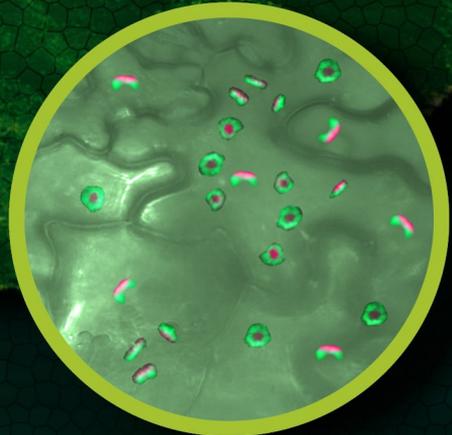# Glyco-engineering in plants:

Fucosyltransferase characterisation and application for the production of immunomodulatory helminth glycoproteins

Kim van Noort

Propositions

1. Expression of *Schistosoma mansoni* glycoproteins in *Nicotiana benthamiana* demonstrates an intimate co-evolution between glycosyltransferases and their substrates.
   (this thesis)

2. Subcellular localisation, micro-environmental conditions and the glycan acceptor determine glycosyltransferase activity.
   (this thesis)

3. An internship should be an integral part of the PhD programme.

4. In order to combat malaria, investments should focus on vector control rather than drug or vaccine development.

5. To ensure comparability in Europe the doctoral degree should be regulated alike the bachelor and master degree in the qualifications framework of the Bologna process.

6. A picture says more than thousand words, but only a movie shows the complete picture.

7. A tennis match is like scientific writing, a mind game of which the outcome does not reflect the journey.

Propositions belonging to the thesis entitled:
Glyco-engineering in plants: Fucosyltransferase characterisation and application for the production of immunomodulatory helminth glycoproteins

Kim van Noort
Wageningen, 27th of March 2020

Glyco-engineering in plants:

Fucosyltransferase characterisation and application for the production of immunomodulatory helminth glycoproteins

Kim van Noort

Glyco-engineering in plants:

Fucosyltransferase characterisation and
application for the production of
immunomodulatory helminth glycoproteins

Kim van Noort

to my family

Table of contents

Chapter 1

General introduction

Kim van Noort

Laboratory of Nematology, Wageningen University and Research, Wageningen, the Netherlands.

The advent of recombinant DNA technology that allows the expression of genes of one organism in another organism, led to the production of pharmaceutical proteins in heterologous systems. Plants and plant cells are among these heterologous systems and have proven to be a suitable host for various proteins of pharmaceutical interest. Over the last two decades plants have been used for the production of (potential) pharmaceutical proteins in relative high amounts. The latter includes cases where isolation from the original source was not possible, laborious and/or resulted in low yields. Currently, almost all biopharmaceutical proteins originate from humans, such as antibodies or hormones, as for instance insulin. Many of these human proteins are glycoproteins implying that they are 'coated' with sugars (glycans). These glycans can be important for protein folding and/or activity, and the composition of the glycans can vary between organisms. Hence, in order to produce a glycoprotein in another organism with (close to) native glycans, the glycosylation pathway has to be comparable or has to be engineered. Various glyco-engineering studies have shown the possibilities to "humanise" the plant glycosylation pathway, enabling the production of human glycoproteins with native glycans in plants. Pharmaceutical glycoproteins of human origin often have a high added value. Therefore, research in the field of plant molecular farming mainly focused on the production of human glycoproteins in plants. Other research fields could benefit from the development of this production platform as well. For example, studies on glycosylated vaccine candidates and the immunomodulatory properties of helminth secretions, show the high potential of these helminth proteins as pharmaceuticals. High amounts of native helminth protein are required for vaccine development and research on the biological and biopharmaceutical properties of helminth secretions. The helminth proteins for such research projects cannot be isolated in sufficient quantities from the helminth or its secretions. Moreover, many of these proteins are glycosylated with glycans that cannot be mimicked in current recombinant production systems. For this purpose, the focus in this thesis is on glyco-engineering of plants to establish a production system for native helminth glycoproteins.

Schistosome infection and secretions

Schistosomiasis is caused by infection with human parasitic trematodes of the genus *Schistosoma* and is the most important parasitic disease after malaria as it affects over 252 million people worldwide [1], [2]. Schistosomiasis belongs to the neglected tropical diseases and is a major health problem in the tropics and sub-tropics [3]. Currently, the only recommended drug against schistosome infection is Praziquantel. School-aged children are at high risk and regular anthelmintic drug treatment from young age onwards is suggested to prevent morbidity by schistosome infection in adulthood [4]. Therefore, the target groups of the World Health Organisation (WHO) are school-aged children and pregnant or lactating women. In 2006, 35.6% of all patients received the preventive chemotherapy required, which included 53.7% of the school-aged-children that required

treatment [2]. The goal of the WHO was to treat 75% of the school-aged-children against schistosome infection before 2020.

One of the schistosome species that can cause schistosomiasis is *Schistosoma mansoni*. This parasite infects humans via free-living cercariae in contaminated water (Figure 1). Cercariae penetrate the human skin and transform into schistosomula, immature adults. The schistosomula mature during their journey through the bloodstream and in the inferior mesenteric vein adult worms mate and start reproduction [5]. A female worm will produce around 350 eggs a day [6], [7]. In five to six days these eggs mature and a miracidium will develop surrounded by the Reynold's layer, the Von Lichtenberg's envelope and the egg shell [8]–[10]. The eggs will either cross the intestinal mucosa and leave the body via the faeces or are trapped in the intestine or liver, where they cause inflammation and organ damage. In contact with fresh water the eggs hatch and miracidia infect their intermediate host, fresh water snails of the species *Biomphalaria glabrata*. In these snails, miracidia develop into sporocysts that reproduce asexually. After four to six weeks cercariae will leave the snail in search for a new host, to restart the life cycle.

The control of schistosomiasis started already in the early 1980's [11]. In 1993 the schistosome genome project was introduced to identify new targets for drugs, vaccines and diagnostics [12]. In 2009, the first full genomes of *S. mansoni* and *S. japonicum* were sequenced and annotated, which were further improved in 2012 and 2019 [13]–[16]. Publication of the *S. mansoni* genome allowed studies that focused on identification of new drug targets based on transcriptional differences between the life stages, among others via transcriptome and proteome analysis [17]–[19].

A group of stage-dependent differentially regulated genes are the soluble egg antigens (SEA). SEA influence the host immune responses by shifting the initial T helper (Th) 1 response against the worm to a Th2 response [20]. The three most abundant SEA are IPSE/alpha-1, omega-1 and kappa-5. IPSE/alpha-1 induces IgE-dependent basophil-derived IL-4 and IL-13 production, thereby inducing alternatively activated macrophages [21]–[23]. Furthermore, IPSE/alpha-1 was shown to induce regulatory B cell development [24]. Omega-1 can condition dendritic cells (DCs) to induce a Th2 response, via a mannose receptor (MR) mediated uptake and subsequent degradation of messenger and ribosomal RNA via its ribonuclease activity [25]–[27]. Furthermore, omega-1 can induce weight loss and improve metabolic homeostasis in obese mice [28]. The induced weight loss and improved metabolic homeostasis additional to the Th2 priming function of omega-1, makes omega-1 a promising candidate for treatment and management of metabolic syndrome and type 2 diabetes. Additionally, omega-1 can be used to prevent allergic airway inflammation [29]. The function of the third most abundant SEA, kappa-5, is still unknown [30].

IPSE/alpha-1, omega-1 and kappa-5 are all glycoproteins with specific glycan compositions. ISPE/alpha-1 and omega-1 carry the same type of N-glycan motifs and are both produced in the same subshell area of *S. mansoni* eggs [31], [32]. In contrast to IPSE/alpha-1 and omega-1, kappa-5 carries different types of N-glycan motifs and is expressed in the eggs by the miracidium [30], [33], [34]. This shows that glycosylation can vary between cell-types. Next to variation in glycosylation within an organism and its various life stages, glycosylation can differ between organisms.
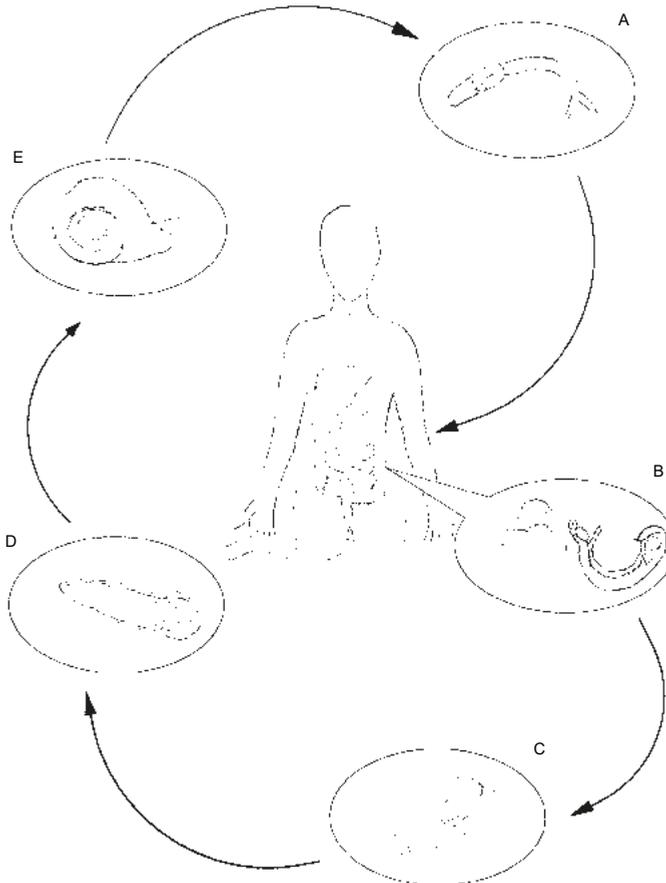
FIGURE 1 | Schematic representation of the *Schistosoma mansoni* life cycle. When humans come in contact with contaminated fresh water, free living cercariae (A) can penetrate the human skin. In the human host the schistosomes mature, reproduce sexually and lay eggs (B, paired female and male worm with eggs). The eggs (C) will either cross the intestinal mucosa and leave the body via the faeces or are trapped in organs such as the intestine or liver. In contact with fresh water the eggs hatch and miracidia (D) infect snails as intermediate host (E) and reproduce asexually. After four to six weeks cercariae will leave the snail in search for a human host, restarting the life cycle. This figure was designed in cooperation with J.M. van Noort-van Dijk and J.C. van Noort.

Glycans are important for correct protein folding, specific interactions, host mimicry or distinguishing self from non-self. For instance, during human blood transfusions, blood from an AB blood donor cannot be used for a patient with blood type O. Antibodies against A and B antigens can be formed and will cause blood clotting. These A and B antigens are sugar epitopes and only vary in the presence or absence of two sugar residues, which shows the impact of a small change in sugar composition. Moreover, blood group type is related to pathogen susceptibility and these host sugars can be used by for instances viruses for cell attachment and entry [35]–[38]. Pathogens and parasites also copy host glycans on their own glycoproteins, host mimicry, to use them as a smoke screen or to interact with host proteins. For instance, for *S. mansoni* infection of *B. glabrata*, the glycoproteins and glycolipids secreted by *S. mansoni* are decorated with glycans that are also found in the snails [39], [40]. Furthermore, a correlation was observed between the expression levels of shared glycan motifs and *B. glabrata* susceptibility towards *S. mansoni* infection [41]. Similarly, the human tissue passing stages of *S. mansoni* secrete glycoproteins and glycolipids with the human blood group type glycan motif LeX [42]–[45] (glycan motifs discussed in this thesis can be found at the end of this chapter in Table 1).

IPSE/Alpha-1 and omega-1 are an example of such LeX carrying glycoproteins secreted during human infection by *S. mansoni* [31]. The uptake of omega-1 by DCs is glycan-dependant and is important for its Th2 skewing function [25]. Moreover, the presence of one antenna with terminal LeX enhances Th2 polarisation in mice [46]. This shows the importance of glycans for receptor binding and subsequent immune response. Also the glycans on kappa-5 seem to be important, since the LDN and LDN-F glycan motifs found on kappa-5 N-glycans are implicated in granuloma formation or immunomodulation, respectively [33], [47]. Kappa-5 is the major component in SEA that carries LDN glycan motifs and antibodies found in sera of infected humans are directed against kappa-5 glycans. This makes kappa-5, next to omega, an interesting protein for host parasite interaction studies or studies to the immunomodulatory properties of helminth glycoproteins and the effect of their glycans.

However, high amounts of protein are required for these kinds of studies. At the moment, it takes 14 to 15 weeks (7 to 8 weeks snail infection and subsequently 7 weeks hamster infection) to isolate approximately 270.000 eggs from the intestines or 80.000 – 200.000 eggs from the liver of one hamster. Approximately 130 mg SEA can be isolated from 40 million intestinal eggs or 30-40 mg SEA from liver eggs. Per mg SEA 18 μg kappa, 15 μg IPSE/alpha-1 or 8 μg omega-1 can be isolated. For one *in vivo* metabolic experiment with kappa-5, requiring 3 mg kappa-5, 160 hamsters need to be infected with *S. mansoni*. This shows that at the moment SEA isolation requires a lot of time and resources, resulting in relative low amounts of protein (personal communication with H.J.P. van der Zande, Dr. A.

van Diepen and Y.C.M Kruize from the Leiden University Medical Centre and G. Schramm from the Forschungszentrum Borstel). In order to enable studies to the pharmaceutical properties of these proteins and the effect of their glycans, a platform is required that produces high amounts of these glycoproteins with native glycans.

Pharmaceutical protein production systems

Production of biopharmaceutical proteins nowadays relies on heterologous gene expression. Several heterologous expression systems exist for the production of biopharmaceuticals, of which the most common platforms are *Escherichia coli* and mammalian cell culture systems, such as Chinese hamster ovary (CHO) cells. Although, also other systems are used such as insect cells, yeast cells and cell free expression systems [48], [49]. *E. coli* and CHO cells are two well established platforms that both meet the most important standards of high quality and high production. *E. coli* is easy to work with, requires simple medium and is cost effective [50]. However, *E. coli* is naturally not able to introduce post-translational modifications, like glycosylation. Moreover, the formation of inclusion bodies due to aggregate formation is a common feature for protein production in *E. coli* and therefore chemical refolding is often required for biological activity. More complex proteins, such as glycoproteins, are expressed in systems that naturally produce such proteins. Yeast for instance, is capable of glycoprotein production and is easy and cheap to culture. However, yeast glycoproteins naturally carry high mannose-enriched N-glycans, which are not found on mammalian glycoproteins. These mannose-enriched N-glycans can bind the MR on antigen presenting cells, which can result in reduced half-life by quick clearance from the blood [51], [52]. Research efforts on glyco-engineering enabled human-like N-glycosylation in yeast [53], [54]. Nevertheless, the most common platform for glycoprotein production is mammalian cell culture that naturally produces glycoproteins with mammalian N-glycans. However, glycosylation in these cultured cells is heterogeneous and therefore maintaining batch-to-batch reproducibility is more difficult [55]–[57]. Moreover, due to the complex endogenous glycome of mammals, glyco-engineering of mammalian cells to increase glycan homogeneity can be difficult.

In the 1990's, plants were introduced as an alternative platform for the production of biopharmaceuticals [58]. Plants are able to correctly fold complex proteins and assemble heteromultimeric protein complexes, such as antibodies [59]. Furthermore, plants are capable of post-translational modifications, such as glycosylation and allow adjustment of their glycosylation pathway. In contrast to mammalian cell lines, plants show remarkable homogeneous glycan profiles [56], [60], [61]. Also, plant cultivation is low in costs, scalable and transient plant expression allows fast production, which is beneficial for the production of emergency vaccines or antibodies. Furthermore, *in planta* production of biopharmaceuticals has a lower risk for contamination with human

pathogens. Despite all these advantages plants still lack behind as a production platform due to low yields, cost inefficient isolation and purification on top of the regulations for plant transformation. However, these drawbacks are hardly of interest for proteins with a higher added value, such as antibodies, vaccines, cytokines and proteins that replace inactive human proteins leading to disease [48], [62], [63]. An example of the latter is plant produced Glucocerebrosidase (product name Elelyso), which is used as medical treatment of type 1 Gaucher's disease with enzyme replacement therapy [64]. In May 2012, Elelyso was the first plant-produced biopharmaceutical approved by the United States Food and Drug Administration (FDA). Furthermore, the plant produced anti-Ebola antibody cocktail ZMapp and the Medicago anti-influenza vaccine were used in emergency outbreaks of Ebola in 2014 and influenza in 2009, respectively [63]. Moreover, multiple other plant-produced biopharmaceuticals are in clinical trials [63], [65], [66]. This shows that plants gradually become accepted as production system for biopharmaceutical proteins. However, for glycoprotein production another limitation to accept plants as a production platform are the differences in glycosylation of plants and humans.

Glycosylation and the glycosylation pathway

Glycosylation is a post-translational modification that covalently links a glycosyl donor to a glycosyl acceptor. The main form of glycosylation discussed in this thesis is protein N-glycosylation in which the common glycan precursor is linked to the nitrogen of an asparagine (Asn) side chain in the consensus glycosylation peptide sequence Asn-X-serine (Ser) / threonine (Thr) (X can be any amino acid except a proline (Pro)). Other common types of glycosylation are protein O-glycosylation and glycosylation of lipids (glycolipids). A high degree of glycan diversity is created by the type of glycosylation and the variation in monosaccharides and linkage in the glycan itself. Each monosaccharide can be linked via an α or a β-linkage to several positions of another monosaccharide or oligosaccharide chain, creating numerous possibilities (monosaccharides, glycans and glycan motifs discussed in this thesis can be found at the end of this chapter in Table 1).

Glycosylation pathways differ between organisms, such as observed for humans and plants. Human N-glycans can carry tri- or tetraantennary N-glycans with a core α1,6-fucose and complex sialylated glycan motifs, such as sialylated LN. In contrast, plant N-glycans carry paucimannosidic or diantennary N-glycans with core α1,3-fucose and β1,2-xylose and sometimes the complex glycan motif LeA. Besides these differences, the human and plant N-glycosylation pathways are quite similar. In the endoplasmic reticulum (ER), the common glycan precursor $Glc_3Man_9GlcNAc_2$ is linked to the protein glycosylation site by the oligosaccharyl transferase (OST) complex. During movement through the ER lumen the glycan is gradually trimmed down to $Man_9GlcNAc_2$ followed by transfer to the Golgi via coat protein complex (COP) II vesicles [67]. The first steps of the N-glycosylation

pathway are important for proper protein folding, which is monitored by ER chaperones, such as the lectins calnexin and calreticulin. α-Glucosidases (GCS) I and GCSII perform the first trimming steps and remove two terminal glucose residues. Proteins that fail to fold correctly are marked with specific truncated glycans by α-mannosidases and are targeted to the cytoplasm for degradation. Correctly folded proteins carry $Man_8GlcNAc_2$ N-glycans and are further processed in the Golgi. The Golgi consists of three sub compartments, the *cis*-, medial and *trans*-Golgi. In the *cis*-Golgi α-mannosidases trim down the glycan to $Man_5GlcNAc_2$ (Figure 2). Subsequently, a GlcNAc is added by the *cis*- to the medial-Golgi localised β-*N*-acetylglucosaminyltransferase (GnT) I. After removal of two additional mannoses from the α1,6-branch by Golgi-α-mannosidase II (GMII), the initiation of more complex N-glycans starts in the medial-Golgi. In this sub Golgi compartment a second GlcNAc is added by GnTII to the α1,6-branch of the $GlcNAcMan_3GlcNAc_2$ acceptor. Without these two steps only hybrid type N-glycans can be formed, carrying one antenna with complex glycans and one antenna with mannoses. After GnTII activity the human glycosylation and the plant glycosylation pathway diverge.

In humans, GnTII activity can be followed by introduction of extra antennae by GnTIII, GnTIV or GnTV activity and/or addition of a core α1,6-fucose by fucosyltransferase 8 (FucT8). In the *trans*-Golgi the terminal GlcNAcs can be further extended with a β1,4-galactose by β1,4-galactosyltransferase (β1,4-HsGalT). Next, α2,3/6-sialic acid can be added by α2,3/6-sialyltransferase (α2,3/6-HsST). In plants, modification of the $GlcNAcMan_5GlcNAc_2$ acceptor by GMII and GnTII is followed by activity of the medial-Golgi localised β1,2-xylosyltransferase (XylT) and/or the medial-to *trans*-Golgi localised core α1,3-fucosyltransferase (core α1,3-FucT, FucT11/12). Alternatively, the $GlcNAcMan_5GlcNAc_2$ acceptor is first modified by XylT before GMII and/or GnTII activity. These alternative processing routes have been proposed based on recombinant enzyme substrate specificity and N-glycan analysis of glycosyltransferase knockout plants [68]–[70]. The last possible step in the plant glycosylation pathway is synthesis of LeA by the *trans*-Golgi localised enzymes β1,3-galactosyltransferase (β1,3-GalT) and α1,4-fucosyltransferase (α1,4-FucT, FucT13). First a galactose is added to the terminal GlcNAc by β1,3-GalT, subsequently the GlcNAc is fucosylated by FucT13 synthesising LeA. At the end of the Golgi proteins are transported to either the vacuole or the apoplastic space. β-*N*-acetylhexosaminidase (HEXO) 1 and 3 present in the vacuole or the apoplast, respectively, can cleave the terminal GlcNAcs resulting in paucimannosidic N-glycan structures [71]–[73].
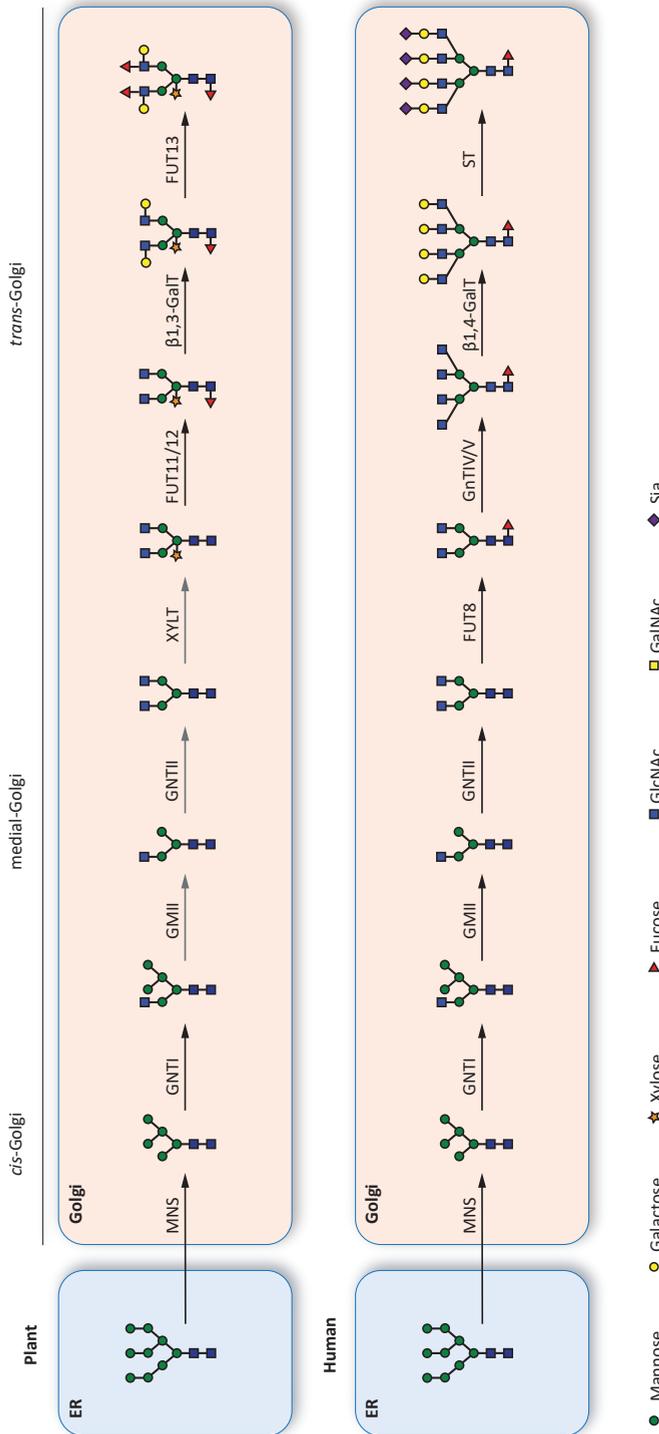
FIGURE 2 | Schematic representation of the N-glycan modifying steps in the plant and human Golgi-system. MNS: Class I mannosidases MNS1, 2 and 3; GnTI: N-acetyl-glucosaminyltransferase I; GMII: Golgi-α-mannosidase II; GnTII: N-acetyl-glucosaminyltransferase II; XYLT: β1,2-xylosyltransferase; FUT11/12: core α1,3-fucosyltransferase; β1,3-GalT: β1,3-galactosyltransferase; FUT13: β1,3-galactosyltransferase; FUT8: α1,6-fucosyltransferase; GnTIV/V: N-acetyl-glucosaminyltransferase IV and V; β1,4-GalT: β1,4-galactosyltransferase; ST: sialyltransferase. Arrows are indicated in grey, when alternative N-glycan processing routes have been postulated.

Adjusted from Wilbers *et al* [46]

1

Plant glyco-engineering

Since plant and human N-glycosylation pathways differ, the plant glycosylation pathway requires adaptation for the production of glycoproteins with human-like N-glycans. Various studies have shown the possibilities to "humanise" the plant glycosylation pathway by introduction of human glycosyltransferases and production of knockdown/knockout plants for plant glycosyltransferases core α1,3-FucT and XylT. The first attempt to "humanise" plants was made by introduction of β1,4-HsGalT in *Nicotiana tabacum* L. cv. Bright Yellow 2 suspension-cultured cells (BY-2 cells) [74]. Stable expression of β1,4-HsGalT in *N. tabacum* plants resulted in galactosylated N-glycans with reduced core α1,3-fucose and β1,2-xylose, depending on β1,4-HsGalT Golgi localisation [74]–[76]. This indicates that depending on its localisation, β1,4-HsGalT hampers the addition of core α1,3-fucose and β1,2-xylose. For complete elimination of core α1,3-fucose and β1,2-xylose, an *Arabidopsis thaliana* core α1,3-FucT and XylT knockout plant was generated via crosses with T-DNA insertion lines [77], [78]. After establishment of a core α1,3-FucT and XylT knockout in *A. thaliana* and glycosyltransferase introduction in *N. tabacum*, studies focussed on *Nicotiana benthamiana*, an important plant species for transient protein production. Strasser and colleagues [79] generated a core α1,3-FucT and XylT down regulated *N. benthamiana* plant by RNA interference (ΔXT/FT *N. benthamiana* plants). Production of a human anti-HIV monoclonal antibody (mAb) 2G12 showed homogenous glycosylation and a lack of core α1,3-fucose and β1,2-xylose. Moreover, the functional properties of the plant-produced 2G12 were comparable to the CHO-derived 2G12. Co-expression of a *trans*-Golgi targeted β1,4-HsGalT in ΔXT/FT *N. benthamiana* plants enabled engineering of the anti-HIV mAb 2G12 with improved homogeneous galactosylation, in comparison to CHO-derived 2G12 [80]. Next, sialylation was achieved in plants by introduction of the mammalian sialylation synthesis pathway [81]. Together with the introduction of multiple antennae by expression of GnTIV and/or GnTV, this enabled production of biologically active recombinant multisialylated human erythropoietin (EPO) in plants [82]–[84]. Next to mammalian N-glycosylation, mammalian mucin-type O-glycosylation was introduced on plant produced EPO [85]. Also, human core α1,6-fucosylation and synthesis of the glycan motif LeX were shown in plants [61], [86], [87]. All these studies reveal that plants are an excellent platform for glycoprotein production with "humanised" N-glycans. This enabled production of biopharmaceuticals and research to the effect of N-glycans on protein function. Thus far, the field of plant molecular farming has focussed on the production of high valued human pharmaceutical proteins and therefore "humanisation" of the plant glycosylation pathway. With "helminthisation" plants could become a promising platform for the production of potentially therapeutic helminth glycoproteins, such as omega-1 of which high amounts are required for *in vivo* studies to its Th2 inducing ability or its effect on metabolic homeostasis. Therefore, next to improvement of the plant as production platform and the production of new human proteins, production of helminth glycoproteins is the next challenge.

Synthesis of fucosylated glycans

Secretions of *S. mansoni* contain highly fucosylated glycans, to which many antibodies are directed. Fucose is a common vertebrate monosaccharide with the chemical formula $C_6H_{12}O_5$. Fucose differs from other six-carbon monosaccharides present in mammals by two characteristics: the lack of a hydroxyl group on the sixth carbon and its L-configuration. The lack of a hydroxyl group makes fucose a deoxy monosaccharide and with its L-configuration it is the only laevorotatory sugar synthesised by mammals. Synthesis of fucosylated glycan motifs requires fucose to be activated to its high-energy donor form guanosine diphosphate β-L-fucose (GDP-fucose). GDP-fucose is synthesised *de novo* or via a salvage pathway in the cytosol, in a 9:1 ratio in mammalian cells [88]. In *de novo* synthesis, GDP-mannose is converted to GDP-fucose in three steps, involving the enzymes GDP-Man 4,6-dehydratase and GDP-keto-6-deoxymannose-3,5-epimerase-4-reductase. In synthesis via the salvage pathway, GDP-fucose is directly derived from free fucose in a two-step reaction. Free fucose is converted to fucose-1-P by a kinase, after which it is linked to GDP to yield GDP-fucose. Concentrations of free fucose are normally low, a few µM at most. Moreover, free fucose seems to be an indicator for disease, such as cancer, diabetes or cardiovascular disease. Since, GDP-fucose is synthesised in the cytosol, N-glycosylation requires GDP-fucose transport over the Golgi membrane into the Golgi lumen via GDP L-fucose transporters. The energy to transport GDP-fucose into the Golgi lumen is provided by simultaneous export of guanosine monophosphate (GMP), converted from GDP after L-fucose transfer [89], [90].

Fucose transfer to *S. mansoni* N-glycans yields terminal modifications or core modifications. These modifications are made by fucosyltransferases (FucTs), which in a stereo- and regiospecific manner transfer a fucose from the GDP-fucose donor to an oligosaccharide acceptor via an α1,2, α1,3 or α1,6 bond. The name of the bond is composed of the type of linkage formed upon fucose transfer and the numbers of the two C-atoms involved in fucose to acceptor binding. FucTs are type II membrane proteins, with at the N-terminus the hypervariable CTS domain, composed of a short cytoplasmic tail (CT), a transmembrane domain (TMD) and a stem region (SR). The C-terminal part is composed of various short sequence motifs involved in GDP-fucose or glycan acceptor binding. These sequence motifs vary depending on the substrate, glycan acceptor or the type of linkage formed upon fucose transfer. For instance, FucTs that couple fucose with an α1,3 linkage (α1,3-FucTs), contain specific sequence motifs, such as the 1st cluster motif and motif V both involved in GDP-fucose binding [91]–[94]. The catalytic domain of the FucT can be found at the C-terminal part of the protein, whereas the N-terminal CTS domain is involved in Golgi localisation.

1

Golgi localisation of glycosyltransferases

The separate parts of the glycosyltransferase CTS domain are important in general Golgi targeting, retention or retrograde transport models, such as adaptor or receptor mediated retrieval, sorting on bilayer thickness or lipid composition and vesicle exclusion by oligomerisation [95]–[100]. It was reported that the CT in some glycosyltransferases is involved in binding to COP coated vesicles, important for Golgi trafficking. The yeast vacuolar protein sorting-associated protein 74 (Vps74p), for instance, binds to a specific amino acid sequence in the CT of yeast Golgi-localised enzymes and COPI vesicles. This suggests that Vps74p binding to the CT and COPI vesicles mediates glycosyltransferase incorporation in COPI vesicles [101], [102]. The mammalian homolog of Vps74p, Golgi phosphoprotein 3 (GOLPH3), shows similar interaction with the CT of human GnTI, α2,6-HsST and Core 2 *N*-acetylglucosaminyltransferase 1 [103], [104]. Interestingly, also direct binding of COPI vesicles to the CT of several mammalian *cis*-glycosyltransferases is observed [105]. In plants, the CT of EMP12 contains two signals, one for ER exit and one for interaction with COPI and COPII vesicles [106], [107]. Similarly, in the CT of *N. tabacum* GnTI (NtGnTI) a signal was found required for COPII vesicle interaction and ER exit [108]. Although, for NtGnTI not the CT but the TMD is important for its Golgi localisation [109], [110].

The TMD is the part of the protein, which anchors the glycosyltransferase in the Golgi membrane. Throughout the secretory pathway the membrane thickness changes, with a thicker membrane at the end of the secretory pathway near the plasma membrane. The bilayer thickness model suggests that the length of the TMD can determine Golgi localisation [97]. This was demonstrated for human lysosome-associated membrane protein-1 (LAMP1) by reduction of the TMD from 23 to either 20 or 17 amino acids resulting in, respectively, plasma membrane, Golgi or ER localisation [111]. Another model based on the membrane thickness is rapid partitioning, which is based on differences in the cell membrane lipid composition [98], [112]. Recently, an key polar residue in *A. thaliana* and *N. benthamiana* GnTIs was identified for Golgi sub-localisation [110]. This polar residue could interact with specific lipids or an adaptor protein regulating Golgi retention or retrograde transport.

For plant GnTI, oligomerisation via homo- or heterodimers is not implicated in its Golgi sub-localisation, whereas for α2,6-HsST oligomerisation it is suggested to be a secondary event that enhances stable Golgi localisation [109], [113]. The kin-recognition model or oligomerisation model suggests that glycosyltransferases homo- and heteromeric protein complexes are retained or concentrated in distinct Golgi regions [95], [96], [114]. However, homo- and heteromeric dimers have been shown for multiple plant glycosyltransferases in the *cis*- and medial-Golgi, such as for plant GnTI that does not require oligomerisation for correct Golgi sub-localisation [115]. Nevertheless, for *Rattus norvegicus* α2,6-sialyltransferase (RnST) no homo- or heterodimers were observed in plants with the tested plant glycosyltransferases, suggesting that in the plant, oligomerisation is not involved

in Golgi localisation of RnST [115]. Remarkably, also heterodimers are formed between successive acting glycosyltransferases, such as GnTI/GMII, GnTI/Golgi α-mannosidase I (MNS1), MNS1/GMII, GMII/β1,3-GalT and GMII/XylT in plants and GnTI/GnTII and β1,4-GalT/α2,3-sialyltransferase in mammals [115]–[117]. This suggests that oligomerisation is important for glycan synthesis, although oligomerisation may still be important for Golgi sub-localisation as well.

Evidence for various Golgi targeting, retention or retrograde transport models have been observed, although the exact mechanisms behind the strict Golgi organisation are still a mystery. Nevertheless, the information gathered over the years enables glyco-engineering in heterologous expression systems, such as plants.

Outline of this thesis

This thesis describes the production of helminth glycoproteins and "helminthisation" of the plant glycosylation pathway, by characterisation of *S. mansoni* and *N. benthamiana* FucTs (SmFucT and NbFucT, respectively). Characterisation of SmFucTs starts with a selection of ten SmFucT sequences of which the Golgi and sub-Golgi localisation is investigated (**Chapter 2**). Sub-Golgi localisation of glycosyltransferases is important for their function, because the correct glycan acceptor and glycan donor should be available for monosaccharide transfer and subsequent glycan synthesis. We show that all SmFucTs are localised in the Golgi, although subtle differences in sub-Golgi localisation are observed. Next, SmFucTs are functionally characterised based on the glycan motif synthesised on the N-glycans of carrier glycoproteins upon *in planta* co-expression (**Chapter 3**). With this approach the core glycan motifs core α1,3- and α1,6-fucose and terminal glycan motifs LeX, LDN-F and F-LDN-F were synthesised. With the same method the activity of putative plant core α1,3-NbFucTs is tested to determine which NbFucT sequence should be targeted to generate a core α1,3-NbFucT knockout plant (**Chapter 4**). Core α1,3-NbFucT sequences are amplified and their activity is checked by *in planta* co-expression with carrier glycoproteins and subsequent analysis of the carrier glycoprotein N-glycan composition. This study shows that only three NbFucTs out of ten putative NbFucT coding sequences lead to an active enzyme. Therefore, time and effort can be saved by targeting only these three NbFucT coding sequences instead of all ten for generation of a core α1,3-NbFucT knockout plant. Characterisation of SmFucTs and NbFucTs enables production of helminth glycoproteins with tailored N-glycans. With the gathered knowledge we produce the helminth glycoproteins omega-1 (**Chapter 5**) and kappa-5 (**Chapter 6**) with tailored N-glycans in *N. benthamiana*. Production of glycoproteins with tailored helminth-like N-glycans enables development of vaccines or diagnostic tests and research to the effect of N-glycan composition on protein function. Moreover, production of helminth proteins, such as omega-1 and kappa-5, enables research to their immunomodulatory properties and their potential use as biopharmaceutical.

TABLE 1 | Monosaccharides, glycans and glycan motifs discussed in this thesis.

Name	Abbreviation	Description	Symbole
fucose	Fuc / F		▲
N-Acetylglucosamine	GlcNAc / Gn		▪
xylose	Xyl / X		★
mannose	Man / M		●
galactose	Gal / A		○
N-Acetylgalactosamine	GalNAc / An		□
sialic acid	Sia		◆
paucimannose	MM	Manα1-6(Manα1-3)Manβ1–4GlcNAcβ1–4GlcNAc	
monoantennary	MGn	GlcNAcβ1–2Manα1-3(Manα1-6)Manβ1–4GlcNAcβ1–4GlcNAc	
diantennary	GnGn	GlcNAcβ1–2Manα1-6(GlcNAcβ1–2Manα1-3)Manβ1–4GlcNAcβ1–4GlcNAc	
triantennary	Gn[GnGn]	GlcNAcβ1–4(GlcNAcβ1–2)Manα1-3(GlcNAcβ1–2Manα1-6)Manβ1–4GlcNAcβ1–4GlcNAc	
triantennary	[GnGn]Gn	GlcNAcβ1–6(GlcNAcβ1–2)Manα1-6(GlcNAcβ1–2Manα1-3)Manβ1–4GlcNAcβ1–4GlcNAc	
tetra-antennary	GnGnGnGn	GlcNAcβ1–6(GlcNAcβ1–2)Manα1-6[GlcNAcβ1–4(GlcNAcβ1–2)Manα1-3]Manβ1–4GlcNAcβ1–4GlcNAc	
Core α1,3-fucose		Manα1-6(Manα1-3)Manβ1–4GlcNAcβ1–4(Fucα1-3)GlcNAc	
Core α1,6-fucose		Manα1-6(Manα1-3)Manβ1–4GlcNAcβ1–4(Fucα1-6)GlcNAc	
Core β1,2-xylose		Manα1-6(Manα1-3)(Xylβ1–2)Manβ1–4GlcNAcβ1–4GlcNAc	
fucosylated GlcNAc	F-GlcNAc	Fucα1–4GlcNAc-R	
LacNAc	LN	Galβ1–4GlcNAc-R	
fucosylated LacNAc	LeA	Galβ1–3(Fucα1–4)GlcNAc-R	
"	LeX	Galβ1–4(Fucα1–3)GlcNAc-R	
"	Pseudo-LeY	Fucα1-3Galβ1-4(Fucα1-3)GlcNAc-R	
sialylated LacNAc	sLN	Sia2-3/6Galβ1–4(Fucα1–3)GlcNAc-R	
double LeX	diLeX	Galβ1–4(Fucα1–3)GlcNAcβ1-3Galβ1–4(Fucα1–3)GlcNAc-R	
LacDiNAc	LDN	GalNAcβ1-4GlcNAc-R	
fucosylated LacDiNAc	LDN-F	GalNAcβ1-4(Fucα1-3)GlcNAc-R	
"	F-LDN	Fucα1-3GalNAcβ1-4GlcNAc-R	
"	F-LDN-F	Fucα1-3GalNAcβ1-4(Fucα1-3)GlcNAc-R	
double fucose	DF	Fucα1-2Fucα1-3-R	
difucosylated GlcNAc	DF-GlcNAc	Fucα1-2Fucα1-3GlcNAc-R	
difucosylated LacDiNAc	LDN-DF	GalNAcβ1-4(Fucα1-2Fucα1-3)GlcNAc-R	
"	DF-LDN-DF	Fucα1-2Fucα1-3GalNAcβ1-4(Fucα1-2Fucα1-3)GlcNAc-R	

References

[1]   P. J. Hotez *et al.*, "The Global Burden of Disease Study 2010: Interpretation and Implications for the Neglected Tropical Diseases," *PLoS Negl. Trop. Dis.*, vol. 8, no. 7, p. e2865, 2014.

[2]   D. Rollinson *et al.*, "Time to set the agenda for schistosomiasis elimination," *Acta Trop.*, vol. 128, no. 2, pp. 423–440, 2013.

[3]   D. P. Mcmanus, D. W. Dunne, M. Sacko, J. Utzinger, B. J. Vennervald, and N. Zhou, "Schistosomiasis," *Nat. Rev. Dis. Prim.*, vol. 4, no. 13, 2018.

[4]   World Health Organization, Preventive chemotherapy in human helminthiasis: coordinated use of anthelminthic drugs in control interventions: a manual for health professionals and programme managers. 2006.

[5]   A. W. Cheever, J. G. Macedonia, J. E. Mosimann, and E. A. Cheever, "Kinetics of egg production and egg excretion by Schistosoma mansoni and S. japonicum in mice infected with a single pair of worms," *Am. J. Trop. Med. Hyg.*, vol. 50, no. 3, pp. 281–295, 1994.

[6]   A. G. Ross, D. P. McManus, J. Farrar, R. J. Hunstman, D. J. Gray, and Y.-S. Li, "Neuroschistosomiasis," *Neglected Trop. Dis. Cond. Nerv. Syst.*, vol. 259, no. 1, pp. 22–32, 2014.

[7]   J. E. H. Pittella, "Neuroschistosomiasis," *Brain Pathol.*, vol. 7, no. 1, pp. 649–662, 1997.

[8]   A. Von Schaewen, A. Sturm, J. O'Neill, and M. J. Chrispeels, "Isolation of a mutant Arabidopsis plant that lacks N-acetyl glucosaminyl transferase I and is unable to synthesize golgi-modified complex N-linked glycans," *Plant Physiol.*, vol. 102, no. 4, pp. 1109–1118, 1993.

[9]   P. D. Ashton, R. Harrop, B. Shah, and R. A. Wilson, "The schistosome egg: Development and secretions," *Parasitology*, vol. 122, no. 3, pp. 329–338, 2001.

[10]   S. deWalick, A. G. M. Tielens, and J. J. van Hellemond, "Schistosoma mansoni: The egg, biosynthesis of the shell and interaction with the host," *Exp. Parasitol.*, vol. 132, no. 1, pp. 7–13, 2012.

[11]   World Health Organization Expert Committee, "The control of schistosomiasis," 1985.

[12]   World Health Organization, "The control of schistosomiasis. Second report of the WHO Expert Committee." *World Health Organization - Technical Report Series*, vol. 830. pp. 1–86, 1993.

[13]   M. Berriman *et al.*, "The genome of the blood fluke Schistosoma mansoni," *Nature*, vol. 460, no. 7253, pp. 352–358, 2009.

[14]   Y. Zhou *et al.*, "The Schistosoma japonicum genome reveals features of host-parasite interplay," *Nature*, vol. 460, no. 7253, pp. 345–351, 2009.

[15]   A. V. Protasio *et al.*, "A systematically improved high quality genome and transcriptome of the human blood fluke Schistosoma mansoni," *PLoS Negl. Trop. Dis.*, vol. 6, no. 1, p. e1455, 2012.

[16]   F. Luo *et al.*, "An improved genome assembly of the fluke Schistosoma japonicum," *PLoS Negl. Trop. Dis.*, vol. 13, no. 8, p. e0007612, 2019.

[17]   J. M. Fitzpatrick *et al.*, "Anti-schistosomal intervention targets identified by lifecycle transcriptomic analyses," *PLoS Negl. Trop. Dis.*, vol. 3, no. 11, p. e543, 2009.

[18]   S. J. Parker-Manuel, A. C. Ivens, G. P. Dillon, and R. A. Wilson, "Gene expression patterns in larval Schistosoma mansoni associated with infection of the mammalian host," *PLoS Negl. Trop. Dis.*, vol. 5, no. 8, p. e1274, 2011.

[19]   L. Anderson *et al.*, "Schistosoma mansoni Egg, Adult Male and Female Comparative Gene Expression Analysis and Identification of Novel Genes by RNA-Seq," *PLoS Negl. Trop. Dis.*, vol. 9, no. 12, p. e0004334, 2015.

[20]   E. J. Pearce, "Priming of the immune response by schistosome eggs," *Parasite Immunol.*, vol. 27, no. 7–8, pp. 265–270, 2005.

[21]   G. Schramm *et al.*, "Molecular characterization of an interleukin-4-inducing factor from Schistosoma mansoni eggs," *J. Biol. Chem.*, vol. 278, no. 20, pp. 18384–18392, 2003.

[22]   K. Knuhr *et al.*, "Schistosoma mansoni egg-released IPSE/alpha-1 dampens inflammatory cytokine responses viabasophil interleukin (IL)-4 and IL-13," *Front. Immunol.*, vol. 9, p. 2293, 2018.

[23]   G. Schramm *et al.*, " Cutting Edge: IPSE/alpha-1, a Glycoprotein from Schistosoma mansoni Eggs, Induces IgE-Dependent, Antigen-Independent IL-4 Production by Murine Basophils In Vivo ," *J. Immunol.*, vol. 178, no. 10, pp. 6023–6027, 2007.

[24]   S. Haeberlein *et al.*, "Schistosome egg antigens, including the glycoprotein IPSE/alpha-1, trigger the development of regulatory B cells," *PLoS Pathog.*, vol. 13, no. 7, pp. 1–28, 2017.

[25]   B. Everts *et al.*, "Schistosome-derived omega-1 drives Th2 polarization by suppressing protein synthesis following internalization by the mannose receptor," *J. Exp. Med.*, vol. 209, no. 10, pp. 1753–1767, 2012.

[26]   B. Everts *et al.*, "Omega-1, a glycoprotein secreted by Schistosoma mansoni eggs, drives Th2 responses," *J. Exp. Med.*, vol. 206, no. 8, pp. 1673–1680, 2009.

[27]   S. Steinfelder *et al.*, "The major component in schistosome eggs responsible for conditioning dendritic cells for Th2 polarization is a T2 ribonuclease (omega-1)," *J. Exp. Med.*, vol. 206, no. 8, pp. 1681–1690, 2009.

[28]   E. Hams *et al.*, "The helminth T2 RNase ω1 promotes metabolic homeostasis in an IL-33- and group 2 innate lymphoid cell-dependent mechanism," *FASEB Journal*, vol. 30, no. 2. pp. 824–835, 2016.

[29]   K. Obieglo, "Immune modulation by schistosomes: mechanisms of regulatory B cell induction and inhibition of allergic asthma," 2019.

[30]   G. Schramm *et al.*, "Molecular characterisation of kappa-5, a major antigenic glycoprotein from Schistosoma mansoni eggs," *Mol. Biochem. Parasitol.*, vol. 166, no. 1, pp. 4–14, 2009.

[31]   M. H. J. Meevissen *et al.*, "Structural characterization of glycans on omega-1, a major schistosoma mansoni egg glycoprotein that drives Th2 responses," *J. Proteome Res.*, vol. 9, no. 5, pp. 2630–2642, 2010.

[32]   M. Wuhrer *et al.*, "IPSE/alpha-1, a major secretory glycoprotein antigen from schistosome eggs, expresses the Lewis X motif on core-difucosylated N-glycans," *FEBS J.*, vol. 273, no. 10, pp. 2276–2292, 2006.

[33]   M. H. J. Meevissen *et al.*, "Targeted glycoproteomic analysis reveals that kappa-5 is a major, uniquely glycosylated component of schistosoma mansoni egg antigens," *Mol. Cell. Proteomics*, vol. 10, no. 5, pp. M110-005710, 2011.

[34]   M. H. J. Meevissen *et al.*, "Specific glycan elements determine differential binding of individual egg glycoproteins of the human parasite Schistosoma mansoni by host C-type lectin receptors," *Int. J. Parasitol.*, vol. 42, no. 3, pp. 269–277, 2012.

[35]   P. Huang *et al.*, "Norovirus and Histo-Blood Group Antigens: Demonstration of a Wide Spectrum of Strain Specificities and Classification of Two Major Binding Groups among Multiple Binding Patterns," *J. Virol.*, vol. 79, no. 11, pp. 6714–6722, 2005.

[36]   E. P. Hennessy, A. D. Green, M. P. Connor, R. Darby, and P. MacDonald, "Norwalk Virus Infection and Disease Is Associated with ABO Histo–Blood Group Type," *J. Infect. Dis.*, vol. 188, no. 1, pp. 176–177, 2003.

[37]   V. R. Graziano, J. Wei, and C. B. Wilen, "Norovirus Attachment and Entry," *Viruses*, vol. 11, no. 6, pp. 1–13, 2019.

[38]   S. Marionneau *et al.*, "Norwalk Virus binds to histo-blood group antigens present on gastroduodenal epithelial cells of secretor individuals," *Gastroenterology*, vol. 122, no. 7, pp. 1967–1977, 2002.

[39]   T. Lehr, H. Geyer, K. Maaß, M. J. Doenhoff, and R. Geyer, "Structural characterization of N-glycans from the freshwater snail Biomphalaria glabrata cross-reacting with Schistosoma mansoni glycoconjugates," *Glycobiology*, vol. 17, no. 1, pp. 82–103, 2007.

[40]   T. Lehr, K. Beuerlein, M. J. Doenhoff, C. G. Grevelding, and R. Geyer, "Localization of carbohydrate determinants common to Biomphalaria glabrata as well as to sporocysts and miracidia of Schistosoma mansoni," *Parasitology*, vol. 135, no. 8, pp. 931–942, 2008.

[41]   T. Lehr *et al.*, "N-Glycosylation patterns of hemolymph glycoproteins from Biomphalaria glabrata strains expressing different susceptibility to Schistosoma mansoni infection," *Exp. Parasitol.*, vol. 126, no. 4, pp. 592–602, 2010.

[42]   C. H. Smit *et al.*, "Glycomic analysis of life stages of the human parasite Schistosoma mansoni reveals developmental expression profiles of functional and antigenic glycan motifs.," *Mol. Cell. Proteomics*, vol. 14, no. 7, pp. 1750–1769, 2015.

[43]   C. H. Hokke and A. van Diepen, "Helminth glycomics – glycan repertoires and host-parasite interactions," *Mol. Biochem. Parasitol.*, vol. 215, pp. 47–57, 2017.

[44]   M. Wuhrer, "Schistosoma mansoni cercarial glycolipids are dominated by Lewis X and pseudo-Lewis Y structures," *Glycobiology*, vol. 10, no. 1, pp. 89–101, 2000.

[45]   A. K. Nyame, T. P. Yoshino, and R. D. Cummings, "Differential Expression of LacdiNAc, Fucosylated LacdiNAc, and Lewis X Glycan Antigens in Intramolluscan Stages of Schistosoma mansoni," *J. Parasitol.*, vol. 88, no. 5, p. 890, 2002.

[46]   R. H. P. Wilbers *et al.*, "Production and glyco-engineering of immunomodulatory helminth glycoproteins in plants," *Sci. Rep.*, vol. 7, p. 45910, 2017.

[47]   K. K. Van de Vijver, A. M. Deelder, W. Jacobs, E. A. Van Marck, and C. H. Hokke, "LacdiNAc- and LacNAc-containing glycans induce granulomas in an in vivo model for schistosome egg-induced hepatic granuloma formation," *Glycobiology*, vol. 16, no. 3, pp. 237–243, 2006.

[48]   S. Schillberg, N. Raven, H. Spiegel, S. Rasche, and M. Buntru, "Critical analysis of the commercial potential of plants for the production of recombinant proteins," *Front. Plant Sci.*, vol. 10, p. 720, 2019.

[49]   G. Walsh, "Biopharmaceutical benchmarks 2018," *Nat. Biotechnol.*, vol. 36, no. 12, pp. 1136–1145, 2018.

[50]   N. A. Baeshen *et al.*, "Cell factories for insulin production," *Microb. Cell Fact.*, vol. 13, p. 141, 2014.

[51]   M. E. Taylor, K. Bezouska, and K. Drickamer, "Contribution to ligand binding by multiple carbohydrate-recognition domains in the macrophage mannose receptor," *J. Biol. Chem.*, vol. 267, no. 3, pp. 1719–1726, 1992.

[52]   S. J. Lee *et al.*, "Mannose receptor-mediated regulation of serum glycoprotein homeostasis," *Science*, vol. 295, no. 5561, pp. 1898–1901, 2002.

[53]   R. Baghban *et al.*, "Yeast Expression Systems: Overview and Recent Advances," *Mol. Biotechnol.*, vol. 61, no. 5, pp. 365–384, 2019.

[54]   C. de Wachter, L. van Landuyt, and N. Callewaert, *Biological Processes for Hydrogen Production*. 2018.

[55]   L. Meuris *et al.*, "GlycoDelete engineering of mammalian cells simplifies N-glycosylation of recombinant proteins," *Nat. Biotechnol.*, vol. 32, no. 5, pp. 485–489, 2014.

[56]   S. A. Brooks, "Strategies for analysis of the glycosylation of proteins: Current status and future perspectives," *Mol. Biotechnol.*, vol. 43, no. 1, pp. 76–88, 2009.

[57]   R. J. Graham, H. Bhatia, and S. Yoon, "Consequences of trace metal variability and supplementation on Chinese hamster ovary (CHO) cell culture performance: A review of key mechanisms and considerations," *Biotechnol. Bioeng.*, vol. 116, no. 12, pp. 3446–3456, 2019.

[58]   J. W. Larrick, L. Yu, J. Chen, S. Jaiswal, and K. Wycoff, "Production of antibodies in transgenic plants," *Res. Immunol.*, vol. 149, no. 6, pp. 603–608, 1998.

[59]   L. B. Westerhof *et al.*, "Transient expression of secretory IgA in Planta is optimal using a multi-gene vector and may be further enhanced by improving joining chain incorporation," *Front. Plant Sci.*, vol. 6, p. 1200, 2016.

[60]   D. Bosch, A. Castilho, A. Loos, A. Schots, and H. Steinkellner, "N-Glycosylation of Plant-produced Recombinant Proteins," *Curr. Pharm. Des.*, vol. 19, no. 31, pp. 5503–5512, 2013.

[61]   R. H. P. Wilbers *et al.*, "The N-glycan on Asn54 affects the atypical N-glycan composition of plant-produced interleukin-22, but does not influence its activity," *Plant Biotechnol. J.*, vol. 14, no. 2, pp. 670–681, 2016.

[62]   E. Stoger, R. Fischer, M. Moloney, and J. K.-C. Ma, "Plant Molecular Pharming for the Treatment of Chronic and Infectious Diseases," *Annu. Rev. Plant Biol.*, vol. 65, no. 1, pp. 743–768, 2014.

[63]   M. Sack, A. Hofbauer, R. Fischer, and E. Stoger, "The increasing value of plant-made proteins," *Curr. Opin. Biotechnol.*, vol. 32, pp. 163–170, 2015.

[64]   J. L. Fox, "First plant-made biologic approved," *Nat. Biotechnol.*, vol. 30, no. 6, pp. 472–472, 2012.

[65]   G. M. Lyon *et al.*, "Clinical care of two patients with Ebola virus disease in the United States," *N. Engl. J. Med.*, vol. 371, no. 25, pp. 2402–2409, 2014.

[66]   G. P. Lomonossoff and M. A. D'Aoust, "Plant-produced biopharmaceuticals: A case of technical developments driving clinical deployment," *Science*, vol. 353, no. 6305, pp. 1237–1240, 2016.

[67]   J. Schoberer *et al.*, "A signal motif retains Arabidopsis ER-α-mannosidase I in the cis-Golgi and prevents enhanced glycoprotein ERAD," *Nat. Commun.*, vol. 10, no. 1, p. 3701, 2019.

[68]   R. Strasser, "Plant protein glycosylation," *Glycobiology*, vol. 26, no. 9, pp. 926–939, 2016.

[69]   P. Bencúr *et al.*, "Arabidopsis thaliana β1,2-xylosyltransferase: An unusual glycosyltransferase with the potential to act at multiple stages of the plant N-glycosylation pathway," *Biochemical Journal*, vol. 388, no. 2. pp. 515–525, 2005.

[70] R. Strasser, J. Schoberer, C. Jin, J. Glössl, L. Mach, and H. Steinkellner, "Molecular cloning and characterization of Arabidopsis thaliana Golgi α-mannosidase II, a key enzyme in the formation of complex N-glycans in plants," *Plant J.*, vol. 45, no. 5, pp. 789–803, 2006.

[71] Y. J. Shin *et al.*, "Reduced paucimannosidic N-glycan formation by suppression of a specific β-hexosaminidase from Nicotiana benthamiana," *Plant Biotechnology J.*, vol. 15, no. 2. pp. 197–206, 2017.

[72] R. Strasser *et al.*, "Enzymatic properties and subcellular localization of arabidopsis β-N-acetylhexosaminidases," *Plant Physiol.*, vol. 145, no. 1, pp. 5–16, 2007.

[73] E. Liebminger *et al.*, "β-N-acetylhexosaminidases HEXO1 and HEXO3 are responsible for the formation of paucimannosidic N-glycans in Arabidopsis thaliana," *J. Biol. Chem.*, vol. 286, no. 12. pp. 10793–10802, 2011.

[74] N. Q. Palacpac *et al.*, "Stable expression of human β1,4-galactosyltransferase in plant cells modifies N-linked glycosylation patterns," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 96, no. 8. pp. 4692–4697, 1999.

[75] H. Bakker *et al.*, "Galactose-extended glycans of antibodies produced by transgenic plants," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 98, no. 5, pp. 2899–2904, 2001.

[76] H. Bakker *et al.*, "An antibody produced in tobacco expressing a hybrid β-1,4- galactosyltransferase is essentially devoid of plant carbohydrate epitopes," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 103, no. 20, pp. 7577–7582, 2006.

[77] R. Strasser, F. Altmann, L. Mach, J. Glössl, and H. Steinkellner, "Generation of Arabidopsis thaliana plants with complex N-glycans lacking β1,2-linked xylose and core α1,3-linked fucose," *FEBS Letters*, vol. 561, no. 1–3. pp. 132–136, 2004.

[78] J. M. Alonso *et al.*, "Genome-wide insertional mutagenesis of Arabidopsis thaliana," *Science*, vol. 301, no. 5633, pp. 653–657, 2003.

[79] R. Strasser *et al.*, "Generation of glyco-engineered Nicotiana benthamiana for the production of monoclonal antibodies with a homogeneous human-like N-glycan structure," *Plant Biotechnology J.*, vol. 6, no. 4. pp. 392–402, 2008.

[80] R. Strasser *et al.*, "Improved virus neutralization by plant-produced anti-HIV antibodies with a homogeneous β1,4-galactosylated N-glycan profile," *J. Biol. Chem.*, vol. 284, no. 31. pp. 20479–20485, 2009.

[81] A. Castilho *et al.*, "In Planta protein sialylation through overexpression of the respective mammalian pathway," *J. Biol. Chem.*, vol. 285, no. 21, pp. 15923–15930, 2010.

[82] A. Castilho *et al.*, "N-Glycosylation engineering of plants for the biosynthesis of glycoproteins with bisected and branched complex N-glycans," *Glycobiology*, vol. 21, no. 6, pp. 813–823, 2011.

[83] A. Castilho *et al.*, "Generation of Biologically Active Multi-Sialylated Recombinant Human EPOFc in Plants," *PLoS One*, vol. 8, no. 1, pp. 1–13, 2013.

[84] J. Jez *et al.*, "Expression of functionally active sialylated human erythropoietin in plants," *Biotechnol. J.*, vol. 8, no. 3, pp. 371–382, 2013.

[85] A. Castilho *et al.*, "Engineering of sialylated mucin-type O-glycosylation in plants," *J. Biol. Chem.*, vol. 287, no. 43, pp. 36518–36526, 2012.

[86] A. Castilho *et al.*, "Rapid high yield production of different glycoforms of ebola virus monoclonal antibody," *PLoS One*, vol. 6, no. 10, p. e26040, 2011.

[87] G. J. A. Rouwendal, D. E. A. Florack, T. Hesselink, J. H. Cordewener, J. P. F. G. Helsper, and D. Bosch, "Synthesis of Lewis X epitopes on plant N-glycans," *Carbohydr. Res.*, vol. 344, no. 12, pp. 1487–1493, 2009.

[88] P. D. Yurchenco and P. H. Atkinson, "Equilibration of Fucosyl Glycoprotein Pools in Hela Cells," *Biochemistry*, vol. 16, no. 5, pp. 944–953, 1977.

[89] C. E. Caffaro and C. B. Hirschberg, "Nucleotide sugar transporters of the Golgi apparatus: From basic science to diseases," *Acc. Chem. Res.*, vol. 39, no. 11, pp. 805–812, 2006.

[90] N. A. Peterson, T. K. Anderson, X. J. Wu, and T. P. Yoshino, "In silico analysis of the fucosylation-associated genome of the human blood fluke Schistosoma mansoni: Cloning and characterization of the enzymes involved in GDP-L-fucose synthesis and Golgi import," *Parasites and Vectors*, vol. 6, no. 1, pp. 1–21, 2013.

[91] P. Both *et al.*, "Distantly related plant and nematode core 1,3-fucosyltransferases display similar trends in structure-function relationships," *Glycobiology*, vol. 21, no. 11, pp. 1401–1415, 2011.

[92] H. Y. Sun *et al.*, "Structure and mechanism of Helicobacter pylori fucosyltransferase: A basis for lipopolysaccharide variation and inhibitor design," *J. Biol. Chem.*, vol. 282, no. 13, pp. 9973–9982, 2007.

[93] S. L. Martin, M. R. Edbrooke, T. C. Hodgman, D. H. Van Den Eijnden, and M. I. Bird, "Lewis X biosynthesis in Helicobacter pylori. Molecular cloning of an α(1,3)-fucosyltransferase gene," *J. Biol. Chem.*, vol. 272, no. 34, pp. 21349–21356, 1997.

[94] C. Breton, R. Oriol, and A. Imberty, "Conserved structural features in eukaryotic and prokaryotic fucosyltransferases," *Glycobiology*, vol. 8, no. 1, pp. 87–94, 1998.

[95] C. E. Machamer, "Golgi retention signals: do membranes hold the key?," *Trends Cell Biol.*, vol. 1, no. 6, pp. 141–144, 1991.

[96] T. Nilsson, P. Slusarewicz, M. H. Hoe, and G. Warren, "Kin recognition. A model for the retention of Golgi enzymes," *FEBS Lett.*, vol. 330, no. 1, pp. 1–4, 1993.

[97] S. Munro, "An investigation of the role of transmembrane domains in Golgi protein retention.," *EMBO J.*, vol. 14, no. 19, pp. 4695–4704, 1995.

[98] G. H. Patterson, K. Hirschberg, R. S. Polishchuk, D. Gerlich, R. D. Phair, and J. Lippincott-Schwartz, "Transport through the Golgi Apparatus by Rapid Partitioning within a Two-Phase Membrane System," *Cell*, vol. 133, no. 6, pp. 1055–1067, 2008.

[99] A. S. Uliana, C. G. Giraudo, and H. J. F. Maccioni, "Cytoplasmic tails of SialT2 and GalNAcT impose their respective proximal and distal Golgi localization," *Traffic*, vol. 7, no. 5, pp. 604–612, 2006.

[100] A. Petrosyan, M. F. Ali, and P. W. Cheng, "Keratin 1 plays a critical role in golgi localization of core 2 N-acetylglucosaminyltransferase M via interaction with its cytoplasmic tail," *J. Biol. Chem.*, vol. 290, no. 10, pp. 6256–6269, 2015.

[101] L. Tu, W. C. S. Tai, L. Chen, and D. K. Banfield, "Signal-mediated dynamic retention of glycosyltransferases in the Golgi," *Science*, vol. 321, no. 5887, pp. 404–407, 2008.

[102] K. R. Schmitz *et al.*, "Golgi Localization of Glycosyltransferases Requires a Vps74p Oligomer," *Dev. Cell*, vol. 14, no. 4, pp. 523–534, 2008.

[103] E. S. P. Eckert *et al.*, "Golgi phosphoprotein 3 triggers signal-mediated incorporation of glycosyltransferases into coatomer-coated (COPI) vesicles," *J. Biol. Chem.*, vol. 289, no. 45, pp. 31319–31329, 2014.

[104] M. F. Ali, V. B. Chachadi, A. Petrosyan, and P. W. Chengs, "Golgi phosphoprotein 3 determines cell binding properties under dynamic flow by controlling golgi localization of core 2 N-acetylglucosaminyltransferase 1," *J. Biol. Chem.*, vol. 287, no. 47, pp. 39564–39577, 2012.

[105] L. Liu, B. Doray, and S. Kornfeld, "Recycling of Golgi glycosyltransferases requires direct binding to coatomer," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 115, no. 36, pp. 8984–8989, 2018.

[106] C. Gao *et al.*, "The Golgi-Localized Arabidopsis endomembrane protein12 contains both endoplasmic reticulum export and golgi retention signals at its C terminus," *Plant Cell*, vol. 24, no. 5, pp. 2086–2104, 2012.

[107] C. Gao *et al.*, "Retention mechanisms for ER and Golgi membrane proteins," *Trends Plant Sci.*, vol. 19, no. 8, pp. 508–515, 2014.

[108] J. Schoberer *et al.*, "Arginine/lysine residues in the cytoplasmic tail promote ER export of plant glycosylation enzyme," *Traffic*, vol. 10, no. 1, pp. 101–115, 2009.

[109] J. Schoberer *et al.*, "The transmembrane domain of N -acetylglucosaminyltransferase i is the key determinant for its Golgi subcompartmentation," *Plant J.*, vol. 80, no. 5, pp. 809–822, 2014.

[110] J. Schoberer *et al.*, "The golgi localization of gnti requires a polar amino acid residue within its transmembrane domain," *Plant Physiol.*, vol. 180, no. 2, pp. 859–873, 2019.

[111] F. Brandizzi, N. Frangne, S. Marc-Martin, C. Hawes, J. M. Neuhaus, and N. Paris, "The destination for single-pass membrane proteins is influenced markedly by the length of the hydrophobic domain," *Plant Cell*, vol. 14, no. 5, pp. 1077–1092, 2002.

[112] G. Van Meer, D. R. Voelker, and G. W. Feigenson, "Membrane lipids: Where they are and how they behave," *Nat. Rev. Mol. Cell Biol.*, vol. 9, no. 2, pp. 112–124, 2008.

[113] F. H. Fenteany and K. J. Colley, "Multiple signals are required for α2,6-sialyltransferase (ST6Gal I) oligomerization and golgi localization," *J. Biol. Chem.*, vol. 280, no. 7, pp. 5423–5429, 2005.

1

[114] K. J. Colley, "Golgi localization of glycosyltransferases: More questions than answers," *Glycobiology*, vol. 7, no. 1, pp. 1–13, 1997.

[115] J. Schoberer, E. Liebminger, S. W. Botchway, R. Strasser, and C. Hawes, "Time-resolved fluorescence imaging reveals differential interactions of N-glycan processing enzymes across the golgi stack in planta," *Plant Physiol.*, vol. 161, no. 4, pp. 1737–1754, 2013.

[116] A. Hassinen *et al.*, "Functional organization of Golgi N- and O-glycosylation pathways involves pH-dependent complex formation that is impaired in cancer cells," *J. Biol. Chem.*, vol. 286, no. 44, pp. 38329–38340, 2011.

[117] S. Kellokumpu, A. Hassinen, and T. Glumoff, "Glycosyltransferase complexes in eukaryotes: Long-known, prevalent but still unrecognized," *Cell. Mol. Life Sci.*, vol. 73, no. 2, pp. 305–325, 2016.

Chapter 2

Sub-Golgi localisation of *Schistosoma mansoni* fucosyltransferases in *Nicotiana benthamiana*

Kim van Noort [1], Annet Lindhout [1], Verena Kriechbaumer [2], Chris Hawes [2], Ruud H.P. Wilbers [1] and Arjen Schots [1]

[1] Laboratory of Nematology, Wageningen University and Research, Wageningen, the Netherlands;
[2] Plant Cell Biology, Oxford Brookes University, Oxford, United Kingdom.

2

Abstract

The human parasite *Schistosoma mansoni* secretes immunomodulatory glycoproteins during its life cycle to dampen the host's immune response. The glycans on the secreted immunomodulatory proteins play a key role in the interaction with the human immune system. The potential of these glycoproteins to dampen allergic reactions and autoimmune disorders has been shown in clinical trials and mouse model studies. To further develop helminth-derived glycoproteins as biopharmaceuticals, a large-scale expression system is required for the production of recombinant glycoproteins with defined and tailored glycosylation. *S. mansoni* synthesises highly fucosylated glycan motifs on the N-glycans of its glycoproteins that cannot be synthesised in current production systems. Production of highly fucosylated N-glycans requires co-expression of specific fucosyltransferases (FucTs) in the expression host. Various putative *S. mansoni* FucT (SmFucT) coding sequences can be found in literature and databases. However, to date only one SmFucT is characterised in chemoenzymatic assays. Sub-Golgi localisation of FucTs determines donor and acceptor availability and may reveal information on FucT function. Therefore, in this study we determined the sub-Golgi localisation of ten SmFucTs by co-localisation with reported Golgi markers. Differences in sub-Golgi localisation were observed and suggested a function in fucosylation of complex glycan motifs or core fucosylation. This information helps us with the characterisation of SmFucTs and creates possibilities to modify the plant glycosylation pathway to produce immunomodulatory helminth glycoproteins. Production of these glycoproteins enables research to their immunomodulatory properties and potential future use as biopharmaceutical for treatment or prevention of allergies and autoimmune disorders.

Introduction

The trematode *Schistosoma mansoni* is a human parasite that causes the disease schistosomiasis, the most important parasitic disease after malaria. *S. mansoni* infects humans through free-living cercariae in fresh water that upon skin penetration lose their tail and transform into schistosomula. The schistosomula travel via the blood stream through the lungs to the mesenteric vein near the intestine, where they pair and start sexual reproduction. The resulting eggs will either leave the human body via the faeces or are trapped in organs, such as the liver or intestine where they can cause severe inflammation. Free-living miracidia will hatch from the eggs upon contact with fresh water and infect fresh-water *Biomphalaria* snails as intermediate hosts. In these snails asexual reproduction takes place via mother and daughter sporocysts. After four to six weeks free-living cercariae leave the snail in search for a human host to continue their life cycle.

Throughout its lifecycle *S. mansoni* secretes glycoproteins and glycolipids with various fucosylated glycan motifs [1]–[3]. These glycoconjugates play a key role in parasitism as they play a prominent role in immunomodulation. Fucosylated N-glycan motifs are synthesised by fucosyltransferases (FucTs), type II membrane proteins that transfer ʟ-fucose to a glycoconjugate acceptor. To synthesise a specific glycan motif, the ʟ-fucose moiety of guanosine diphosphate β-ʟ-fucose (GDP-fucose) is transferred to the corresponding glycoconjugate acceptor in an α1,2, α1,3, α1,4 or α1,6 confirmation. To ensure correct linkage FucTs have GDP-fucose and glycoconjugate acceptor binding pockets. These binding pockets show substrate specificity and enable synthesis of specific glycan motifs, which give FucTs different functions. FucTs are functionally conserved in various organisms, although depending on which FucT is expressed glycan motifs can be organism-, stage-, tissue- or cell-specific. In order to understand how the fucosylated glycan motifs produced by *S. mansoni* are synthesised and how these motifs influence parasite biology, we studied the function of the FucTs of *S. mansoni* (SmFucTs).

S. mansoni has many different fucosylated glycan motifs, therefore it is not surprising that within the genome of *S. mansoni* different putative SmFucT coding sequences have been found. The first SmFucT was found by Trottein and colleagues [4] based on the sequence of *Schistosoma japonicum* FucTs. After the genome of *S. mansoni* was published in 2009 more putative SmFucT genes were found, resulting in 14 full-length SmFucTs that could be picked up from cDNA [5]–[9]. Based on sequence comparison the SmFucTs can be grouped according to the type of linkages formed upon fucose transfer. Two SmFucTs were characterised as putative O-FucTs involved in direct fucosylation of serine or threonine. The other 12 SmFucTs were characterised as putative α1,3-SmFucT (denoted SmFucTA to F) or putative α1,6-SmFucT (denoted SmFucT H to M) involved in synthesis of N-glycans, O-glycans and/or glycolipids [9]. Further functional characterisation was

done by Mickum and colleagues [10], who reported that SmFucTF synthesises the glycan motif LeX in chemoenzymatic assays with synthetic glycoconjugate acceptors. The function of the other five α1,3-SmFucTs was predicted based on phylogenetic and sequence analysis.

The *in vivo* function of glycosyltransferases, however, not only depends on their sequence but also on their Golgi localisation and corresponding donor and acceptor availability. For instance, in *S. mansoni*, fucosylation of the glycan core requires a GlcNAc residue on the α1,3-mannose antenna of the glycoconjugate acceptor. α1,6-Fucosylation of the glycan core is inhibited upon addition of a terminal galactose to this GlcNAc residue, whereas addition of an α1,3 fucose is not inhibited [11]. Furthermore, a core α1,3-fucosylated glycoconjugate acceptor cannot receive a core α1,6-fucose, whereas a core α1,6-fucosylated glycoconjugate acceptor can receive a core α1,3-fucose. This acceptor specificity indicates that a core α1,6-SmFucT will probably localise earlier in the Golgi than a core α1,3-SmFucT, although both SmFucTs will probably localise after the β1,2-*N*-acetylglucosaminyltransferase (GnT) I and before the β1,4-galactosyltransferase (GalT). Thus, sub-Golgi localisation is important for acceptor availability and FucT functioning. Therefore, sub-Golgi localisation can reveal information on SmFucT function.

Golgi localisation involves the CTS region of glycosyltransferases, consisting of a N-terminal short cytoplasmic tail (CT), a transmembrane domain (TMD) and a stem region (S). Four models are described in which the CTS region of glycosyltransferases is important for Golgi targeting and either Golgi retrieval or retention: the receptor-mediated retrieval model; the bilayer thickness model; the kin-recognition or oligomerisation model; and the lipid based partitioning model [12]–[17]. Golgi retrieval or retention can rely on different and multiple parts of the CTS domain, such as reported for human FucT3 [18]–[21]. Moreover, no consensus amino acid sequences were found that serve as Golgi retrieval or retention signal for a group of glycosyltransferases, although for single glycosyltransferases important amino acid motifs were found [19], [22], [23]. Therefore, the Golgi position of the SmFucTs cannot be determined based on sequence analysis only. Since, Golgi localisation studies in *S. mansoni* are nearly impossible, plants are used as an alternative platform in this study. In plants various studies have been done on Golgi retrieval or retention mechanisms of plant glycosyltransferases and markers are available for co-localisation studies [20], [22], [24]–[26]. Furthermore, endogenous glycosyltransferases can be introduced by transient expression. Synthesis of endogenous N-glycans by introduction of endogenous glycosyltransferases shows that these glycosyltransferases are active in the plant glycosylation pathway.

In this study, we determined the sub-Golgi localisation of ten SmFucTs. Four α1,6-SmFucTs and six α1,3-SmFucTs were selected based on earlier research and sequence analysis [9]. The sub-Golgi localisation of these SmFucTs was determined in *Nicotiana*

benthamiana, by co-localisation studies of C-terminally GFP-tagged SmFucTs and mRFP-tagged reference markers. Differences in sub-Golgi localisation were observed between the α1,3-SmFucTs (except SmFucTC) and the α1,6-SmFucTs. Furthermore, differences between the sub-Golgi localisation of the α1,3-SmFucTs suggests a function in core α1,3-fucosylation or fucosylation of terminal glycan motifs.

Results

Putative *Schistosoma mansoni* fucosyltransferases

Many putative SmFucTs are described in literature and in the databases UniProtKB and GeneDB [4]–[7], [9], [10], [27]–[29]. To create an overview, compare and group the reported putative SmFucT coding sequences, a table was created (Table 1). In this table, the protein sequences of 46 out of 50 putative SmFucT coding sequences were compared to the protein sequences of SmFucTA to M, previously amplified from cDNA by Peterson and colleagues [9]. Since FucTs are type II membrane proteins, the TMD of the putative SmFucTs was predicted by the TMHMM server v2.0 [30]. Based on this prediction the length of the CT and the TMD were calculated. For many of the shorter putative SmFucT coding sequences no TMD was predicted, which suggested that these sequences were incomplete. The predicted TMDs were all around 23 amino acids except for sequences mostly related to SmFucTA, K and L, which show a TMD length of 18 to 20 amino acids. The length of the CT was more variable. The percentage of identity between the putative SmFucT coding amino acid sequences and the protein sequences of SmFucTA to M varied and various mutations, deletions and insertions were observed.

Since only SmFucTA to M have been amplified from cDNA, our analysis continued with these SmFucT protein sequences [9]. SmFucTG was also excluded as it encodes a premature stop codon and therefore lacks a large part of the catalytic domain. With SmFucTH as exception, remarkably the percentage of identity between the α1,6-SmFucTs (H to M) was higher compared to the α1,3-SmFucTs (A to F). Since, the percentage of identity between SmFucTI, J and M was above 89% and the observed mutations were not present in conserved protein regions shared by α1,2- and α1,6-FucTs, SmFucTI and SmFucTM were excluded from further analysis [9]. We continued our study with the remaining SmFucTs (A to F, H and J to L).

TABLE 1 | Putative *Schistosoma mansoni* fucosyltransferases (SmFucTs). This table describes the putative SmFucT coding sequences reported in literature and found in the databases UniProtKB and GeneDB. The transmembrane domain (TMD) was predicted by the TMHMM server v2.0. Based on these predictions the length of the cytoplasmic tail (CT) and the TMD were calculated. x indicates that no sequence was found or no TMD was predicted by the TMHMM server v2.0. SmFucTA to M, amplified from cDNA by Peterson and colleagues [9], were used as reference sequences. The percentage of identity between the putative SmFucT coding amino acid (aa) sequences and the most similar protein reference sequences were determined by clustal Omega alignment. If the percentage of identity exceeded 70% differences are indicated according to the protein mutation nomenclature of den Dunnen and Antonarakis (2000). The differences between the SmFucTs amplified from cDNA are only indicated when the percentage of identity exceeded 90%.

Protein	Length (aa)	TMD (length CT/ TMD in aa)	Percentage of identity	Sequence diversity and sequence related information	Published in
SmFucTA	426aa	20-37 (19/18)	47% with SmFucTD		[4], [9], [10], [27], [28]
smp_148850	x			on ncbi linked with SmFucTA described by Trottein and colleagues [4]	[5], [6], [9]
smp_211180	643aa	x	92.88% with SmFucTA	p.M1_K61del, p.K62M, p.N398Mfs*308	[28], [29]
smp_214370.1	426aa	20-37 (19/18)	100% with SmFucTA		[27]
smp_129730	49aa	x	30.43% with SmFucTA		[28]
smp_214380.1	338aa	x	21.55% with SmFucTA		[27]
SmFucTB	416aa	7-29 (6/23)	50.72% with SmFucTF		[9], [10], [27], [28]
smp_099090	347aa	x	21.09% with SmFucTB		[5], [6], [9], [28]
smp_209060	277aa	7-29 (6/23)	99.28% with SmFucTB	previously smp_109500, p.V162I, p.I181_E319del	[27]–[29]
A0A3Q0KJK9	416aa	7-29 (6/23)	98.08% with SmFucTB	p.V162I, p.V156E, p.P157A, p.K170S, p.T263I, p.S306G, p.A316V, p.C320G	[28]
A0A146MI48	416aa	7-29 (6/23)	97.6% with SmFucTB	p.Y48H, p.V162I, p.V166E, p.P167A, p.K170S, p.I254V, p.G378E, p.S306G, p.A316V, p.E320G	[28]
smp_109500	x			see smp_209060	[6]
SmFucTC	463aa	13-35 (12/23)	41.38% with SmFucTA		[9], [10], [27], [28]
smp_154410	311aa	x	100% with SmFucTC	p.M1_L152del	[5], [9], [27]–[29]
A0A3Q0KQD9	148aa	13-35 (12/23)	97.3% with SmFucTC	p.N148Lfs*2	[28]
SmFucTD	398aa	13-35 (12/23)	47% with SmFucTA		[9], [10], [27], [28]
smp_053400	394aa	13-35 (12/23)	100% with SmFucTD	p.N395*	[5]–[7], [9], [27]–[29]
A0A3Q0KG08	324aa	x	99.69% with SmFucTD	p.M1_V74del, p.I346K	[28]
smp_129750	304aa	x	50.99% with SmFucTD		[5], [9], [27]–[29]

Protein	Length (aa)	TMD (length CT/TMD in aa)	Percentage of identity	Sequence diversity and sequence related information	Published in
SmFucTE	420aa	17-39 (16/23)	69.63% with SmFucTF		[9], [10], [27]–[29]
smp_028910	101aa	x	100% with SmFucTE	previously smp_205640, p.M1_V326del	[27]–[29]
smp_137740	271aa	17-39 (16/23)	94.46% with SmFucTE	p.V190_K355delins15, p.N420K, p.G421D, p.R422H	[5], [6], [9], [27]–[29]
smp_205640	x			see smp_028910	[27]–[29]
A0A3Q0KMV7	355aa	17-39 (16/23)	98.59% with SmFucTE	p.Y44H, p.V190-S262del, p.N420K, p.G421D, p.R422H	[28]
SmFucTF	434aa	21-43 (20/23)	69.63% with SmFucTE		[9], [10], [27], [28]
smp_137730	31aa	x	93.55% with SmFucTF	p.M1_K295del, p.T317A, p.N326Kfs*2	[6], [7], [9], [27]–[29]
smp_142860	139aa	x	100% with SmFucTF	p.M1_K295del	[5], [27]–[29]
A0A3Q0KNN7	434aa	21-43 (20/23)	99.77% with SmFucTF	p.V36I	[28]
smp_193620	92aa	x	72.83% with SmFucTF	p.M1_P131del, p.I196Vfs*29	[29]
smp_193870	101aa	x	98.02% with SmFucTF	p.M1_V333del, p.Y348H, p.P395L	[27]–[29]
smp_194990	118aa	21-43 (20/23)	96.61% with SmFucTF	p.S25Q, p.V36I, p.H91Q, p.K118Sfs*2	[6], [9], [27]–[29]
smp_199790	96aa	x	85.42% with SmFucTF	p.M1_A187del, p.L104S, p.D270Sfs*14	[27]–[29]
SmFucTG	x			pseudogene (Peterson and colleagues [9])	[9], [27]
SmFucTH	599aa	9-31 (8/23)	33.87% with SmFucTJ and SmFucTM		[9], [27], [28]
smp_175120.1	882aa	21-43 (20/23)	83.6% with SmFucTH	p.M1YextM-13, p.V3L, p.T4I, p.S5G, p.K7E, p.V31Nfs390, p.V31_G481ins379	[5], [9], [27], [29]
smp_175120.2	990aa	21-43 (20/23)	97.33% with SmFucTH	p.M1YextM-13, p.V3L, p.T4I, p.S5G, p.K7E, p.K78_D87del, p.R88K, p.L181_N194del, p.I281T, p.K398_E399V, p.D428E, p.G450E, p.491C_Ifs404*	[5], [9], [27], [29]
smp_211170	475aa	x	20.05% with SmFucTH		[28]
SmFucTI	592aa	x	98.99% with SmFucTJ	p.G398R, p.L399S, p.S405R, p.E502K, p.G547R, p.K581T	[9], [27], [28]

2

Protein	Length (aa)	TMD (length CT/ TMD in aa)	Percentage of identity	Sequence diversity and sequence related information	Published in
SmFucTJ	592aa	x	98.99% with SmFucTI	p.R398G, p.S399L, p.R405S, p.K502E, p.R547G, p.T581K	[9], [27], [28]
smp_138730	474aa	x	93.88% with SmFucTJ	p.M1_S77del, p.K79R, p.T80L, p.T81M, p.R82I, p.N83D, p.V84G, p.Q85A, p.I86E, p.T87V, p.N88D, p.G258_S260del	[5], [9], [27]–[29]
SmFucTK	579aa	12-30 (11/19)	87.65% with SmFucTL		[9], [27], [28]
smp_138750	502aa	x	100% with SmFucTK	p.M1_S77del	[5], [9], [27]–[29]
A0A3Q0KMV6	579aa	12-30 (11/19)	97.06% with SmFucTK	p.T80I, p.Q85K, p.E89V, p.S101R, p.L127F, p.Y157H, p.G146D, p.K175R, p.M176I, p.N180D, p.E187Q, p.N296D, p.Y309F, p.A318S, p.Q378R, p.F383L, p.T557A	[9]
SmFucTL	588aa	7-26 (6/20)	87.65% with SmFucTK		[5], [9], [27]–[29]
smp_030650	484aa	x	94.21% with SmFucTL	p.M1_K82del, p.I339M, p.I355_R406delin30, p.M452I	[27], [28]
SmFucTM	592aa	x	98.14% with SmFucTJ	p.S35C, p.S50P, p.L118S, p.I122L, p.N281H, p.H283L, p.R324K, p.E336K, p.V342G, p. G398R, p.L399S	[9], [28]
smp_212520	147aa	x	97.96% with SmFucTM	p.M1_K200del p.V345Cfs*4	[27]–[29]
smp_185720	x				[5]–[7]
smp_189280	x				[5]
smp_138740	x			found on same scaffold as SmFucTJ, K and L, but could not be amplified by Peterson and colleagues [9]	[5], [9]

Schistosoma mansoni fucosyltransferases localise in the Golgi

To determine expression of SmFucTs (A to F, H and J to L) in plants, C-terminally GFP-tagged SmFucTs were transiently co-expressed with P19 in *N. benthamiana* leaves. Leaves were harvested three days post infiltration (dpi) and leaf tissue was homogenised. Isolated proteins were analysed on a western blot using anti-GFP antibodies (Figure 1). The western blot revealed full-length GFP-tagged SmFucTA to F and multimers detected around 75 kDa and above 150 kDa, respectively. These bands indicate that GFP tagged SmFucTA to F were expressed. Although, for SmFucTH and J to L no bands were observed around 75 kDa, the smear above 250 kDa and the faint bands around 100 kDa indicated multimerisation. Since these bands were absent in the GFP control this suggests expression of SmFucTH and J to L at a lower concentration and/or poor isolation of these GFP tagged SmFucTs. The band around 25 kDa corresponded to free GFP, which indicated partial cleavage of GFP.
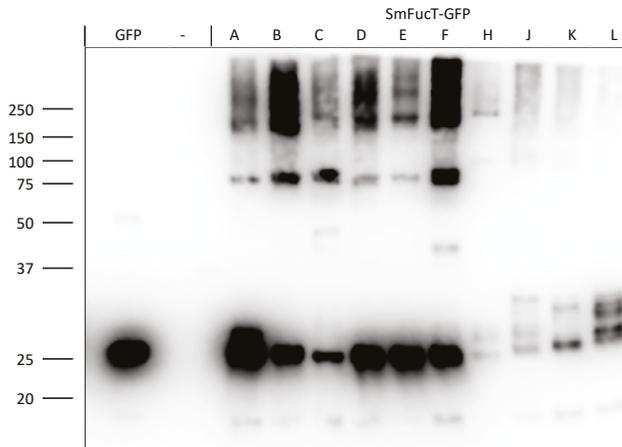
FIGURE 1 | Analysis of *Schistosoma mansoni* fucosyltransferase (SmFucT) expression in *Nicotiana benthamiana*. To analyse expression of SmFucTs, GFP-tagged SmFucTs were transiently expressed in *N. benthamiana*. Subsequently, 1 µg of total protein from crude plant extracts was analysed under reducing conditions on a western blot using anti-GFP antibodies. The western blot reveals cleavage of GFP and expression of the GFP-tagged SmFucTs.

After confirmed expression, the subcellular localisation of the GFP-tagged SmFucTs was investigated by transient co-expression with P19 in *N. benthamiana* leaves. Leaves were harvested three dpi and the fluorescently labelled proteins were visualised by confocal microscopy. Since Golgi bodies move, latrunculin B was used to inhibit Golgi stack mobility. For SmFucTA to F, H and J punctuated structures were observed corresponding to Golgi localisation (Figure 2). In addition, localisation in different cellular organelles was observed, such as localisation in the nucleus, in the ring around the nucleus corresponding to the nuclear envelop and in network-like structures corresponding to the ER and the

cytoplasm (Figure 2 and Supplemental Figure 1). The localisation in the nucleus and the cytoplasm was comparable to the localisation of free GFP and can probably be attributed to cleaved GFP (Figure 2). The localisation in the nuclear envelope and the network-like structures were comparable to the localisation of cSP-GFP and probably correspond to ER
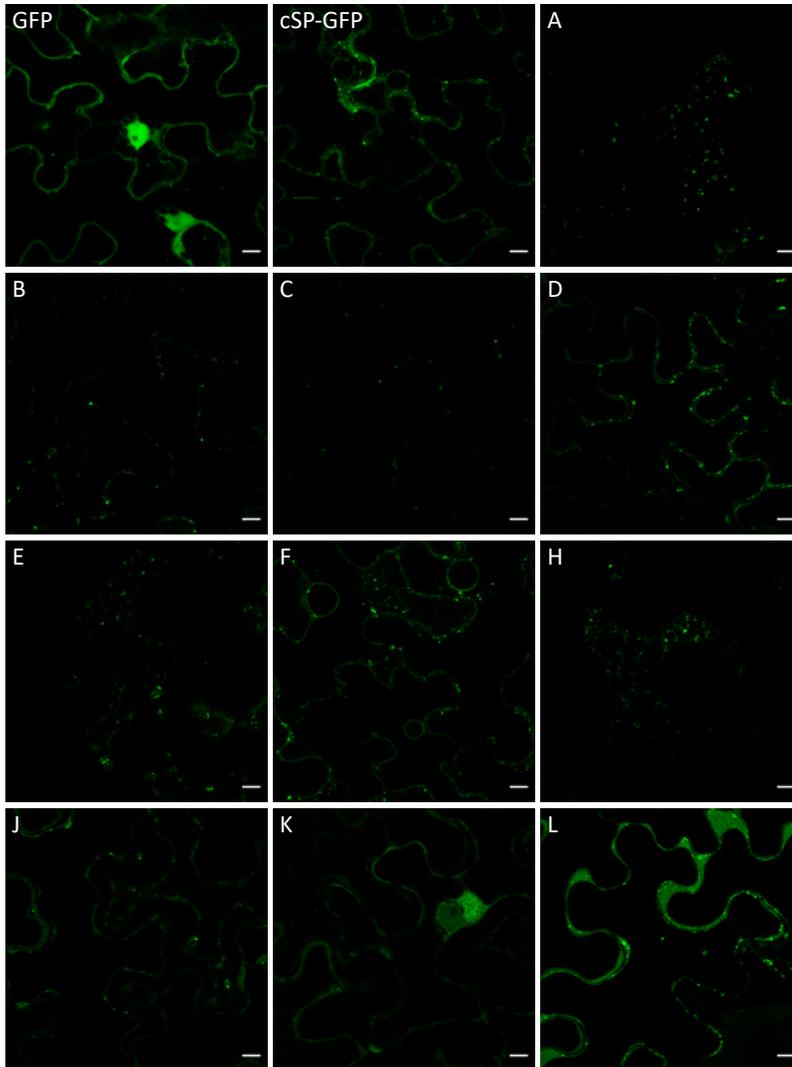
FIGURE 2 | Subcellular localisation of *Schistosoma mansoni* fucosyltransferases (SmFucTs). The subcellular localisation of GFP-tagged SmFucTs (in green) was investigated by transient expression in *Nicotiana benthamiana* leaves. Leaves were harvested and the fluorescently labelled proteins were visualised by confocal microscopy. GFP and cSP-GFP were taken along as controls. GFP showed localisation in the nucleus and cytoplasm, whereas cSP-GFP showed localisation in the nuclear envelope, the ER and the Golgi. The letters in the upper left corner of the panels correspond to the SmFucT analysed. The punctuated structures observed for SmFucTA to F, H and J indicate Golgi localisation. Localisation of SmFucTK and SmFucTL in the Golgi cannot be confirmed. The scale bar indicates 10 μm.

localisation (Figure 2 and Supplemental Figure 1). Next to localisation in these organelles, aggregate formation or storage in bodies adjacent to the nucleus (resembling virus-like-induced inclusion bodies) was observed for SmFucTJ, K and L (Figure 2 and Supplemental Figure 1). Since the same experimental set-up without latrunculin B showed identical results, aggregates could be formed due to P19 overexpression (data not shown). Altogether, we revealed expression and Golgi localisation for SmFucTA to F, H and J in plants.

Sub-Golgi localisation of *Schistosoma mansoni* fucosyltransferases

Next, SmFucT localisation was determined inside the Golgi. Therefore, GFP-tagged SmFucTs were co-expressed with mRFP-tagged Golgi reference markers GnTI-mRFP, XylT-mRFP and ST-mRFP. GnTI-mRFP localises more to the *cis*-Golgi in comparison to XylT-mRFP and the *trans*-Golgi localising ST-mRFP. XylT-mRFP localises in between GnTI-mRFP and ST-mRFP. Since, P19 co-expression was suggested to result in aggregate formation, P19 was not co-expressed in this experiment. To capture the fast-moving Golgi bodies without fixation or use of latrunculin B movies were made and individual frames were analysed. It was observed that all SmFucTs localised in the Golgi. Moreover, aggregate formation was reduced.

Golgi bodies tumble, therefore two different views were distinguished, the "donut-shaped Golgi body" and the "side view Golgi body" (Figure 3). These tumbling Golgi bodies were mostly observed in cytoplasmic streams. Donut-shaped Golgi bodies were mostly observed against the cell periphery on the epidermal side, whereas side Golgi bodies were mostly observed in cytoplasmic streams or against the cell periphery to the side of the cell.

Co-localisation of the SmFucTs with the three different mRFP-tagged reference makers was monitored. SmFucTA to F seemed to co-localise more with ST-mRFP than with GnTI-mRFP and XylT-mRFP (Supplemental Figure 2). SmFucTH and J to L seemed to co-localise with all three reference markers (Supplemental Figure 2). Striking was the different appearance of the three markers. ST-mRFP could be observed as donut, whereas GnTI-mRFP and XylT-mRFP rather seemed to form a smaller ball than a donut. The donut-shape observed for ST-mRFP was also observed for SmFucTA to F. For SmFucTH and J to L the donut-shape was difficult to distinguish, due to lower expression and subsequently more background noise. Interestingly, in donut-shape on some occasions two ring structures were observed for SmFucTA to E (Supplemental Figure 3). Furthermore, in cytoplasmic streams at times tubule-like structures were observed (Supplemental Figure 3).

In order to visualise the degree of separation, fluorescent intensity profiles were made from the path drawn by a white arrow across the different Golgi stacks (Figure 4). Since, a clearer degree of separation between the two colours could be seen for the side Golgi bodies, side-on view Golgi bodies were selected for analysis. Three merged pictures of SmFucTE and the different reference markers were analysed (Figure 4). The fluorescent intensity profiles of SmFucTE and ST-mRFP overlapped, whereas a shift to the right was observed for the profile of SmFucTE relative to the profiles of GnTI-mRFP and XylT-mRFP.

Furthermore, a difference in intensity was observed for the profile of SmFucTE relative to the profiles of GnTI-mRFP and XylT-mRFP. The intensity difference was largest for SmFucTE relative to XylT-mRFP, suggesting differences in expression level or a horizontal as well as a vertical separation of SmFucTE and XylT-mRFP.
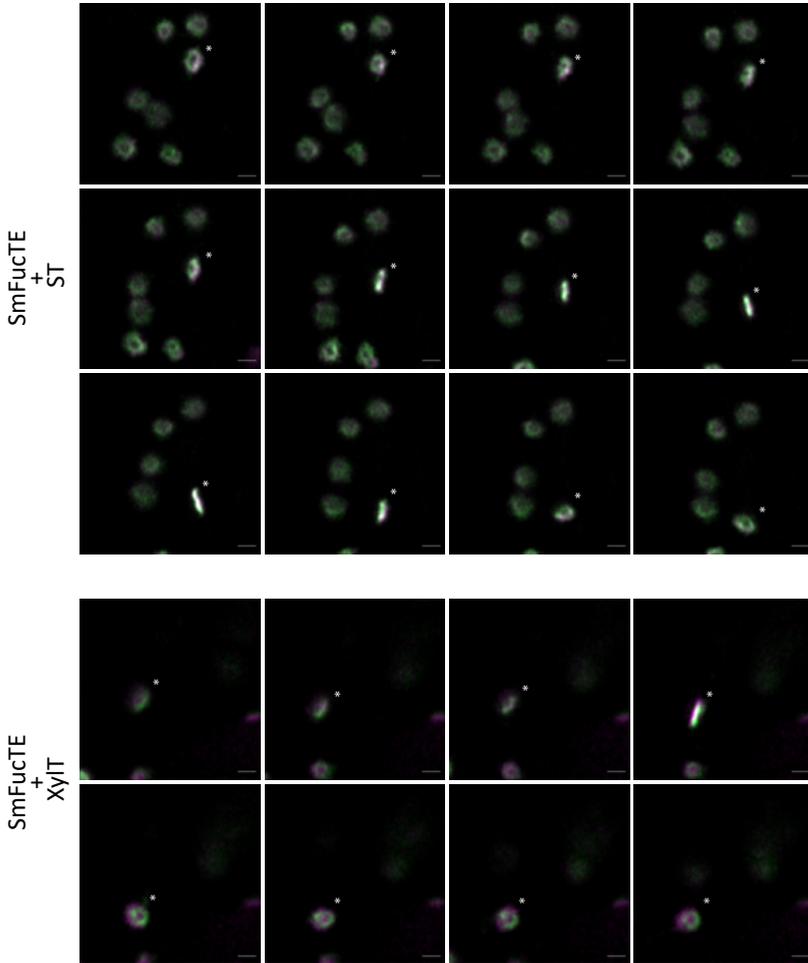
FIGURE 3 | Tumbling Golgi bodies in "donut-shape" and "side view". Co-infiltrated GFP-tagged SmFucTE (in green) and mRFP-tagged reference markers ST and XylT (in violet) in *Nicotiana benthamiana* leaves were visualised by confocal microscopy via movies of ~50 frames taken 1.27sec apart from each other. 12 and 8 frames were selected that show a tumbling Golgi body (indicated by *) turning from "donut-shape" to "side view" back to "donut-shape" while moving downwards. Co-localisation is shown in white. Merged representative pictures of co-infiltrated SmFucTE and ST reveal that these proteins localise together during tumbling. Merged representative pictures of co-infiltrated SmFucTE and XylT reveal that these proteins show a degree of co-localisation during tumbling. The scale bar indicates 1 μm.
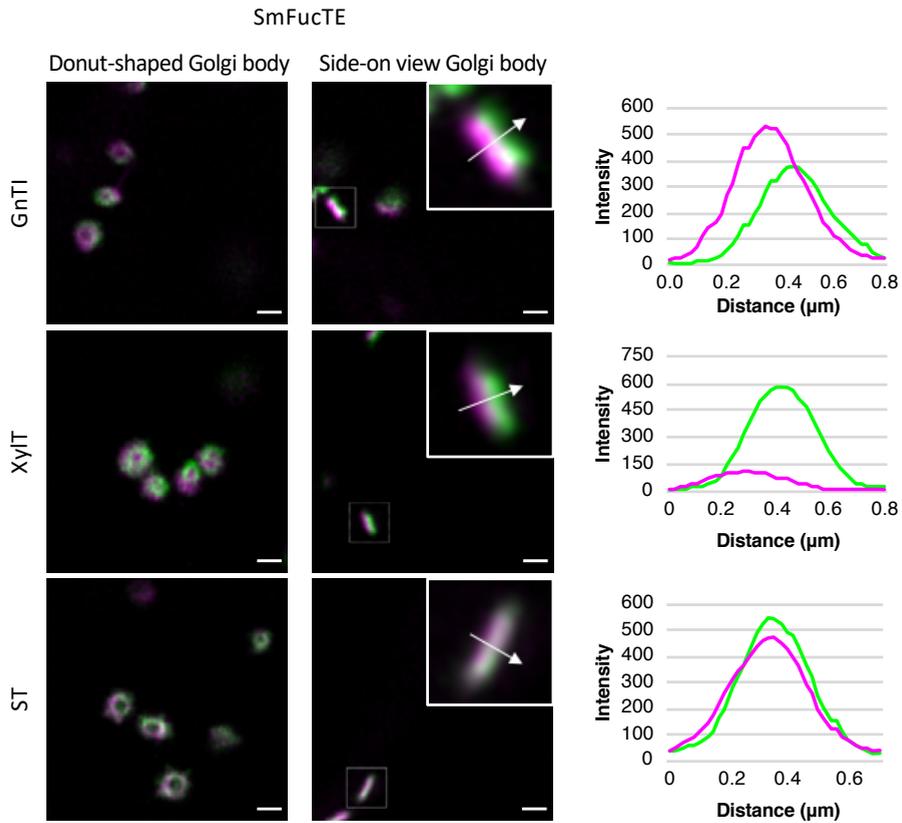
FIGURE 4 | Co-localisation fluorescent intensity profiles. Fluorescently tagged proteins were co-expressed in *Nicotiana benthamiana* leaves and co-localisation was analysed with confocal microscopy. Merged representative pictures are shown of GFP-tagged SmFucTE in green and the three mRFP-tagged reference markers GnTI, XylT or ST in violet. Co-localisation is seen in white. From left to right: a merged picture of a donut-shaped Golgi body ; a merged picture of a side-on view Golgi body ; a magnification of the boxed area in the side-on view Golgi body picture ; a fluorescent intensity profile generated from the path of the white arrow drawn across the centre of the side-on view Golgi stack. The fluorescent intensity profiles reveal better co-localisation of SmFucTE with ST in relation to GnTI and XylT. The scale bar indicates 1 μm.

Since the amount of co-localisation with the individual reference markers gave a more accurate indication of the SmFucT sub-Golgi localisation, the Pearson's correlation coefficient was determined for co-expression of the SmFucTs and the three reference markers (Figure 5). Analysis of the Pearson's correlation coefficient revealed that all α1,3-SmFucTs, except SmFucTC, co-localised significantly more with ST-mRFP than with GnTI-mRFP or XylT-mRFP, which indicated localisation towards the *trans* side of the Golgi. Interestingly, a difference in co-localisation was observed for SmFucTC compared to the other α1,3-SmFucTs, since SmFucTC did not co-localise significantly more with ST-mRFP than with XylT-mRFP. Furthermore, in comparison to the other α1,3-SmFucTs, SmFucTC

co-localised significantly more with XylT-mRFP. This suggested that SmFucTC probably localised more to the medial-Golgi than the other α1,3-SmFucTs. The α1,6-SmFucTs showed no significant difference in co-localisation with the three reference markers. In conclusion, differences in sub-Golgi localisation were observed between the α1,3-SmFucTA, B, D and E, SmFucTC and the α1,6-SmFucTs, which suggested a function in fucosylation of complex glycan motifs or core fucosylation.
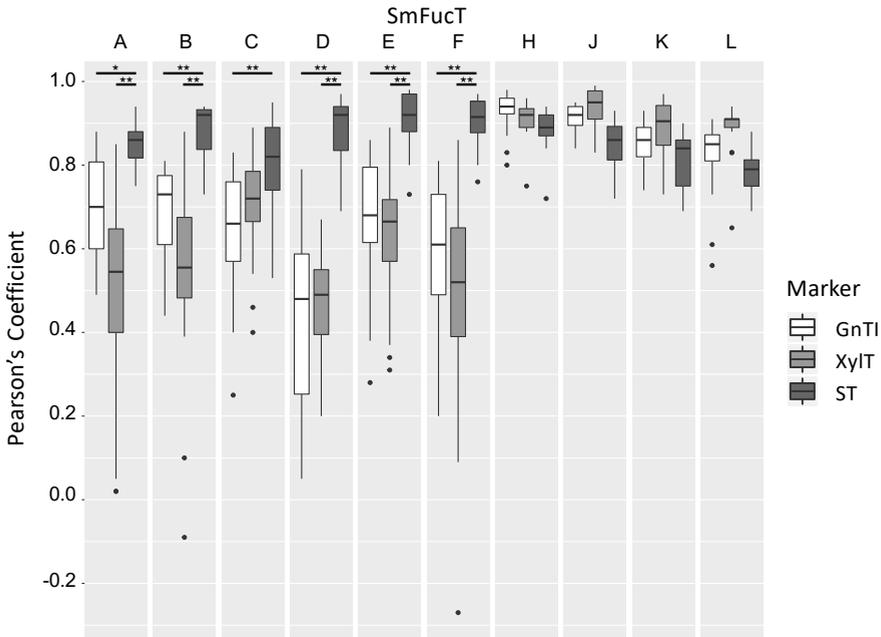
FIGURE 5 | Analysis of *Schistosoma mansoni* fucosyltransferase (SmFucT) sub-Golgi localisation. GFP-tagged SmFucTs were co-expressed with mRFP-tagged Golgi markers GnTI, XylT or ST in *Nicotiana benthamiana* leaves. Co-localisation was analysed by confocal microscopy. Subsequently, the Pearson's coefficient of side-on-view Golgi bodies was determined and plotted. P-values were calculated by TukeyHSD, * corresponds to $P < 5.0*10^{-3}$ and ** corresponds to $P < 5.0*10^{-7}$.

Discussion

After sequencing of the *S. mansoni* genome in 2009 various putative SmFucT coding sequences were found, of which 14 full-length proteins could be amplified from cDNA [5], [9]. These expressed SmFucTs were characterised as putative O-FucTs, involved in O-fucosylation, or putative α1,3- or α1,6-SmFucT involved in synthesis of N-glycans, O-glycans and/or glycolipids [9]. Only for SmFucTF a LeX synthesising function has been shown in chemoenzymatic assays [10]. Since, sub-Golgi localisation of a FucT is important for donor and acceptor availability and therefore its function, sub-Golgi localisation may

provide information on FucT function. In this study, we determined Golgi localisation and sub-Golgi distribution of ten SmFucTs involved in synthesis of N-glycans, O-glycans and/or glycolipids.

In comparison to the 12 expressed full length SmFucTA to M protein sequences, many of the putative SmFucT protein sequences seemed to code for incomplete proteins, lacking for instance a predicted TMD or containing a deletion at the C-terminal part. Important domains, such as the GDP-fucose binding domain, can be found at the C-terminal part of the FucTs [9]. For instance, for *Arabidopsis thaliana* Fut11 a deletion of the last 113 amino acids leads to a 99% reduction in enzyme activity [31]. Many of the putative SmFucT protein sequences were mentioned and used in research papers [6], [7], [9], [29]. Results of these incomplete and/or not expressed putative SmFucT coding sequences should therefore be interpreted with much care.

For the 12 full length expressed SmFucTA to M proteins, an amino acid sequence identity below 50% was observed for the putative α1,3-SmFucTs (SmFucTA to F), except for SmFucTE and F that had an identity of almost 70%. This protein sequence identity below 50% suggests that multiple functions can be attributed to the α1,3-SmFucTs, which is in agreement with the many different glycan motifs reported to have one or more α1,3-linked fucose(s) [1]–[3]. The relatively high similarity between SmFucTE and F suggests for SmFucTE a function comparable to the LeX synthesising function described for SmFucTF based on phylogeny and sequence analysis [10]. FucT9a from *Tetraodon nigroviridis* has shown the ability to synthesis both LeX and LDN-F [32]. Since SmFucTF cannot synthesise LDN-F this may indicate a LDN-F synthesising function for SmFucTE, although LDN-F synthesis was attributed to SmFucTA, B or D by Mickum and colleagues [10]. The protein sequence identity of the putative α1,6-SmFucTs SmFucTJ to L is at least 87%, which is high in comparison to the percentage of identity found among the α1,3-SmFucTs. The relatively low identity of SmFucTH (34%) with the other putative α1,6-SmFucTs suggests that SmFucTH has a different function than the other putative α1,6-SmFucTs, although core α1,6-fucosylation is the only known glycan motif in *S. mansoni* with an α1,6-linked fucose [1], [2].

Cellular localisation of the SmFucTs revealed Golgi localisation, although P19 overexpression seemed to lead to an SmFucT excess and subsequent storage in the ER or around the nucleus in virus-like-induced inclusion bodies [33]. Furthermore, localisation was observed in the cytoplasm and nucleus, which corresponded to free GFP as the same localisation pattern was observed for expressed GFP. Free GFP localises in the cytoplasm and is small enough to accumulate in the nucleus via passive diffusion, whereas proteins larger than 40 kDa, such as GFP-tagged SmFucTs, cannot pass the membrane passively [34], [35].

When latrunculin B was used to immobilise Golgi bodies, Golgi bodies were mostly observed in donut-shape and localised more at the outer edges of the cell. Latrunculin B blocks actin polymerisation and thereby influences ER shape and sheet versus tubule formation [36]. Since Golgi bodies are suggested to be linked to the ER, ER shape and movement can influence Golgi body localisation [37]. Furthermore, Golgi bodies move faster in a cytoplasmic stream than at the outer edges of the cell, therefore they stay relatively long at the outer edges of the cell and probably localise here when movement is abolished [38].

Sub-Golgi localisation of SmFucTs revealed differences in the degree of co-localisation with each of the Golgi markers GnTI-mRFP, XylT-mRFP and ST-mRFP. GnTI-mRFP localises more to the *cis*-Golgi in comparison to the more medial-Golgi localising XylT-mRFP and the *trans*-Golgi localising ST-mRFP. Therefore, from here on GnTI-mRFP, XylT-mRFP and ST-mRFP are referred to as the *cis*/medial-, medial- and *trans*-Golgi marker, respectively. All ten SmFucTs showed a degree of co-localisation with the *trans*-Golgi marker, this indicates that all SmFucTs localised at least partly to the *trans*-Golgi. The degree of co-localisation with the *trans*-Golgi marker significantly differed for SmFucTA, B, D, E and F in comparison to co-localisation with the *cis*/medial- and medial-Golgi markers. Therefore, SmFucTA, B, D, E and F localised more to the *trans*-Golgi. SmFucTC's co-localisation with the medial-Golgi marker was not significantly different from co-localisation with the *cis*/medial- or *trans*-Golgi marker. This suggested that SmFucTC localised more towards the medial-Golgi. Co-localisation of the α1,6-SmFucTs with the three markers showed that SmFucTH and J to L co-localised with all three markers as no significant difference was observed in co-localisation with each of the markers. This indicates that SmFucTH and J to L localised throughout the Golgi or the observed background noise hampered co-localisation analysis.

In the glycosylation pathway of *S. mansoni* a strict order of fucosylation is shown starting with core α1,6-fucosylation, followed by core α1,3-fucosylation and then fucosylation of terminal glycan motifs [11]. Regarding the localisation of SmFucTs, a core α1,6-SmFucT is, therefore, expected to localise before α1,3-SmFucTs. Since, SmFucTH and J to L localised more to the *cis*/medial-Golgi than to the other SmFucTs, SmFucTH and J to L are likely involved in core α1,6-fucosylation. After core α1,6-fucosylation, core α1,3-SmFucT activity is expected. Of the six analysed α1,3-SmFucTs, SmFucTC was more localised to the medial-Golgi, which corroborates its role in core α1,3-fucosylation. This is in agreement with medial/*trans*-Golgi localisation observed for the plant core α1,3-FucT and the functional prediction made by Mickum and colleagues [10], [39]–[41]. SmFucTA, B, D, E and F were more localised to the *trans*-Golgi and are therefore likely involved in the synthesis of more complex glycan motifs. The complex glycan motifs LeA and LeX in the antennae of respectively plants and mammals are also synthesised more to the *trans*-Golgi, corroborating our observations regarding the localisation and function of SmFucTA, B, D,

E and F [40]. These α1,3-SmFucTs are therefore expected to be involved in the synthesis of more complex glycan motifs such as LDN-F, LeX and DF as predicted by Mickum and colleagues [10].

Differences observed in donut-shape for the three markers suggests that not all Golgi compartments have the same shape and size. Where the *trans*-Golgi marker showed a donut-shape, the *cis*/medial- and medial-Golgi markers showed a smaller ball. The observed size in the donut-shape corresponds to the expected localisation, since the Golgi compartment grows from *cis* to late medial and becomes smaller again from the late medial to the *trans*-Golgi Network [42]. However, no obvious differences in size were observed for the α1,3-SmFucTs and the α1,6-SmFucTs, although background noise made it harder to draw this conclusion for the latter. Two ring structures, which were sometimes observed in donut shape, can correspond to maturation of a second Golgi stack (Supplemental Figure 3) [43]. Furthermore, the observed tubule-like structures in the cytoplasmic stream can link the Golgi to the ER or are formed by a continues transport of coat protein complex (COP) vesicles between the ER and the Golgi (Supplemental Figure 3) [44].

During sub-Golgi localisation experiments with the α1,6-SmFucTs, much background noise was observed, which at first glance could influence our observations. However, the Pearson's correlation coefficient is not sensitive towards differences in mean signal intensities and background noise makes the Pearson's correlation coefficient closer to 0 than to 1, whereas we observed coefficient values closer to 1 than to 0 in our co-localisations studies with the α1,6-SmFucTs [45].

A high degree of variation was observed for the Pearson's correlation coefficient in co-localisation studies with the α1,3-SmFucTs and the *cis*/medial- and medial-Golgi markers. This variation was caused by the movement of the Golgi bodies. The low coefficient values corresponded to fast moving Golgi bodies in a cytoplasmic stream, which were partly pulled apart by the high velocity. This means that proteins that do not co-localise were pulled apart even further. Therefore, the degree of variation in the coefficient values gives an indication of the degree of co-localisation. This indicates that the α1,3-SmFucTs co-localise less with the *cis*/medial- and medial-Golgi markers than the α1,6-SmFucTs, as also indicated by the median of the Pearson's correlation coefficient.

In conclusion, this study shows the Golgi localisation of ten SmFucTs involved in synthesis of N-glycans, O-glycans and/or glycolipids and gives an indication of their function based on sub-Golgi localisation. With the knowledge that transiently expressed SmFucTs localise in the Golgi, the indication of their function can be tested in *N. benthamiana*.

Materials and Methods

Construction of expression vectors

Ten full length open reading frames of SmFucT sequences (A to F, H and J to L) were codon optimised in house and flanked by SacI and NcoI/BspHI restriction sites at the '5-end and a KpnI restriction site at the '3-end (Supplemental Figure 4) [46]. Subsequently, these sequences were synthetically constructed at GeneArt. For SmFucTA an extra alanine residue was introduced after the start codon to introduce a NcoI restriction site. The synthetically constructed sequences were cloned into the plant expression vector pHYG via the NcoI/BspHI and KpnI restriction sites [47].

For confocal studies C-terminally GFP-tagged SmFucTs were constructed in the pHYG plant expression vector. The SmFucT sequences were reamplified by PCR to introduce a GGGGS-linker and a NheI restriction site at the 3' end in order to clone the SmFucTs in frame with a C-terminal GFP fragment in pHYG (primers depicted in Supplemental Table 1). The GGGGS-linker was introduced to prevent cleavage and intervention of GFP with protein localisation. In the localisation studies free GFP (GFP) and secreted GFP (cSP-GFP), behind the *Arabidopsis thaliana* chitinase signal peptide, were taken along as controls.

For sub-Golgi localisation we used three mRFP Golgi reference markers: the CTS domain of β1,2-*N*-acetylglucosaminyltransferase I (GnTI) from *Nicotiana tabacum* fused to mRFP (GnTI-mRFP); β1,2-XylT from *A. thaliana* fused to mRFP (XylT-mRFP); and the CTS domain of α2,6-sialyltransferase from *Rattus norvegicus* fused to the catalytic domain of *N. tabacum* GnTI and mRFP (ST-mRFP) [23], [26], [48]. These mRFP-tagged Golgi reference markers were kindly provided by Dr Richard Strasser from the University of Natural Resources and Life Sciences, Vienna.

Unless stated differently, all constructs were driven by the Cauliflower mosaic virus 35S promoter with duplicated enhancer (d35S) and the *Agrobacterium tumefaciens* nopaline synthase transcription terminator (Tnos) [47]. Furthermore, to boost translation a 5' leader sequence of the alfalfa mosaic virus RNA 4 (AlMV) was included between the promoter and the construct. To enhance expression, the P19 silencing suppressor from tomato bushy stunt virus pBIN61 was used in all experiments, unless stated differently [49]. For plant expression, all constructs were transformed into *A. tumefaciens* strain MOG101 except GnTI-mRFP, XylT-mRFP and ST-mRFP that were transformed into GV3101.

Protein expression

For subcellular localisation studies and protein extraction, *A. tumefaciens* clones were cultured at 28 °C/250 rpm in LB medium (10 g/L peptone140, 10 g/L NaCl, 5 g/L yeast, pH 7.0), with 20 µM acetosyringone and 50 µg/mL kanamycin. After 16 hrs, bacteria were centrifuged for 15 min/2880 x*g* and subsequently resuspended in MMA (1.95 g/L MES, 20

g/L sucrose, 5 g/L MS-salts, pH 5.6) with 200 µM acetosyringone to increase *A. tumefaciens* transformation efficiency. *A. tumefaciens* cultures of different constructs were mixed for co-expression. The final optical density (OD) of each *A. tumefaciens* culture in the mixture varied between 0.1 and 0.5, depending on the construct. The two youngest fully expanded leaves of four to six week old wild type *N. benthamiana* plants were infiltrated at the abaxial side. *N. benthamiana* plants were grown in a controlled greenhouse compartment (UNIFARM, Wageningen). Three dpi leaves were harvested for confocal microscopy or protein extraction.

For sub-Golgi localisation studies, *A. tumefaciens* clones were cultured at 28 °C/250 rpm in LB medium (10 g/L peptone140, 10 g/L NaCl, 5 g/L yeast, pH 7.0), with 20 µM acetosyringone and 50 µg/mL kanamycin for MOG101 or 100 µg/mL spectinomycin for GV3101. After o/n incubation 2 mL bacteria culture was centrifuged for 5 min/4000 rpm. Subsequently, the pellet was washed twice in 1 mL infiltration medium (50 mM MES, 2 mM $Na_3PO_4 \cdot 12H_2O$, 0.1 mM acetosyringone and 5 mg/ml glucose) and resuspended in infiltration medium. *A. tumefaciens* cultures of different constructs were mixed for co-expression. The final OD of each *A. tumefaciens* culture in the mixture varied between 0.1 and 0.5, depending on the construct. Fully expanded leaves of five to six week old wild type *N. benthamiana* plants were infiltrated at the abaxial side. *N. benthamiana* plants were grown in a controlled greenhouse compartment and after infiltration grown in incubators at 21 °C, 14 hrs light, 10 hrs dark. Three dpi leaf epidermal cells were studied with confocal microscopy.

Protein extraction and visualisation on SDS-PAGE and western blot

In order to confirm GFP-tagged SmFucT expression, *N. benthamiana* leaves infiltrated with GFP-tagged SmFucTs were harvested three dpi. Subsequently, proteins were extracted via crude extraction. 8 leaf discs per infiltrated leaf were collected in 2 mL tubes, snap frozen and homogenised with 2 mL/g fresh weight of ice-cold extraction buffer (0.5 M PBS, 0.1 M NaCl, 2 % w/v PVP and 0.1 % v/v Triton 100, pH 6) using a Mixer Mill MM 400 for 2 times 30 sec with 30 r/sec. Crude extracts were clarified by centrifugation for 5 min/16.000 ×g at 4 °C. Total protein content was analysed using the Pierce Bicinchoninic Acid Protein Assay (BCA, Fisher Scientific). Total soluble plant proteins were separated under reducing conditions by SDS-PAGE on a 12% Bis-Tris gel (Invitrogen) and subsequently transferred to a PVDF membrane by semi dry blotting. After blotting the membrane was blocked with 5 % w/v milk powder in PBST (PBS containing 0.1 % v/v Tween-20). Subsequent the membrane was incubated with 1:2000 diluted anti-GFP-HRP antibodies (Miltenyi Biotec). The anti-HRP labelled antibodies were detected with a 1:1 SuperSignal West Femto:Dura substrate (Fisher Scientific) in the G:BOX Chemi System (Syngene).

2

Confocal imaging

To determine SmFucT subcellular and sub-Golgi localisation, *N. benthamiana* leaves infiltrated with GFP-tagged SmFucTs and mRFP-tagged Golgi reference markers for sub-Golgi localisation were harvested three dpi. To observe fluorescent protein expression and subsequent subcellular localisation, pictures were acquired with the 63x/1.4 Oil DIC objective on a Zeiss LSM 510 laser scanning microscope (Zeiss) using the ZEN-2012 software with scanning speed 9 and average of 16. GFP was excited with a 488 nm argon laser line and detected between 505-530 nm.

To observe fluorescent protein expression and potential co-localisation with mRFP-tagged Golgi reference markers for sub-Golgi localisation studies, high resolution movies were acquired with the Zeiss time series software on a Zeiss LSM 880 confocal microscope with Airyscan detectors (Zeiss) using a 100x 1.46NA lens. Images were acquired with 8x digital zoom, dual emission filters 500-550BP and 565LP, chlorophyll auto fluorescence blocking filter 620SP and simultaneous GFP-RFP laser excitation at 488 nm and 561 nm, respectively. In the Zeiss software, Airyscan acquired movies were processed to one picture per frame. From these movies with integration time ~0.8 sec, frames with side-on view Golgi bodies were selected for analysis with ImageJ plugin Coloc 2. Per combination the Pearson's correlation coefficient was determined of minimal 16 side-on view Golgi bodies. The Pearson correlation coefficient of analysed Golgi bodies was visualized in R (v 3.5.0; win x64) using ggplot [50]. To determine which variables were capturing variance in the data, an ANOVA was performed using R. The model used was Yi ~ SmFucT + marker + ε. Where Y is the calculated Pearson correlation coefficient of the analysed Golgi body i (1, 2, ..., 879), which was explained over GFP tagged-SmFucT (A, B, C, D, E, F, H, J, K or L), mRFP-tagged Golgi reference marker (either GnTI, XylT or ST), and error term ε. Subsequently, a post-hoc analysis was performed using a Tukey's HSD test. The significances reported were corrected for multiple testing.

For graphical display, fluorescent intensity profiles were generated across side-on view Golgi stacks using the Profile tool in the ZEN Blue software by drawing a line across dual-labelled Golgi stacks.
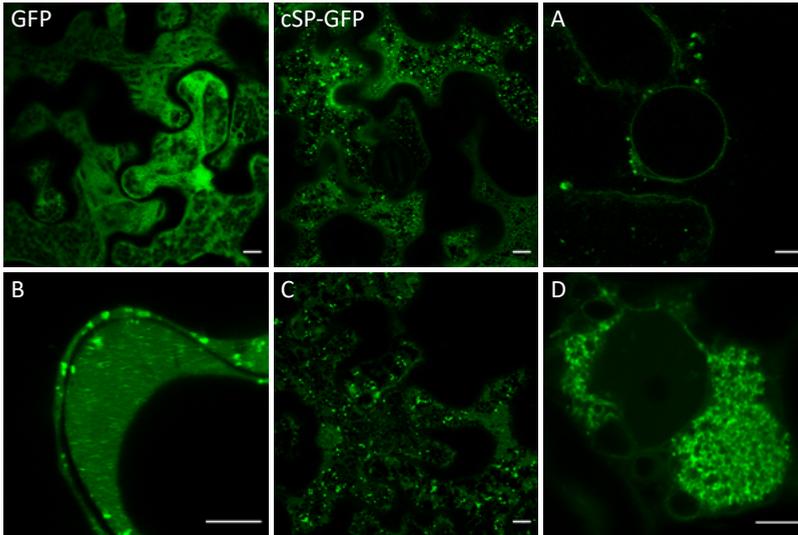
Acknowledgements

References

[1]   C. H. Hokke and A. van Diepen, "Helminth glycomics – glycan repertoires and host-parasite interactions," *Mol. Biochem. Parasitol.*, vol. 215, pp. 47–57, 2017.

[2]   C. H. Smit *et al.*, "Glycomic analysis of life stages of the human parasite Schistosoma mansoni reveals developmental expression profiles of functional and antigenic glycan motifs.," *Mol. Cell. Proteomics*, vol. 14, no. 7, pp. 1750–1769, 2015.

[3]   N. A. Peterson, C. H. Hokke, A. M. Deelder, and T. P. Yoshino, "Glycotope analysis in miracidia and primary sporocysts of Schistosoma mansoni: Differential expression during the miracidium-to-sporocyst transformation," *Int. J. Parasitol.*, vol. 39, no. 12, pp. 1331–1344, 2009.

[4]   F. Trottein *et al.*, "Molecular cloning of a putative α3-fucosyltransferase from Schistosoma mansoni," *Molecular and Biochemical Parasitology*, vol. 107, no. 2. pp. 279–287, 2000.

[5]   M. Berriman *et al.*, "The genome of the blood fluke Schistosoma mansoni," *Nature*, vol. 460, no. 7253, pp. 352–358, 2009.

[6]   J. M. Fitzpatrick *et al.*, "Anti-schistosomal intervention targets identified by lifecycle transcriptomic analyses," *PLoS Negl. Trop. Dis.*, vol. 3, no. 11, p. e543, 2009.

[7]   S. J. Parker-Manuel, A. C. Ivens, G. P. Dillon, and R. A. Wilson, "Gene expression patterns in larval Schistosoma mansoni associated with infection of the mammalian host," *PLoS Negl. Trop. Dis.*, vol. 5, no. 8, p. e1274, 2011.

[8]   A. V. Protasio *et al.*, "A systematically improved high quality genome and transcriptome of the human blood fluke Schistosoma mansoni," *PLoS Negl. Trop. Dis.*, vol. 6, no. 1, p. e1455, 2012.

[9]   N. A. Peterson, T. K. Anderson, and T. P. Yoshino, "In Silico Analysis of the Fucosylation-Associated Genome of the Human Blood Fluke Schistosoma mansoni: Cloning and Characterization of the Fucosyltransferase Multigene Family," *PLoS One*, vol. 8, no. 5, pp. 1–15, 2013.

[10]   M. L. Mickum, T. Rojsajjakul, Y. Yu, and R. D. Cummings, "Schistosoma mansoni α1,3-fucosyltransferase-F generates the Lewis X antigen," *Glycobiology*, vol. 26, no. 3. Glycobiology, pp. 270–285, 2015.

[11]   K. Paschinger, E. Staudacher, U. Stemmer, G. Fabini, and I. B. H. Wilson, "Fucosyltransferase substrate specificity and the order of fucosylation in invertebrates," *Glycobiology*, vol. 15, no. 5, pp. 463–474, 2005.

[12]   C. E. Machamer, "Golgi retention signals: do membranes hold the key?," *Trends Cell Biol.*, vol. 1, no. 6, pp. 141–144, 1991.

[13]   T. Nilsson, P. Slusarewicz, M. H. Hoe, and G. Warren, "Kin recognition. A model for the retention of Golgi enzymes," *FEBS Lett.*, vol. 330, no. 1, pp. 1–4, 1993.

[14]   S. Munro, "An investigation of the role of transmembrane domains in Golgi protein retention.," *EMBO J.*, vol. 14, no. 19, pp. 4695–4704, 1995.

[15]   G. H. Patterson, K. Hirschberg, R. S. Polishchuk, D. Gerlich, R. D. Phair, and J. Lippincott-Schwartz, "Transport through the Golgi Apparatus by Rapid Partitioning within a Two-Phase Membrane System," *Cell*, vol. 133, no. 6, pp. 1055–1067, 2008.

[16]   A. S. Uliana, C. G. Giraudo, and H. J. F. Maccioni, "Cytoplasmic tails of SialT2 and GalNAcT impose their respective proximal and distal Golgi localization," *Traffic*, vol. 7, no. 5, pp. 604–612, 2006.

[17]   A. Petrosyan, M. F. Ali, and P. W. Cheng, "Keratin 1 plays a critical role in golgi localization of core 2 N-acetylglucosaminyltransferase M via interaction with its cytoplasmic tail," *J. Biol. Chem.*, vol. 290, no. 10, pp. 6256–6269, 2015.

[18]   V. L. Sousa, C. Brito, and J. Costa, "Deletion of the cytoplasmic domain of human α3/4 fucosyltransferase III causes the shift of the enzyme to early Golgi compartments," *Biochim. Biophys. Acta - Gen. Subj.*, vol. 1675, no. 1–3, pp. 95–104, 2004.

[19]   V. L. Sousa, C. Brito, T. Costa, J. Lanoix, T. Nilsson, and J. Costa, "Importance of Cys, Gln, and Tyr from the transmembrane domain of human α3/4 fucosyltransferase III for its localization and sorting in the Golgi of baby hamster kidney cells," *J. Biol. Chem.*, vol. 278, no. 9, pp. 7624–7629, 2003.

[20]   C. Saint-Jore-Dupas *et al.*, "Plant N-glycan processing enzymes employ different targeting mechanisms for their spatial arrangement along the secretory pathway," *Plant Cell*, vol. 18, no. 11, pp. 3182–3200, 2006.

2

[21] J. Schoberer and R. Strasser, "Sub-compartmental organization of golgi-resident N-glycan processing enzymes in plants," *Mol. Plant*, vol. 4, no. 2, pp. 220–228, 2011.

[22] J. Schoberer *et al.*, "The golgi localization of gnti requires a polar amino acid residue within its transmembrane domain," *Plant Physiol.*, vol. 180, no. 2, pp. 859–873, 2019.

[23] J. Schoberer *et al.*, "Arginine/lysine residues in the cytoplasmic tail promote ER export of plant glycosylation enzyme," *Traffic*, vol. 10, no. 1, pp. 101–115, 2009.

[24] J. Schoberer, J. Runions, H. Steinkellner, R. Strasser, C. Hawes, and A. Osterrieder, "Sequential depletion and acquisition of proteins during Golgi stack disassembly and reformation," *Traffic*, vol. 11, no. 11, pp. 1429–1444, 2010.

[25] J. Schoberer, E. Liebminger, S. W. Botchway, R. Strasser, and C. Hawes, "Time-resolved fluorescence imaging reveals differential interactions of N-glycan processing enzymes across the golgi stack in planta," *Plant Physiol.*, vol. 161, no. 4, pp. 1737–1754, 2013.

[26] J. Schoberer *et al.*, "The transmembrane domain of N -acetylglucosaminyltransferase i is the key determinant for its Golgi subcompartmentation," *Plant J.*, vol. 80, no. 5, pp. 809–822, 2014.

[27] F. J. Logan-Klumpler *et al.*, "GeneDB-an annotation database for pathogens," *Nucleic Acids Res.*, vol. 40, no. D1, pp. 98–108, 2012.

[28] A. Bateman, "UniProt: A worldwide hub of protein knowledge," *Nucleic Acids Res.*, vol. 47, no. D1, pp. D506–D515, 2019.

[29] M. L. Mickum, N. S. Prasanphanich, J. Heimburg-Molinaro, K. E. Leon, and R. D. Cummings, "Deciphering the glycogenome of schistosomes," *Front. Genet.*, vol. 5, p. 262, 2014.

[30] E. L. Sonnhammer, G. von Heijne, and A. Krogh, "A hidden Markov model for predicting transmembrane helices in protein sequences." *Int. Conf. Intell. Syst. Mol. Biol.*, vol. 6, pp. 175–182, 1998.

[31] P. Both *et al.*, "Distantly related plant and nematode core 1,3-fucosyltransferases display similar trends in structure-function relationships," *Glycobiology*, vol. 21, no. 11, pp. 1401–1415, 2011.

[32] R. H. P. Wilbers *et al.*, "Production and glyco-engineering of immunomodulatory helminth glycoproteins in plants," *Sci. Rep.*, vol. 7, p. 45910, 2017.

[33] P. Boevink, S. Santa Cruz, C. Hawes, N. Harris, and K. J. Oparka, "Virus-mediated delivery of the green fluorescent protein to the endoplasmic reticulum of plant cells," *Plant J.*, vol. 10, no. 5, pp. 935–941, 1996.

[34] E. Slootweg *et al.*, "Nucleocytoplasmic distribution is required for activation of resistance by the potato NB-LRR receptor Rx1 and is balanced by its functional domains," *Plant Cell*, vol. 22, no. 12, pp. 4195–4215, 2010.

[35] T. Merkle, "Nucleo-cytoplasmic partitioning of proteins in plants: Implications for the regulation of environmental and developmental signalling," *Curr. Genet.*, vol. 44, no. 5, pp. 231–260, 2003.

[36] C. Pain, V. Kriechbaumer, M. Kittelmann, C. Hawes, and M. Fricker, "Quantitative analysis of plant ER architecture and dynamics," *Nat. Commun.*, vol. 10, no. 1, pp. 1–15, 2019.

[37] J. F. Mckenna, S. E. D. Webb, V. Kriechbaumer, and C. Hawes, "Is Actin Filament Sliding Responsible for Endoplasmic Reticulum and Golgi Movement?," *Biorxiv*, p. 347187, 2018.

[38] P. Boevink, K. Oparka, S. S. Cruz, B. Martin, A. Betteridge, and C. Hawes, "Stacks on tracks: The plant Golgi apparatus traffics on an actin/ER network," *Plant J.*, vol. 15, no. 3, pp. 441–447, 1998.

[39] R. Strasser, J. Schoberer, C. Jin, J. Glössl, L. Mach, and H. Steinkellner, "Molecular cloning and characterization of Arabidopsis thaliana Golgi α-mannosidase II, a key enzyme in the formation of complex N-glycans in plants," *Plant J.*, vol. 45, no. 5, pp. 789–803, 2006.

[40] A. -C Fitchette-Lainé, V. Gomord, A. Chekkafi, and L. Faye, "Distribution of xylosylation and fucosylation in the plant Golgi apparatus," *Plant J.*, vol. 5, no. 5, pp. 673–682, 1994.

[41] P. Lerouge, M. Cabanes-Macheteau, C. Rayon, A. C. Fischette-Lainé, V. Gomord, and L. Faye, "N-glycoprotein biosynthesis in plants: Recent developments and future trends," *Plant Mol. Biol.*, vol. 38, no. 1–2, pp. 31–48, 1998.

[42] L. A. Staehelin and B. H. Kang, "Nanoscale architecture of endoplasmic reticulum export sites and of Golgi membranes as determined by electron tomography," *Plant Physiol.*, vol. 147, no. 4, pp. 1454–1468, 2008.

[43] C. Hawes, J. Schoberer, E. Hummel, and A. Osterrieder, "Biogenesis of the plant Golgi apparatus," *Biochem. Soc. Trans.*, vol. 38, no. 3, pp. 761–767, 2010.

[44] I. A. Sparkes, T. Ketelaar, N. C. A. De Ruijter, and C. Hawes, "Grab a golgi: Laser trapping of golgi bodies reveals in vivo interactions with the endoplasmic reticulum," *Traffic*, vol. 10, no. 5, pp. 567–571, 2009.

[45] K. W. Dunn, M. M. Kamocka, and J. H. McDonald, "A practical guide to evaluating colocalization in biological microscopy," *Am. J. Physiol. - Cell Physiol.*, vol. 300, no. 4, pp. C723–C742, 2011.

[46] L. B. Westerhof, "Plant as a production platform for high-value proteins," 2014.

[47] L. B. Westerhof *et al.*, "3D Domain Swapping Causes Extensive Multimerisation of Human Interleukin-10 When Expressed In Planta," *PLoS One*, vol. 7, no. 10, pp. 1–10, 2012.

[48] L. Renna, S. L. Hanton, G. Stefano, L. Bortolotti, V. Misra, and F. Brandizzi, "Identification and characterization of AtCASP, a plant transmembrane Golgi matrix protein," *Plant Mol. Biol.*, vol. 58, no. 1, pp. 109–122, 2005.

[49] O. Voinnet, S. Rivas, P. Mestre, and D. Baulcombe, "An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus," *Plant J.*, vol. 33, no. 5, pp. 949–956, 2003.

[50] H. Wickham, "Ggplot2: Elegant Graphics for Data Analysis," *Springer-Verslag*, 2016.

2

Supplemental Figures and Tables

SUPPLEMENTAL TABLE 1 | Primers. Primers used for *Schistosoma mansoni* fucosyltransferase (SmFucT) reamplification. The part of the sequences that corresponds to the original SmFucT sequence is written in capital letters.

Primer name	Sequence 5' to 3'
SmFucTA - *Forward*	caccCCATGGCCAACATCCCTTG
SmFucTA-GGGGS-Nhe - *Reverse*	tgctagcggaaccaccaccaccGTTCATCTGAGAGTTAGTAGAAGTAGTAC
SmFucTB - *Forward*	caccTCATGAAGCGTACTCGTCAC
SmFucTB-GGGGS-Nhe - *Reverse*	tgctagcggaaccaccaccaccGTGGATGAAGTTCCAACGG
SmFucTC - *Forward*	caccCCATGGCCATGTTCTACC
SmFucTC-GGGGS-Nhe - *Reverse*	tgctagcggaaccaccaccaccGATAGTAGGGAACTTACGGTTC
SmFucTD - *Forward*	caccCCATGGAGCCTCTCCGTG
SmFucTD-GGGGS-Nhe - *Reverse*	tgctagcggaaccaccaccaccAACAGCCTTGTTCCAACGG
SmFucTE - *Forward*	caccTCATGATTGTTCAGCGTGGTAAG
SmFucTE-GGGGS-Nhe - *Reverse*	tgctagcggaaccaccaccaccAGGGTTCCAACGGAGCTTAC
SmFucTF - *Forward*	caccCCATGGCCATGTGCGAGG
SmFucTF-GGGGS-Nhe - *Reverse*	tgctagcggaaccaccaccaccAGGGTTCCAACGGAGCTTAC
SmFucTH - *Forward*	caccCCATGGTTGTTACTTCTTTCAAG
SmFucTH-GGGGS-Nhe - *Reverse*	tgctagcggaaccaccaccaccCTTGTAGTCGAAAGTGAACTC
SmFucTJ - *Forward*	caccCCATGGCCATGCGTTACTC
SmFucTJ-GGGGS-Nhe - *Reverse*	tgctagcggaaccaccaccaccGTTAGACCAAGTGATGTTGATACG
SmFucTK - *Forward*	caccCCATGGCCATGCGTTACTC
SmFucTK-GGGGS-Nhe - *Reverse*	tgctagcggaaccaccaccaccGTTGAACCAAGTCATGTTGATAGTAG
SmFucTL - *Forward*	caccCCATGGCCATGCGTTACCTC
SmFucTL-GGGGS-Nhe - *Reverse*	tgctagcggaaccaccaccaccAGTGATGTTGATACGACCACGG

**SUPPLEMENTAL FIGURE 1 | Non-Golgi cellular localisation of *Schistosoma mansoni* fucosyltransferases (SmFucTs).** In addition to Golgi localisation, localisation in different cellular organelles was observed, in the nuclear envelope (A), in the cytoplasm (GFP and B), in network-like structures (cSP-GFP and C) and in or around the nucleus (D). The scale bar indicates 10 μm (GFP, cSP-GFP and C) or 5 μm (A, B and D).
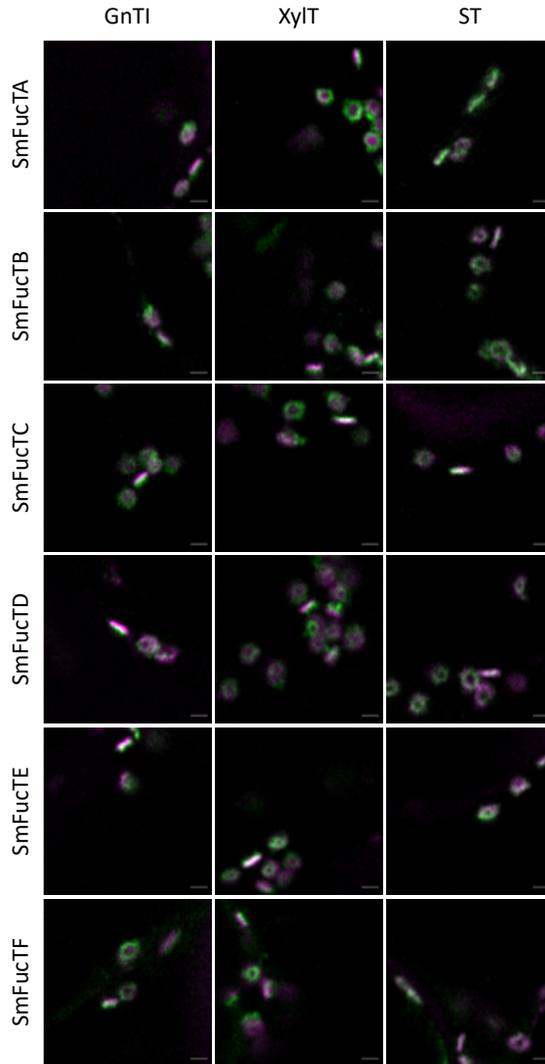
2

SUPPLEMENTAL FIGURE 2 | Sub-Golgi localisation of *Schistosoma mansoni* fucosyltransferases (SmFucTs). GFP-tagged SmFucTs and mRFP-tagged Golgi markers GnTI, XylT and ST were co-expressed in *Nicotiana benthamiana* leaves and co-localisation was analysed three days post infiltration with confocal microscopy. Merged representative pictures of each SmFucT (in green) with the three reference markers (in violet) are depicted. Co-localisation is seen in white. The scale bar indicates 1 μm.
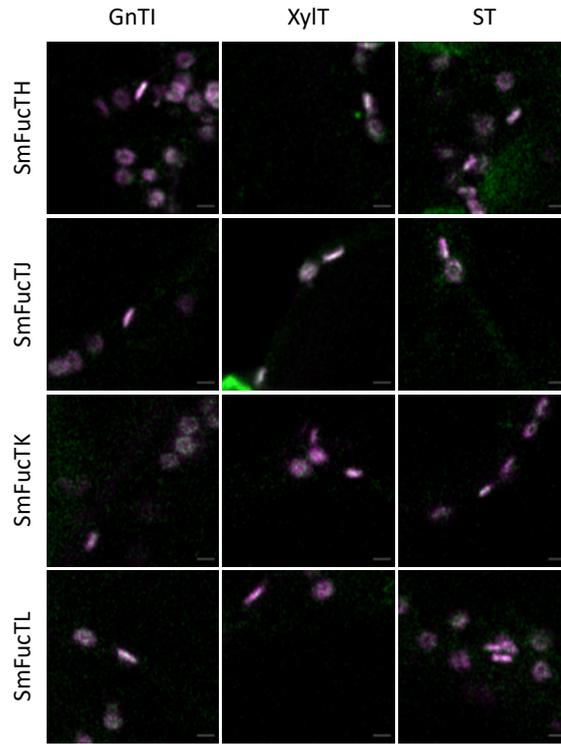
SUPPLEMENTAL FIGURE 2 | Sub-Golgi localisation of *Schistosoma mansoni* fucosyltransferases (SmFucTs).
Continued
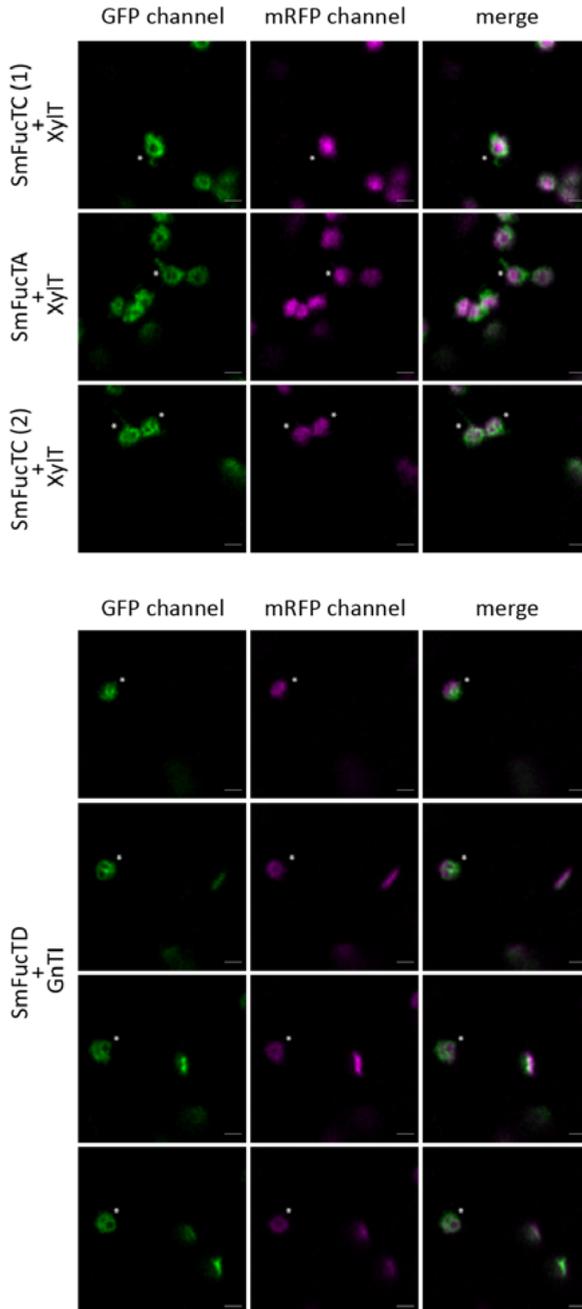
SUPPLEMENTAL FIGURE 3 | Two ring structures and tubule-like structures. SmFucTA + XylT and SmFucTC (1) + XylT exemplify the "tubule-like" structures (indicated by *) observed during sub-Golgi localisation studies with GFP-tagged SmFucTs and mRFP-tagged Golgi markers GnTI, XylT and ST. SmFucTC (2) + XylT and SmFucTD + GnTI exemplify the "two ring" structures (indicated by *) observed during sub-Golgi localisation studies. The four frames for the combination of SmFucTD + GnTI were selected from the same movie. The scale bar indicates 1 μm.

SmFucTA

ATGGCCAACATCCCTTGGATCCAGTTCAACCACCACTTCATCTACCGTCAGCGTTCTTACTACTACATCATCCTCTCTATCCTCCTCCT
CCCTTTCATCTTCACTTTCATCCAGAAGTACCACCTCTACAAGTACAACGAGATCGACGAGGTTTACAACTCTTACTCTACTTTCCAC
GAGCTTAAGAAGGACTTCGTTAACGGTAAGTTCAAGCTCAACACTGAGAAGATCAAGGAGATCGTTATCTACGGTAAGACTACT
GTTCCTATCAACACTGACCTCCGTGGTTGCCTCGTTTCTAACTGCGTTCTCCACAAGACTAAGTGGCGTTCTTACTACGCTGAGATG
GTTATCATCACTAACGGTAACTTCCCTAAGGTTAAGAACCACCAGAACCAGGCTTGGCTCGCTTACCACTACGAGGCTCCTCCT
CACTCTCGTCTCAACAAGCTCCTCGGTGAGAAGATCAACTTCACTGCTACTTACCGTCTCGACTCTACTATCCACCTCCCT
TACGGTAAGGTTATCGAGAACCTCGAGTACCACCAGAAGCTCAACAACAACGACGACTCTTCTCGTTCTTCTATCTCTACT
GACAACTTCTTCAAGCCTCGTAACATCTCTGCTCTCAAGAAGAAGACTGTTGCTTGGATCGTTTCTAACTGCCACCCTGACTCTCCTC
GTAACCTCTACGCTTACGAGCTTTCTAAGTACATCACTGTTGACGTTTACGGTAAGTGCTCTCAGCGTGAGTGCCAGGACTC
TAAGTGCCACAAGATGCTCAAGGAGCAGTACAAGTTCTACCTCTCTTTCGAGAACTCTCTCTGCCAGGACTACATCACTGAGA
AGTTCTTCAAGAACGCTCTCATGAACGACGTTATCCCTGTTGTTATGGGTGCTTCTATCGAGGAGTACAAGTCTGTTGCTCCTCCTA
ACTCTTTCATCCACGTTGACCAGTTCTCTTCTCCTCGTCAGCTCGCTAAGTACCTCCACTACCTCGACAACAACCACACTGCTTTCAAC
GAGTACTTCATCTGGCAGAACAAGTGGAAGGTTCTCCTCTTCCCTGAGCGTCCTGAGTGCGACTTCTGCCTCCTCGCTAAC
GCTCTCCCTAACCTCAAGCCTTCTTGGTACATCGACGTTAACTCTTGGTTCGACCGTTCTTGCCACGGTCGTCGTCTCAAGTGGAAGGG
TAACCAGAAGGACTTCTCTGCTGCTATCTGGTTCTCTAAGCTCAAGCAGCAGAACCCTGTTACTACTTCTACTAACTCTCAGATGAACTAA

SmFucTB

ATGAAGCGTACTCGTCACCTCGTTGTTCGTGTTGTTCAGCTCGCTGTTATCATCTTCATCTCTTACGAGCTTACTCGTCTCCTCTTCCAG
TACAAGTGGGACGAGAAGTTCGGTATCTCTCGTAACAACGGTGACTTCTTCTACAACATCCTCCTCCACGAGTCTTACTCTTCTCTCACTG
GTAACCTCAACCAGACTTACTACGTTTCTTCTCAGATGGAGAAGTACGAGAACAAGATCACTCTCGGTTTCACTGAGAAGGACCTC
GAGGAGCTTTACAAGGTTAAGTCTGCTCCTGCTCTGCTGCTGTAAGATCATCCTCAACTACGGTGACCTCTACATCAACATCATCCGTA
ACTTCTCTTTCTGCTTCGCTTGCGACTGCGAGCTTATCTTCGACCGTTCTCGTTGGCTCGAGGCTGACGTTATCATCCTCACTGACCACCTC
TACCCTAAGGGTCCTCGTCCTCCTAACCAGCTCTGGTTCGTTTTCGTTCACGTTCCTCCTCTCAAGATCCGTATCGCTGACGGTCTC
GAGAACAAGATCAACTACACTATCACTTACCGTATGGACTCTACTATCTACCTCCCTTACCACAACTACGTTCCTTCTGTTGCTTCTCACG
GTCCTGACACTAAGTACAAGCTCCCTTCTCGTAACTACGCTACTGGTAAGTCTAAGATGGTTGCTTGGTTCGTTTCTAACTGCCAGCCTA
AGTCTCCTCGTATGATCTACGTTCGTGAGCTTTCTCGTTACATCCAGATCGACATCTACGGTACTTGCGGTACTCTCACTTGCCCTCGT
GAGGTTGAGGACCAGTGCTTCTCTCTCCTCGGTAAGCACTACAAGTTCTACCTCTCTTTCGAGAACTCTCTCTGCGGTGACTACATCACT
GAGAAGCTCTTCCGTAACGCTTTCTCTCAGAACGTTATCCCTGTTGTTATGGGTGCTTCTATCGAGGAGTGCAAGCGTGTTGCTCCTCCT
CACTCTTTCATCCACGTTGACCAGTTCGAGTCTCCTGCTAAGCTCGCTGACTACCTCAAGTACCTCGACCGTAACGAGACTGCTTACAAC
GAGTACTTCGCTTGGCACGAGCACGGTACTATCGACGCTTGGTCTCCTATGCCTCAGTGCGTTATCTGCCTCCTCGCTCACTGCTCA
CAAGCTCAAGCCTTACACTTTCCCTAACGTTTCTAAGTGGTGGAACGACGGTTGCGTTGGTCGTAAGATCCGTTGGAACTTCATCCACTAA

SmFucTC

ATGTTCTACCCTCGTACTTACATCAACCGTAAGCGTCGTTCTCTCAAGAAGATCACTTTCGCTTTCATCCTCTTCCTCATCATCTTCCTCAT
CATCAACTACTCTAAGAAGTCTTTCCTCTCTGTTAACCTCACTGACATCAACAACTACAACCTCATCCGTAAGTACGCTCAGTCTTTCTCTG
GTCACGAGAACTCTATCTCTGACATCGACTACGTTACTCGTTCTACTGAGAAGCGTAAGAACGTTATCATCAACTCTAAGCGTCACCTC
GGTTCTATCGACACTAAGCTCTTCGACGAGCAGGACTTCATCCTCAAGAAGAAGTTCAACATGAACAACCTCATCCTCTTCCCTTTC
GAGACTGACAACCCTTACGAGGACGACCGTATCCTCGCTCAGATGTCTTACGTTCCTGTTCAGGTTCTCGAGGCTCGTCGTAACGGT
CAGATCCTCATGAAGCGTATCTACTACCACGGTAAGTTCTCTAACTCTCTCCCTGACCTCTCTAAGTGCCCTGTTTTACGAGTGCAAC
GTTTCTGGTAACGGTGACAACATCCAGCAGGCTGACCTCGTTGTTTTCGAGTCTAACCCTGGTCTCACTGACCTCAACAAGCCTC
GTAACCAGTCTTGGCTCATCTACCACATCGAGTCTCCTCGTCACTACTCTTTCGCTATGCCTAAGAACCTCATCAACTTCACTGC
TACTTACGCTGTTGACTCTACTATCGTTACTCCTTACTACAAGTACGTTCGTTTCTCTGACGTTCAGCCTCGTTCTCTCAACAACCT
CATCCAGCCTGACAAGGACTTCTCTATCGGTAAGACTAAGATGATCGCTTGGTTCGTTTCTAACTGCGTTGGTTCTTCTCCTCGTAT
GCAGTACGCTAAGGAGCTTATGAAGTACATCAAGGTTGACATGTACGGTGGTTGCGGTAACCTCAACTGCGACCGTCACGACTC
TAAGTGCTTCGAGCTTCTCAAGAACGACTACAAGTTCTACCTCTCTTTCGAGAACTCTATCTGCCAGGACTACATCACTGAGA
AGTTCTTCCAACGCTCTCAAGAACAACGTTGTTCCTATCGTTATGGGTGGTTCTAAGTCTGACTACGTTCGTTCTGCTCCTTCTA
ACTCTTTCATCCACGTTGACGACTTCACTTCTGTTAAGCAGCTCTCTGACTACCTCTACTACCTCGACCGTAACAAGACTGCTTACAAC
GAGTACTTCAAGTGGCACAACCACGGTACTGTTGACTTCAACACTTTCTCTGACTGCCGTCTCTGCATGCTCGCTCACGAGATCGA
CAACATGGAGCGTTCTTACTGGTACGAGGACGTTGACCAGTGGCGTATGTCTTCTTGCGAGAACCGTAAGTTCCCTACTATCTAA

SmFucTD

ATGGAGCCTCTCCGTGACTTCATCCGTATCTCTTCTGTTGTTCGTCTCAAGGTTATCATCATCCTCGCTGTTTGCACTGCTTCTTTCG
GTGTTTGGTACATCTGGGGTCTCTCTCTCAACGAGATCGGTCAGTACATCTCTAACTACCAGAACCTCGCTATCTGGAACCTCGTTC
GTCAGCCTGTTTACTACCCTAACTTCAAGGACTTCCGTAACGACGTTATGATCGGTCGTCTCCTCCTCAACAAGACTAAGAAG
TACGAGATCATCATCTACGGTCAGATGCTCGCTGGTATGAACAAGGACGTTACTACTTGCCCTGTTAACAACTGCGTTATCCA
CACTGACACTACTCGTTGGATCAACTCTGACCTCATCCTCATCCCTAACCGTCTCTTCCCTTCTGGTAAGCGTCCTCACCAGCAG
GCTTGGGTTGCTTTCGAGTACGAGTCTGCTCTCCACACTCGTTTCTCTGACGAGCTTAACGACAAGATCAACTTCACTGCTTCT
TACCGTTTCGACTCTACTATCCGTACTCCTTACGGTATGTACACTCCTGACGAGCCTAAGACTGACATCAACAAGACTATCCAGCT

CATCAAGCTCGAGGACATCGCTAAGGGTAAGGACCGTGCTGTTGCTTGGATCGTTTCTAACTGCAACCCTAAGTCTCCTCGTAAC
GCTTACGCTGACGAGCTTTCTAAGTACATCACTGTTGACGTTTACGGTCGTTGCGGTCGTATGACTTGCTACGGTTCTCAGTGCCTC
GACCTCGTTAAGCGTCACTACAAGTTCTACCTCTCTTTCGAGAACTCTCTCTGCCAGGACTACATCACTGAGAAGTTCTTCCTCAAC
GCTCTCATGAACAACGCTCTCCCTATCGTTATGGGTGCTTCTGTTGAGGAGTACCACAAGGTTGCTCCTCCTCACTCTTTCATC
CACGTTGACCAGTTCGAGAACCCTAAGGAGCTTGCTAAGTACCTCAAGTACCTCGACAAGAACGACACTGCTTACAACGAG
TACTTCACTTGGCACATCAAGGGTGTTGTTACTGTTTGGTCTTTCAAGCCTGAGTGCGAGTTCTGCATCCTCGCTAACGCTCTCCCT
TACTTCAAGCCTACTATCCACGAGAACTTCATGTTCTGGTGGAAGAACGGTTGCAAGAACCGTACTCTCCGTTGGAACAAGGCTGTTTAA

SmFucTE
ATGATTGTTCAGCGTGGTAAGTGGACTCTCCGTATCAAGGCTATCGGTTTCGTTCTCATCGCTCTCTTCATCGTTTGGCTCATCAA
CATCTCTCCTTCTCTCCTCTACGTTGGTATCCACAACGAGTCTTACACTTACCTCTCTGACAACATGGGTCTCATCCACAAGGA
CATCTTCGACTCTAACTCTTCTGAGCTTGAGCTTGCTATGATCCTCCACAAGTCTTACTCTACTATCTACGAGACTTCTGTTCAGAAG
TACTCTCTCCGTTTCGACGTTGAGGAGAACCGTCTCAAGCGTCTCTTCGACTCTCTCTCTAAGAACACTTCTGTTAAGACTTTCCT
CAACTACGGTTCTCAGAAGATCTACACTCCTATGGACTTCTCTAAGTGCTACGCTTCTAAGTGCACTGTTATCACTGAGATGGAGC
GTTGGCACGAGGCTGACGCTCTCATCCTCACTAAGAACGTTAAGCCTGACGGTATCCGTCCTCCGGTCAGCTCTGGTTCACTCTCATC
CACGAGTCTCCTGAGCACATCTCTCTCGCTAACACTCTCGACAACGAGGTTAACTTCACTATCTCTTTCCGTCTCGACTCTACTATC
TACTCTCCTTACGGTTACTACGAGCCTCACTCTAAGTACCACGAGCCTAAGACTCGTTACCCTCTCCCTTCTCGTAACTACGCTGCTG
GTAAGTCTAAGCAGGTTGCTTGGTTCGTTTCTAACTGCGGTGCTAAGTCTCCTCGTCTCCAGTACGCTAACGAGCTTTCTCGTTA
CATCTCTGTTGACATCTACGGTAAGTGCGGTACTAAGATCTGCCCTCGTAACCTCGTTAACGACCCTCTCCTCCCTGACTGCTTCAAGCT
CATCCGTGAGAACTACAAGTTCTACCTCTCTTTCGAGAACTCTCTCTGCCGTTGGTACATCACTGAGAAGCTCTTCAAGAAC
GCTCTCAAGAACGACATCCTCCCTATCGTTATGGGTGCTTCTATCGAGGAGTACGAGCGTGTTGCTCCTCCTTACTCTTTCATCCAC
GTTGACCAGTTCGAGTCTCCTGCTAAGCTCGCTGACTACCTCAAGTACCTCGACAAGAACGACACTGCTTACAACGAGTACTTC
GCTTGGCACGGTCACGGTATCATCCACGACTGGGACTCTCAGCCTCAGTGCGCTATGTGCCTCCTCGCTCACACTTCTCAG
GCTTTCGGTCCTTACTGGGTTCCTAAGGTTGCTCGTTGGTGGAACGACGGTTGCAACGGTCGTAAGTCCGTTGGAACCCTTAA

SmFucTF
ATGTGCGAGGTTTTCGGTATCGTTGTTCGTAACGACAAGTACTCTCTCAAGGTTAAGCTCCTCGTTTCTGTTTCTATCATCCTCCT
CATCATCTGGCTCGTTAACGTTGTTCCTTCTCTCATCCAGTGGAAGATCAACCAGGTTGACCACGACCGTATGCCTGAGACTATC
GAGGTTCTCTGCTCTAACATCTCTATCACTAACCAGTCTGAGTCTGACCTCGCTTCTATCCTCCACAAGATCTACTCTCCTATCTC
TAAGACTTCTGTTCACAAGTACCCTCTCCGTTTCTCTCTCGAGGAGAACCGTCTCGAGAAGCTCTACGACTCTCTCAACATGAA
CAAGACTGTTAAGACTTTCCTCAACTACGGTCACCGTACTATCTTCACTCCTATGGACTTCTCTAAGTGCCACGCTTCTAAGTGCGCT
GTTATCACTGACATGAACCGTTGGCGTGAGGCTGACGCTCTCATCCTCACTGAGGACAAGGTTCCTAAGGGTACTCGTCCTCCTGAG
CAGCTCTGGTTCTCTCTCATCCAGGAGTCTCCTGTTCACATCGCTATGGCTGGTTTCCTGAGAACGAGATCAACTTCACTATCTCTTTC
CGTCTCGACTCTACTATCTACTCTCCTTACGGTTCTTACGAGCCTTACATCAAGCACCACGGTCCTGAGACTCGTTACCCTCTCCCTTCTC
GTAACATCGCTTCTGGTAAGTCTAAGAAGGTTGCTTGGTTCGTTTCTAACTGCATCCCTAAGTCTCCTCGTATGCAGTACGCTA
AGGAGCTTTCTCGTCACATCTCTGTTGACATCTACGGTGAGTGCGGTACTCTCTCTTGCCCTAAGTACATCCCTACTTTCACTTCTTCT
GTTGGTTGCCTCAAGATGATGCGTGAGAACTACAAGTTCTACCTCTCTTTCGAGAACTCTCTCTGCTCTGAGTACGTTACTGAGA
AGCTCTACAAGAACGCTCTCAACAACGAGCTTCTCCCTATCGTTATGGGTGCTTCTATCGAGGAGTACGAGCGTGTTGCTCCTCCT
TACTCTTTCATCCACGTTGACCAGTTCGAGTCTCCTGCTAAGCTCGCTGACTACCTCAAGTACCTCGACAAGAACGACACTGCTTACAAC
GAGTACTTCGCTTGGCACGGTCACGGTATCATCCACGACTACGACGCTCAGCCTCAGTGCGCTATGTGCCTCCTCGCTCACACTTCT
CAGGCTTTCGGTCCTTACTGGGTTCCTAAGGTTGCTCGTTGGTGGAACGACGGTTGCAACGGTCGTAAGCTCCGTTGGAACCCTTAA

SmFucTH
ATGGTTGTTACTTCTTTCAAGAACCTCTTCCACGTTCGTCCTCTCATCACTACTGCTGTTGCTCTCTGGATCTGCGTTATCATC
TACCTCTCTTTCAACTTCTTCAACCTCGAGTCTTCTGACCAGCAGTCTTCTGCTTGCCAGTCTCTCCTCGACATCATCACTGGTCT
CAAGACTCAGGTTGACATCGGTTCTGCTAACCTCCGTATCTTCCTCACTTCTTACAACGCTAAGAACAAGGAGGAGAACACTATC
GTTGACCGTGTTGGTATCCACGGTATGTCTTACGAGATCCTCCGTCGTCGTGCTCTCCAGTACGCTCGTGAGATCAAGTACTCTCTC
GTTCAGCTCGACTCTGTTAACGAGAAGATCAAGACTATCAAGAACTACTCTCTCATCCGTCCTCTCAAGGAGGTTAAGGTTTCT
GTTGAGCAGCTCAAGTCTCAGCTCCAGGACATCTCTCTCCACCTCCACGTTGACATCGACGGTATCGGTCGTGTTGACGGTGTTCTC
GAGTCTCGTATCAACTACCTCGAGTACCTCTCTAACGAGCTTCAGTCTGAGCTTTTCCAGCTCCAGAACCCTTCTTCTTGCACTA
AGGCTAAGTACGTTGTTGCTTCTCTCAACCGTCCTTGCGCTTTCGGTTGCAACGCTCACCACCTCATGCACTGCTTCCAGATGGCT
TACGCTACTGGTCGTACTCTCATCCTCGACCCTACTGACGGTGAGGGTTACACTCAGTGGTGGATGAAGAACTTCGTTCCTCTCTCT
CAGAAGTGCTCTATCAACGACATCCAGTCTAACATCCACTCTGACTTCTCTGCTGGTGAGGCTTCTAACATCTACCAGGCTATCGAGT
GCCCTCGTATCGACGCTGTTTCTTCTTCTTTCGACTGGGTTCCTCCTGCTGTTCCTTCTCACCTCTCTAAGATCCTCACTCGTCTCCAC
GGTGCTCCTTTCGTTTGGTTCATCGGTCAGCTCGGTAAGTTCCTCATGCGTCCTACTTTCAACCTCACTGAGGAGTTCAAGATCTTC
GACAACCAGTCTGACAACCCTATCGTTGGTATCCACGTTCGTCGTACTGACAAGATCAACACTGAGGCTTCTTTCCACTCTCTCGAG
GAGTACATGACTGAGGTTGAGTCTTACTTCCAGTTCATCGACGCTAAGCGTCAGATGATGTCTCGTACTAAGGAGTGGCGTAAC
GACATCAAGTCTCCTTTCCGTCACAACGTTTACCACCAGCTCAAGCCTGTTAAGCGTCGTGTTTACATCGCTACTGACGACCCTAA
CATCTTCGACGAGGCTAAGTCTAAGTACCCTAACTACACTTTCTACGGTGACCGTGGTCGTGCTGACTCTGCTTCTCTCTTCCGTC

GTAAGAACGAGGACTCTATCATGGGTGTTGTTACTGACGTTTTCGCTCTCTCTCGTACTAACTACCTCGTTTGCACTTTCTCTTCT
CAGGTTTGCCGTCTCGCTTACGAGCTTATGCAGTCTAACCACCTCGAGCTTGGTGACGCTTCTCAGCAGTTCCGTTCTCTC
GACGACATCTACTACTTCGGTGGTCAGCAGTCTTCTCCTTACGAGACTATCATCTCTGACAAGGAGAACGACCTCTCTCCTGGT
GACCTCGTTCACTTCCACGGTAACCACTGGAACGGTTACGCTCGTGTTGAGAAGCTCAACACTAACCGTAAGGTTATGGCTCCT
GCTTTCAAGTTCTCTTCTCGTCTCCTCTCTGCTCCTATGATCGGTGCTCACGGTAACCGTTCTGAGTTCACTTTCGACTACAAGTAA

SmFucTJ

ATGCGTTACTCTCGTAAGGAGTTCAACTTCACTTTCGTTCTCTGGCTCCCTCTCCTCCCTGTTCTCATGCTCTGCATCGAGGACTACTTCCG
TAAGAAGACTTCTTCTAAGACTGAGCTTTGCTGCAAGGGTCTCATCAACAACGTTTCTGTTGCTAAGGAGGGTGTTGACTGGGTTCCTACT
CAGATCTCTCGTCTCCGTCGTCTCTTCGACTTCTCTGAGCGTATCTCTATGAAGACTACTCGTAACGTTCAGATCACTAACGAGAACCG
TATCGGTATCTCTCACGAGACTCTCCGTCGTCGTGCTCTCAACTACATCCACGAGGCTCAGAACACTCTCAACAACCTCAACTCTATCTC
TACTCAGCTCCGTGGTCTCAACCACAAGGACCTCGACACTAAGACTGTTTCTTCTATGCAGGCTTCTCTCATCCGTTACCGTCAGAAC
GTTGAGGACATCATCCTCCACCTCGAGATCATCCTCAACAAGCTCGGTCAGGTTGACGGTTTCTCTCAGATCCGTGTTATCGGTCTCAA
CAACCTCGCTCAGCGTCTCCAGAAGAAGATCCAGGAGCAGCAGAACCCTTCTGACTGCAAGATGGCTAAGTACGTTACTTTCGACAT
CACTAACATGTGCGGTTTCGGTTGCCTCATGCACCAGATCATGTTCTGCCTCCAGCTCGCTTTCGAGAAGGAGCGTGTTCTCATCCTCAC
TAACTACCGTTCTCGTGACATCTTCCACGAGTGGCTCCACCAGAACACTCTCCCTCTCTCTGAGAAGTGCTCTCACGAGGACACTAACG
GTCGTTCTGACAACATCGAGTGCCCTTACAACATCAACGACCACGTTGCTCACAACTGGCTCCCTCACGTTCTCCCTATCGACATGAAC
GAGGAGCTTCTCCGTCTCCACGAGGCTCCTTACGTTTGGTTCGCTGGTCAGCTCGCTGCTTACATCCTCCGTCCTCGTCCTAAGTTCTCT
GAGCTTATCAACGCTACTCTCGAGACTTTCAAGGCTTCTGTTGACCCTGTTGTTGGTGTTCACATCCGTCGTACTGACAAGATCATCCTC
GAGGCTAAGTTCCACAACCTCTCTGAGTACATGGAGCAGGTTGACAAGTTCTTCGACCGTCAGAACGCTAACCACCTCGTTATGTCTA
AGCGTCACGACCTCAACGAGGAGAAGCTCGGTCTCATCTCTGAGAAGTCTATCGGTCGTAAGAACGCTCTCCGTGCTAAGCGTTCT
GTTTACCTCGCTACTGACGACCCTAACGTTTTCACTCAGTTCTCTTACTCTCACCCTAACTACATCCTCTACGGTTCTCCTGGTCGTTCT
CACTCTGCTGACGTTGAGTTCCGTATGACTTACAACTCTATGAACAACGCTGTTATCGACGTTATCGCTCTCTCTATGACTGACTTCCTC
GTTTGCACTTTCTCTTCTAACATCTGCCGTCTCGCTTACGAGCTTATGCAGACTCGTCACATGGAGCTTGGTGACGCTACTCAGCTCTGC
CACTCTATCGACTTCGCTTTCCACGAGGAGGACTACAAGCGTCTCGAGTTCGACGTTATCATCCCTGACAAGAAGGGTCACCTCAAG
TACGGTGACATCGTTGACACTTTCAAGAACCACGGTAACGGTCTCGCTTCTGTTAAGCTCGACCAGTCTTCTAAGAAGAAGGAGAA
CAACAACCAGATGTACATGGAGAACGAGCTTTTCATCGTTCCTGCTTACAAGTTCCGTCCTCGTATCAACATCACTTGGTCTAACTAA

SmFucTK

ATGCGTTACTCTCGTAAGGAGTTCAACTTCTCTGTTGTTCTCTGGCTCCCTCTCCTCCCTGTTCTCATGCTCTGCATCGGTGACTACTTC
CGTAAGAAGACTTCTTCTAAGACTGACCTCTGCTGGAAGGACCTCATCAACAACGCTTCTGTTGCTAAGGAGGGTGTTGACTG
GTTCTCTACTCGTATCTCTCGTCTCCGTCGTCTCTTCGACTTCTCTGAGTCTATCTCTATGAAGACTATCAAGATGGACCAGATCACTAAC
GAGAACCGTATCGGTATCTCTCACGAGACTCTCCGTTCTCGTGCTCTCAACTACGTTCACGAGGCTCAGAACACTCTCACTAACCTC
CACTCTATCTCTACTCAGCTCCAGGGTCTCCACCACAAGGACCTCAACAACGAGACTGTTTCTTCTATCCAGGCTTCTCTCATCCGT
TACCGTCAGAACGTTGAGGAGATCATCCTCTACCTCGAGATCATCCTCAACAAGCTCGGTCAGCTCGAGGGTTTCTCTCAGATCAA
GATGATCGGTCTCAACAAGCTCTCTCAGCGTCTCGAGAAGAAGATCCAGGAGCAGCAGAACCCTTCTGACTGCAAGATGGCTAAG
TACGTTACTTTCGACTTCACTAACCTCTGCGGTTTCGGTTGCTCTATGCACCAGCTCATGTTCTGCTTCCAGCTCGCTTTCGAGTCT
GAGCGTGTTCTCATCGTTATGAACCACCCTCCTGGTGACGTTTTCCGTGAGTGGCTCCACCAGAACACTCTCCCTCTCTCTGAGAAGT
GCTCTCACGTTGACACTAACGGTCGTTCTGACAACATCCGTTGCCCTTTCATCCGTTACGGTCACATCGCTCACAACTGGCTCCCT
CACGTTCTCCCTATCAACATGAACGCTGAGCTTCTCCGTCTCCACGAGGCTCCTTACGTTTGGTTCGCTGGTCAGCTCGCTGCT
TACATCCTCCGTCCTCGTCCTCAGTTCGCTCAGCTCATCAACGCTACTCTCAAGTCTTTCAAGGCTTCTGGTAACCCTGTTGTTGGT
GTTCACATCCGTCGTACTGACAAGATCATCTACGAGGCTAAGTTCCACAACCTCTCTGAGTACATGAAGCAGGTTGACAAGTTC
GACCAGCAGAACGCTTCTTTCCTCACTATGTCTAAGCGTCACGACATCGAGGAGAAGAAGTTCGGTTCTATCTTCGGTCGTACTATC
CGTAAGAACAACTCTGCTAAGCGTTCTGTTTACCTCGCTACTGACGACCCTAACGTTTTCACTCAGTTCCCTTACTCTCACCCTAACTA
CATCCTCTACGGTTCTCCTGGTCGTTCTCACTCTGCTGACCTCCGTATGCGTAAGACTTACAACTCTATCAACAACGCTATCATCGAC
GTTATCGCTCTCTCTATGACTGACTTCCTCGTTTGCACTTTCTCTTCTAACATCTGCCGTCTCGCTTACGAGCTTATGCAGACTCGTCA
CATGGAGTTCGGTGACGCTACTCAGCTCTGCCACTCTATCGACTTCGCTTTCCACCAGGAGGACTACTCTCGTCTCAAGTTCGACGT
TATCATCCCTGACATGAAGGGTCACCTCAAGTACGGTGACGTTGTTGACAACTTCATCAACCACGGTAACGGTTCTGCTTCTGTTA
AGCTCGAGAAGAACACTAAGAACGAGTTCTTCATCGCTCCTGCTTACAAGTTCCGTCCTACTATCAACATGACTTGGTTCAACTAA

SmFucTL

ATGCGTTACCTCAAGAACGAGTTCAACCTCTCTATCGTTCTCTGGCTCCCTCTCCTCCCTGTTCTCATGTTCTGCATCGAGGACTACTTCC
GTAAGAAGACTTTCTCTAAGACTGACCTCTGCTGGAAGGGTCTCATCAACAACGCTTCTGTTGCTAAGGAGGGTGTTGACTGGGTTTCT
GCTGAGATCCTCCACCTCCGTCGTCTCTTCGACTTCTCTGAGTCTATCTCTATGAAGACTATCAAGATGGACCAGATCACTAACGTTAACC
GTATCGGTATCTCTCACGAGACTCTCCGTCGTCGTGCTCTCAACTACCTCCACGAGGCTCAGAACACTCTCACTAACCTCCACTCTATCTC
TATCCAGCTCCAGGGTCTCCACCACAAGGACCTCAACAACGAGACTGTTTCTTCTATCCAGGCTTCTCTCATCCGTTACCGTCAGAAC

GTTGAGGAGATCATCCTCCACCTCGAGATCATCCTCAACAAGCTCGGTCAGCTCGACGGTTTCTCTCAGATCCGTATGATCGGTCTCAA
CAAGCTCTCTCAGCGTCTCCAGAAGAAGATCCAGGAGCAGCAGAACCCTGCTGACTGCAAGATGGCTAAGTACGTTACTTTCGAC
TACACTAACCTCTGCGGTTTCGGTTGCTCTATGCACCAGCTCATGTACTGCCTCCAGCTCGCTTTCGAGAACGACCGTGTTCTCATCGT
TAAGAACCACCCTCCTGGTGACATCTTCCGTAAGTGGCTCCACCAGAACACTCTCCCTCTCTCTGAGAAGTGCTCTCACCTCGACACTA
ACGGTCGTTCTGACAACATCCAGTGCCCTCGTCTCCGTTACGGTCACATCCACAACCTCCTCCCTAACGTTCTCCCTATCAACATGAAC
GCTGAGCTTCTCCGTCTCCACGAGGCTCCTTACGTTTGGTTCGCTGGTCAGCTCGCTGCTTACATCCTCCGTCCTCGTCCTCAGTTC
GCTCAGCTCATCAACGCTACTCTCAAGTCTTTCAAGATCTCTGGTGACCCTGTTGTTGGTGTTCACATCCGTCGTACTGACAAGATCATC
CGTGAGGCTAAGTTCCACAACCTCTCTGAGTACATGGAGCAGGTTGACAAGTTCTTCGACCGTCAGAACGCTAACCACCTCATCAT
GTCTAAGCGTCACAACAAGAACCGTAAGAAGCTCGGTTCTATCTCTGAGCGTACTATGCGTCGTAACAACTCTATCCAGATCAAGC
GTTCTATCTACCTCGCTACTGACGACCCTAACGTTTTCACTCAGTTCCCTAACTCTCACCCTAACTACATCCTCTACGGTTCTCCTG
GTCGTTCTCACTCTGCTGACCTCAAGATGCGTAAGACTTACAACTCTTGGAACAACGCTATCATCGACGTTATCGCTCTCTCTATGACT
GACTTCCTCGTTTGCACTTTCTCTTCTAACATCTGCCGTCTCGCTTACGAGCTTATGCAGACTCGTCAGATGGAGCTTGGTGACGCTACT
CAGCTCTGCCACTCTATCGACATCGCTTTCCACGAGGAGGACTACTCTCGTCTCAAGTTCGACGTTATCATCCCTGACATGAAGGAG
CAGCTCAAGTACGGTGACGTTGTTGACATCTTCATCAACCACGGTAACGGTTCTGCTTCTGTTAAGCTCGACCAGTCTTCTCGTAA
CATCCACAACAAGTCTCAGAAGGACACTCGTAACGAGTTCTTCATGGCTCCTGCTTACAAGTTCCGTCCTCGTATCAACATCACTTAA

SUPPLEMENTAL FIGURE 4 | Sequences. SmFucTA to F, H and J to L codon optimised nucleotide sequences.

Chapter 3

Functional characterisation of *Schistosoma mansoni* fucosyltransferase in *Nicotiana benthamiana*

Kim van Noort [1], Kasper ter Horst [1], Annet Lindhout [1], Linh Nguyen [2],
Lotte B. Westerhof [1], Cornelis H. Hokke [2], Arjen Schots [1] and Ruud H.P. Wilbers [1]

[1] Laboratory of Nematology, Wageningen University and Research, Wageningen, the Netherlands;
[2] Department of Parasitology, Leiden University Medical Centre, Leiden, the Netherlands.

3

Abstract

Clinical trials with live parasites and mouse model studies have shown the potential of helminths and their excretory/secretory (ES) proteins to treat allergies and autoimmune disorders. Glycan-dependent mechanisms have shown to be essential in the *modus operandi* of several ES proteins. To further develop helminth-derived ES glycoproteins as biopharmaceuticals, a large-scale expression system is required for the production of recombinant glycoproteins with defined and tailored glycans. The trematode *Schistosoma mansoni* produces ES proteins that have the potential to function as biopharmaceutical. These *S. mansoni* ES proteins often possess highly fucosylated N-glycans that cannot be synthesised in current production systems. Co-expression of specific fucosyltransferases (FucTs) in the expression host is required to introduce helminth-like N-glycan modifications. In the GeneDB database 20 different *S. mansoni* FucT (SmFucT) genes can be found that are possibly involved in the synthesis of N-glycans, O-glycans and glycolipids. To date, only one SmFucT has been characterised in chemoenzymatic assays with glycan acceptors. Therefore, in this study the function of ten SmFucTs was examined by transient co-expression with carrier glycoproteins in *Nicotiana benthamiana* plants. With this method we have identified SmFucTs that are involved in N-glycan core α1,3- or α1,6-fucosylation or the synthesis of antennary LeX, LDN-F or F-LDN-F. These functionally characterised SmFucTs can directly be applied to synthesise complex helminth N-glycan motifs on recombinant glycoproteins, in order to study their contribution to immunomodulation. Characterisation of SmFucTs, other glycosyltransferases and combinations of different glycosyltransferases will expand the glyco-engineering toolbox and offers perspectives for large scale production of glycoproteins with tailored N-glycan structures in plants.

Introduction

Schistosoma mansoni is a human parasitic trematode, which together with other *Schistosoma* species infects 252 million people worldwide [1]. During host infection *S. mansoni* secretes various glycoproteins and glycolipids that affect the host immune system [2], [3]. Notable among the secretions of *S. mansoni* are highly fucosylated glycan motifs that differ in composition in different life stages. Fucosylated glycans can be highly immunogenic and many anti-*S. mansoni* antibodies in the human host are directed against fucose containing glycan motifs [4]. N-glycans on proteins of *S. mansoni* can be fucosylated at either the core, the antennae or both. At the N-glycan core a fucose can be linked via an α1,3 and/or α1,6 bond to the innermost GlcNAc residue. On the N-glycan antennae complex fucose containing motifs can be found, such as those depicted in Figure 1 [2], [5]. Notable among these motifs is the double fucose (DF) motif, found on the N-glycans of glycoprotein SmVAL9, glycolipids and O-glycans [2], [3], [6].
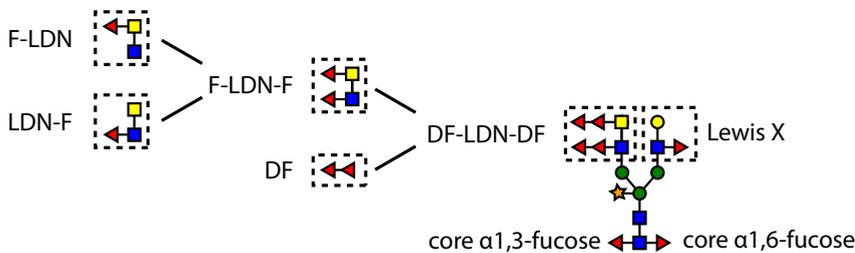

FIGURE 1 | Graphical representation of fucosylated N-glycan motifs synthesised by *Schistosoma mansoni*.

Glycans attached to glycoproteins and glycolipids are important in parasite biology [7]. For instance, the T-helper (Th) 2 inducing capacity of the *S. mansoni* soluble egg antigen omega-1 is glycosylation dependent. LeX on the N-glycans of omega-1 binds to the mannose receptor and the dendritic cell-specific ICAM-grabbing non-integrin (DC-SIGN) receptor on dendritic cells (DCs). Both receptors mediate uptake and allow omega-1 to subsequently prime DCs for a Th2 response via its capacities as RNase [8]. The presence of terminal LeX on omega-1 N-glycans enhances the Th2 inducing ability of omega-1 in comparison to omega-1 with paucimannosidic N-glycans [9]. Binding of LeX to DC-SIGN is fucose dependent and seems to depend on fucose linkage and/or surrounding monosaccharides [10]–[12]. Interestingly, next to fucose containing glycans DC-SIGN can also bind to mannose containing glycans. A different immunological response is observed upon binding of mannose containing glycans in comparison to fucose containing glycans [13]. This different immunological response indicates that glycans are important for receptor binding as well as for glycoprotein functioning and parasite biology.

In order to further investigate the importance of fucosylated glycan motifs on parasite secreted proteins and their role in parasite biology, it is crucial to know how these fucosylated motifs are synthesised. Fucosylated glycan motifs are synthesised by specific fucosyltransferases (FucTs), which link L-fucose from guanosine diphosphate β-L-fucose (GDP-fucose) to the glycoconjugate acceptor via an α1,2, α1,3, α1,4 or α1,6 linkage. However, most of the *S. mansoni* FucTs (SmFucTs) are not yet functionally characterised. To date one putative α1,3-SmFucT, SmFucTF, is functionally characterised to synthesise LeX in chemoenzymatic assays with various glycan acceptors [14]. For the other five putative α1,3-SmFucTs only functional predictions, based on phylogenetic and sequence analysis are currently available [14], [15]. The function of other putative SmFucTs is as yet unknown.

Functional predictions and characterisation in chemoenzymatic assays can differ from the *in vivo* function, since the characterisation is performed outside the context of a cell. *In vivo* the Golgi localisation of the glycosyltransferase is regulated by the CTS domain, consisting of a N-terminal short cytoplasmic tail (C), a transmembrane domain (T) and a stem region (S). The Golgi localisation determines for instance the pH in a sub-Golgi compartment, dimer formation and/or availability and composition of glycan acceptors and nucleotide sugars. Glycosyltransferase activity is therefore highly dependent on its specific Golgi localisation, which is difficult to mimic *in vitro*. *In vivo* SmFucT characterisation requires a platform that allows SmFucT Golgi transfer and Golgi retention or retrieval. Furthermore, the correct glycan acceptor and donor should be present. *Nicotiana benthamiana* can be a promising platform for *in vivo* characterisation. Plants are tolerant to glyco-engineering as their glycosylation pathway can be adjusted without influencing plant growth or development. Furthermore, so-called ΔXT/FT *N. benthamiana* plants are available, wherein endogenous plant FucTs are silenced and will not hamper SmFucT characterisation [16]. Lastly, transiently expressed exogenous glycosyltransferases can modify glycans on plant glycoproteins to generate glycan acceptors LN and LDN, for characterisation of SmFucTs involved in formation of LeX and LDN-F or F-LDN, respectively [9], [17].

In this study, we used *N. benthamiana* to functionally characterise the previously selected SmFucTs (A to F, H and J to L, Chapter 2) *in vivo*. We functionally characterised SmFucTs by transient co-expression with *S. mansoni* carrier glycoproteins kappa-5 and omega-1 in *N. benthamiana* plants and subsequent glycan analysis (Figure 2). With this plant-based expression platform we identify SmFucTs that are involved in N-glycan core α1,3- or α1,6-fucosylation or the synthesis of antennary LeX, LDN-F or F-LDN-F. Furthermore, we attributed a novel biological function to SmFucTF and were able to synthesise a new glycan motif in our plant expression platform.
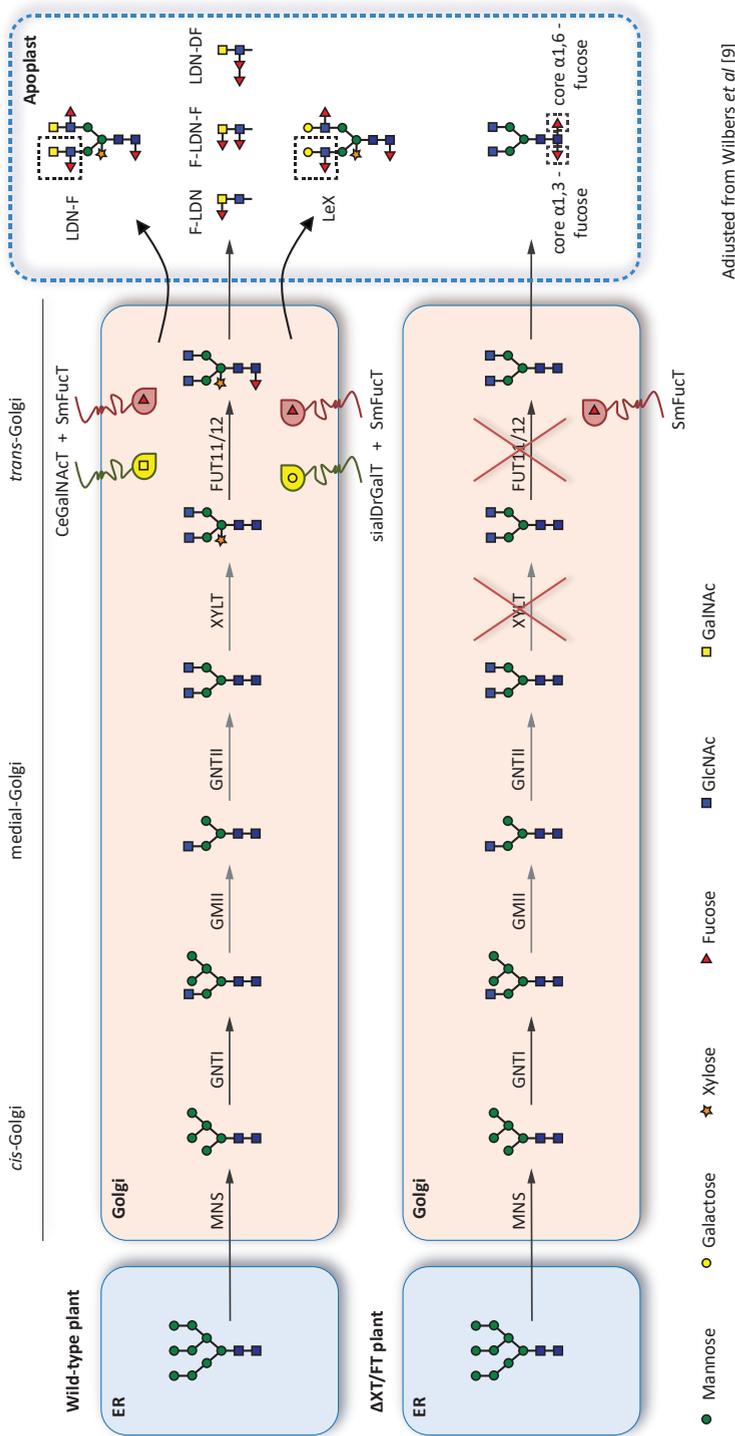
FIGURE 2 | Engineering of the plant glycosylation pathway for the characterisation of *Schistosoma mansoni* fucosyltransferases (SmFucTs). A schematic overview of the N-glycan modifying steps in the plant Golgi-system of *Nicotiana benthamiana* plants. The plant N-glycosylation machinery was engineered by introducing (hybrid) glycosyltransferases that allow synthesis of LN (sialDrGalT) or LDN (CeGalNAcT) for the synthesis of LeX, LDN-F or F-LDN by SmFucTs. Characterisation of core α1,3- and core α1,6-SmFucTs was performed in ΔXT/FT *N. benthamiana* plants, wherein the plant β1,2-xylosyltransferases and core α1,3-FucTs were down regulated by RNA interference. MNS: Class I mannosidases MNS1, 2 and 3; GnTI: N-acetyl-glucosaminyltransferase I; GMII: Golgi-α-mannosidase II; GnTII: N-acetyl-glucosaminyltransferase II; XYLT: β1,2-xylosyltransferase; FUT11/12: core α1,3-fucosyltransferase. Arrows are indicated in grey, when alternative N-glycan processing routes have been postulated.

Adjusted from Wilbers *et al* [9]

3

65

Results

Core fucosylation by *Schistosoma mansoni* fucosyltransferases

To investigate which of the previously selected SmFucTs (A to F, H and J to L, Chapter 2) can add fucose to the N-glycan core, each of the SmFucTs or the α1,6-FucT DmFucT8 was co-expressed with carrier glycoprotein omega-1 in ΔXT/FT *N. benthamiana* plants. Additionally, omega-1 was expressed in wild type (wt) plants as a control for core α1,3-fucosylation. Omega-1 was purified from the leaf apoplast and fucosylation of the omega-1 N-glycans was analysed with the fucose binding lectin AAL. As expected, only weak binding of AAL was observed to the N-glycans of omega-1 produced in ΔXT/FT plants and was therefore considered background (Figure 3A). On the contrary, strong binding of AAL was observed to the N-glycans of omega-1 produced in wt plants and omega-1 produced in ΔXT/FT plants upon co-expression of DmFucT8. Upon co-expression of SmFucTs, clear AAL binding was only observed to omega-1 N-glycans upon co-expression of SmFucTC and H, which suggested that SmFucTC and SmFucTH add a core fucose to the N-glycans of omega-1.

To distinguish whether SmFucTC and SmFucTH added a core α1,3- or α1,6-linked fucose to the N-glycan core, a lectin assay was performed with *Pholiota squarrosa* lectin (PhoSL). PhoSL specifically binds to core α1,6-fucose. PhoSL binding to the omega-1 N-glycans upon co-expression of SmFucTH indicated that SmFucTH added a core α1,6-fucose to omega-1 N-glycans, although less efficient than the positive control DmFucT8 (Figure 3B). Analysis of the N-glycan composition of omega-1 upon co-expression of DmFucT8 or SmFucTH with MALDI-TOF MS confirmed addition of a core fucose (Figure 3F and G). The major N-glycan found on omega-1 produced in ΔXT/FT plants (Figure 3D) is paucimannosidic (peak 1030 m/z), whereas both DmFucT8 and SmFucTH induced a shift towards a fucosylated paucimannosidic N-glycan (peak 1176 m/z).

To confirm that SmFucTC added an α1,3-linked fucose to the core of omega-1 N-glycans, purified omega-1 was treated with PNGase F and visualised on an Oriole stained SDS-PAGE gel. PNGase F releases N-glycans only in the absence of core α1,3-fucose. The release of N-glycans from omega-1 resulted in a shift in protein size from 30 kDa to 27 kDa, which is shown when comparing omega-1 samples produced in wt and ΔXT/FT plants (Figure 3C). Upon co-expression of SmFucTs, only SmFucTC prevented N-glycan release by PNGase F. This confirms that SmFucTC adds a core α1,3-fucose to the omega-1 N-glycans. The similar band pattern observed for omega-1 expressed in ΔXT/FT plants upon co-expression of the α1,6-SmFucTs, probably corresponded to background core α1,3-fucosylation by endogenous plant FucTs. In ΔXT/FT plants the *N. benthamiana* fucosyltransferases are knocked down and not knocked out. MALDI-TOF MS analysis confirmed fucosylation by SmFucTC (Figure 3E), comparable to the degree of fucosylation of omega-1 N-glycans produced in wt plants [9]. MALDI-TOF MS analysis on omega-1 N-glycans upon co-

expression of SmFucTF and SmFucTJ revealed predominantly unfucosylated core N-glycans (Supplemental Figure 1A and B). Since, P19 co-expression could influence SmFucTK and L localisation (Chapter 2), also omega-1 N-glycans upon co-expression of SmFucTK or L without P19 co-expression were analysed by MALDI-TOF MS and lacked a core fucose (Supplemental Figure 1C and D). Taken together, the N-glycan profiles, the analysis with PNGase F, and the lectin binding assays with AAL and PhoSL showed that SmFucTC is a core α1,3-SmFucT and SmFucTH is a core α1,6-SmFucT.
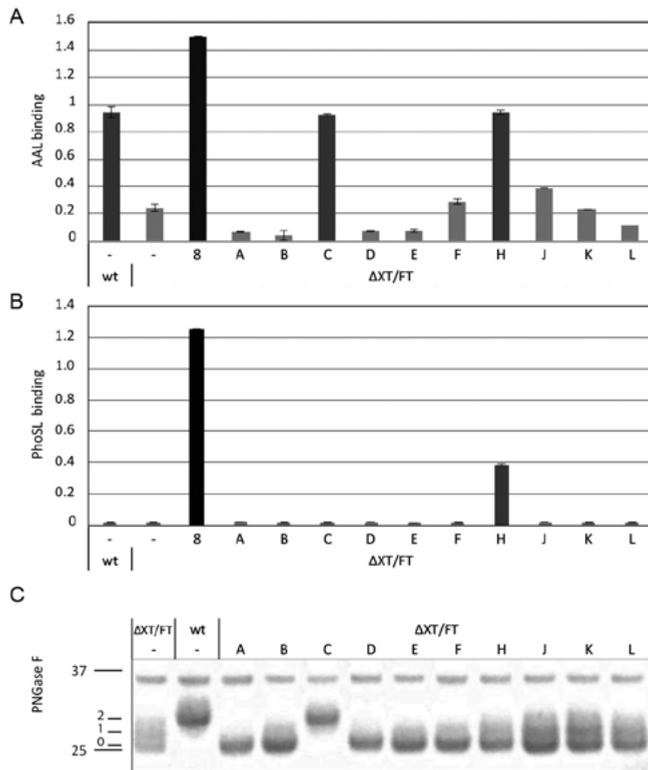

FIGURE 3 | Core fucosylation of omega-1 N-glycans by fucosyltransferases of *Schistosoma mansoni* (SmFucTs). To screen for core fucosylation omega-1 was co-expressed with each SmFucT (A to F, H and J to L) or DmFucT8 in ΔXT/FT *Nicotiana benthamiana* plants or wild type (wt) plants as positive control. After extraction and subsequent purification from the apoplast fluid the N-glycan composition of omega-1 was analysed. (A) *Aleuria aurantia* lectin (AAL) binding assay reveals that SmFucTC and SmFucTH add a core fucose. (B) *Pholiota squarrosa* lectin (PhoSL) binding assay for the specific detection of α1,6-linked fucose, reveals that SmFucTH adds a core α1,6-fucose. (C) Oriole stained SDS-PAGE gel with 200ng PNGase F treated omega-1, reveals that SmFucTC is a core α1,3-FucT. The observed 35kD band corresponds to a co-purified plant protein. The amount of fucosylated glycosylation sites (2, 1, 0) is indicated on the left. (D-G) MALDI-TOF MS N-glycan profile of omega-1 expressed in ΔXT/FT plants (D) upon co-expression of SmFucTC (E), DmFucT8 (F) or SmFucTH (G). When a peak represents multiple N-glycan structures of identical mass, the number of sugar residues of which the position on the N-glycan is not clear is indicated above the N-glycan structure. The possible positions of these sugar residues on the N-glycan are indicated between brackets.
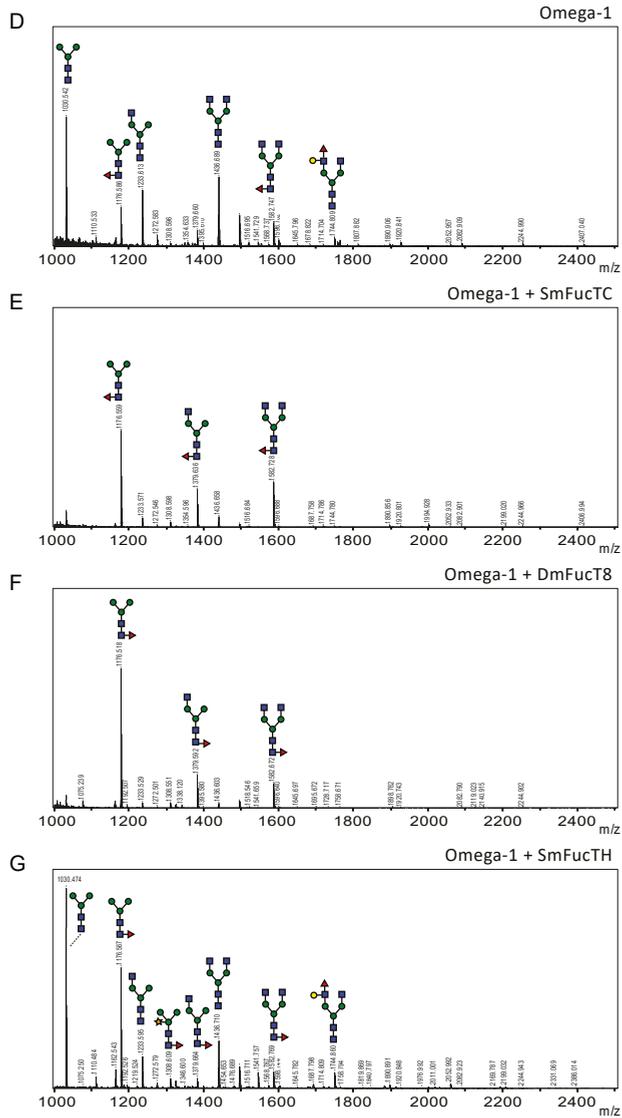
FIGURE 3 | Core fucosylation of omega-1 N-glycans by fucosyltransferases of *Schistosoma mansoni* (SmFucTs). Continued

Lewis X synthesis by *Schistosoma mansoni* fucosyltransferases

To identify which of the SmFucTs is able to synthesise LeX, carrier glycoprotein kappa-5 was co-expressed with the β1,4-galactosyltransferase (GalT) sialDrGalT for LN synthesis in ΔXT/FT *N. benthamiana* plants. Subsequently, each of the SmFucTs or positive control sialTnFucT9a was included to screen for LeX synthesis. Proteins were isolated from the leaf apoplast and screened for fucosylated N-glycans with the lectin AAL. In the lectin binding assay, AAL bound to N-glycans when LeX was synthesised upon co-expression of sialDrGalT and sialTnFucT9a (Figure 4A). Upon co-expression of SmFucTs, AAL bound to N-glycans upon co-expression of SmFucTA, D, E or F, although AAL binding was weak upon co-expression of SmFucTA (Figure 4A). Strikingly, no AAL binding was observed upon co-expression of SmFucTC or H, which suggests that AAL was unable to bind the core fucoses on the most abundant glycoprotein in the apoplast kappa-5 or SmFucTC and H were unable to core fucosylate kappa-5.

To check LeX synthesis, a western blot was performed on apoplast samples using anti-LeX antibodies. Upon co-expression of sialDrGalT with sialTnFucT9a, SmFucTD or E, the western blot revealed two bands around 45 kDa and 90 kDa for the kappa-5 monomer and dimer, respectively (Figure 4B). A fainter band around 45 kDa was observed upon co-expression of SmFucTF, indicating that SmFucTF can synthesis LeX, although less efficient than SmFucTD or E. No bands were observed upon co-expression of SmFucTA.

MALDI-TOF MS analysis was performed on PNGase A released N-glycans of purified kappa-5 to confirm LeX synthesis. As expected co-expression of sialDrGalT with sialTnFucT9a resulted in synthesis of a single branched LeX (peak 1541 m/z) (Figure 4C) [9]. Upon co-expression of SmFucTD and E also single branched LeX was detected (Figure 4D and E and Supplemental Figure 3 A-F). For SmFucTF only minor peaks corresponding to single branched LeX were observed, confirming the lower signal observed in the western blot (Figure 4F and Supplemental Figure 3G-I). Also, for SmFucTA minor peaks corresponding to a single galactosylated and α1,3-fucosylated terminal glycan motif were observed (Supplemental Figure 2 and Supplemental Figure 3J-L). Since, on the western blot no bands were observed upon co-expression of SmFucTA, the released N-glycans of kappa-5 upon SmFucTA co-expression were treated with α1,2/4/6-fucosidase O from *Omnitrophica to confirm LeX synthesis.* After α1,2/4/6-fucosidase O treatment the remaining minor peaks correspond to LeX carrying N-glycans (Supplemental Figure 3M). The ratio of LeX vs LeA synthesised glycan motifs upon SmFucTA, D, E or F co-expression was determined by digestions with α1,3/4-fucosidase from *Xanthomonas* sp. and β1,3/6-galactosidase from *Xanthomonas manihotis* or β1,4/6-galactosidase from Jack Bean, respectively. Altogether, we conclude that SmFucTD, E and to a lesser extent SmFucTF and SmFucTA can synthesise LeX on kappa-5 N-glycans.
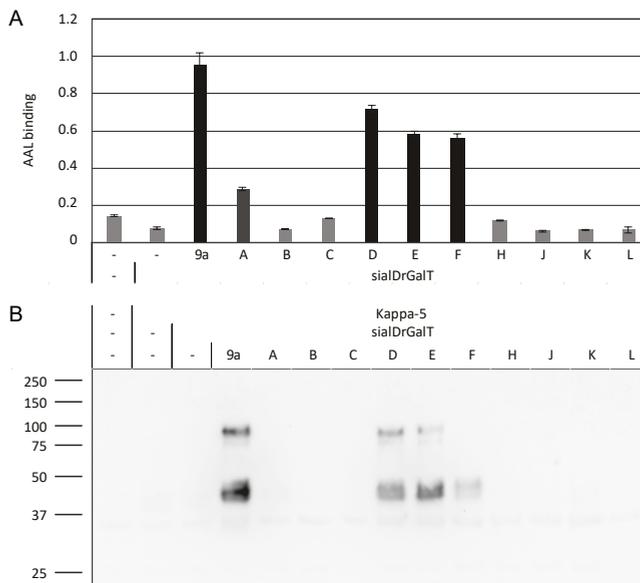
3

FIGURE 4 | LeX synthesis on kappa-5 N-glycans by fucosyltransferases of *Schistosoma mansoni* (SmFucTs).
Kappa-5 was co-expressed with sialDrGalT and each of the selected SmFucTs (A to F, H and J to L, Chapter 2) or sialTnFucT9a in ΔXT/FT *Nicotiana benthamiana* plants. After extraction the N-glycan composition of total apoplast proteins was analysed. (A) *Aleuria aurantia* lectin (AAL) binding assay reveals that SmFucTD, SmFucTE and SmFucTF add a fucose. (B) LeX western blot with 4ug apoplast protein reveals that SmFucTD, SmFucTE and SmFucTF synthesise LeX. (C-F) N-glycan profiles are shown for purified kappa-5 upon co-expression of sialDrGalT with sialTnFucT9a (C), SmFucTD (D), SmFucTE (E) or SmFucTF (F). When a peak represents multiple N-glycan structures of identical mass, the peak indicates the major N-glycan present based on enzymatic digestions (for details see Supplemental Figure 3).
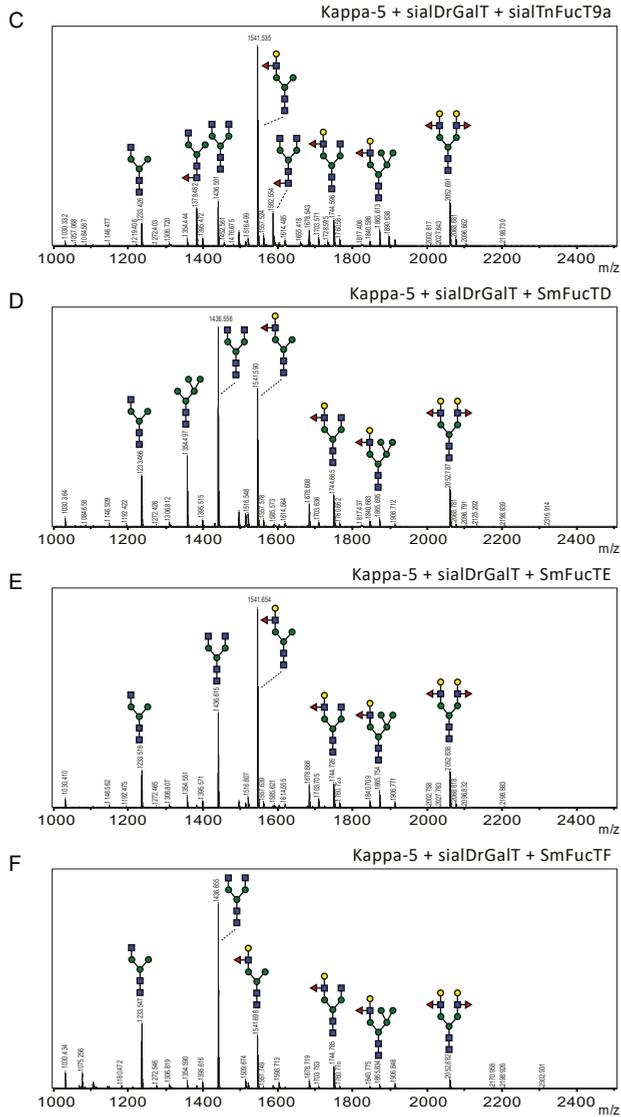
FIGURE 4 | LeX synthesis on kappa-5 N-glycans by fucosyltransferases of *Schistosoma mansoni* (SmFucTs).
Continued

Fucosylation of LDN by *Schistosoma mansoni* fucosyltransferases

To identify which of the SmFucTs is able to fucosylate LDN, carrier glycoprotein kappa-5 was co-expressed with the β1,4-*N*-acetylgalactosaminyltransferase (GalNAcT) CeGalNAcT for LDN synthesis in *N. benthamiana* plants. Subsequently, each of the SmFucTs or positive control sialTnFucT9a was co-expressed to screen for LDN fucosylation. Proteins were isolated from the leaf apoplast and screened for fucosylation of LDN glycan motifs with the lectin soybean agglutinin (SBA). SBA binding to the GalNAc residue in LDN is inhibited upon fucosylation of LDN. As expected SBA bound to N-glycans when LDN was synthesised upon co-expression of CeGalNAcT, whereas LDN fucosylation upon co-expression of sialTnFucT9a abolished SBA binding (Figure 5A). Similarly, SBA binding to N-glycans was abolished upon co-expression of SmFucTD and E, indicating LDN fucosylation by SmFucTD and E.

To distinguish whether SmFucTD and E synthesised LDN-F and/or F-LDN a western blot was performed on apoplast samples using anti-LDN-F and anti-F-LDN antibodies. Upon co-expression of CeGalNAcT with sialTnFucT9a, SmFucTD or E, the western blot treated with anti-LDN-F antibodies revealed two bands around 45 kDa and 90 kDa corresponding with the kappa-5 monomer and dimer, respectively (Figure 5B). This indicates that all three enzymes synthesised LDN-F. Next, a western blot treated with anti-F-LDN antibodies was performed. Remarkably, this western blot revealed bands only upon co-expression of SmFucTF, suggesting that SmFucTF synthesises F-LDN (Figure 5C and Supplemental Figure 4 and 5A-D). This was surprising, since co-expression of SmFucTF did not abolish binding of SBA to the kappa-5 N-glycans (Figure 5A).

MALDI-TOF MS analysis was performed on released N-glycans of purified kappa-5 to confirm LDN-F and F-LDN synthesis. As expected co-expression of CeGalNAcT with sialTnFucT9a resulted in synthesis of single branched LDN-F (peak 1713 m/z) (Figure 5D). Additionally, we detected a significant proportion of N-glycan structures with a fucosylated terminal GlcNAc residue (1657 m/z; Figure 5D and Supplemental Figure 5A-C). Upon co-expression of SmFucTD, E and F single branched fucosylated LDN was observed (1713 m/z) (Figure 5E-G). However, enzymatic digestion on released N-glycans of purified kappa-5 upon co-expression of SmFucTF revealed no fucosylation of LDN (Supplemental Figure 5M-P). This suggested that F-LDN is not formed upon co-expression of SmFucTF, although the western blot suggested otherwise (Figure 5C). Interestingly, enzymatic digestions on released N-glycans of purified kappa-5 upon co-expression of sialTnFucT9a, SmFucTD and E revealed more efficient LDN-F formation upon co-expression of SmFucTD and E than sialTnFucT9a (Supplemental Figure 5A-L). Furthermore, no fucosylated terminal GlcNAc residues without terminal GalNAc were detected upon co-expression of SmFucTD, E and F. In conclusion, LDN-F can be synthesised by SmFucTD and E.
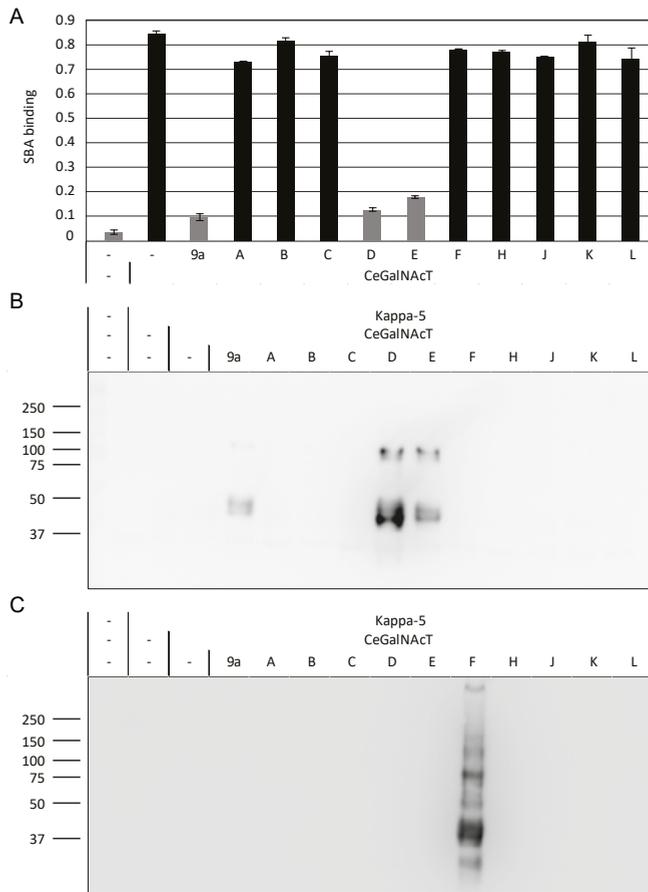
FIGURE 5 | LDN fucosylation on the N-glycans of kappa-5 by fucosyltransferases of *Schistosoma mansoni* (SmFucTs). Kappa-5 was co-expressed with CeGalNAcT and each of the selected SmFucTs (A to F, H and J to L, Chapter 2) or sialTnFucT9a in *Nicotiana benthamiana* plants. After extraction the N-glycan composition of total apoplast proteins was analysed. (A) Soybean agglutinin (SBA) lectin binding assay reveals that SmFucTD and SmFucTE inhibited SBA binding, suggesting LDN fucosylation. (B) LDN-F western blot with 2ug apoplast protein reveals that SmFucTD and SmFucTE synthesise LDN-F. (C) F-LDN western blot with 2ug apoplast protein reveals that SmFucTF synthesises F-LDN. (D-G) After purification of kappa-5 from the apoplast fluid, the composition of PNGase A released N-glycans was analysed by MALDI-TOF MS. N-glycan profiles are shown for kappa-5 upon co-expression of CeGalNAcT with sialTnFucT9a (D), SmFucTD (E), SmFucTE (F) or SmFucTF (G). When a peak represents multiple N-glycan structures of identical mass, the peak indicates the major N-glycan present based on enzymatic digestions (for details see Supplemental Figure 4 and 5). When the sugar position is still unclear, the number of sugar residues of which the position on the N-glycan is not clear is indicated above the N-glycan structure. The possible positions of these sugar residues on the N-glycan are indicated between brackets.
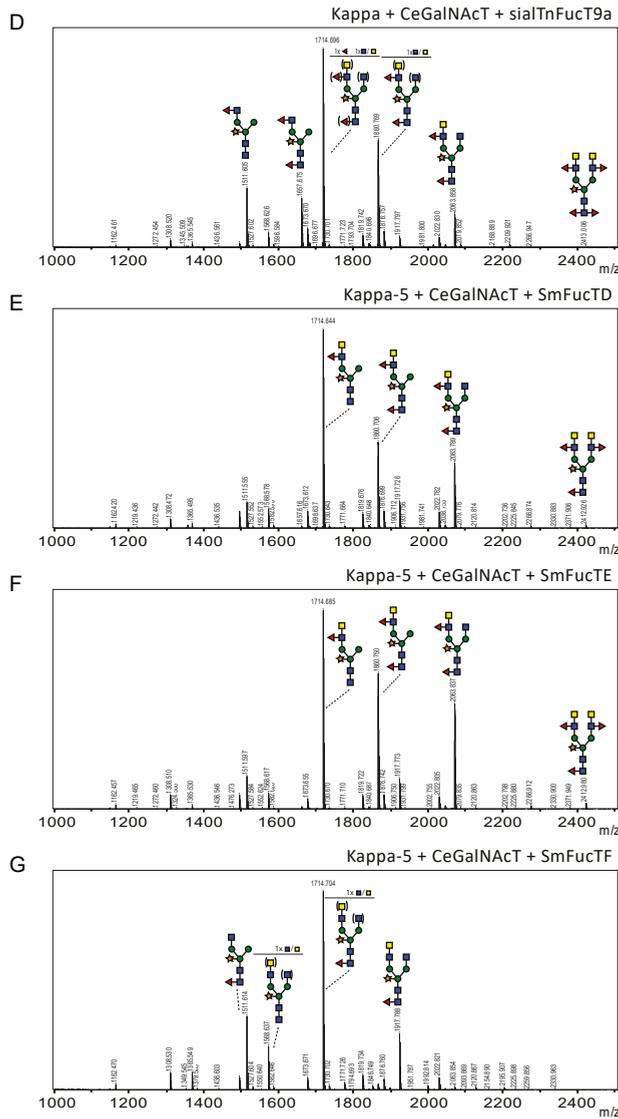
FIGURE 5 | LDN fucosylation on the N-glycans of kappa-5 by fucosyltransferases of *Schistosoma mansoni* (SmFucTs). Continued

Synthesis of F-LDN-F by *Schistosoma mansoni* fucosyltransferases

The identification of SmFucTD and E as LDN-F synthesising SmFucTs and the suggestion of F-LDN synthesis by SmFucTF allowed us to examine whether the glycan motif F-LDN-F could be synthesised in plants. Therefore, kappa-5 was co-expressed with CeGalNAcT, SmFucTF and SmFucTD or E in *N. benthamiana* plants. After purification of kappa-5 the presence of F-LDN-F was analysed by a western blot treated with anti-F-LDN-F antibodies. This blot revealed multiple bands around 45 kDa, 55 kDa and 100 kDa upon co-expression of SmFucTF with SmFucTD or E, indicating synthesis of F-LDN-F (Figure 6A). To confirm synthesis of F-LDN-F, MALDI-TOF MS analysis was performed on released N-glycans of purified kappa-5 (Figure 6B and C). F-LDN-F carrying N-glycans were detected upon co-expression of SmFucTF with SmFucTD or E (2006 and 2209 m/z). Presence of F-LDN-F glycan motifs was confirmed after enzymatic digestions (Supplemental Figure 6). To our knowledge this is the first time that F-LDN-F has been synthesised in plants.

No indication for synthesis of the double fucose glycan motif

To investigate synthesis of the glycan motif DF by one SmFucT or a combination of SmFucTs, each of the SmFucTs was co-expressed in wt or ΔXT/FT *N. benthamiana* plants with carrier glycoprotein omega-1 or kappa-5 in combination with different glycosyltransferases and *N. benthamiana* β-*N*-hexosaminidase (NbHEXO) to create the N-glycan acceptors, LDN-F, F-LDN-F and F-GlcNAc. After extraction the N-glycan composition of glycoproteins in the apoplast fluid was analysed by a western blot treated with anti-LDN-DF or anti-DF-LDN-DF antibodies. Unfortunately, none of the performed western blots indicated DF synthesis. Furthermore, MALDI-TOF MS analysis on purified kappa-5 N-glycans upon co-expression of all ten SmFucTs and CeGalNAcT did not indicate DF carrying N-glycans (unpublished data).
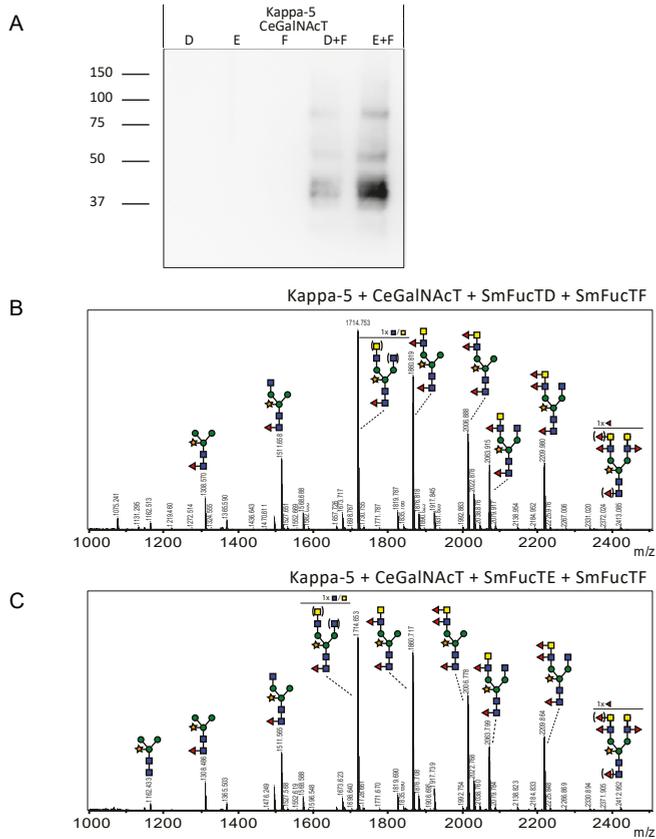
3

FIGURE 6 | F-LDN-F synthesis on kappa-5 N-glycans by fucosyltransferases D, E and F of *Schistosoma mansoni* (SmFucTD, E and F). Kappa-5 was co-expressed with CeGalNAcT, SmFucTF and SmFucTD or SmFucTE in *Nicotiana benthamiana* plants. After extraction and subsequent purification from the apoplast fluid the N-glycan composition of kappa-5 was analysed. (A) F-LDN-F western blot with 50ng purified kappa-5 reveals synthesis of F-LDN-F. (B-C) MALDI-TOF MS N-glycan profile for kappa-5 upon co-expression of CeGalNAcT, SmFucTF and SmFucTD (B) or SmFucTE (C). When a peak represents multiple N-glycan structures of identical mass, the peak indicates the major N-glycan present based on enzymatic digestions (for details see Supplemental Figure 6). When the sugar position is still unclear, the number of sugar residues of which the position on the N-glycan is not clear is indicated above the N-glycan structure. The possible positions of these sugar residues on the N-glycan are indicated between brackets.

Discussion

Currently, all expressed SmFucTs are characterised based on sequence analysis [15]. SmFucTF is the only SmFucT that was also functionally characterised in chemoenzymatic assays [14]. Characterisation based on chemoenzymatic assays and sequence-analysis can differ from the biological function *in vivo*. Here, previously selected SmFucTs (A to F, H and J to K, Chapter 2) were characterised *in vivo*, to determine their function and to understand how certain fucosylated N-glycan motifs are synthesised. Depending on the expected glycan motif SmFucTs were either co-expressed with carrier glycoprotein omega-1 or kappa-5 in *N. benthamiana*. Subsequently, the N-glycan composition of the carrier glycoprotein was analysed for fucosylated glycan motifs. With this approach we have identified SmFucTs that are involved in N-glycan core α1,3- or α1,6-fucosylation or the synthesis of antennary LeX, LDN-F or F-LDN-F (Figure 7). Furthermore, we were able to attribute a new function to SmFucTF, which allows formation of F-LDN-F in plants.
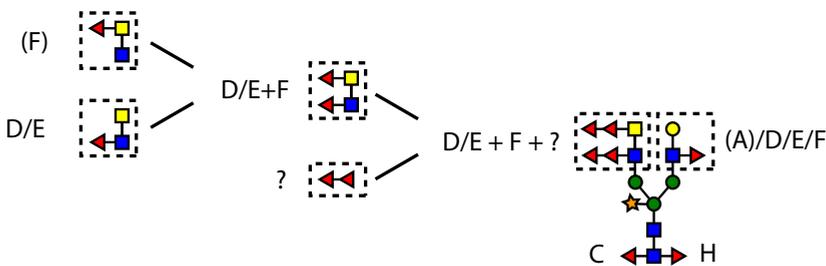

FIGURE 7 | Graphical representation of N-glycan motifs synthesised by characterised fucosyltransferases of *Schistosoma mansoni* (SmFucTs). In this figure, A is placed between brackets, because only a minority of kappa-5 N-glycans carried LeX upon co-expression of SmFucTA. Since, F-LDN synthesis on N-glycans by SmFucTF was not confirmed in the absence of SmFucTD or E by MALDI-TOF MS, F is place between brackets.

One of the fucosylated motifs on the N-glycans of *S. mansoni* is core fucosylation, either with an α1,3 or α1,6 linkage. We investigated core fucosylation on carrier glycoprotein omega-1 and characterised SmFucTC and H as core α1,3-FucT and core α1,6-FucT, respectively. An α1,6 linked fucose is only known to be present on the N-glycan core, whereas four of the previously selected SmFucTs (Chapter 2) were characterised as α1,6-SmFucTs based on sequence analysis [15]. Based on our analysis the function of the three α1,6-SmFucTs J, K and L remains unknown. Depending on the protein environment the α1,6-SmFucTs may differ in their capacity to add an α1,6-fucose. Therefore, SmFucTJ, K and L may be involved in the core α1,6-fucosylation of other glycoproteins than omega-1. Protein environment is thought to influence glycosylation, although protein specific activity of glycosyltransferases with the same function has never been shown [18]–[20].

Furthermore, in other organisms this singular function is linked to one α1,6-FucT instead of six. Peterson and colleagues [15] suggested that a secondary function may explain the large number of α1,6-SmFucTs. Another possibility is stage specific expression of different α1,6-SmFucTs as it has been shown that α1,6 linked fucoses are present in different developmental stages of *S. mansoni* [2].

LeX is another fucosylated glycan motif found on glycans of *S. mansoni*. In *N. benthamiana* LeX can be synthesised on N-glycans by co-expression of sialDrGalT and sialTnFucT9a [9]. For *S. mansoni*, Mickum and colleagues [14] showed that in chemoenzymatic assays SmFucTF fucosylated LN to synthesise LeX on glycan acceptors. In this study, we showed that SmFucTF can synthesise LeX *in vivo* as well. Next to SmFucTF, we found that also SmFucTA, D and E are able to synthesise LeX on N-glycans. Although, upon SmFucTA co-expression only a minority of the N-glycans on kappa-5 carried LeX. This suggests that N-glycans may not be the most optimal glycan acceptors for LeX synthesis by SmFucTA, whereas glycolipids or O-glycans could be more suitable. Furthermore, next to LeX the glycan motifs diLeX and pseudo LeY are synthesised by *S. mansoni* [2], [3]. Therefore, it could be that SmFucTA synthesises these β1,4-galactosylated and α1,3-fucosylated glycan motifs, found on O-glycans and/or glycolipids, more efficiently.

Sequence comparison of SmFucTF vs SmFucTD and SmFucTF vs SmFucTE show only 41.6 and 69.6% similarity, respectively, although all these SmFucTs can fucosylate LN to synthesise LeX on N-glycans. In an earlier study of Mickum and colleagues [14], SmFucTF showed specificity towards different glycan acceptors. In our experiments we observe differences in efficiency, which suggests that although SmFucTD, E and F show the same function, they can have specificity for N-glycans, O-glycans or glycolipids. Furthermore, these SmFucTs may also be active in different life stages or tissues. Transcriptomic life cycle analysis performed by Fitzpatrick and colleagues [21] showed differential expression of smp_053400, smp_137740 and smp_137730. These sequences correspond or relate to SmFucTD, E and F, respectively (Chapter 2). All three SmFucTs show quite high expression in the miracidium and sporocyst stages. Smp_137740 and smp_137730 are clearly expressed in all life stages. Smp_053400 is scarcely expressed in cercariae, schistosomula and young worms. Smp_137740 expression is high in eggs, whereas smp_053400 and smp_137730 expression levels are low to moderate in eggs. These differences in expression during various life stages combined with glycan acceptor specificities may explain why *S. mansoni* has multiple SmFucTs for LeX synthesis.

Next to fucosylation of LN to synthesise LeX, we showed that SmFucTD and E can fucosylate LDN on the GlcNAc residue synthesising LDN-F. SmFucTF can also fucosylate LDN although fucosylation occurs on the GalNAc residue in LDN synthesising F-LDN. N-Glycan profiles show that fucosylation of GalNAc by SmFucTF appears not to be present. Yet, fucosylation of GalNAc is observed upon co-expression of SmFucTF with SmFucTD or E, which results

in synthesis of F-LDN-F. Moreover, glycan analysis performed by Smit and colleagues [2] showed that fucosylation of GalNAc is mostly seen in combination with fucosylation of GlcNAc, synthesising F-LDN-F motifs. This suggests that SmFucTF prefers LDN-F as N-glycan acceptor over LDN. The importance of the glycan acceptor can also explain why Mickum and colleagues [14] were unable to show fucosylation of LDN by SmFucTF in their experiments with glycan acceptor GalNAcβ1–4GlcNAcβ1–3Galβ1–4Glc-AEAB. This glycan acceptor is not a natural glycan acceptor, whereas the plant paucimannosidic N-glycan with LDN(-F) antennae was found in *S. mansoni* [2]. Furthermore, the presence of SmFucTD or E seems to be required for proper fucosylation by SmFucTF on N-glycans.

LDN-F and F-LDN-F can be double fucosylated to LDN-DF and DF-LDN-DF wherein the DF motif is synthesised by an α1,2-SmFucT. None of the putative SmFucTs was initially characterised as α1,2-SmFucT. Therefore, the capability to link fucose with an α1,2 bond to another fucose was expected among the characterised α1,3 or α1,6 SmFucTs. Unfortunately, co-expression of SmFucTs (A to F, H and J to L) with different combinations of glycosyltransferases, glycan cutting enzymes and glycoproteins, in order to generate different N-glycan acceptor structures, did not result in detectable DF formation. The DF glycan motif was probably not detected due to co-expression of the wrong SmFucT or incorrect environmental conditions, such as glycan acceptor or donor accessibility. Since in *S. mansoni*, only a minority of the DF motif is found on N-glycans [2], [6].

In our study, we were not able to characterise SmFucTB, J, K and L. Therefore, one of the uncharacterised SmFucTs could also be involved in formation of DF. α1,6-FucT glycan motif I and II, as described by Breton and colleagues [22], are shared between the α1,6-FucTs and the known α1,2-FucTs adding an α1,2-fucose to a galactose [23]. This suggests that one of non-characterised putative α1,6 SmFucTs could be involved in DF synthesis. On the other hand, sequence comparison of Mickum and colleagues [14] suggests that SmFucTB could be involved in synthesis of the DF motif. Another option could be that none of the selected SmFucTs can synthesise DF and one of the many splice variants found among the SmFucTs encodes the missing α1,2-SmFucT [15].

In this study, we revealed synthesis of fucosylated N-glycan motifs by six SmFucTs. This knowledge can be used in studies on the effects of fucosylated glycans in parasite biology. Characterised transferases can be knocked out to study the loss of specific glycan motifs during parasite infection. Additionally, we established a promising new *in vivo* platform for the functional characterisation of *S. mansoni* glycosyltransferases using *N. benthamiana* plants. The functionally characterised SmFucTs can directly be applied to synthesise helminth glycoproteins with native N-glycans in order to study their role in parasite biology. This includes motifs such as the newly synthesised F-LDN-F. Furthermore, these native glycoproteins can be used in immunological studies to investigate applications as vaccine or biopharmaceutical drug.

3

Materials and Methods

Construction of expression vectors

To determine SmFucT activity, selected SmFucT constructs (SmFucT A to F, H and J to L, Chapter 2) were co-expressed with carrier glycoprotein omega-1 or kappa-5 [9]. For screening purposes, LN was synthesised in plants by co-expression of a hybrid *Danio rerio* β1,4-galactosyltransferase with the CTS domain of *Rattus norvegicus* α2,6-sialyltransferase (sialDrGalT) under the control of the Gpa2 promoter and LDN was synthesised in plants by co-expression of *Caenorhabditis elegans* β1,4-*N*-acetylgalactosaminyltransferase (CeGalNAcT) in the pBINPLUS plant expression vector [9], [24]. As positive controls we used a *Drosophila melanogaster* core α1,6-FucT (DmFucT8) and a hybrid *Tetraodon nigriviridus* α1,3-FucT IXa, containing the CTS domain of *R. norvegicus* α2,6-sialyltransferase (sialTnFucT9a). Both genes were inserted in the pBINPLUS plant expression vector to allow synthesis of, respectively, LeX and LDN-F [9], [24], [25]. *Homo sapiens* β1,2-*N*-acetylglucosaminyltransferase (HsGnT) II and *N. benthamiana* β-*N*-hexosaminidases (NbHEXO) were used to yield various glycan acceptors for DF formation. HsGnTII was amplified using illustra PuReTaq Ready-To-Go PCR Beads (GE Healthcare) from human MGAT2 gene cDNA clone (Sino Biological) flanking BspHI and BsrGI restriction sites and cloned into the plant expression vector (Supplemental Table 1). NbHEXO3 was reamplified with flanking BspHI and KpnI restriction sites, from a NbHEXO3-RFP construct (kindly provided by Dr. Richard Strasser from the University of Natural Resources and Life Sciences, Vienna) [26] (Supplemental Table 1). Reamplified NbHEXO3 was cloned into the plant expression vector by flanking restriction sites.

Unless stated differently, all constructs were driven by the Cauliflower mosaic virus 35S promoter with duplicated enhancer (d35S) and the *Agrobacterium tumefaciens* nopaline synthase transcription terminator (Tnos) in the plant expression vector pHYG [27]. Furthermore, to boost translation a 5' leader sequence of the alfalfa mosaic virus RNA 4 (AIMV) was included between the promoter and the construct. To enhance expression, the P19 silencing suppressor from tomato bushy stunt virus pBIN61 was co-infiltrated in all experiments, unless stated differently [28]. All constructs were transformed into *A. tumefaciens* strain MOG101 for plant expression.

Recombinant protein production and isolation

A. tumefaciens clones were cultured at 28 °C/250 rpm in LB medium (10 g/L peptone140, 10 g/L NaCl, 5 g/L yeast, pH 7.0), with 50 μg/mL kanamycin and 20 μM acetosyringone. After 16 hrs, bacteria were centrifuged for 15 min/2880 x*g* and subsequently resuspended in MMA (1.95 g/L MES, 20 g/L sucrose, 5 g/L MS-salts, pH 5.6) with 200 μM acetosyringone to increase *A. tumefaciens* transformation efficiency. *A. tumefaciens* cultures of different constructs were mixed for co-expression. The final optical density (OD) of each *A.*

tumefaciens culture in the mixture varied between 0.1 and 0.5 depending on the construct. The two youngest fully expanded leaves of 4 to 6 week old WT or ΔXT/FT *N. benthamiana* plants were infiltrated at the abaxial side [16]. *N. benthamiana* plants were grown in a controlled greenhouse compartment (UNIFARM, Wageningen). 5 to 6 days post infiltration (dpi) leaves were harvested and total apoplast proteins were isolated via apoplast wash. Harvested leaves were submerged in extraction buffer (50 mM phosphate buffer, pH 8.0, containing 0.1 M NaCl and 0,1 % v/v Tween-20), a vacuum was applied and after 5 min slowly released to infiltrate the leaves with buffer. The apoplast fluid was extracted from the leaves by centrifugation for 10 min/2000 x*g*. The protein concentration of the apoplast fluid was determined by the Pierce Bicinchoninic Acid Protein Assay (BCA, Fisher Scientific).

Protein purification

Prior to purification, extracted apoplast fluids were passed through Sephadex G25 chromatography columns to exchange extraction buffer for binding buffer (10 mM Sørensen's phosphate buffer, pH 6.0, containing 0.1 M NaCl) and subsequently clarified by centrifugation for 5 min/16.000 x*g* at 4 °C. Proteins were then bound to Pierce Strong Cation Exchange Mini Spin Columns (Fisher Scientific). Omega-1 was eluted with a 50 mM Tris-HCl buffer, pH 9.0, containing 2 M NaCl, whereas kappa-5 was eluted with binding buffer containing 2 M NaCl. Column loading, washing and elution was done by centrifugation for 5 min/2000 x*g*. After elution samples were dialysed against PBS. The protein concentration of the purified proteins was determined by the BCA (Fisher Scientific). Aliquots of different purification steps and eluted omega-1 or kappa-5 were separated under reducing conditions by SDS-PAGE on a 12% Bis-Tris gel (Invitrogen) and subsequently stained with Coomassie brilliant blue staining.

Lectin binding assay

Fucose binding *Aleuria aurantia* lectin (AAL, Bio-Connet) was used to screen for fucosylation in general, whereas *Pholiota squarrosa* lectin (PhoSL) was used to screening for core α1,6-fucosylation. PhoSL was kindly provided by Dr. A. Varrot (Université Grenoble Alpes, Grenoble, France) and was biotinylated in house following manufactures protocol (Pierce) [29]. Soybean agglutinin (SBA, Bio-Connect) was used to determine fucosylation of LDN, since SBA binding to the GalNAc residue in LDN is inhibited by fucosylation of LDN. Microtiter plates were coated overnight (o/n) with apoplast fluids or purified protein in PBS at a protein concentration of 1 µg/mL for binding by AAL and 10 µg/mL for binding by PhoSL or SBA. Plates were blocked with carbohydrate-free blocking buffer (Vector Laboratories) for 1 hr/room temperature (RT). Plates were then incubated for 1 hr/RT with biotinylated lectin at a concentration of 2 µg/mL AAL, 1.8 µg/mL PhoSL or 5 µg/mL SBA. Subsequently, plates were incubated with avidin-HRP (eBioscience) for 30 min/

RT. After every incubation step the microtiter plate was washed 5 times with PBST (PBS containing 0.05 % v/v Tween-20). Lectin binding was visualised by TMB substrate (Fisher Scientific) and absorbance was measured at a wavelength of 450 nm while using 655 nm as reference filter.

SDS-PAGE and western blot

For screening of core α1,3-fucosylation, purified omega-1 samples were treated with Peptide:N-glycosidase F (PNGase F, New England Biolabs) according to manufacturer's protocol. Deglycosylation by PNGase F was then analysed on a 12 % Bis-Tris SDS-PAGE gel (Fisher Scientific) followed by Oriole staining (BioRad). For glycan detection by western blot, apoplast fluids or purified kappa-5 samples were run on a 12 % Bis-Tris gel under reducing conditions and subsequently transferred to a PVDF membrane by wet blotting. After blotting the membrane was blocked for 1 hr/RT or o/n at 4 $^\circ$C with 5 % w/v bovine serum albumin in PBST (PBS containing 0.1 % v/v Tween-20). Next, the membrane was incubated for 1 hr/RT or o/n at 4 $^\circ$C with primary antibodies: mouse IgM 290-2E6 for LDN-F detection [30], mouse IgM 291-5D5 for F-LDN detection [31] or mouse IgM 128-1E7 for F-LDN-F [31] each diluted 1:500 or rat IgM 5750 for LeX detection [32] diluted 1:1000. Subsequently, the membrane was incubated for 1 hr/RT with 0.75 μg/mL secondary antibody, HRP labelled donkey anti-mouse IgM (Jackson ImmunoResearch laboratories). After every incubation step the membrane was washed 5 times with PBST. The anti-HRP labelled antibodies were detected with a 1:1 SuperSignal West Femto:Dura substrate (Fisher Scientific) in the G:BOX Chemi System (Syngene).

Matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry

The glycan composition was analysed with matrix-assisted laser desorption/ionisation time-of-flight (MALDI-TOF) mass spectrometry (MS). Thereto, 5 μg purified omega-1 or kappa-5 was denatured by incubation with 1.3 % w/v SDS and 0.1 % v/v β-mercaptoethanol for 10 min/95 $^\circ$C. SDS was neutralised by adding 1.3 % v/v NP-40, after which proteins were digested with trypsin (Sigma-Aldrich) linked to NHS-activated Sepharose (GE Healthcare) o/n at 37 $^\circ$C. Trypsin beads were removed by centrifugation for 3 min/400 rpm and the supernatant was transferred and dried under vacuum. Dried samples were dissolved in 1 M sodium acetate pH 4.5 and sonicated for 5 min. Subsequently the N-glycans were released by incubation with PNGase A (Roche Diagnostics) 24-48 hrs/37 $^\circ$C. The released N-glycans were separated from peptides using C18 Bakerbond™ SPE cartridges (JT Baker) and subsequent binding of the N-glycans to Extract Clean™ Carbo SPE columns. Eluates were dried o/n under vacuum and reconstituted in MQ. N-glycans were then labelled with anthranilic acid (Sigma-Aldrich), by incubation for 2 hrs/65 $^\circ$C with labelling mix (DMSO:Acetic acid (10:3) with 48 mg/mL 2-aminobenzoic acid (Sigma-Aldrich) and 107.23 mg/mL 2-picoline-borane complex (Sigma-Aldrich)). Samples were cooled to RT

and subsequently desalted by hydrophilic interaction chromatography on Biogel P10 (BioRad). Samples were cleaned using C18 ZipTip (Millipore) and eluted by 2 µL of matrix solution (20 mg/mL 2,5-dihydroxybenzoic acid in 50 % acetonitrile containing 0.1 % v/v TFA) and subsequently pipetted on a polished steel target plate. MALDI-TOF mass spectra were obtained using an Ultraflex II mass spectrometer (Bruker Daltonics) in negative-ion reflection mode as previously described [33].

To confirm the presence of specific glycan motifs, N-glycans were treated with glycosidases, prior to ZipTip C18 clean-up (Supplemental Table 1). The following glycosidases were used according to the suppliers protocols: α1,3/4-fucosidase from *Xanthomonas* sp. (Sigma-Aldrich), β1,3/6-galactosidase 80120 from *Xanthomonas manihotis* (Prozyme)*,* β1,4/6-galactosidase 5012 from Jack bean (Prozyme), α1,2/4/6-fucosidase O from *Omnitrophica* (New England Biolabs), β-N-acetyl-hexosaminidase from *Streptomyces plicatus* (New England Biolabs) and/or β-N-acetyl-glucosaminidase from *Streptococcus pneumoniae* (New England Biolabs).

Acknowledgements

3

References

[1] P. J. Hotez *et al.*, "The Global Burden of Disease Study 2010: Interpretation and Implications for the Neglected Tropical Diseases," *PLoS Negl. Trop. Dis.*, vol. 8, no. 7, p. e2865, 2014.

[2] C. H. Smit *et al.*, "Glycomic analysis of life stages of the human parasite Schistosoma mansoni reveals developmental expression profiles of functional and antigenic glycan motifs.," *Mol. Cell. Proteomics*, vol. 14, no. 7, pp. 1750–1769, 2015.

[3] C. H. Hokke and A. van Diepen, "Helminth glycomics – glycan repertoires and host-parasite interactions," *Mol. Biochem. Parasitol.*, vol. 215, pp. 47–57, 2017.

[4] A. van Diepen, N. S. J. Van Der Velden, C. H. Smit, M. H. J. Meevissen, and C. H. Hokke, "Parasite glycans and antibody-mediated immune responses in Schistosoma infection," *Parasitology*, vol. 139, no. 9, pp. 1219–1230, 2012.

[5] N. S. Prasanphanich, M. L. Mickum, J. Heimburg-Molinaro, and R. D. Cummings, "Glycoconjugates in host-helminth interactions," *Front. Immunol.*, vol. 4, no. 240, pp. 1–22, 2013.

[6] T. P. Yoshino *et al.*, "Excreted/secreted Schistosoma mansoni venom allergen-like 9 (SmVAL9) modulates host extracellular matrix remodelling gene expression," *Int. J. Parasitol.*, vol. 44, no. 8, pp. 551–563, 2014.

[7] D. van der Kleij *et al.*, "Triggering of Innate Immune Responses by Schistosome Egg Glycolipids and Their Carbohydrate Epitope GalNAcβ1-4(Fucα1-2Fucα1-3)GlcNAc," *J. Infect. Dis.*, vol. 185, no. 4, pp. 531–539, 2002.

[8] B. Everts *et al.*, "Schistosome-derived omega-1 drives Th2 polarization by suppressing protein synthesis following internalization by the mannose receptor," *J. Exp. Med.*, vol. 209, no. 10, pp. 1753–1767, 2012.

[9] R. H. P. Wilbers *et al.*, "Production and glyco-engineering of immunomodulatory helminth glycoproteins in plants," *Sci. Rep.*, vol. 7, p. 45910, 2017.

[10] M. H. J. Meevissen *et al.*, "Specific glycan elements determine differential binding of individual egg glycoproteins of the human parasite Schistosoma mansoni by host C-type lectin receptors," *Int. J. Parasitol.*, vol. 42, no. 3, pp. 269–277, 2012.

[11] E. van Liempt *et al.*, "Specificity of DC-SIGN for mannose- and fucose-containing glycans," *FEBS Letters*, vol. 580, no. 26. pp. 6123–6131, 2006.

[12] I. van Die *et al.*, "The dendritic cell-specific C-type lectin DC-SIGN is a receptor for Schistosoma mansoni egg antigens and recognizes the glycan antigen Lewis x," *Glycobiology*, vol. 13, no. 6, pp. 471–478, 2003.

[13] S. I. Gringhuis, J. den Dunnen, M. Litjens, M. van der Vlist, and T. B. H. Geijtenbeek, "Carbohydrate-specific signaling through the DC-SIGN signalosome tailors immunity to Mycobacterium tuberculosis, HIV-1 and Helicobacter pylori," *Nat. Immunol.*, vol. 10, no. 10, pp. 1081–1088, 2009.

[14] M. L. Mickum, T. Rojsajjakul, Y. Yu, and R. D. Cummings, "Schistosoma mansoni α1,3-fucosyltransferase-F generates the Lewis X antigen," *Glycobiology*, vol. 26, no. 3. Glycobiology, pp. 270–285, 2015.

[15] N. A. Peterson, T. K. Anderson, and T. P. Yoshino, "In Silico Analysis of the Fucosylation-Associated Genome of the Human Blood Fluke Schistosoma mansoni: Cloning and Characterization of the Fucosyltransferase Multigene Family," *PLoS One*, vol. 8, no. 5, pp. 1–15, 2013.

[16] R. Strasser *et al.*, "Generation of glyco-engineered Nicotiana benthamiana for the production of monoclonal antibodies with a homogeneous human-like N-glycan structure," *Plant Biotechnology J.*, vol. 6, no. 4. pp. 392–402, 2008.

[17] H. Bakker *et al.*, "Galactose-extended glycans of antibodies produced by transgenic plants," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 98, no. 5, pp. 2899–2904, 2001.

[18] R. B. Parekh, "Site-specific protein glycosylation," *Adv. Drug Deliv. Rev.*, vol. 13, no. 3, pp. 251–266, 1994.

[19] J. Wilhelm *et al.*, "Alterations in the domain structure of ttssue-type p1asminogen activator change the nature of asparagine olycosyution," *Bio/Technology*, vol. 8, no. 4, pp. 321–325, 1990.

[20] S. O. Lee, J. M. Connolly, D. Ramirez-Soto, and R. D. Poretz, "The polypeptide of immunoglobulin G influences its galactosylation in vivo," *J. Biol. Chem.*, vol. 265, no. 10, pp. 5833–5839, 1990.

[21] J. M. Fitzpatrick *et al.*, "Anti-schistosomal intervention targets identified by lifecycle transcriptomic analyses," *PLoS Negl. Trop. Dis.*, vol. 3, no. 11, p. e543, 2009.

[22] C. Breton, R. Oriol, and A. Imberty, "Conserved structural features in eukaryotic and prokaryotic fucosyltransferases," *Glycobiology*, vol. 8, no. 1, pp. 87–94, 1998.

[23] R. Oriol, R. Mollicone, A. Cailleau, L. Balanzino, and C. Breton, "Divergent evolution of fucosyltransferase genes from vertebrates, invertebrates, and bacteria," *Glycobiology*, vol. 9, no. 4, pp. 323–334, 1999.

[24] W. J. van Engelen, F. A., Molthoff, J. W., Conner, A. J., Nap, J. P., Pereira, A., & Stiekema, "pBINPLUS/ARS: An improved plant transformation vector based on pBIN19," *Transgenic Res.*, vol. 4, no. 4, pp. 288–290, 1995.

[25] R. H. P. Wilbers *et al.*, "The N-glycan on Asn54 affects the atypical N-glycan composition of plant-produced interleukin-22, but does not influence its activity," *Plant Biotechnol. J.*, vol. 14, no. 2, pp. 670–681, 2016.

[26] Y. J. Shin *et al.*, "Reduced paucimannosidic N-glycan formation by suppression of a specific β-hexosaminidase from Nicotiana benthamiana," *Plant Biotechnology J.*, vol. 15, no. 2. pp. 197–206, 2017.

[27] L. B. Westerhof *et al.*, "3D Domain Swapping Causes Extensive Multimerisation of Human Interleukin-10 When Expressed In Planta," *PLoS One*, vol. 7, no. 10, pp. 1–10, 2012.

[28] O. Voinnet, S. Rivas, P. Mestre, and D. Baulcombe, "An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus," *Plant J.*, vol. 33, no. 5, pp. 949–956, 2003.

[29] A. Cabanettes, L. Perkams, C. Spies, C. Unverzagt, and A. Varrot, "Recognition of Complex Core-Fucosylated N-Glycans by a Mini Lectin," *Angew. Chemie - Int. Ed.*, vol. 57, no. 32, pp. 10178–10181, 2018.

[30] A. v. Remoortere, C. H.Hokke, G. J. v. Dam, I. v. Die, A. M.Deelder, and D. H. v. d. Eijnden, "Various stages of Schistosoma express Lewisx, LacdiNAc, GalNAc 1-4 (Fuc 1-3)GlcNAc and GalNAc 1-4(Fuc 1-2Fuc 1-3)GlcNAc carbohydrate epitopes: detection with monoclonal antibodies that are characterized by enzymatically synthesized neoglycoproteins," *Glycobiology*, vol. 10, no. 6, pp. 601–609, 2000.

[31] M. L. M. Robijn, M. Wuhrer, D. Kornelis, A. M. Deelder, R. Geyer, and C. H. Hokke, "Mapping fucosylated epitopes on glycoproteins and glycolipids of Schistosoma mansoni cercariae, adult worms and eggs," *Parasitology*, vol. 130, no. 1, pp. 67–77, 2005.

[32] E. Hennen, T. Czopka, and A. Faissner, "Structurally distinct LewisX glycans distinguish subpopulations of neural stem/progenitor cells," *J. Biol. Chem.*, vol. 286, no. 18, pp. 16321–16331, 2011.

[33] M. H. J. Meevissen *et al.*, "Targeted glycoproteomic analysis reveals that kappa-5 is a major, uniquely glycosylated component of schistosoma mansoni egg antigens," *Mol. Cell. Proteomics*, vol. 10, no. 5, pp. M110-005710, 2011.

3

Supplemental Figures and Tables

SUPPLEMENTAL TABLE 1 | Primers. This table depicts primers used for amplification of HsGnTII and NbHEXO3. The part of the sequences that corresponds to the open reading frame is written in capital letters.
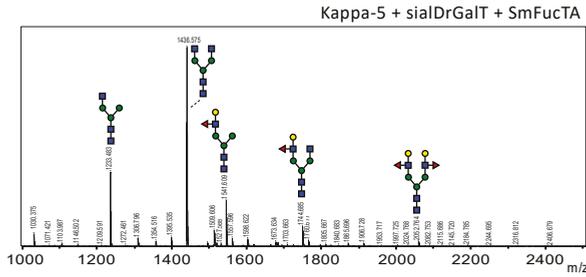
Primer name	Sequence 5' to 3'
HsGnTII – *Forward*	acggccgccagtgtgctgtcATGAGGTTCCGCATCTACAAACG
HsGnTII – *Reverse*	cgccagtgtgatggatatctgcatgtacaTCACTGCAGTCTTCTATAACTTTTACAGAG
NbHEXO3 – *Forward*	ggtcatgagcATGGGGAAGTTAGGATTCCG
NbHEXO3 – *Reverse*	gggtaccTTATTGCTGATAGCAAGAACCTG

SUPPLEMENTAL TABLE 2 | Glycosidases. This table depicts glycosidases used to screen for specific N-glycan motifs prior to MALDI-TOF MS analysis. The for this chapter relevant glycan specificity is indicated, in which the arrow indicates the site of cleavage.

Name of the glycosidase	Abbreviation used in this thesis	Used in screening for	Specificity
β1,4/6-galactosidase	β1,4-gal	LeX vs LeA	Galβ1↓4GlcNAcβ-R
β1,3/6-galactosidase	β1,3-gal	LeX vs LeA	Galβ1↓3GlcNAcβ-R
α1,3/4-fucosidase	fuc	LeX/LeA vs core	Galβ1-4(Fucα1↓3)GlcNAcβ-R Galβ1-3(Fucα1↓4)GlcNAcβ-R
		LDN-F vs core	GalNAcβ1-4(Fucα1↓3)GlcNAcβ-R
α1,2/4/6-fucosidase O	fuc-o	LeX/LeA vs core	R-GlcNAcβ1-4(Fucα1↓3)GlcNAcβ-AA
		LDN-F vs core (or F-LDN)	Fucα1↓3GalNAcβ1-4GlcNAcβ-R R-GlcNAcβ1-4(Fucα1↓3)GlcNAcβ-AA
β-N-acetyl-hexosaminidase	hexo	presence of fucosylated LDN	GalNAcβ1↓4GlcNAcβ-R
β-N-acetyl-glucosaminidase	gluc	presence of LDN or fucosylated LDN	GlcNAcβ1↓2Manα1-3,6-R
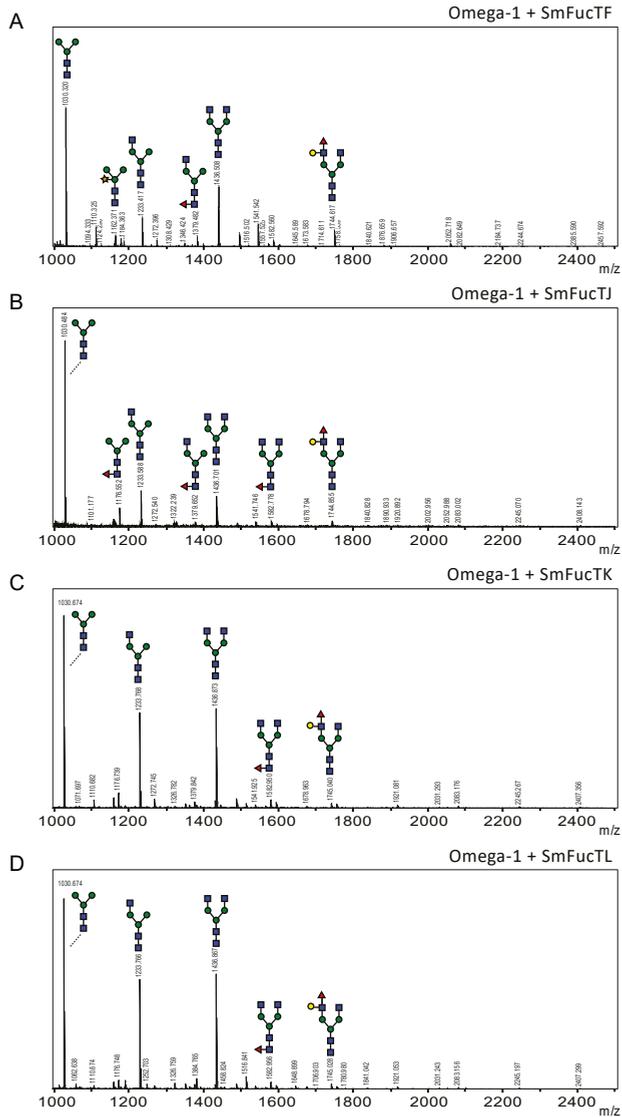
SUPPLEMENTAL FIGURE 1 | MALDI-TOF MS N-glycan analysis for core fucosylation of omega-1 N-glycans by fucosyltransferases of *Schistosoma mansoni* (SmFucTs). (A-B) N-glycan profile of purified omega-1 expressed in ΔXT/FT plants upon co-expression of P19 and SmFucTF (A) or SmFucTJ (B). (C-D) N-glycan profile of purified omega-1 expressed in ΔXT/FT plants upon co-expression of SmFucTK (C) or SmFucTL (D) without P19 co-expression.
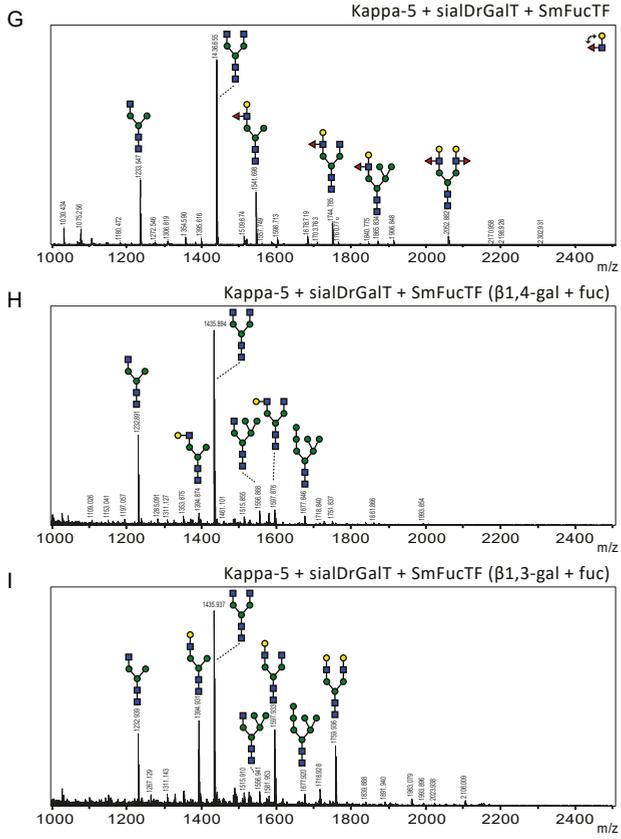
SUPPLEMENTAL FIGURE 2 | MALDI-TOF MS N-glycan analysis for LeX synthesis on omega-1 N-glycans by fucosyltransferase A of *Schistosoma mansoni* (SmFucTA). MALDI-TOF MS N-glycan profile of purified kappa-5 expressed in in ΔXT/FT *Nicotiana benthamiana* plants upon co-expression of SmFucTA. When a peak represents multiple N-glycan structures of identical mass, the peak indicates the major N-glycan present based on enzymatic digestions (for details see Supplemental Figure 3).

SUPPLEMENTAL FIGURE 3 | LeX vs LeA synthesis. N-glycan profiles of purified kappa-5 upon co-expression of sialDrGalT and SmFucTD (A-C), SmFucTE (D-F), SmFucTF (G-I) or SmFucTA (J-M). Released N-glycans were treated with the exoglycosidases α1,2/4/6-fucosidase O from *Omnitrophica* (fuc-o), α1,3/4-fucosidase from *Xanthomonas* sp. (fuc), β1,3/6-galactosidase from *Xanthomonas manihotis* (β1,3-gal) or β1,4/6-galactosidase 5012 from Jack bean (β1,4-gal). If in the left corner of the N-glycan profile a LeX glycan motif with an arrow is indicated, then the LeX in the profile could also be LeA.

SUPPLEMENTAL FIGURE 3 | LeX vs LeA synthesis. Continued

SUPPLEMENTAL FIGURE 3 | LeX vs LeA synthesis. Continued

SUPPLEMENTAL FIGURE 3 | LeX vs LeA synthesis. Continued

SUPPLEMENTAL FIGURE 4 | MALDI-TOF MS N-glycan analysis of kappa-5. (A-B) MALDI-TOF MS N-glycan profile of purified kappa-5 expressed in *Nicotiana benthamiana* plants (A) and upon co-expression of CeGalNAcT (B). When a peak represents multiple N-glycan structures of identical mass, the number of sugar residues of which the position on the N-glycan is not clear is indicated above the glycan structure. The possible positions of these sugar residues on the N-glycan are indicated between brackets.
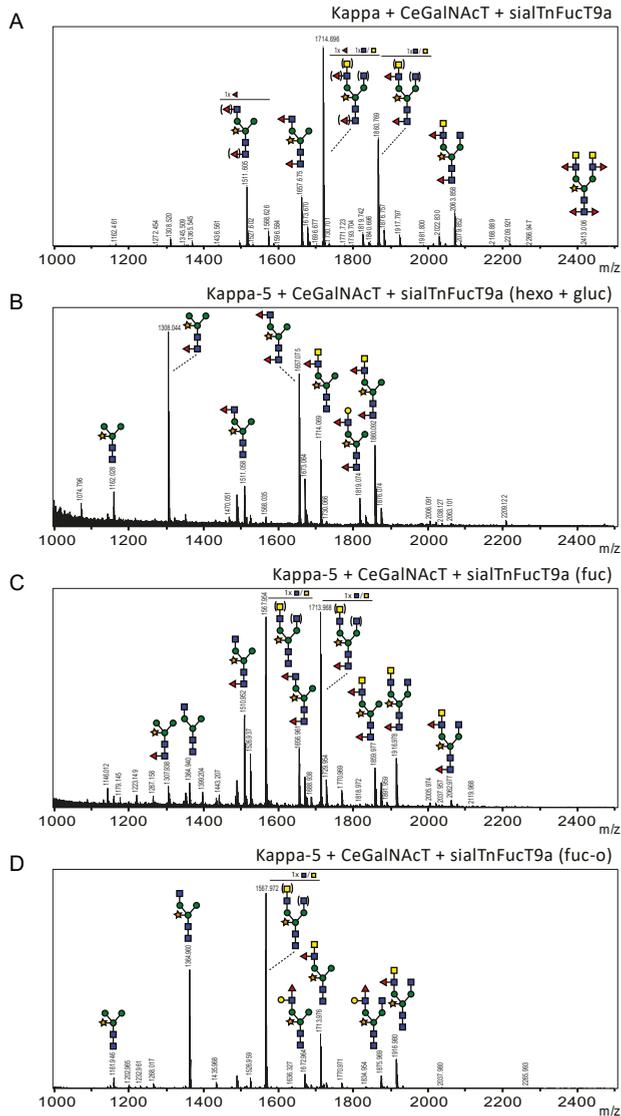
SUPPLEMENTAL FIGURE 5 | MALDI-TOF MS N-glycan analysis of enzyme digested samples for detection of LDN fucosylation by fucosyltransferases of *Schistosoma mansoni* (SmFucTs). (A-D) MALDI-TOF MS N-glycan profiles of purified kappa-5 upon co-expression of CeGalNAcT and sialTnFucT9a (A-D), SmFucTD (E-H), SmFucTE (I-L) or SmFucTF (M-P). Released N-glycans were treated with the exoglycosidases β-N-acetyl-hexosaminidase from *Streptomyces plicatus* (hexo) and β-N-acetyl-glucosaminidase from *Streptococcus pneumoniae* glucosaminidase (gluc) or α1,3/4-fucosidase from *Xanthomonas* sp. (fuc) or α1,2/4/6-fucosidase O from *Omnitrophica* (fuc-o). Sugar residues for which the position is not clear (prior to enzymatic digestion) are placed between brackets. When a peak represents multiple N-glycan structures of identical mass, the number of sugar residues of which the position on the N-glycan is not clear is indicated above the glycan structure. The possible positions of these sugar residues on the N-glycan are indicated between brackets.

SUPPLEMENTAL FIGURE 5 | MALDI-TOF MS N-glycan analysis of enzyme digested samples for detection of LDN fucosylation by fucosyltransferases of *Schistosoma mansoni* (SmFucTs). Continued

SUPPLEMENTAL FIGURE 5 | MALDI-TOF MS N-glycan analysis of enzyme digested samples for detection of LDN fucosylation by fucosyltransferases of *Schistosoma mansoni* (SmFucTs). Continued
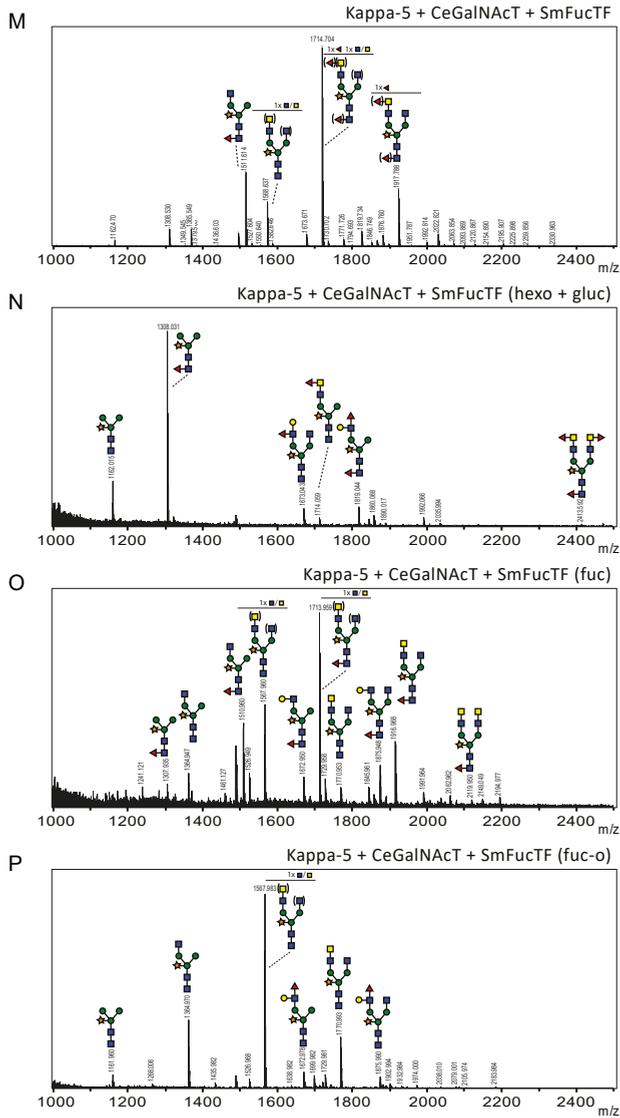
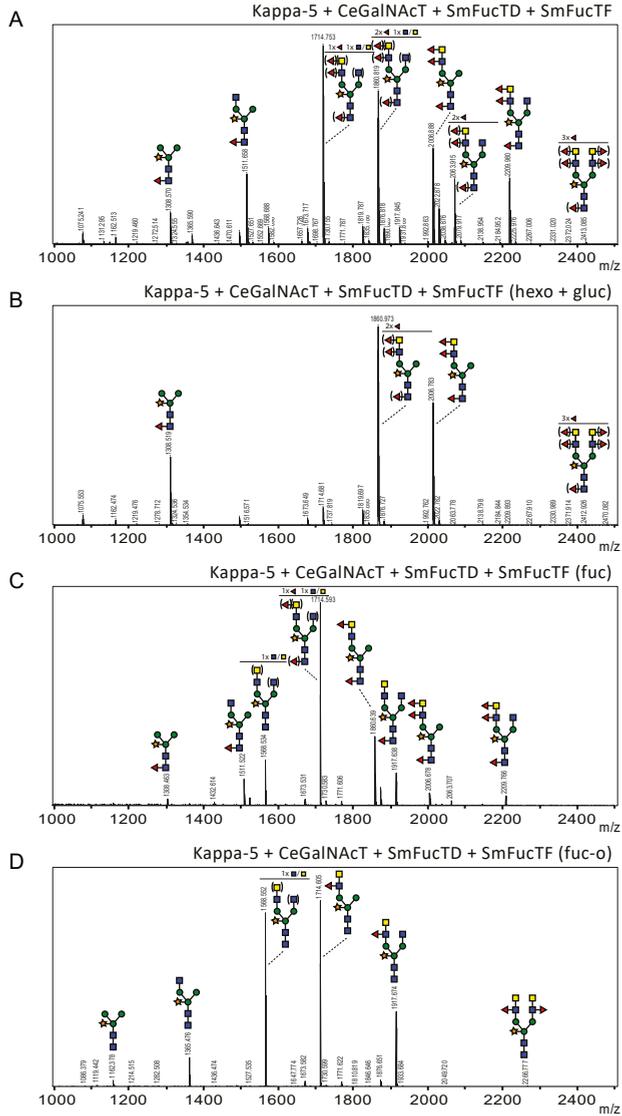**SUPPLEMENTAL FIGURE 5 | MALDI-TOF MS N-glycan analysis of enzyme digested samples for detection of LDN fucosylation by fucosyltransferases of *Schistosoma mansoni* (SmFucTs).** Continued

SUPPLEMENTAL FIGURE 6 | MALDI-TOF MS N-glycan analysis of enzyme digested samples for detection of F-LDN-F synthesis by fucosyltransferases of *Schistosoma mansoni* (SmFucTs). *(A-D)* MALDI-TOF MS N-glycan profiles of purified kappa-5 upon co-expression of CeGalNAcT, SmFucTF and SmFucTD (A-D) or SmFucTE (E-H). Released N-glycans were treated with the exoglycosidases β-N-acetyl-hexosaminidase from *Streptomyces plicatus* (hexo) and β-N-acetyl-glucosaminidase from *Streptococcus pneumoniae* glucosaminidase (gluc) or α1,3/4-fucosidase from *Xanthomonas* sp. (fuc) or α1,2/4/6-fucosidase O from *Omnitrophica* (fuc-o). Sugar residues for which the position is not clear (prior to enzymatic digestion) are placed between brackets. When a peak represents multiple N-glycan structures of identical mass, the number of sugar residues of which the position on the N-glycan is not clear is indicated above the glycan structure. The possible positions of these sugar residues on the N-glycan are indicated between brackets.
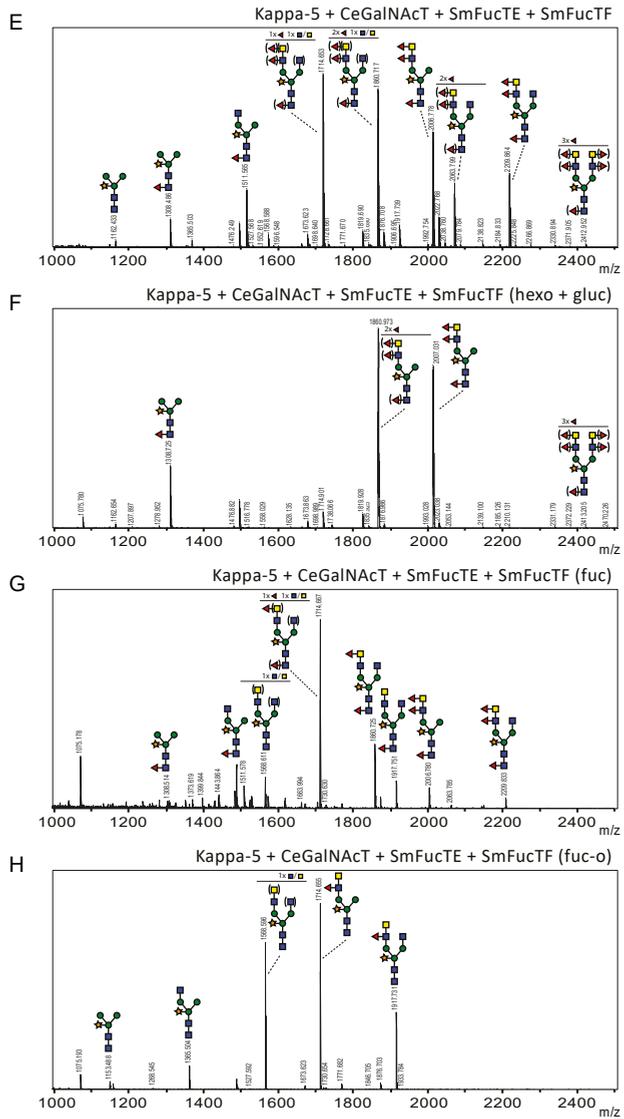
SUPPLEMENTAL FIGURE 6 | MALDI-TOF MS N-glycan analysis of enzyme digested samples for detection of F-LDN-F synthesis by fucosyltransferases of *Schistosoma mansoni* (SmFucTs). *Continued*

Chapter 4

Identification and functional characterisation of *Nicotiana benthamiana* fucosyltransferases

Kim van Noort, Billy Nugraha, Martijn Holterman,
Arjen Schots and Ruud H. P. Wilbers

Laboratory of Nematology, Wageningen University and Research, Wageningen, the Netherlands.

Abstract

Nicotiana benthamiana is a promising platform for the production of biopharmaceutical proteins, with many advantages over well-established production platforms. One of the reasons plants still lack behind as a production platform is that their glycosylation pathway differs from humans. Plant N-glycans carry the sugar residues core α1,3-fucose and β1,2-xylose, which are not found on human N-glycans. Antibodies against these plant sugar residues may circulate in the human blood and although these antibodies seem clinically irrelevant, they are a public cause of concern. Therefore, plants have been developed without these plant sugar residues with available techniques, such as RNA interference, transcription activator-like effector nucleases (TALENs) and random mutagenesis. The drawback of these studies is that all or just a few reported core fucosyltransferase (FucT) and/ or xylosyltransferase (XylT) encoding genes are targeted, before these genes were characterised. *N. benthamiana* is an allotetraploid plant, consequently multiple allelic variants of a gene can be present. Knowledge of the endogenous XylT and FucT encoding genes and their alleles could save time and effort during genome editing. In this study, we identified *N. benthamiana* FucT (NbFucT) sequences and assessed the activity of all retrieved full-length sequences. Ten different NbFucT sequences were retrieved, of which six encode full-length proteins. Of these six proteins only three were able to introduce a core α1,3-fucose. This knowledge allows the design of specific guide RNAs, which need to target only these three sequences for the production of knockout plants with CRISPR-Cas9 or similar technologies. Thereby saving time, effort and resources. Therefore, we anticipate our study to be an example for future research on generating genome-edited plants.

Introduction

Biopharmaceuticals are drugs produced in biological systems, such as mammalian and bacterial cell cultures, or extracted from tissue, such as insulin from bovine or porcine pancreas. Already for 30 years recombinant proteins are produced in plants. Plants have emerged as a viable alternative to well-established biopharmaceutical protein production systems, such as mammalian and bacterial cell cultures [1]. The use of plants over these well-established protein production systems has several advantages; plant-based protein production is low in costs, easy to set up, scale up and adjust [2]. Furthermore, plants are not likely to carry mammalian pathogens, can perform post-translational modifications and introduce homogenous glycosylation patterns. Despite all these advantages it took until May 2012 to get the first plant-produced biopharmaceutical, Elelyso (taliglucerase alfa), approved by the United States Food and Drug Administration (FDA). This enzyme is produced in carrot cells and can treat type 1 Gaucher's disease by replacing its non-functional human equivalent [3]. Next, in 2014 the tobacco produced anti-Ebola antibody cocktail ZMapp was used as treatment for Ebola infected patients. Currently, several plant-produced biopharmaceuticals are in clinical trials [4], [5]. This shows that plants are gradually becoming accepted as production system for biopharmaceutical proteins.

One of the reasons plants still lack behind as a production system is the difference in plant and human N-glycosylation. N-glycosylation is the linkage of an oligosaccharide chain (glycan) composed of different sugar residues, to a protein on the nitrogen atom of asparagine in the protein sequence asparagine-X-serine/threonine. In the plant glycosylation pathway, core fucosyltransferases (FucTs) and xylosyltransferases (XylTs) transfer respectively an α1,3-fucose and a β1,2-xylose to the core of the N-glycans on glycoproteins. These typical plant sugar residues are not found in human N-glycans. Moreover, IgE antibodies against these plant sugar residues can circulate in the human blood. These cross-reactive anti-carbohydrate antibodies seem to be clinically irrelevant due to low affinity binding of IgE to the plant N-glycan residues, however they do reflect negatively on plants as a protein production system [6].

Next to the presence of plant sugar residues, typical residues such as terminal β1,4-galactose and sialic acid that are often found in human N-glycans, are absent on plant N-glycans. Since plants are tolerant to changes in their glycosylation pathway, the plant glycosylation pathway can be altered by introduction of human glycosyltransferases, such as β1,4-galactosyltransferase and α2,3/6-sialyltransferase [7]–[11]. Furthermore, the human core α1,6-fucose and multiple N-glycan branches can be synthesised in plants [12], [13]. While various exogenous glycosyltransferases can be introduced to humanise the plant glycosylation pathway, endogenous core FucTs and XylTs are still present. To produce glycoproteins with human-like N-glycans these endogenous glycosyltransferases have to be knocked out.

4

Via random T-DNA insertion the first *Arabidopsis thaliana* endogenous core FucTs (AtFucTs) and XylT (AtXylT) mutant lines were created [14]. Based on these lines Strasser and colleagues [15]generated the first AtFucTs and AtXylT knockout lines by generation of homozygous plants with T-DNA insertions in AtFucTA, AtFucTB and AtXylT. The single, double and triple knockout lines were viable and did not show a phenotype under normal growth conditions [15]. Antibody production in the triple knockout plants resulted in correct assembly and unaltered antigen binding specificity of the 2G12 anti-HIV IgG monoclonal antibody containing complex N-glycan structures lacking plant core α1,3-fucose and β1,2-xylose [16]. After establishment of recombinant protein production with humanised N-glycans in *A. thaliana*, the focus shifted to *Nicotiana benthamiana*, an important plant species for transient protein production. ΔXT/FT *N. benthamiana* plants were generated in which the endogenous *N. benthamiana* core FucTs (NbFucTs) and XylTs (NbXylTs) were targeted by RNA interference [17]. Although, NbFucTs and NbXylTs are not completely knocked out in ΔXT/FT plants, N-glycans on the monoclonal antibody 2G12 produced in these plants lacked core α1,3-fucose and β1,2-xylose. To generate full *N. benthamiana* knockout plants, point mutations were introduced in the five NbFucT and two NbXylT genes by random mutagenesis with ethylmethanesulfonate. These lines were back-crossed with the ΔXT/FT line of Strasser and colleagues [17], creating a full knockout plant, patented in 2015 [18]. In another approach a knockout line was created by transcription activator-like effector nucleases (TALENs), targeting two NbFucT and two NbXylT genes. Transient expression of the recombinant rituximab antibody still resulted in a considerable amount of core α1,3-fucose containing N-glycans showing that creation of a full knockout was unsuccessful [19]. In all these studies different methods were used to abolish NbFucT and NbXylT activity in *N. benthamiana* plants.

In most knockout studies all known NbFucT sequences were targeted, whereas not all reported NbFucT sequences have to encode an active NbFucT protein. However, next to the available *N. benthamiana* sequences in the Sol Genomics Network (SGN) database nothing is known about NbFucT activity or enzyme specificity [20]. In this study, we focussed on a combined analysis of NbFucT sequences and the functionality of the encoded enzymes. Furthermore, an overview was created of the different names used in papers, patents and databases for the reported NbFucT sequences (Table 1). Based on the available sequence data in the SGN database, primers were designed for NbFucT amplification from cDNA. The six retrieved sequences encoding full-length proteins were transiently co-expressed in ΔXT/FT plants with carrier glycoproteins. Subsequently, N-glycans on the carrier glycoproteins were analysed for addition of a core α1,3-fucose, indicating NbFucT activity. Our data suggest that only three of the five reported NbFucT sequences have to be targeted to create a *N. benthamiana* plant without core α1,3-fucosylated N-glycans. With this knowledge time and effort can be saved in the generation of NbFucT knockout plants. This study can function as an example for future research on generating genome-edited plants.

TABLE 1 | Names used for reported *Nicotiana benthamiana* core α1,3-fucosyltransferases (NbFucTs). NbFucTs are described in different databases, patents and literature. This table gives an overview of the different names used in the database Sol Genomics Network (SGN), patents (patent 1: Mathis and colleagues [21]; patent 2: Weterings and van Eldik [18]) and literature. ~ indicates that FucT2, EF562631 is almost similar to NbFucT7, minus the mutations: p.M1LextM-17, p.A2P, p.Q3L, p.L4F, p.S7L, p.R9V, p.C10I, p.P11A, p.K12E, p.A86S (described following the protein mutation nomenclature of den Dunnen and Antonarakis [22]).

Sol Genomics Network (SGN)	Abbreviation in this thesis	Patent 1 and Li and colleagues [19]	Patent 2 (NbFucT, NbFucT with mutation)
Niben101Scf17626g00001.1	NbFucT1		FucTD, FucT003
Niben101Scf02631g00007.1	NbFucT7	~FucT2, EF562631.1	FucTB, FucT006
Niben101Scf05447g03009.1	NbFucT9		FucTE, FucT009
Niben101Scf05494g01011.1	NbFucT11		FucTC, FucT007
Niben101Scf01272g00014.1	NbFucT14	FucT1, EF562630.1	FucTA, FucT004

Results

Sequence analysis of *Nicotiana benthamiana* core α1,3-fucosyltransferases

To analyse the differences between the five NbFucT protein sequences reported in the SGN database (NbFucT1: Niben101Scf17626g00001.1; NbFucT7: Niben101Scf02631g00007.1; NbFucT9: Niben101Scf05447g03009.1; NbFucT11: Niben101Scf05494g01011.1; and NbFucT14: Niben101Scf01272g00014.1) an amino acid alignment was made (Figure 1). In this alignment FucT sequence motifs were indicated, such as the fucose binding motif V (following renumbering of Mollicone and colleagues [23]) and two specific plant motifs (the plant core FucT motif and the 1st cluster motif) [24]–[26]. Characteristics and differences between the NbFucT amino acid sequences are indicated in Table 2.

Since, FucTs are type II membrane proteins, the transmembrane domain (TMD) of the putative NbFucT coding sequences was predicted by the TMHMM server v2.0 [27]. Based on this prediction the length of the cytoplasmic tail (CT) and the TMD were calculated (Table 2). Remarkable was the apparent lack of a TMD for three of the five SGN NbFucT protein sequences. For NbFucT1 and NbFucT7 this was probably due to a deletion observed at the N-terminus of both protein sequences in comparison to NbFucT9 and NbFucT14, respectively. These deletions and the apparent absence of the TMD indicated incomplete protein sequences. The 4 nucleotide (nt) deletion observed for NbFucT11 in comparison to NbFucT9 results in a frameshift and introduction of an early stop codon. Since, this frameshift could be introduced by sequencing or prediction errors, both parts of NbFucT11 were compared to NbFucT9. For the first part of NbFucT11 no TMD was predicted due to a mutation from isoleucine to leucine at the N-terminus of the protein in comparison to NbFucT9. For the second part of NbFucT11, no remarkable differences

4

were observed compared to NbFucT9 except for single nucleotide polymorphisms (SNPs). Remarkable for NbFucT14 was the 33 amino acids (aa) duplication from methionine at position 342 to lysine at position 376 in comparison to the other NbFucTs in the alignment. This duplication seems to be induced during sequencing, since the insertion starts with amino acid "X" that is introduced when an ambiguity is present (represented by n) in the nucleotide sequence. At last a 40 aa deletion was observed for NbFucT7 at the C-terminus.

In the alignment the first ~100 aa at the N-terminus of the NbFucT amino acid sequences seemed to be most variable (Figure 1). This variable part codes for the CTS domain (<u>C</u>T, <u>T</u>MD and <u>S</u>tem region), which is known to be hypervariable between different glycosyltransferases. After the CTS domain the sequences become more similar and various FucT sequence motifs were indicated. The first motif indicated is the DxD motif, in plants present as SSD, expected to be involved in binding of Mn^{2+} [28]. After this the plant core α1,3-FucT specific motif (in short pcf motif) SNCG(A/G)ARN is indicated. This pcf motif is found within the 1st cluster motif, which combined with motif V is involved in guanosine diphosphate β-L-fucose (GDP-fucose) binding [26], [29]. Within these two motifs, mutations were observed in the NbFucT amino acid sequences at nine positions. Furthermore, an aspartic acid, tyrosine and valine deletion was observed for NbFucT9 in motif V. The four cysteine residues (indicated with *) that are conserved in most of the α1,3-FucTs were also conserved in the NbFucTs [30], [31].

To compare the NbFucT sequences with FucTs of other organisms a Bayesian phylogenetic tree was made based on an alignment with all the α1,3-NbFucT amino acid sequences (incl. glycan motif LeA synthesising plant α1,4-FucTs) and α1,3-FucT amino acid sequences from *A. thaliana*, *Homo sapiens*, *Mus musculus*, *Drosophila melanogaster*, *Schistosoma mansoni* (Chapter 2 and 3) and *Caenorhabditis elegans* (Figure 2). For this phylogenetic tree the variable CTS domain was excluded from the alignment. In the phylogenetic tree a clear division is shown between the animal and the plant FucTs. Within the plant clade the FucTs are separated by function, the core α1,3-FucTs vs the LeA synthesising FucTs. In the core α1,3-FucT group the NbFucTs group separately from the two core α1,3-FucTs from *A. thaliana*, FUT11 and FUT12. The NbFucTs group in two groups, where group 1 consists of NbFucT1, 9 and 11 and group 2 constitutes NbFucT7 and 14. Within the animal clade *H. sapiens* and *M. musculus* group together based on function. Furthermore, the α1,3-FucTs from *D. melanogaster* group among different clades within the animal cluster.
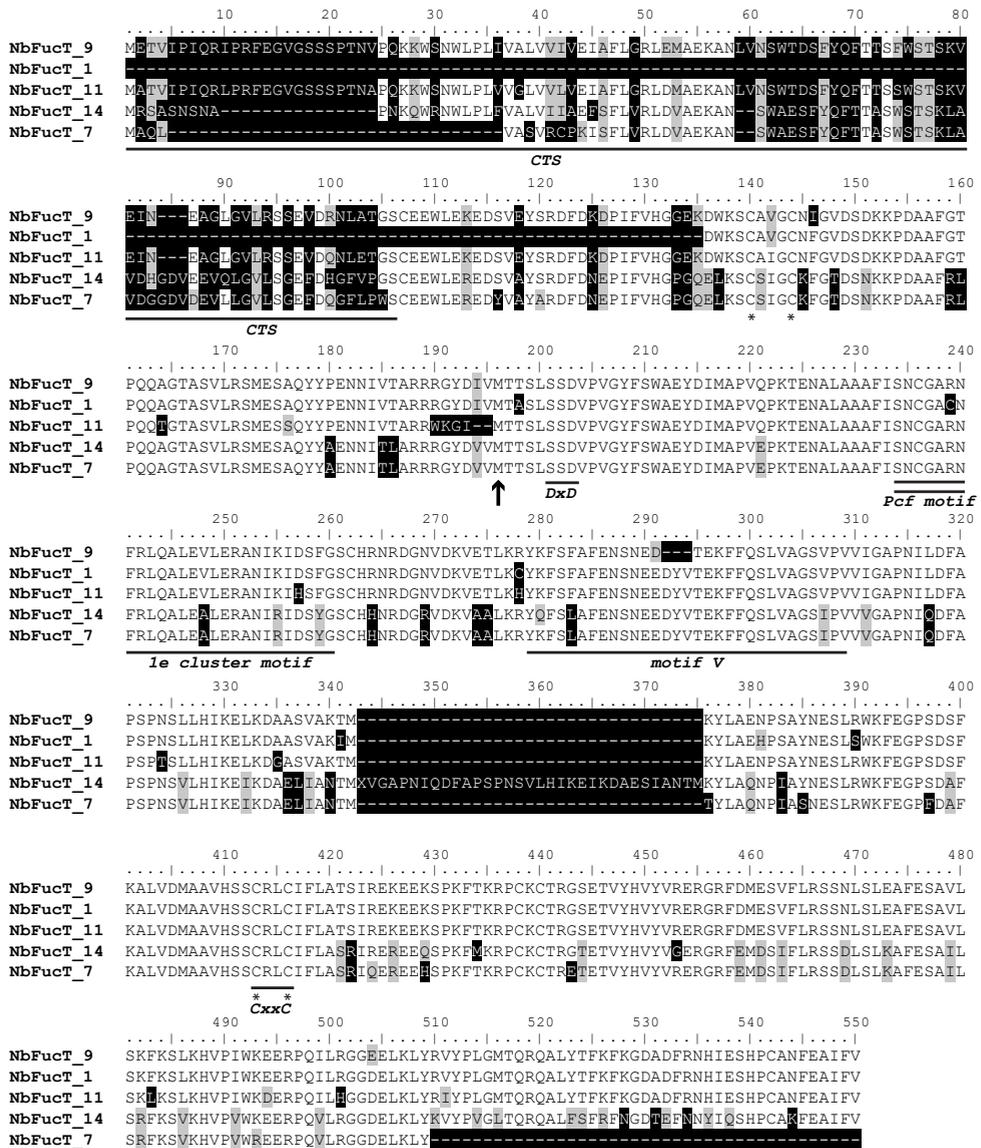
FIGURE 1 | Amino acid alignment of *Nicotiana benthamiana* core α1,3-fucosyltransferase (NbFucT). The NbFucT protein sequences from the sol genomics network were compared in an amino acid alignment. Positions with less than 60% conservation are highlighted in black. Positions that are not similar, but with more than 60% conservation are highlighted in grey. The α1,3-FucT sequence motifs: DxD (SSD in plants), pcf motif, 1st cluster motif, motif V and CxxC are indicated under the sequences [26], [28]–[31]. CTS indicates the cytoplasmic tail-transmembrane-stem region. * Indicates the four cysteine residues conserved in most of the α1,3-FucTs. The arrow indicates the start of the part of NbFucT11 after the stop codon, which was introduced by a 4 nucleotide deletion and subsequent frameshift.

TABLE 2 | Characteristics and differences between the *Nicotiana benthamiana* core α1,3-fucosyltransferases (NbFucTs). NbFucT amino acid sequences reported in the Sol Genomics Network database were compared. The transmembrane domain (TMD) was predicted by the TMHMM server v2.0. Based on these predictions the length of the cytoplasmic tail (CT) and the TMD were calculated. x indicates that no TMD was predicted by the TMHMM server v2.0. The percentage of identity between the NbFucT amino acid sequences and the most similar NbFucT amino acid sequence were determined by Clustal Omega alignment. Differences are indicated following the protein mutation nomenclature of den Dunnen and Antonarakis [22], based on the alignment in Figure 1.

Protein		TMD (length CT/TMD)	Percentage of identity	Sequence diversity
NbFucT1	382aa	x	97.63% with NbFucT9	p.M1_D136del, p.I146F, p.T198A, p.R239C, p.R278C, p.D291_T295ins3,p.T341I, p.N381H, p.R390S, p.E504D
NbFucT7	479aa	x	93.74% with NbFucT14	p.R2A, p.S3Q, p.A4L, p.S5_A9del, p.P25-V37del, p.L39S, p.I41R, p.I42C, p.A43P, p.E44K, p.F45I, p.H85G, p.E87D, p.Q90L, p.H100Q, p.V103L, p.G105W, p.S116Y, p.S120A, p.Q280K, p.X343_K376del, p.K376T, p.Y385S, p.S397F, p.R424Q, p.Q429H, p.M434T, p.G4443E, p.G453R, p.K493R, p.K510*
NbFucT9	511aa	29-48 (28/20)	97.63% with NbFucT1	-
NbFucT11	191aa	x	90.53% with NbFucT9	p.E2A, p.I10L, p.V24A, p.I36V, p.A38G, p.I42L p.E52D, p.F74S, p.R100Q, p.A103E, p.V142I, p.I146F, p.A164T, p.A176S, p.R190Wfs*5
	323aa	x	96.57% with NbFucT9	p.D257H, p.R278H, p.D291E, p.D291_T295ins3, p.N324T, p.A335G, p.F483L, p.E494D, p.R501H, p.E504D, p.V411I
NbFucT14	533aa	17-39 (16/23)	93.44% with NbFucT7	-
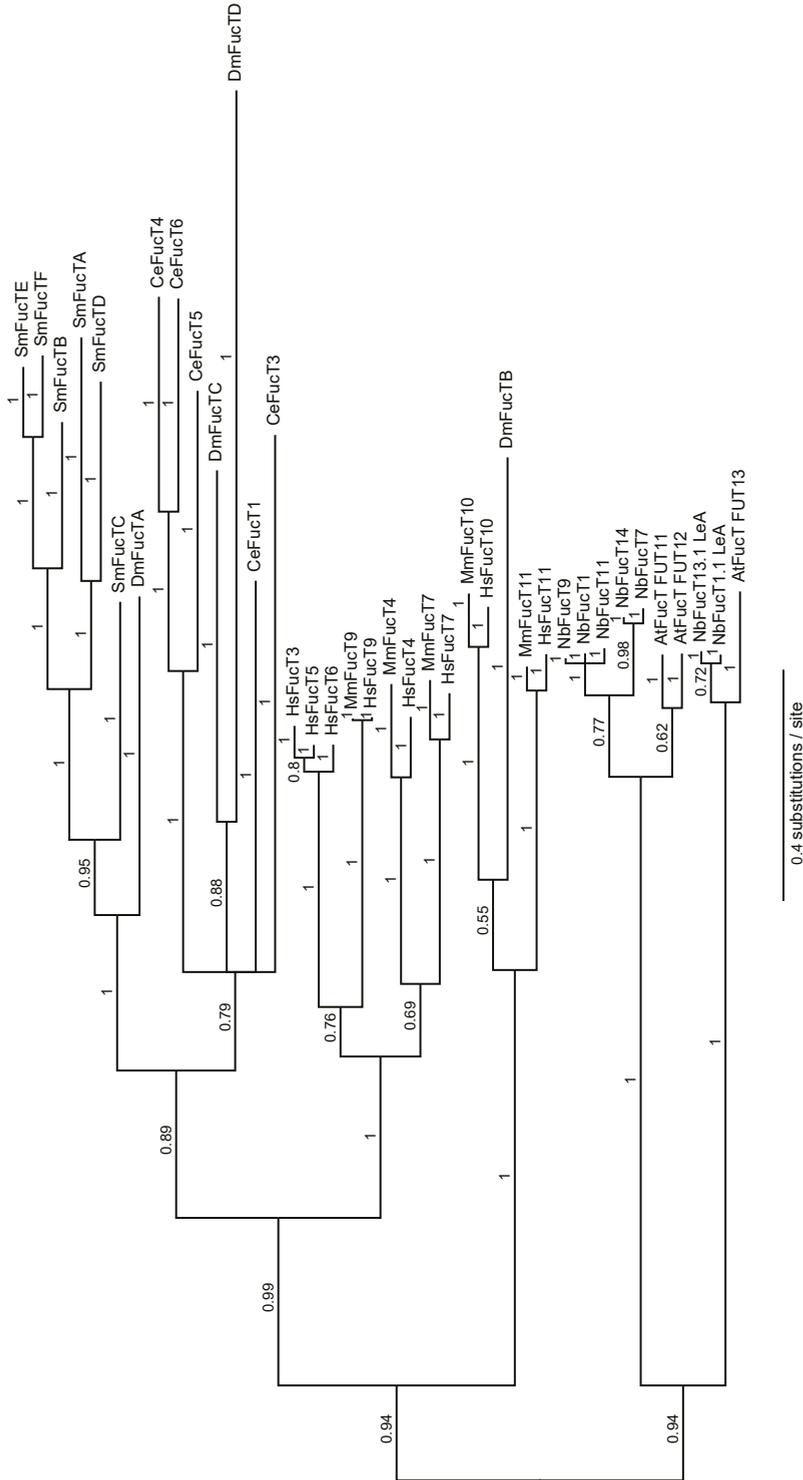
FIGURE 2 | Phylogenetic tree of fucosyltransferases (FucTs). A Bayesian phylogenetic tree of FucTs from *Nicotiana benthamiana* core α1,3-FucTs (NbFucT1, NbFucT7, NbFucT9, NbFucT11, NbFucT14), LeA FucTs (NbFucT1.1 LeA, NbFucT13.1 LeA) and FucTs of other organisms (*A. thaliana* (AtFucT FUT11, AtFucT FUT12 and AtFucT FUT13), *Homo sapiens* (HsFucT3 to 7 and 9 to 11), *Mus musculus* (MmFucT4,7,9 to 11), *Drosophila melanogaster* (DmFucTA to D), *Schistosoma mansoni* (SmFucTA to F) (Chapter 2 and 3) and *Caenorhabditis elegans* (CeFucT1 and 3 to 6)). Numbers above the branches indicate the posterior probability. Genetic distance is indicated by the scale bar.

4

Identification of *Nicotiana benthamiana* core α1,3-fucosyltransferase sequences

To identify the NbFucT sequences in our *N. benthamiana* plants primer pairs were designed based on the full-length sequences of NbFucT9, 11 and 14 from SGN (Supplemental Table 1). In total 74 sequences were collected by PCR amplification and subsequent sequencing. Among these sequences full-length sequences resembling NbFucT1 and NbFucT7 were also found. The sequences were grouped based on their SNP profiles (Figure 3A). Ten sequences, corresponding to the ten consensus sequences, were selected and named after their initial bacterial colony number. Next to the SNP profile, the 45 nt deletion at the beginning of the sequences in group 2 is worth noting as well as the characteristic 9 nt deletion and 12 nt insertion in the first 218 nt in comparison to group 1 (Figure 3B).

Group 1 contained five selected sequences resembling NbFucT1, 9 and 11. Of these five sequences NbFucTA1, A4 and B2 corresponded to the SGN protein sequences of NbFucT1, 9 and 11, respectively. For both NbFucT1 and 11 an additional sequence variant was identified in comparison to the SGN sequence (NbFucTC5 and D5, respectively). NbFucTC5 encoded an incomplete protein due to a deletion-introduced frameshift. Since, the SGN NbFucT11 protein sequence was incomplete, it was surprising to find 26 sequences (selected sequence NbFucTD5) that resemble NbFucT11. These sequences did encode a full-length protein due to a 4 nt insertion in comparison to NbFucT11.

Group 2 contained five selected sequences resembling NbFucT7 and 14. Of these five sequences NbFucTE10 and F11 corresponded to the SGN protein sequences of NbFucT7 and 14, respectively. For NbFucT7 three extra variants were identified in comparison to the SGN sequence, one full-length sequence and two sequences encoding an incomplete protein due to an early stop codon. The first variant, NbFucTG10, had a 3 nt deletion and encoded a full-length protein. The second variant, NbFucTF9, had a 4 nt insertion which introduced a frameshift resulting in an early stop codon and an incomplete protein. The third variant, NbFucTA10, had the same 3 nt deletion as NbFucTG10 as well as the same 4 nt insertion as NbFucTF9, which resulted in an early stop codon and an incomplete protein. For NbFucT14 a full-length sequence was found without the remarkable 33 aa duplication, suggesting a sequencing or annotation mistake in the SGN sequence.

All six full-length sequences were compared with the SGN sequence of the most similar NbFucT (Table 3). Next to the already mentioned observations only a few mutations were found in comparison to the SGN sequences.
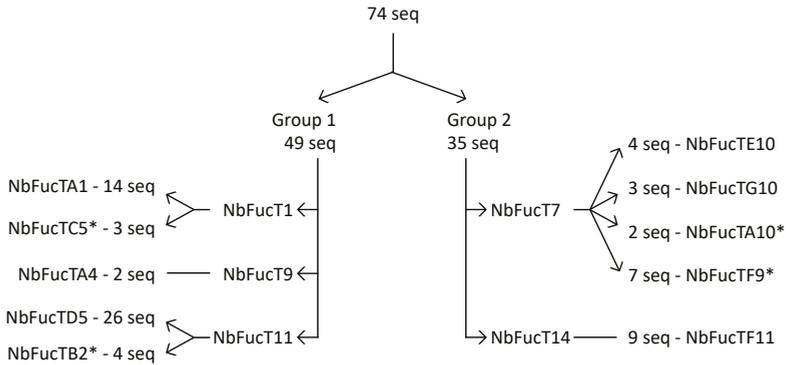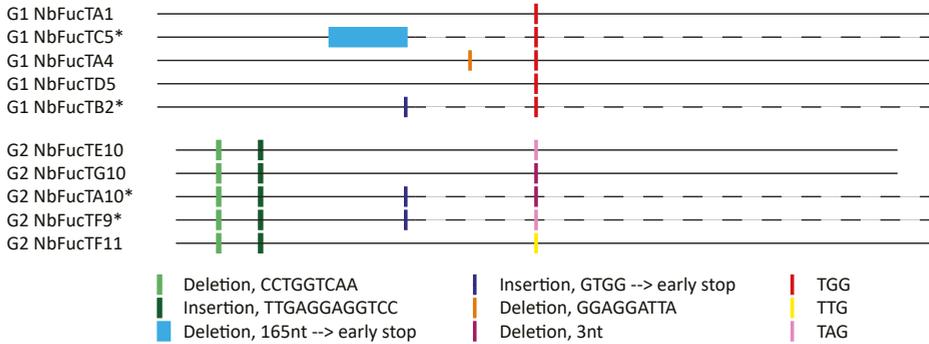
FIGURE 3 | Grouping and characteristics of retrieved *Nicotiana benthamiana* core α1,3-fucosyltransferase (NbFucT) sequences. (A) 74 sequences, collected by PCR amplification, were grouped based on SNP profile. Two initial groups as found in the phylogenetic tree (Figure 2) could be distinguished. Group 1, consisting of NbFucT1, 9 and 11 was made up of 49 sequences, which were represented by five sequences. Of these sequences two resembled NbFucT1, one resembled NbFucT9 and two resembled NbFucT11. Group 2, consisting of NbFucT7 and 14 was made up of 25 sequences, which were also represented by five sequences. Of these sequences four resembled NbFucT7 and one resembled NbFucT14. B) Schematic overview of sequence specific characteristic deletions, insertions or succeeded nucleotides (nt). Corresponding alignment can be found in Supplemental Figure 1. Sequences preceded by G1 belong to Group 1 and sequences preceded by G2 belong to Group 2. * indicates a sequence that encodes a protein with an early stop codon. The horizontal continues line resembles the length of the protein coding sequence, whereas the dotted line in combination with the continues line resembles the retrieved sequence.

4

TABLE 3 | Characteristics and differences of retrieved full length *Nicotiana benthamiana* core α1,3-fucosyltransferases (NbFucTs). NbFucT nucleotide and amino acid sequences retrieved from *N. benthamiana* cDNA were analysed and compared with the reported NbFucT sequences in the Sol Genomics Network database. The transmembrane domain (TMD) was predicted by the TMHMM server v.2.0. Based on these predictions the length of the cytoplasmic tail (CT) and the TMD were calculated. x indicates that no TMD was predicted. The percentage of identity between the retrieved NbFucT sequences (nucleotide (c.) and amino acid (p.)) and the most similar NbFucT sequence were determined by Clustal Omega alignment. Differences are indicated following the nucleotide (c.) and protein (p.) mutation nomenclature of den Dunnen and Antonarakis [22].

Protein		TMD (length CT/TMD)	Percentage of identity	Sequence diversity
NbFucTA1	1545nt -514aa	29-48 (28/20)	c.100% / p.100% with NbFucT1	c.0_1ins336, p.D1DM-132
NbFucTA4	1536nt -511aa	29-48 (28/20)	c.99.93% / p.99.80% with NbFucT9	c.1270G>A p.V424I
NbFucTD5	1545nt -514aa	x	c.99.94% / p.99.22% with NbFucT11	c.557_558ins4, p.W187R, p.K188G, p.G189Y, p.I190D, p.*191del
NbFucTE10	1380nt -459aa	17-39 (16/22)	c.98.72% / p.97.96% with NbFucT7	c.0-1ins51, c.1A>T, c.4G>C, c.6G>T, c.8A>T, c.9A>G, c.12G>C, c.19T>C, c.20C>T, c.21T>A, c.25C>G, c.26G>T, c.28T>A, c.29G>T, c.30C>A, c.31C>G, c.33T>A c.34A>G, p.M1LM-17, p.A2P, p.Q3L, p.L4F, p.S7L, p.R9V, pC10I, p.P11A, p.K12E
NbFucTF11	1503nt -500aa	17-39 (16/22)	c.100% / p.100% with NbFucT14	c.930_1028del99, c.1017T>C, p.313_345del33, p.L320S
NbFucTG10	1377nt -458aa	17-39 (16/22)	c.98.72% / p.97.96% with NbFucT7	c.0-1ins51, c.1A>T, c.4G>C, c.6G>T, c.8A>T, c.9A>G, c.12G>C, c.19T>C, c.20C>T, c.21T>A, c.25C>G, c.26G>T, c.28T>A, c.29G>T, c.30C>A, c.31C>G, c.33T>A c.34A>G, c.705_707del, p.M1LM-17, p.A2P, p.Q3L, p.L4F, p.S7L, p.R9V, pC10I, p.P11A, p.K12E, p.V236del

Screening for *Nicotiana benthamiana* core α1,3-fucosyltransferases activity

To investigate whether or not the retrieved six full-length sequences encode active FucTs, each NbFucT sequence was co-expressed with carrier glycoproteins in wild type (wt) or ΔXT/FT *Nicotiana benthamiana* plants. Three different carrier glycoproteins (kappa-5, omega-1 or IL-22) were tested to reveal protein specific core α1,3-fucosylation. Upon expression in wt plants the N-glycans on kappa-5 and omega-1 are core α1,3-fucosylated, whereas the N-glycans on IL-22 are not [32], [33]. Upon co-expression of the various NbFucT sequences in ΔXT/FT plants, the carrier glycoproteins were isolated from the leaf apoplast and fucosylation of the N-glycans was analysed by PNGase F digestion and subsequent visualised on western blot using anti-FLAG antibodies. PNGase F releases N-glycans only in the absence of core α1,3-fucose. The release of N-glycans from kappa-5 resulted in a shift in protein size from around 45 kDa to 37 kDa, which is observed when comparing kappa-5 samples produced in wt and ΔXT/FT plants (Figure 4). Upon co-expression, NbFucTA1, D5 and F11 prevented N-glycan release by PNGase F, which indicated the presence of a core α1,3-fucose. A similar pattern for N-glycan cleavage was observed for the carrier glycoproteins omega-1 and IL-22 upon co-expression of NbFucTA1, D5 and F11 (Supplemental Figure 2). The faint protein band around 30 kDa (just below deglycosylated kappa-5) probably corresponds to a plant protein as was shown in Figure 3C of Chapter 3.

In conclusion, NbFucTA1, D5 and F11 showed activity towards the N-glycans on all three carrier glycoproteins, whereas NbFucTA4, E10 and G10 did not. This suggests that NbFucTA4, E10 and G10 are inactive, indicating that in order to create a *N. benthamiana* NbFucT knockout plant, only the variants NbFucTA1, D5 and F11 have to be knocked-out.
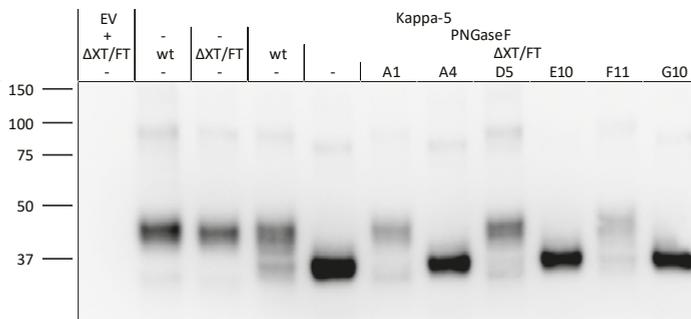
FIGURE 4 | *Nicotiana benthamiana* core α1,3-fucosyltransferase (NbFucT) activity on the N-glycans of kappa-5. Kappa-5 was co-expressed in wild type (wt) or in ΔXT/FT *N. benthamiana* plants with NbFucTs (NbFucTA1, A4, D5, E10, F11 and G10). After extraction from the apoplast fluid the N-glycan composition of kappa-5 was analysed by PNGase F treatment. To assess the presence of core α1,3-fucose 200 ng PNGase F treated sample was visualised on western blot using anti-FLAG antibodies. The western blot reveals that NbFucTA1, D5 and F11 prevented N-glycan release by PNGase F.

Discussion

Different techniques have been used for the production of NbFucT and NbXylT knockout plants [18], [19]. The drawback of these studies is that they target all or only a few reported NbFucT genes, whereas not all reported NbFucT genes have to encode active proteins. Moreover, different allelic variants of one gene can be present and targeting all these variants and subsequent screening for introduced mutations is time consuming. Therefore, in this study we first analysed reported NbFucT genes *in silico*. Next, we retrieved NbFucT sequences to possibly correct for mistakes in sequences in databases or different sequences found in our *N. benthamiana* line. Finally, we tested the activity of six full length sequences that were obtained.

In the amino acid alignment of the SGN NbFucT protein sequences with the described FucT sequence motifs an aspartic acid, tyrosine and valine deletion was observed for NbFucT9 in motif V. Jost and colleagues [34] have shown that mutation of the tyrosine to an alanine in motif V results in an inactive protein. Furthermore, Jansing and colleagues [35] reported that the NbFucT9 gene does not contain introns and is not preceded by a TATA-box. Together this suggest that NbFucT9 is an inactive pseudogene. Other differences in motif V observed between the NbFucTs are not suggested to result in an inactive protein, but can influence the FucT activity level [29], [34]. The plant core α1,3-FucT motif, found in all plant α1,3-FucTs, overlaps with the 1st cluster motif [26], [28]. This is not remarkable, since the serine and asparagine at the beginning of the 1st cluster motif are reported to be important for GDP-fucose binding and will therefore most likely be present in all active plant α1,3-FucTs [29]. The mutations observed in the plant core α1,3-FucT motif and the 1st cluster motif are not in conserved residues involved in GDP-fucose binding, besides similar substitutions are reported in other FucT protein sequences [26], [28].

The phylogenetic tree shows that the plant α1,3-FucTs form a separate clade, in which the core α1,3-FucTs are separated from LeA synthesising FucTs. Within the core α1,3-FucT group the two AtFucT genes group separately from the five NbFucT genes reported in the SGN database. Minus the inactive pseudogene NbFucT9, the 2 to 4 ratio suggests that the allotetraploid *N. benthamiana* retained the FucT genes of both parent species. This is supported by the two groups in which the NbFucT genes are divided, although these two groups do not group together with the AtFucTs. On the other hand, the two core α1,3-FucTs found in *Nicotiana attenuata*, NaFucT11 and NaFucT12, do group with another NbFucT group each [20]. This suggests that the increase in FucT genes is probably due to hybridisations among *Nicotiana* polyploid ancestors [36].

Amplification of NbFucT genes from *N. benthamiana* cDNA resulted in ten NbFucT sequences. Among these ten sequences, variants of the five reported NbFucT gene sequences were found. These variants probably correspond to allelic variation in the allotetraploid *N. benthamiana*. Four of the ten NbFucT sequences encoded an incomplete

protein by deletion- or insertion-induced frameshifts resulting in early stop codons and subsequent absence of GDP-fucose binding motifs. The other six sequences encoded a full-length protein, of which the activity was tested. Only, three out of the six sequences, NbFucTA1, D5 and F11, showed activity towards the N-glycans of the co-expressed carrier glycoproteins. Furthermore, no specificity was observed between the NbFucTs for one of the three carrier glycoproteins, since co-expression of respectively NbFucTA1, D5 and F11 with kappa-5, omega-1 or IL-22 resulted in core α1,3-fucosylation of the N-glycans of all three glycoproteins.

NbFucTA4 encodes the pseudogene NbFucT9 and showed in our activity studies that the absence of the tyrosine and two surrounding amino acids in motif V is likely important for FucT function as shown by Jost and colleagues [34] for human FucTVI. NbFucTE10 and NbFucTG10, the other two sequences encoding inactive proteins, do not show mutations in reported FucT sequence motifs. However, these two sequences do show mutations at the beginning of the sequences and are 41 aa shorter in comparison to the active protein encoding sequence NbFucTF11. For *A. thaliana* FUT11 it was previously shown that a deletion of the last 113 aa leads to a 99% less active protein [26]. This suggests that these two sequences encode an inactive protein, due to a deletion at the C-terminal part.

The five NbFucT sequences used to produce a *N. benthamiana* NbFucT and NbXylT knockout plant by Weterings and van Eldik [18] recently became available on lens.org. Comparison of these NbFucT sequences to our retrieved full-length NbFucT sequences showed that only a few mutations occur. A recently published paper of Schiavinato and colleagues [37] showed that most *N. benthamiana* laboratory accessions are derived from a single source. Together this suggests that within the different laboratory accessions few minor differences can be expected in the NbFucT sequences. Therefore, our sequences can also be used by others for guide RNA design and subsequent production of NbFucT knockout plants by CRISPR-Cas9.

Recently, CRISPR-Cas9 NbFucT and NbXylT knockout plants became available [35]. Jansing and colleauges [35] knocked out eight NbFucT variants, excluding NbFucT9 marked as pseudogene. With activity tests we have now shown that only three of the five reported NbFucT genes encode proteins with activity towards N-glycans and four of their sequence variants encode incomplete proteins. Therefore, only three out of the eight NbFucT variants targeted by Jansing and colleagues [35] had to be knocked out to create a NbFucT knockout plant for the production of proteins without core α1,3-fucosylated N-glycans. With knowledge of the fact that only three NbFucT genes encode active variants, time, effort and resources could have been saved. Less screenings for introduced mutations and less crosses would have been required to end up with a full NbFucT knockout plant. Moreover, targeting less variants by CRISPR-Cas9 will decrease the chance of modifying off target sites. Therefore, we anticipate our study to be an example for future research on generating genome-edited plants.

4

Materials and Methods

Characterisation of *Nicotiana benthamiana* fucosyltransferase sequences

To find NbFucT sequences, the sequence of FUT11, a known core α1,3-FucT from *A. thaliana* (accession number Q9LJK1), was used in a BLAST search against the predicted *N. benthamiana* cDNA sequences in the Sol Genomics Network (SGN) database [20]. The BLAST search resulted in five FucT sequences coding for core α1,3-NbFucTs. Of each sequence the largest predicted open reading frame was used as predicted protein sequence. Since FucTs are type II membrane proteins, the transmembrane domain (TMD) of the proteins was predicted by TMHMM v2.0. To obtain the sequences of the NbFucT proteins in our *N. benthamiana* plant line primers were developed based on the SGN sequences of NbFucT9, 11 and 14, since NbFucT1 and 7 seemed to encode incomplete protein sequences (Supplemental Table 1 and Figure 1). *N. benthamiana* leaf mRNA was isolated with the Maxwell® 16 System RNA Purification Kit (Promega) and subsequently cDNA was synthesised with GoScript™ Reverse Transcriptase (Promega). NbFucT sequences were amplified from cDNA in a PCR reaction with Expand High Fidelity DNA polymerase (Roche). PCR products were cloned into the pCR2.1 vector (Invitrogen) and cloned in *Escherichia coli*. *E. coli* colonies equally originating from the primer combinations of NbFucT9, 11 or 14 were sequenced. The resulting 74 sequences were aligned by Clustal Omega and the alignment was manually improved in BioEdit. Groups were formed of which ten sequences (corresponding to the consensus sequence) were selected for further analysis. Of these ten sequences four sequences showed the introduction of an early stop codon and were not taken along in subsequent steps.

Phylogenetic tree

A Bayesian phylogenetic tree of FucTs was created based on an alignment with *N. benthamiana* core α1,3-FucTs (NbFucT1, NbFucT7, NbFucT9, NbFucT11, NbFucT14), LeA FucTs (NbFucT1.1 LeA, NbFucT13.1 LeA) and FucTs of other organisms (*A. thaliana* (AtFucT FUT11, AtFucT FUT12 and AtFucT FUT13), *Homo sapiens* (HsFucT3 to 7 and 9 to 11), *Mus musculus* (MmFucT4,7,9 to 11), *Drosophila melanogaster* (DmFucTA to D), *Schistosoma mansoni* (SmFucTA to F) (Chapter 2 and 3) and *Caenorhabditis elegans* (CeFucT1,3 to 6)). Corresponding accession numbers can be found in Supplemental Table 2. The alignment was created in BioEdit v.7.2.6 [38]. The variable CTS domain of the FucTs was excluded from the alignment. The phylogenetic tree was created in MrBayes v.3.2.6 using an mixed amino acid substitution model [39]. The tree was run for 1 million generations using 4 parallel runs with 4 chains each. The burn in was 10,000 generations.

Construction and description of expression vectors

In order to overexpress the NbFucT sequences encoding full-length proteins, the sequences were cloned in the pHYG plant expression vector [40]. Thereto, primers were developed for the amplification of the NbFucT sequences and the introduction of flanking restriction sites NcoI or BspHI at the 5'-end and KpnI at the 3'-end (Supplemental Table 1). Via these restriction sites the NbFucT sequences were cloned in the pHYG expression vector. To determine the NbFucT activity, constructs with FLAG-tagged carrier glycoprotein kappa-5, omega-1 or IL-22 were co-expressed from pHYG [32], [33]. All constructs were driven by the Cauliflower mosaic virus 35S promoter with duplicated enhancer (d35S) and the *Agrobacterium tumefaciens* nopaline synthase transcription terminator (Tnos) [40]. To boost translation a 5' leader sequence of the alfalfa mosaic virus RNA 4 (AIMV) was included between the promoter and the construct. To enhance expression, the P19 silencing suppressor from tomato bushy stunt virus pBIN61 was co-infiltrated in all experiments, unless stated differently [41]. All constructs were transformed into *A. tumefaciens* strain MOG101 for plant expression.

Recombinant protein production and isolation

A. tumefaciens cultures were grown at 28 °C/250 rpm in LB medium (10 g/L peptone140, 10 g/L NaCl, 5 g/L yeast, pH 7.0), with 50 µg/mL kanamycin and 20 µM acetosyringone. After 16 hrs, bacteria were centrifuged for 15 min/2880 xg and subsequently resuspended in MMA (1.95 g/L MES, 20 g/L sucrose, 5 g/L MS-salts, pH 5.6) with 200 µM acetosyringone to increase *A. tumefaciens* transformation efficiency. *A. tumefaciens* cultures of different constructs were mixed for co-expression. The final optical density (OD) of each *A. tumefaciens* culture in the mixture was 0.5. The two youngest fully expanded leaves of four to six week old wt or ΔXT/FT *N. benthamiana* plants were infiltrated at the abaxial side [17]. *N. benthamiana* plants were grown in a controlled greenhouse compartment (UNIFARM, Wageningen). five to six days post infiltration (dpi) leaves were harvested and apoplastic proteins were isolated via apoplast wash. Harvested leaves were submerged in extraction buffer (50 mM phosphate buffer, pH 8.0, containing 0.1 M NaCl and 0,1 % v/v Tween-20), a vacuum was applied and after 5 min slowly released to infiltrate the leaves with buffer. The apoplast fluid was extracted from the leaves by centrifugation for 10 min/2000 xg. The protein concentration of the apoplast fluid was determined by the Pierce Bicinchoninic Acid Protein Assay (BCA, Fisher Scientific).

SDS-PAGE and western blot

For screening of N-glycan core α1,3-fucosylation, samples were denatured and subsequent incubated with Peptide:N-glycosidase (PNGase) F for 1 h/37 $^{\circ}$C . After PNGase F treatment samples were analysed on a 12 % Bis-Tris SDS-PAGE gel (Fisher Scientific) and subsequently transferred to a PVDF membrane by wet blotting. After blotting the membrane was blocked for 1 hr at RT or o/n at 4 $^{\circ}$C with 5 % w/v milk powder in PBST (PBS

4

containing 0.1 % v/v Tween-20). Subsequent the membrane was incubated for 1 hr at RT or o/n at 4 ^{0}C with antibody anti-FLAG diluted 1:5000. After every incubation step the membrane was washed 5 times with PBST. The anti-HRP labelled antibody was detected with a 1:1 SuperSignal West Femto:Dura substrate (Fisher Scientific) in the G:BOX Chemi System (Syngene).

Acknowledgements

4

References

[1] J. W. Larrick, L. Yu, J. Chen, S. Jaiswal, and K. Wycoff, "Production of antibodies in transgenic plants," *Res. Immunol.*, vol. 149, no. 6, pp. 603–608, 1998.

[2] E. Stoger, R. Fischer, M. Moloney, and J. K.-C. Ma, "Plant Molecular Pharming for the Treatment of Chronic and Infectious Diseases," *Annu. Rev. Plant Biol.*, vol. 65, no. 1, pp. 743–768, 2014.

[3] J. L. Fox, "First plant-made biologic approved," *Nat. Biotechnol.*, vol. 30, no. 6, pp. 472–472, 2012.

[4] G. M. Lyon *et al.*, "Clinical care of two patients with Ebola virus disease in the United States," *N. Engl. J. Med.*, vol. 371, no. 25, pp. 2402–2409, 2014.

[5] G. P. Lomonossoff and M. A. D'Aoust, "Plant-produced biopharmaceuticals: A case of technical developments driving clinical deployment," *Science*, vol. 353, no. 6305, pp. 1237–1240, 2016.

[6] F. Altmann, "The role of protein glycosylation in allergy," *Int. Arch. Allergy Immunol.*, vol. 142, no. 2, pp. 99–115, 2007.

[7] S. Kallolimath *et al.*, "Engineering of complex protein sialylation in plants (Proceedings of the National Academy of Sciences of the United States of America (2016) 113 (9498-9503) DOI: 10.1073/pnas.1604371113)," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 113, no. 34, pp. 9498–9503, 2016.

[8] A. Castilho *et al.*, "Engineering of sialylated mucin-type O-glycosylation in plants," *J. Biol. Chem.*, vol. 287, no. 43, pp. 36518–36526, 2012.

[9] N. Q. Palacpac *et al.*, "Stable expression of human β1,4-galactosyltransferase in plant cells modifies N-linked glycosylation patterns," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 96, no. 8. pp. 4692–4697, 1999.

[10] H. Bakker *et al.*, "An antibody produced in tobacco expressing a hybrid β-1,4- galactosyltransferase is essentially devoid of plant carbohydrate epitopes," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 103, no. 20, pp. 7577–7582, 2006.

[11] H. Bakker *et al.*, "Galactose-extended glycans of antibodies produced by transgenic plants," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 98, no. 5, pp. 2899–2904, 2001.

[12] A. Castilho *et al.*, "N-Glycosylation engineering of plants for the biosynthesis of glycoproteins with bisected and branched complex N-glycans," *Glycobiology*, vol. 21, no. 6, pp. 813–823, 2011.

[13] D. N. Forthal *et al.*, "Fc-Glycosylation Influences Fcγ Receptor Binding and Cell-Mediated Anti-HIV Activity of Monoclonal Antibody 2G12," *The Journal of Immunology*, vol. 185, no. 11. pp. 6876–6882, 2010.

[14] J. M. Alonso *et al.*, "Genome-wide insertional mutagenesis of Arabidopsis thaliana," *Science*, vol. 301, no. 5633, pp. 653–657, 2003.

[15] R. Strasser, F. Altmann, L. Mach, J. Glössl, and H. Steinkellner, "Generation of Arabidopsis thaliana plants with complex N-glycans lacking β1,2-linked xylose and core α1,3-linked fucose," *FEBS Letters*, vol. 561, no. 1–3. pp. 132–136, 2004.

[16] M. Schähs, R. Strasser, J. Stadlmann, R. Kunert, T. Rademacher, and H. Steinkellner, "Production of a monoclonal antibody in plants with a humanized N-glycosylation pattern," *Plant Biotechnol. J.*, vol. 5, no. 5, pp. 657–663, 2007.

[17] R. Strasser *et al.*, "Generation of glyco-engineered Nicotiana benthamiana for the production of monoclonal antibodies with a homogeneous human-like N-glycan structure," *Plant Biotechnology J.*, vol. 6, no. 4. pp. 392–402, 2008.

[18] K. Weterings and G. van Eldik, "NICOTIANA BENTHAMIANA PLANTS DEFICIENT IN FUCOSYLTRANSFERASE ACTIVITY," 2015.

[19] J. Li *et al.*, "Multiplexed, targeted gene editing in Nicotiana benthamiana for glyco-engineering and monoclonal antibody production," *Plant Biotechnol. J.*, vol. 14, no. 2, pp. 533–542, 2016.

[20] N. Fernandez-Pozo *et al.*, "The Sol Genomics Network (SGN)-from genotype to phenotype to breeding," *Nucleic Acids Res.*, vol. 43, no. D1, pp. D1036–D1041, 2015.

[21] L. Mathis, D. F. Voytas, J. Li, F. Zhang, T. Stoddard, and M. A. D'Aoust, "Plants for production of therapeutic proteins," 2014.

[22] J. T. den Dunnen and S. E. Antonarakis, "Mutation nomenclature extensions and suggestions to describe complex mutations: A discussion," *Hum. Mutat.*, vol. 15, no. 1, pp. 7–12, 2000.

4

[23] R. Mollicone *et al.*, "Activity, splice variants, conserved peptide motifs, and phylogeny of two New α1,3-fucosyltransferase families (FUT10 and FUT11)," *J. Biol. Chem.*, vol. 284, no. 7, pp. 4723–4738, 2009.

[24] S. L. Martin, M. R. Edbrooke, T. C. Hodgman, D. H. Van Den Eijnden, and M. I. Bird, "Lewis X biosynthesis in Helicobacter pylori. Molecular cloning of an α(1,3)-fucosyltransferase gene," *J. Biol. Chem.*, vol. 272, no. 34, pp. 21349–21356, 1997.

[25] C. Breton, R. Oriol, and A. Imberty, "Conserved structural features in eukaryotic and prokaryotic fucosyltransferases," *Glycobiology*, vol. 8, no. 1, pp. 87–94, 1998.

[26] P. Both *et al.*, "Distantly related plant and nematode core 1,3-fucosyltransferases display similar trends in structure-function relationships," *Glycobiology*, vol. 21, no. 11, pp. 1401–1415, 2011.

[27] E. L. Sonnhammer, G. von Heijne, and A. Krogh, "A hidden Markov model for predicting transmembrane helices in protein sequences." *Int. Conf. Intell. Syst. Mol. Biol. ; ISMB. Int. Conf. Intell. Syst. Mol. Biol.*, vol. 6, pp. 175–182, 1998.

[28] I. B. H. Wilson *et al.*, "Cloning and expression of cDNAs encoding α1,3-fucosyltransferase homologues from Arabidopsis thaliana," *Biochimica et Biophysica Acta - General Subjects*, vol. 1527, no. 1–2. pp. 88–96, 2001.

[29] H. Y. Sun *et al.*, "Structure and mechanism of Helicobacter pylori fucosyltransferase: A basis for lipopolysaccharide variation and inhibitor design," *J. Biol. Chem.*, vol. 282, no. 13, pp. 9973–9982, 2007.

[30] T. de Vries, R. M. A. Knegtel, E. H. Holmes, and B. A. Macher, "Fucosyltransferases: Structure/function studies," *Glycobiology*, vol. 11, no. 10, pp. 119–128, 2001.

[31] T. de Vries *et al.*, "Neighboring cysteine residues in human fucosyltransferase VII are engaged in disulfide bridges, forming small loop structures," *Glycobiology*, vol. 11, no. 5, pp. 423–432, 2001.

[32] R. H. P. Wilbers *et al.*, "The N-glycan on Asn54 affects the atypical N-glycan composition of plant-produced interleukin-22, but does not influence its activity," *Plant Biotechnol. J.*, vol. 14, no. 2, pp. 670–681, 2016.

[33] R. H. P. Wilbers *et al.*, "Production and glyco-engineering of immunomodulatory helminth glycoproteins in plants," *Sci. Rep.*, vol. 7, p. 45910, 2017.

[34] F. Jost, T. de Vries, R. M. A. Knegtel, and B. A. Macher, "Mutation of amino acids in the alpha 1,3-fucosyltransferase motif affects enzyme activity and Km for donor and acceptor substrates," *Glycobiology*, vol. 15, no. 2, pp. 165–175, 2005.

[35] J. Jansing, M. Sack, S. M. Augustine, R. Fischer, and L. Bortesi, "CRISPR/Cas9-mediated knockout of six glycosyltransferase genes in Nicotiana benthamiana for the production of recombinant proteins lacking β-1,2-xylose and core α-1,3-fucose," *Plant Biotechnology J.*, vol. 17, no. 2. pp. 350–361, 2019.

[36] S. Knapp, M. W. Chase, J. J. Clarkson, S. Knapp, M. W. Chase, and J. J. Clarkson, "Nomenclatural Changes and a New Sectional Classification in Nicotiana (Solanaceae) Published by : International Association for Plant Taxonomy (IAPT) ," vol. 53, no. 1, pp. 73–82, 2017.

[37] M. Schiavinato, R. Strasser, L. Mach, J. C. Dohm, and H. Himmelbauer, "Genome and transcriptome characterization of the glycoengineered Nicotiana benthamiana line ΔXT/FT," *BMC Genomics*, vol. 20, no. 1, pp. 1–16, 2019.

[38] T. A. Hall, "BioEdit, a user-friendly biological sequence alignment editor and analysis program for Windows 95 and 98.pdf," *Nucleic Acids Symp. Ser.*, vol. 41, no. 41, pp. 95–98, 1999.

[39] F. Ronquist *et al.*, "MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space," *Syst. Biol.*, vol. 61, no. 3, pp. 539–542, 2012.

[40] L. B. Westerhof *et al.*, "3D Domain Swapping Causes Extensive Multimerisation of Human Interleukin-10 When Expressed In Planta," *PLoS One*, vol. 7, no. 10, pp. 1–10, 2012.

[41] O. Voinnet, S. Rivas, P. Mestre, and D. Baulcombe, "An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus," *Plant J.*, vol. 33, no. 5, pp. 949–956, 2003.

[42] A. Bateman, "UniProt: A worldwide hub of protein knowledge," *Nucleic Acids Res.*, vol. 47, no. D1, pp. D506–D515, 2019.
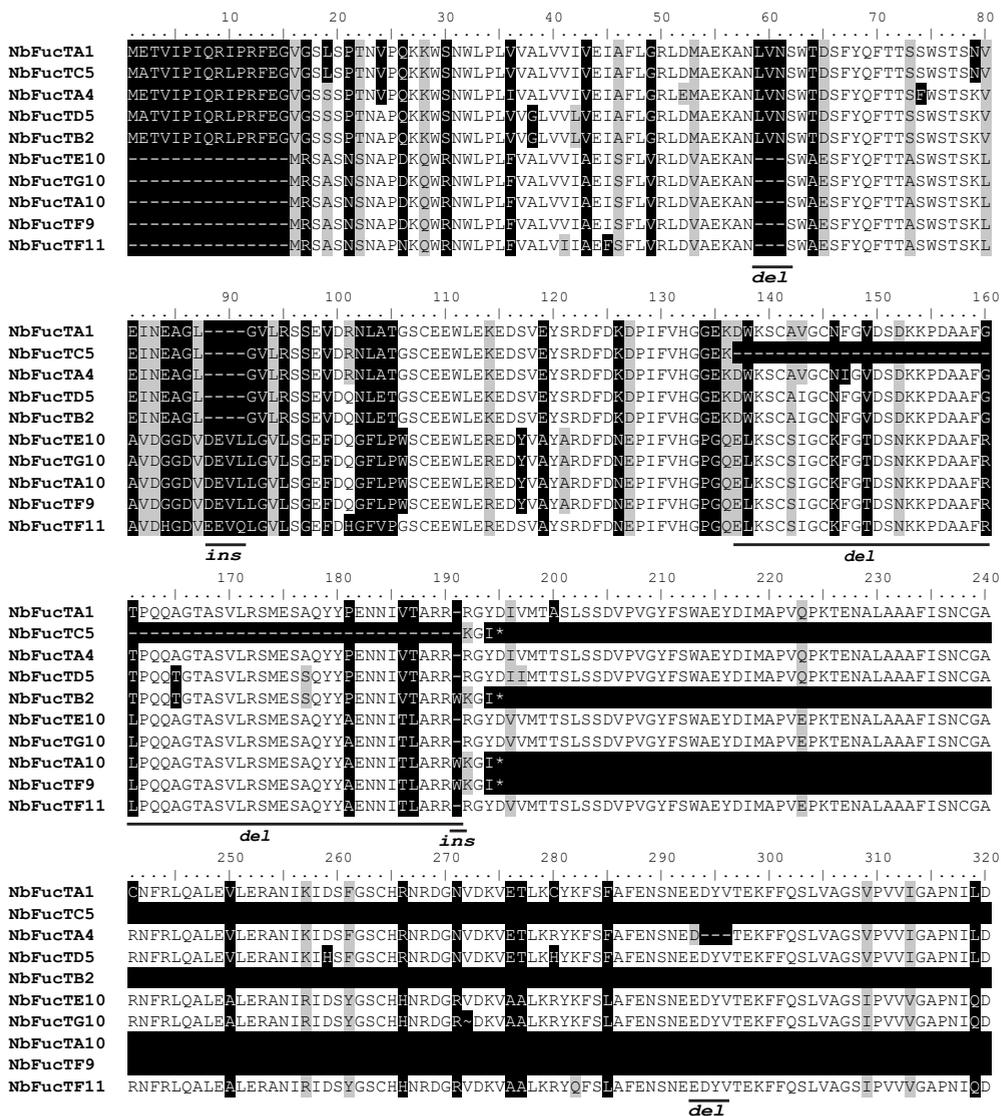
Supplemental Figures and Tables

SUPPLEMENTAL TABLE 1 | Primers for *Nicotiana benthamiana* core α1,3-fucosyltransferase (NbFucT) amplification and cloning. Primers used for *Nicotiana benthamiana* core α1,3-fucosyltransferase (NbFucT) amplification from cDNA or reamplification for cloning in the pHYG plant expression vector. The part of the sequences that corresponds to the open reading frame is written in capital letters.

Primer name	Sequence 5' to 3'
Amplification NbFucT9 - *Forward*	ATGGAAACAGTTATTCCAATTCAAAGAA
Amplification NbFucT11 - *Forward*	ATGGCAACAGTTATTCCAATTCAAAGAT
Amplification NbFucT9 and 11 - *Reverse*	CTATACAAATATGGCTTCAAAGTTTGC
Amplification NbFucT 14 - *Forward*	ATGAGATCGGCGTCAAATTCAAAC
Amplification NbFucT14 - *Reverse*	CTATACGAAGATGGCTTCAAATTTTG
Cloning NbFucT9 - *Forward*	acggccgccagtgtgctgccATGGAAACAGTTATTCCAATTC
Cloning NbFucT11 - *Forward*	acggccgccagtgtgctgccATGGCAACAGTTATTCCAATTC
Cloning NbFucT9 and 11 - *Reverse*	cgccagtgtgatggatatctgcaggtacCTATACAAATATGGCTTCAA AGTTTG
Cloning NbFucT14 - *Forward*	acggccgccagtgtgctgtcATGAGATCGGCGTCAAATTC
Cloning NbFucT14 - *Reverse*	cgccagtgtgatggatatctgcaggtacCTATACGAAGATGGCTTCAA ATTTTG

SUPPLEMENTAL TABLE 2 | Accession numbers of fucosyltransferases (FucTs) used in the phylogenetic tree. FucT amino acid sequences used for the Bayesian phylogenetic tree were gathered from UniProtKB or the Sol Genomics Network [20], [42]. Accession numbers of the FucT sequences are indicated.

Fucosyltransferase	Accession number	Fucosyltransferase	Accession number
SmFucTA	Q9NJ24	HsFucT4	P22083
SmFucTB	E2EAI5	HsFucT5	Q11128
SmFucTC	E2EAI6	HsFucT6	P51993
SmFucTD	E2EAI7	HsFucT7	Q11130
SmFucTE	E2EAI8	HsFucT9	Q9Y231
SmFucTF	E2EAI9	HsFucT10	Q6P4F1
AtFucT11	Q9LJK1	HsFucT11	Q495W5
AtFucT12	Q9FX97	DmFucTA	Q9VUL9
AtFucT13	Q9C8W3	DmFucTB	Q9VLC1
NbFucT13.1 LeA	Niben101Scf05896g00013.1	DmFucTC	P83088
NbFucT1.1 LeA	Niben101Scf02272g00001.1	DmFucTD	Q9W0F6
CeFucT1	G5EDR5	MmFucT4	Q11127
CeFucT3	G5EFP5	MmFucT7	Q11131
CeFucT4	G5EEL7	MmFucT9	O88819
CeFucT5	G5EE06	MmFucT10	Q5F2L2
CeFucT6	G5EEE1	MmFucT11	Q8BHC9
HsFucT3	P21217		

4

4

```
                10        20        30        40        50        60        70        80
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
NbFucTA1  METVIPIQRIPRFEGVGSLSPTNVPQKKWSNWLPLVVALVVIVEIAFLGRLDMAEKANLVNSWTDSFYQFTTSSWSTSNV
NbFucTC5  MATVIPIQRLPRFEGVGSLSPTNVPQKKWSNWLPLVVALVVIVEIAFLGRLDMAEKANLVNSWTDSFYQFTTSSWSTSNV
NbFucTA4  METVIPIQRIPRFEGVGSSSPTNVPQKKWSNWLPLIVALVVIVEIAFLGRLEMAEKANLVNSWTDSFYQFTTSEWSTSKV
NbFucTD5  MATVIPIQRLPRFEGVGSSSPTNAPQKKWSNWLPLVVGLVVIVEIAFLGRLDMAEKANLVNSWTDSFYQFTTSSWSTSKV
NbFucTB2  METVIPIQRLPRFEGVGSSSPTNAPQKKWSNWLPLVVGLVVIVEIAFLGRLDMAEKANLVNSWTDSFYQFTTSSWSTSKV
NbFucTE10 ---------------MRSASNSNAPDKQWRNWLPLFVALVVIAEISFLVRLDVAEKAN---SWAESFYQFTTASWSTSKL
NbFucTG10 ---------------MRSASNSNAPDKQWRNWLPLFVALVVIAEISFLVRLDVAEKAN---SWAESFYQFTTASWSTSKL
NbFucTA10 ---------------MRSASNSNAPDKQWRNWLPLFVALVVIAEISFLVRLDVAEKAN---SWAESFYQFTTASWSTSKL
NbFucTF9  ---------------MRSASNSNAPDKQWRNWLPLFVALVVIAEISFLVRLDVAEKAN---SWAESFYQFTTASWSTSKL
NbFucTF11 ---------------MRSASNSNAPNKQWRNWLPLFVALVIIAEFSFLVRLDVAEKAN---SWAESFYQFTTASWSTSKL
                                                               del

                90       100       110       120       130       140       150       160
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
NbFucTA1  EINEAGL----GVLRSSEVDRNLATGSCEEWLEKEDSVEYSRDFDKDPIFVHGGEKDWKSCAVGCNFGVDSDKKPDAAFG
NbFucTC5  EINEAGL----GVLRSSEVDRNLATGSCEEWLEKEDSVEYSRDFDKDPIFVHGGEK-----------------------
NbFucTA4  EINEAGL----GVLRSSEVDRNLATGSCEEWLEKEDSVEYSRDFDKDPIFVHGGEKDWKSCAVGCNIGVDSDKKPDAAFG
NbFucTD5  EINEAGL----GVLRSSEVDQNLETGSCEEWLEKEDSVEYSRDFDKDPIFVHGGEKDWKSCAIGCNFGVDSDKKPDAAFG
NbFucTB2  EINEAGL----GVLRSSEVDQNLETGSCEEWLEKEDSVEYSRDFDKDPIFVHGGEKDWKSCAIGCNFGVDSDKKPDAAFG
NbFucTE10 AVDGGDVDEVLLGVLSGEFDQGFLPWSCEEWLEREDYVAYARDFDNEPIFVHGPGQELKSCSIGCKFGTDSNKKPDAAFR
NbFucTG10 AVDGGDVDEVLLGVLSGEFDQGFLPWSCEEWLEREDYVAYARDFDNEPIFVHGPGQELKSCSIGCKFGTDSNKKPDAAFR
NbFucTA10 AVDGGDVDEVLLGVLSGEFDQGFLPWSCEEWLEREDYVAYARDFDNEPIFVHGPGQELKSCSIGCKFGTDSNKKPDAAFR
NbFucTF9  AVDGGDVDEVLLGVLSGEFDQGFLPWSCEEWLEREDYVAYARDFDNEPIFVHGPGQELKSCSIGCKFGTDSNKKPDAAFR
NbFucTF11 AVDHGDVEEVQLGVLSGEFDHGFVEGSCEEWLEREDSVAYSRDFDNEPIFVHGPGQELKSCSIGCKFGTDSNKKPDAAFR
              ins                                                          del

               170       180       190       200       210       220       230       240
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
NbFucTA1  TPQQAGTASVLRSMESAQYYPENNIVTARR-RGYDIVMTASLSSDPVGYFSWAEYDIMAPVQPKTENALAAAFISNCGA
NbFucTC5  ----------------------------KGI*
NbFucTA4  TPQQAGTASVLRSMESAQYYPENNIVTARR-RGYDIVMTTSLSSDPVGYFSWAEYDIMAPVQPKTENALAAAFISNCGA
NbFucTD5  TPQQTGTASVLRSMESSQYYPENNIVTARR-RGYDIIMTTSLSSDPVGYFSWAEYDIMAPVQPKTENALAAAFISNCGA
NbFucTB2  TPQQTGTASVLRSMESSQYYPENNIVTARRWKGI*
NbFucTE10 LPQQAGTASVLRSMESAQYYAENNITLARR-RGYDVVMTTSLSSDPVGYFSWAEYDIMAPVEPKTENALAAAFISNCGA
NbFucTG10 LPQQAGTASVLRSMESAQYYAENNITLARR-RGYDVVMTTSLSSDPVGYFSWAEYDIMAPVEPKTENALAAAFISNCGA
NbFucTA10 LPQQAGTASVLRSMESAQYYAENNITLARRWKGI*
NbFucTF9  LPQQAGTASVLRSMESAQYYAENNITLARRWKGI*
NbFucTF11 LPQQAGTASVLRSMESAQYYAENNITLARR-RGYDVVMTTSLSSDPVGYFSWAEYDIMAPVEPKTENALAAAFISNCGA
               del                       ins

               250       260       270       280       290       300       310       320
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
NbFucTA1  CNFRLQALEVLERANIKIDSFGSCHRNRDGNVDKVETLKCYKFSFAFENSNEEDYVTEKFFQSLVAGSVPVVIGAPNILD
NbFucTC5
NbFucTA4  RNFRLQALEVLERANIKIDSFGSCHRNRDGNVDKVETLKRYKFSFAFENSNED---TEKFFQSLVAGSVPVVIGAPNILD
NbFucTD5  RNFRLQALEVLERANIKIHSFGSCHRNRDGNVDKVETLKHYKFSFAFENSNEEDYVTEKFFQSLVAGSVPVVIGAPNILD
NbFucTB2
NbFucTE10 RNFRLQALEALERANIRIDSYGSCHHNRDGRVDKVAALKRYKFSLAFENSNEEDYVTEKFFQSLVAGSIPVVGAPNIQD
NbFucTG10 RNFRLQALEALERANIRIDSYGSCHHNRDGR~DKVAALKRYKFSLAFENSNEEDYVTEKFFQSLVAGSIPVVGAPNIQD
NbFucTA10
NbFucTF9
NbFucTF11 RNFRLQALEALERANIRIDSYGSCHHNRDGRVDKVAALKRYQFSLAFENSNEEDYVTEKFFQSLVAGSIPVVGAPNIQD
                                                                        del
```

SUPPLEMENTAL FIGURE 1 | Amino acid alignment of the *Nicotiana benthamiana* core α1,3-fucosyltransferases (NbFucTs) retrieved sequences. The ten retrieved NbFucT protein sequences are compared in an amino acid alignment. Positions that are less than 60% conservation are highlighted in black. Positions that are not similar, but show a higher than 60% conservation are highlighted in grey. A line under the sequences indicates the sequence specific characteristic deletions, insertions or succeeded nucleotides (mut) indicated in Figure 3.

```
                330       340       350       360       370       380       390       400
       ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
NbFucTA1   FAPSPNSLLHIKELKDAASVAKIMKYLAEHPSAYNESLSWKFEGPSDSFKALVDMAAVHSSCRLCIFLATSIREKEEKSP
NbFucTC5
NbFucTA4   FAPSPNSLLHIKELKDAASVAKTMKYLAENPSAYNESLRWKFEGPSDSFKALVDMAAVHSSCRLCIFLATSIREKEEKSP
NbFucTD5   FAPSPTSLLHIKELKDGASVAKTMKYLAENPSAYNESLRWKFEGPSDSFKALVDMAAVHSSCRLCIFLATSIREKEEKSP
NbFucTB2
NbFucTE10  FAPSPNSVLHIKEIKDAELIANTMTYLAQNPIASNESLRWKFEGPEDAFKALVDMAAVHSSCRLCIFLASRIQEREEHSP
NbFucTG10  FAPSPNSVLHIKEIKDAELIANTMTYLAQNPIASNESLRWKFEGPEDAFKALVDMAAVHSSCRLCIFLASRIQEREEHSP
NbFucTA10
NbFucTF9
NbFucTF11  FAPSPNSVLHIKEIKDAESIANTMKYLAQNPIAYNESLRWKFEGPSDAFKALVDMAAVHSSCRLCIFLASRIREREEQSP
                                                      mut

                410       420       430       440       450       460       470       480
       ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
NbFucTA1   KFTKRPCKCTRGSETVYHVYVRERGRFDMESVFLRSSNLSLEAFESAVLSKFKSLKHVPIWKEERPQILRGGDELKLYRV
NbFucTC5
NbFucTA4   KFTKRPCKCTRGSETVYHVYVRERGRFDMESIFLRSSNLSLEAFESAVLSKFKSLKHVPIWKEERPQILRGGEELKLYRV
NbFucTD5   KFTKRPCKCTRGSETVYHVYVRERGRFDMESVFLRSSNLSLEAFESAVLSKLKSLKHVPIWKDERPQILHGGDELKLYRI
NbFucTB2
NbFucTE10  KFTKRPCKCTRETETVYHVYVRERGRFEMDSIFLRSSDLSLKAFESAILSRFKSVKHVPVWREERPQVLRGGDELKLY*
NbFucTG10  KFTKRPCKCTRETETVYHVYVRERGRFEMDSIFLRSSDLSLKAFESAILSRFKSVKHVPVWREERPQVLRGGDELKLY*
NbFucTA10
NbFucTF9
NbFucTF11  KFMKRPCKCTRGTETVYHVYVGERGRFEMDSIFLRSSDLSLKAFESAILSRFKSVKHVPVWKEERPQVLRGGDELKLYKV

                490       500       510       520
       ....|....|....|....|....|....|....|....|
NbFucTA1   YPLGMTQRQALYTFKFKGDADFRNHIESHPCANFEAIFV*
NbFucTC5
NbFucTA4   YPLGMTQRQALYTFKFKGDADFRNHIESHPCANFEAIFV*
NbFucTD5   YPLGMTQRQALYTFKFKGDADFRNHIESHPCANFEAIFV*
NbFucTB2
NbFucTE10
NbFucTG10
NbFucTA10
NbFucTF9
NbFucTF11  YPVGLTQRQALFSFRFNGDTEFNNYIQSHPCAKFEAIFV*
```

SUPPLEMENTAL FIGURE 1 | Amino acid alignment of the *Nicotiana benthamiana* core α1,3-fucosyltransferases (NbFucTs) retrieved sequences. Continued
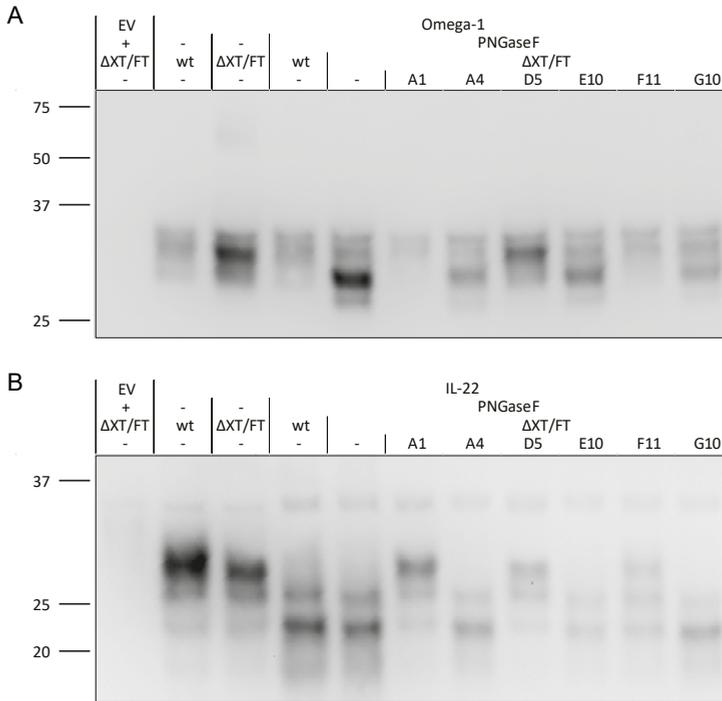
SUPPLEMENTAL FIGURE 2 | *Nicotiana benthamiana* **core α1,3-fucosyltransferase (NbFucT) activity on the N-glycans of omega-1 and IL-22.** Omega-1 (A) or IL-22 (B) was co-expressed in wild type (wt) or in ΔXT/FT *Nicotiana benthamiana* plants with NbFucTs (A1, A4, D5, E10, F11 or G10). After extraction from the apoplast fluid the N-glycan composition was analysed by PNGase F treatment. 500 ng omega-1 or IL-22 PNGase F treated sample was visualised on western blot with anti-FLAG antibodies to assess the presence of core α1,3-fucose.

Chapter 5

Production and glyco-engineering of *Schistosoma mansoni* omega-1 in plants

Kim van Noort [1], Linh Nguyen [2], Lotte B. Westerhof [1], Ben Excell [1], Cornelis H. Hokke [2], Arjen Schots [1] and Ruud H.P. Wilbers [1]

[1] Laboratory of Nematology, Wageningen University and Research, Wageningen, the Netherlands;
[2] Department of Parasitology, Leiden University Medical Centre, Leiden, The Netherlands.

5

Abstract

The human blood fluke *Schistosoma mansoni* controls host-immune responses by secreting immunomodulatory proteins, such as omega-1 a major constituent of *S. mansoni* soluble egg antigens. Mouse model studies have demonstrated the potential of omega-1 for the treatment of immune-related diseases, such as type-2 diabetes. Studies are however hampered by the limited availability of native parasite-derived omega-1. Moreover, recombinant protein production systems have thus far been unable to reconstitute helminth-like N-glycosylation essential for the uptake of omega-1 and C-type lectin mediated signalling. Here we show the production of omega-1 and exploited the flexibility of the N-glycosylation machinery of plants to reconstruct native helminth-like N-glycans on omega-1. Fine-tuning transient co-expression of specific glycosyltransferases in *Nicotiana benthamiana* enabled the synthesis of diantennary LeX carrying N-glycans with a fucosylated N-glycan core, similar to native *S. mansoni* omega-1. These data indicate that mimicking the complex carbohydrate structures of helminths in plants is a promising strategy to create essential tools for the targeted evaluation of therapeutic glycoproteins for the treatment of inflammatory disorders.

Introduction

Parasitic helminths can cause long lasting infections in humans and affect over 2 billion people worldwide. The capacity of helminths to establish long lasting infections can be attributed to their ability to modulate the immune system in the direction of a modified type 2 immune response. This immune response leads to a regulatory network that prevents exacerbated inflammatory immune responses. Remarkable is the inverse correlation between helminth infection and diabetes [1]. For example, a lower prevalence of diabetes was found in Chinese people that were previously infected with schistosomes [2], [3]. Similarly, a study on Flores island in Indonesia has shown that people infected with soil-transmitted helminths (STH) are more sensitive to insulin in comparison to non-infected people [4]. During infection, helminth secretions modulate the host immune system and skew the immune response to a modified Th2 response. This modified Th2 response may be beneficial for the low-grade inflammatory environment in obese patients. Experiments with helminth-infected animals show promising improvement of metabolic homeostasis and type-2 diabetes. For example, the lymph-dwelling filariae *Wuchereria bancrofti* decreases the pro-inflammatory cytokine response related to type-2 diabetes [5]. Diabetic mice infected with the nematode *Nippostrongylus brasiliensis* or injected with *Schistosoma mansoni* soluble egg antigens (SEA) showed improved insulin sensitivity and glucose tolerance [6], [7].

Although helminth infections revealed possibilities to treat or prevent diabetes, the use of helminths in therapy has clinical and pathological downsides. Therefore, research has focussed on the use of helminth derived molecules or molecular mixtures such as soluble egg antigens (SEA) of *Schistosoma mansoni*. Mice that were put on a high-fat-diet and treated with SEA showed a reduction in pro-inflammatory M1 macrophages as well as improved glucose tolerance and insulin sensitivity [7]. Also mice treated with the neo-glycoconjugate pentasaccharide lacto-N-fucopentaose III (LNFPIII) that contains LeX glycan motifs, showed improved insulin sensitivity [8]. During the egg stage of *S. mansoni* infection, the glycan motif LeX is mainly associated with the egg's secretions [9]. LeX carrying omega-1, can skew a Th2 response via dendritic cells (DCs) and has recently been shown to improve metabolic homeostasis in mice [10]–[14]. The main N-glycan structure on omega-1 contains α1,3 and α1,6 linked core-fucose and two LeX carrying antennae [14]. In a human *in vitro* DC model, the N-glycans on omega-1 are necessary for uptake by the mannose receptor (MR) on DCs [10], [11]. Once inside DCs, omega-1 skews a Th2 immune response through its ribonuclease (RNase) activity. Also, omega-1 can induce weight loss and improve metabolic homeostasis via binding to the MR on adipocytes [12]. Furthermore, obese mice treated with omega-1 show a decrease in white adipose tissue and improved glucose homeostasis.

For these reasons omega-1 is a promising helminth-derived molecule for treatment of type 2 diabetes. However, at the moment it is impossible to generate reasonable amounts of

native omega-1 required for research and possible future clinical studies. Native omega-1 can only be extracted from *S. mansoni* eggs in a time consuming and inefficient manner. Therefore, omega-1 was produced recombinantly in mammalian cells, although without native LeX containing N-glycans that could be crucial for efficient uptake of omega-1 [10]. To obtain omega-1 in quantities that are needed in research and therapy, a production system is required that produces large amounts of omega-1 carrying native helminth N-glycans. Over the years plants have shown to be a promising production platform for biopharmaceutical proteins. Moreover, plant N-glycans show limited variation that together with the tolerance to adaptions in their glycosylation machinery allows synthesis of custom-engineered N-glycans on glycoproteins such as omega-1. Glyco-engineering of the plant glycosylation pathway is necessary for production of omega-1 with native helminth N-glycans, since glycan motifs LeX and core α1,6-fucose are not present on native plant N-glycans. Furthermore, the typical plant N-glycan motif β1,2-xylose is not present on omega-1 N-glycans For production of glycoproteins without the typical plant N-glycan motifs, core α1,3-fucose and β1,2-xylose, *Nicotiana benthamiana* plants were generated wherein endogenous xylosyltransferases (NbXylTs) and fucosyltransferases (NbFucTs) were down regulated by RNA interference (ΔXT/FT *N. benthamiana* plants) [15].

LeX was introduced in *Nicotiana tabacum* by introduction of two hybrid glycosyltransferases, *Homo sapiens* β1,4-galactosyltransferase (HsGalT) and *Tetraodon nigroviridis* α1,3-FucT IXa (TnFucT9a), both targeted to the medial-Golgi compartment by the first 53 amino acids of *Arabidopsis thaliana* XylT (XylTHsGalT and XylTTnFucT9a) [16], [17]. However, introduction of the LeX motif resulted in hybrid N-glycan structures with only one LeX antenna, whereas the native N-glycans on *S. mansoni* omega-1 carry two LeX antennae. Diantennary galactose was introduced by targeting the HsGalT or a *Danio rerio* β1,4-GalT (DrGalT) to the *trans*-Golgi by the CTS domain (C, cytoplasmic tail; T, transmembrane domain; S, stem region) of *Rattus norvegicus* α2,6-sialyltransferase (sialHsGalT and sialDrGalT, respectively) [18], [19]. These studies show that targeting the β1,4-GalT to the *trans*-Golgi compartment is of uttermost importance.

In this chapter, we show that the production of large amounts of pure recombinant omega-1 with *S. mansoni* N-glycosylation in *N. benthamiana* plants is feasible. Omega-1 can be produced with diantennary LeX carrying N-glycans. Furthermore, core α1,3- and α1,6-fucosylation can be introduced on omega-1 N-glycans. Although, optimisation steps are required for homogeneous N-glycosylation, omega-1 can be produced with tailored N-glycans to study the effect of glycosylation on omega-1 function. Furthermore, the production of omega-1 enables studies to the immunomodulatory properties of omega-1 and the development of omega-1 as biopharmaceutical.

Results

Efficient production and purification of *Schistosoma mansoni* glycoprotein omega-1

To achieve high expression levels of omega-1 in plants, in-house codon optimised omega-1 was expressed by agroinfiltration in *N. benthamiana* plants. Production of omega-1 (~30 kDa) was analysed in crude extracts and apoplast fluids by SDS-PAGE (Figure 1A). Approximately 90% of omega-1 was recovered from the apoplast fluid. Secretion to the apoplast fluid with a low abundance of endogenous plant proteins enabled single-step purification of >0.5 mg omega-1 per plant (3–4 gram fresh leaf material) by cation exchange chromatography (CEX) (Figure 1B). Since omega-1 is a RNase, the RNase activity of plant produced omega-1 was assessed by incubation with total RNA from mouse bone marrow-derived DCs (Figure 1C). This assay revealed RNA breakdown, although to a lesser extent than RNA breakdown by RNase A. The N-glycan composition of purified omega-1 was assessed by Endo H / PNGase F and MALDI-TOF MS analysis of released N-glycans by PNGase A (Figure 1D). This analysis revealed that the predominant N-glycan on plant produced omega-1 is a paucimannosidic N-glycan with typical plant core α1,3-fucose and β1,2-xylose residues (peak 1308 m/z). Taken together, our data show that active glycosylated omega-1 can be produced in plants, which is easily purified from the apoplast.

Synthesis of LeX glycan motifs on omega-1

The function of omega-1 as Th2-polarising compound depends on its glycan-dependent uptake by DCs. In an attempt to synthesise the diantennary LeX glycan motifs of native omega-1 in plants, we introduced the two hybrid glycosyltransferases sialTnFucT9a and sialDrGalT. When a double CaMV 35S promoter was used to drive the expression of sialDrGalT, hybrid LeX-type N-glycans were obtained on omega-1 (peak 1865 m/z) (Figure 2A). This suggests that transient overexpression of sialDrGalT interferes with the activity of endogenous glycan-modifying enzymes. To reduce sialDrGalT expression, sialDrGalT was placed under the control of the weak constitutive promoter of the potato resistance gene Gpa2 in two different plant expression vectors (pHYG and pBIN).

Upon co-expression of sialTnFucT9a and *Gpa2*:sialDrGalT from pBIN, the predominant N-glycan type found on omega-1 is the typical plant paucimannosidic N-glycan (peak 1308 m/z) (Figure 2B). However, significant proportions of monoantennary and diantennary LeX carrying N-glycans were also found, while hybrid LeX-type N-glycans were absent. Co-expression of sialTnFucT9a and *Gpa2*:sialDrGalT from pHYG enabled synthesis of N-glycans carrying monoantennary LeX (peak 1673 m/z) (Figure 2C). The presence of LeX was confirmed by combined treatment of the PNGase A released N-glycans with α1,3/4-fucosidase from *Xanthomonas* sp. and β1,4/6-galactosidase from Jack bean and subsequent

5

MALDI-TOF MS N-glycan analysis (Supplemental Figure 1B). In conclusion, relatively homogeneous monoantennary LeX containing N-glycans can be synthesised on plant produced recombinant omega-1 by co-expression of sialTnFucT9a and *Gpa2*:sialDrGalT from pHYG.
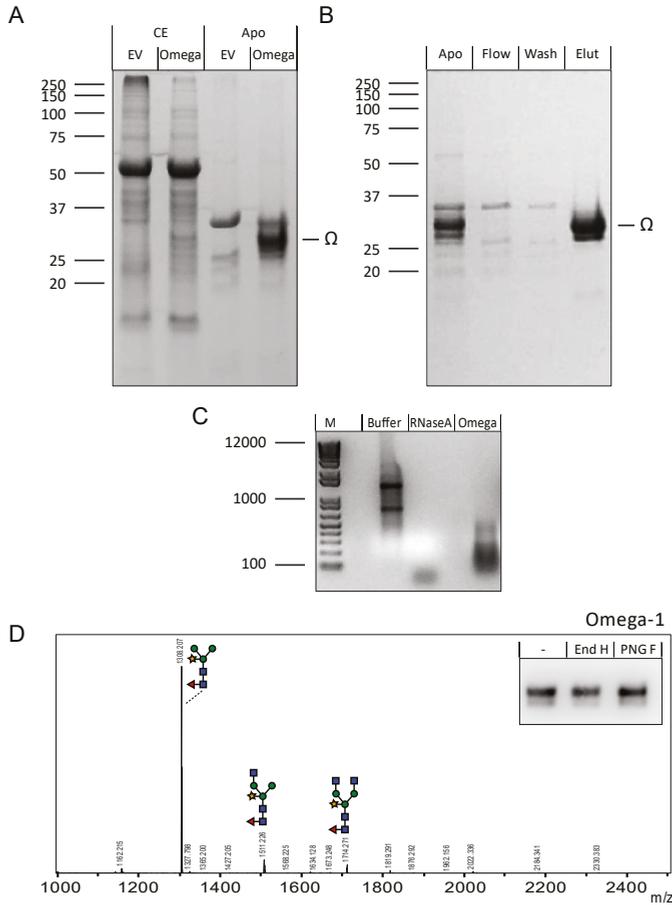
FIGURE 1 | Plant-based production of omega-1. Omega-1 was expressed in *Nicotiana benthamiana* plants. (A) Coomassie stained SDS-PAGE gel with crude extracts (CE) and apoplast fluids (Apo) from omega-1 (Ω) or empty vector (EV) infiltrated *N. benthamiana* plants reveals efficient secretion of omega-1. (B) Coomassie stained SDS-PAGE gel of different purification steps shows single-step cation exchange chromatography omega-1 (Ω) purification from leaf apoplast fluid. (C) The ability of purified plant produced omega-1 to degrade total RNA from mouse bone marrow-derived dendritic cells was tested by using 500ng purified omega-1 or RNase A (RNaseA) as positive control. RNA breakdown was assessed by agarose gel electrophoresis. (D) N-glycan composition of purified omega-1 was analysed by MALDI-TOF MS and N-glycan releasing enzymes Endo H (End H) and PNGase F (PNG F). The predominant N-glycans on omega-1 are paucimannosidic with β1,2-xylose and core α1,3-fucose.
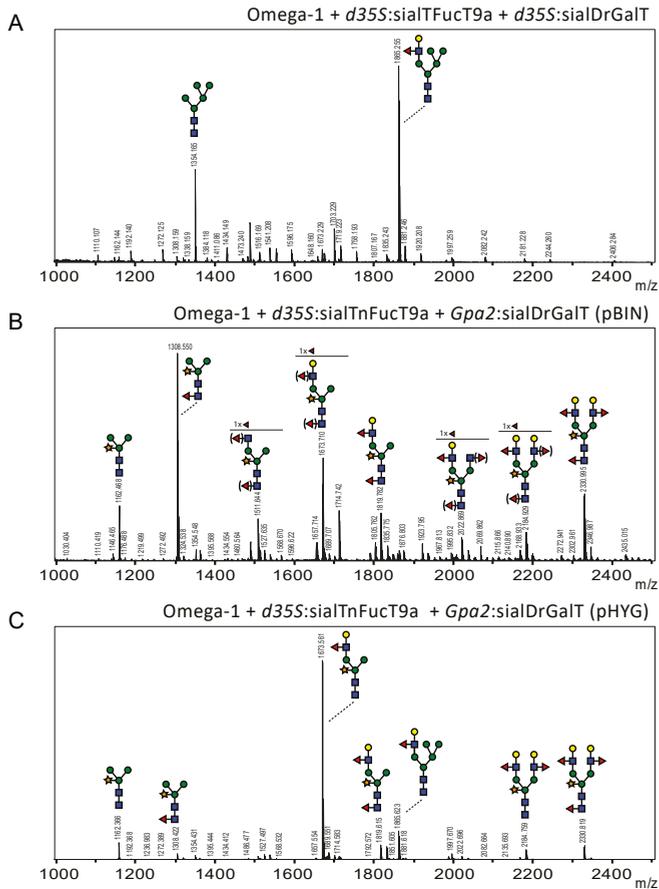
FIGURE 2 | Controlled expression of sialDrGalT enables synthesis of non-hybrid monoantennary N-glycans with LeX glycan motifs. Engineering of LeX on omega-1 was attempted with different sialDrGalT expression strategies. The N-glycan composition on purified omega-1 was analysed by MALDI-TOF MS. (A) N-glycan profile upon co-expression of *d35S:*sialTnFucT9a and *d35S:*sialDrGalT both in the pBIN vector. (B) N-glycan profile upon co-expression of *d35S:*sialTnFucT9a and *Gpa2:*sialDrGalT both in the pBIN vector. The weaker constitutive *Gpa2* promoter was chosen to reduce sialDrGalT expression. (C) N-glycan profile upon co-expression of *d35S:*sialTnFucT9a and *Gpa2:*sialDrGalT, but the latter being expressed using the pHYG vector. The pHYG vector was chosen as it generally yields more protein compared to pBIN. When a peak represents multiple N-glycan structures of identical mass, the peak indicates the major N-glycan present based on enzymatic digestions (for details see Supplemental Figure 1). When no enzymatic digestions are performed or the position is still unclear, the number of sugar residues of which the position on the N-glycan is not clear, this is indicated above the glycan structure. The possible positions of these sugar residues on the N-glycan are indicated between brackets.

5

Synthesis of diantennary LeX glycan motifs on omega-1

To synthesise N-glycans with diantennary LeX on omega-1, we introduced either one of the three exogenous GnTIIs from *Arabidopsis thaliana* (AtGnTII), *Caenorhabditis elegans* (CeGnTII) or *Homo sapiens* (HsGnTII). After extraction and purification of omega-1, its N-glycan composition was analysed on a western blot using anti-LeX antibodies and a Coomassie stained SDS-PAGE gel as sample loading control (Figure 3A). On the western blot bands representing omega-1 run slightly higher upon co-expression of AtGnTII or HsGnTII, suggesting synthesis of diantennary LeX. MALDI-TOF analysis of PNGase A released N-glycans suggested introduction of diantennary LeX on omega-1 upon co-expression of AtGnTII or HsGnTII by a shift of the predominant peak (from 1673 m/z to 2185 m/z) (Figure 3B, C and Supplemental Figure 2A). Upon co-expression of CeGnTII this shift was not observed, suggesting no introduction of diantennary LeX on omega-1 N-glycans (Supplemental Figure 2B). Since co-expression of sialDrGalT inhibits core α1,3-fucosylation on omega-1 N-glycans, we conclude that AtGnTII or HsGnTII co-expression increases the synthesis of diantennary LeX carrying N-glycans upon co-expression of omega-1, sialTnFucT9a and *Gpa2*:sialDrGalT from pHYG.

Core fucosylation of omega-1 N-glycans

To ultimately introduce native core fucosylation on omega-1 N-glycans carrying diantennary LN, co-expression of HsGnTII, *Gpa2*:sialDrGalT in pHYG and either one of the core FucTs, MmFucT8, NbFucTA1 (Chapter 4) or SmFucTC (Chapter 3) was investigated in ΔXT/FT *N. benthamiana* plants. The motif LN was expressed to unmistakably assess the presence of fucose in the core of the N-glycan and not in the antennae. MmFucT8 was introduced to obtain core α1,6-fucosylation of omega-1 N-glycans, whereas NbFucTA1 and SmFucTC were introduced to obtain core α1,3-fucosylation. After extraction and purification, N-glycans on omega-1 were analysed with the fucose binding *Aleuria aurantia* lectin (AAL). The lectin binding assay revealed that AAL bound to omega-1 N-glycans upon co-expression of each of the three core FucTs (Figure 4A and B). This indicates that all three FucTs were able to core fucosylate the N-glycans on omega-1 upon co-expression of glycosyltransferases that synthesise diantennary LN carrying N-glycans on omega-1.

Core α1,6-fucosylation of omega-1 N-glycans by MmFucT8 was confirmed by binding of the core α1,6-fucose binding *Pholiota squarrosa* lectin (PhoSL) in a lectin binding assay upon co-expression of MmFucT8 (Figure 4C). Core α1,3-fucosylation was assessed by PNGase F treatment and subsequent visualisation on a western blot using anti-FLAG antibodies (Figure 4D). PNGase F releases N-glycans only in the absence of a core α1,3-fucose. Therefore, the size of omega-1 is reduced by 1.5 kDa for release of one N-glycan or by 3 kDa for the release of both its N-glycans. Upon co-expression of NbFucTA1 or SmFucTC the N-glycans on omega-1 were partly protected from release by PNGase F, indicating the presence of core α1,3-fucose.

To confirm core fucosylation, N-glycans were analysed by MALDI-TOF MS upon release by PNGase A. Almost complete core α1,6-fucosylation was observed upon co-expression of MmFucT8 by a shift of the predominant peak (from 1030 m/z to1176 m/z) (Supplemental Figure 3A and B). In contrast, co-expression of NbFucTA1 or SmFucTC resulted in incomplete core α1,3-fucosylation (Supplemental Figure 3C and D). Remarkably, almost no diantennary LN carrying N-glycans were observed in these MALDI-TOF MS spectra. In conclusion, core fucosylation can be introduced on omega-1 N-glycans upon co-expression of MmFucT8, NbFucTA1 or SmFucTC.

FIGURE 3 | Co-expression of *Homo sapiens* GnTII for synthesis of diantennary LeX carrying N-glycans on omega-1 in plants. Engineering of diantennary LeX carrying N-glycans on purified omega-1 was attempted by transient co-expression of Gpa2:sialDrGalT in pHYG, sialTnFucT9a and HsGnTII. (A) The N-glycan composition on purified omega-1 was analysed by Coomassie stained SDS-PAGE gel (left) and a western blot treated with anti-LeX antibodies (right). The SDS-PAGE gel and western blot suggest differences in glycosylation based on variations in protein size. *(B-C)* The N-glycan composition of omega-1 was analysed by MALDI-TOF MS upon co-expression of *Gpa2*:sialDrGalT in pHYG with sialTnFucT9a (B) or sialTnFucT9a and HsGnTII (C). When a peak represents multiple N-glycan structures of identical mass, the number of sugar residues of which the position on the N-glycan is not clear is indicated above the glycan structure. The possible positions of these sugar residues on the N-glycan are indicated between brackets.

Double core fucosylation of omega-1 N-glycans

Next, synthesis of omega-1 with a double fucosylated core and diantennary LN carrying N-glycans was investigated. Therefore, MmFucT8 and NbFucTA1 or SmFucTC were co-expressed with omega-1, HsGnTII and *Gpa2*:sialDrGalT in pHYG. After extraction and purification, N-glycans on omega-1 were analysed with PhoSL and PNGase F. A lectin assay with PhoSL revealed that the amount of core α1,6-fucose declines upon co-expression of NbFucTA1 or SmFucTC, suggesting that NbFucTA1 and SmFucTC both interfere with MmFucT8 activity (Figure 4C). PNGase F treated samples were analysed on a western blot using anti-FLAG antibodies (Figure 4D). On the western blot a 27 Da and 28.5 kDa band were observed for omega-1 upon co-expression of MmFucT8 and NbFucTA1, indicating that MmFucT8 interferes with core α1,3-fucosylation by NbFucTA1. Upon co-expression of SmFucTC also a 30 kDa band was observed for omega-1, although interference of MmFucT8 remained this band indicated that on a part of the produced omega-1 both N-glycans carry a core α1,3-fucose. MALDI-TOF MS N-glycan profiles of omega-1 upon co-expression of MmFucT8 and NbFucTA1 or SmFucTC revealed the presence of double core fucosylated N-glycans, indicating that MmFucT8 and the core α1,3-FucTs do not completely inhibit each other (Figure 4E and Supplemental Figure 3E). Remarkably, also in this experiment almost no diantennary LN carrying N-glycans were observed in the MALDI-TOF MS N-glycans profiles. Altogether, the core α1,6-FucT and the α1,3-FucTs seem to inhibit one another upon co-expression.

5
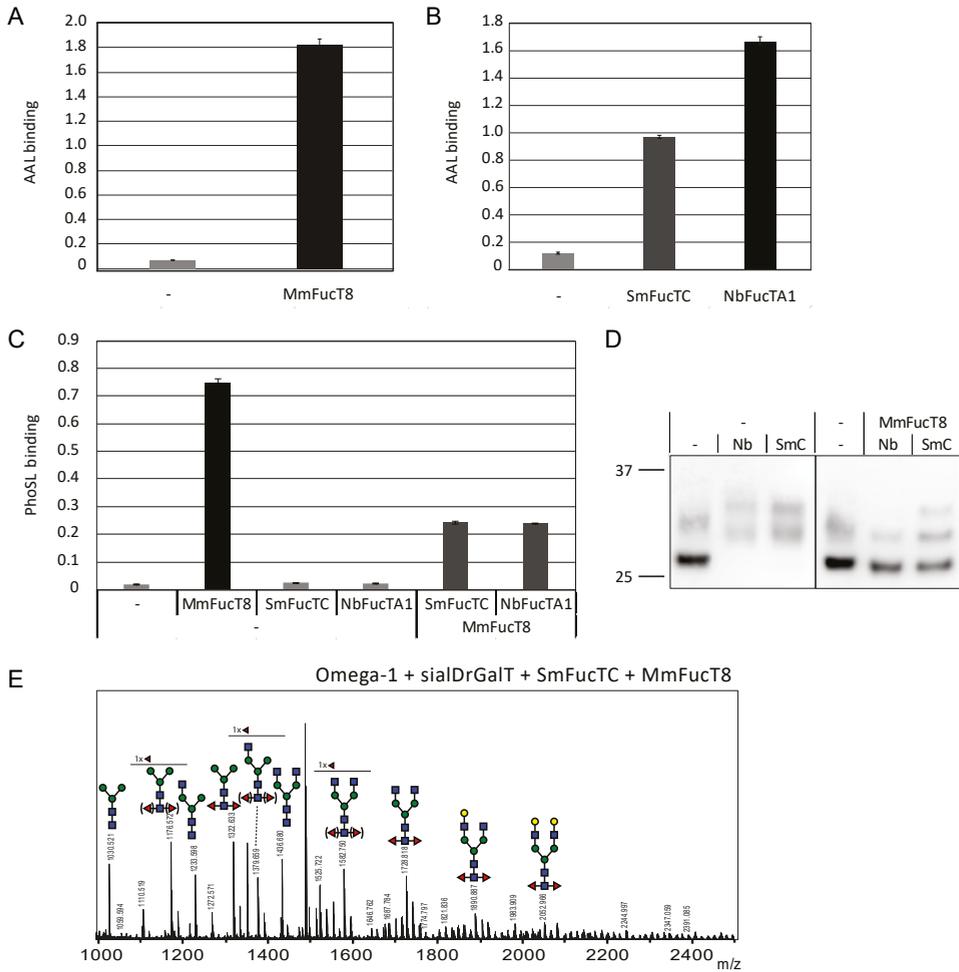
FIGURE 4 | Introduction of core fucosylation on diantennary LN carrying N-glycans on omega-1. Omega-1 was co-expressed with HsGnTII, *Gpa2*:sialDrGalT in pHYG and the core FucTs, MmFucT8, NbFucTA1 or SmFucTC in ΔXT/FT *Nicotiana benthamiana* plants. After purification the N-glycan composition of omega-1 was analysed for core fucosylation. (A) *Aleuria aurantia* lectin (AAL) binding assay reveals that MmFucT8 adds a fucose. (B) AAL binding assay reveals that NbFucTA1 and SmFucTC add a fucose. (C) *Pholiota squarrosa* lectin (PhoSL) binding assay for the specific detection of α1,6-linked fucose, reveals that MmFucT8 adds an α1,6-fucose, which is inhibited upon co-expression of SmFucTC or NbFucTA1. (D) 700ng omega-1 treated with PNGase F was visualised on a western blot using anti-FLAG antibodies. This western blot reveals that NbFucTA1 (Nb) and SmFucTC (SmC) add a core α1,3-fucose, which is partly inhibited upon co-expression of MmFucT8. (E) MALDI-TOF MS N-glycan profile of omega-1 upon co-expression of HsGnTII, *Gpa2*:sialDrGalT in pHYG, MmFucT8 and SmFucTC. When a peak represents multiple N-glycan structures of identical mass, the number of sugar residues of which the position on the N-glycan is not clear is indicated above the glycan structure. The possible positions of these sugar residues on the N-glycan are indicated between brackets.

Core fucosylation of omega-1 N-glycans carrying diantennary LeX

In a final experiment we attempted to introduce both core fucoses on omega-1 carrying diantennary LeX glycans. Therefore, omega-1 was co-expressed with HsGnTII, *Gpa2*:sialDrGalT in pHYG, sialTnFut9a, MmFucT8 and NbFucTA1 or SmFucTC. Again, in order to determine core α1,3- or α1,6-fucosylation samples were analysed with PhoSL and PNGase F. The PhoSL Lectin assay showed a slight increase in PhoSL binding in comparison to co-expression with LN synthesising glycosyltransferases (Figure 5A). This indicates that interference with core α1,6-fucosylation is slightly reduced upon co-expression of sialTnFucT9a. The western blot treated with anti-FLAG antibodies showed no band at 27 kDa for omega-1 upon co-expression of MmFucT8, glycosyltransferases that synthesise diantennary LeX N-glycans and NbFucTA1 or SmFucTC (Figure 5B). This indicated that interference of MmFucT8 with core α1,3-fucosylation is also reduced upon co-expression of sialTnFucT9a.

N-glycans of omega-1 were analysed by MALDI-TOF MS upon release by PNGase A. The N-glycan profiles revealed presence of core fucosylation and LeX synthesis (Figure 5C and Supplemental Figure 4). To determine the proportion of LeX carrying N-glycans the released N-glycans of omega-1 upon co-expression of SmFucTC were treated with β1,4/6-galactosidase from Jack bean (Figure 5D). To determine the proportion of core α1,6-fucose the released N-glycans of omega-1 upon co-expression of SmFucTC were treated with hydrofluoric acid (Figure 5E). Together these treatments revealed that some core α1,6-fucose was present, although most of the fucoses were α1,3 linked of which the majority composes LeX. In conclusion, co-expression of omega-1 with HsGnTII, *Gpa2*:sialDrGalT, sialTnFucT9a, MmFucT8 and SmFucTC resulted in synthesis of diantennary LeX carrying N-glycans with a fucosylated N-glycan core on omega-1 (Figure 6).
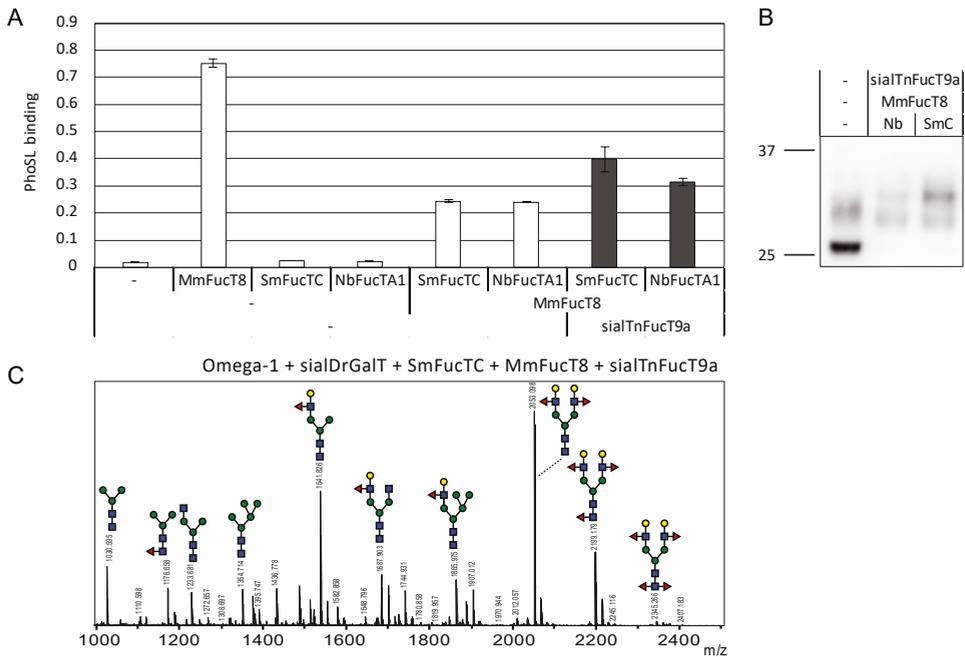
5

FIGURE 5 | Synthesis of double core fucosylated diantennary LeX carrying N-glycans on omega-1.
Omega-1 was co-expressed with HsGnTII, *Gpa2*:sialDrGalT in pHYG, MmFucT8 and each of the core α1,3-FucTs
NbFucTA1 or SmFucTC in ΔXT/FT *Nicotiana benthamiana* plants. After purification, the N-glycan composition of
omega-1 was analysed for core fucosylation. (A) *Pholiota squarrosa* lectin (PhoSL) binding assay for the specific
detection of α1,6-linked fucose, reveals a slight increase in α1,6-fucosylation by MmFucT8 upon introduction of
sialTnFucT9a (in grey) in comparison to the earlier presented data without introduction of sialTnFucT9a (Figure
4C, here indicated in white). (B) 700ng omega-1 treated with PNGase F was visualised on a western blot using
anti-FLAG antibodies. This western blot reveals core α1,3-fucosylation by NbFucTA1 (Nb) and SmFucTC (SmC). (C)
MALDI-TOF MS N-glycan profile of omega-1 upon co-expression of HsGnTII, *Gpa2*:sialDrGalT in pHYG, MmFucT8
and SmFucTC. When a peak represents multiple N-glycan structures of identical mass, the peak indicates the
major N-glycan present based on results in B, D and E. (D-E) Profile of the same N-glycans upon treatment with
β1,4/6-galactosidase from Jack bean (β1,4-gal) (D) or hydrofluoric acid (HF) (E).
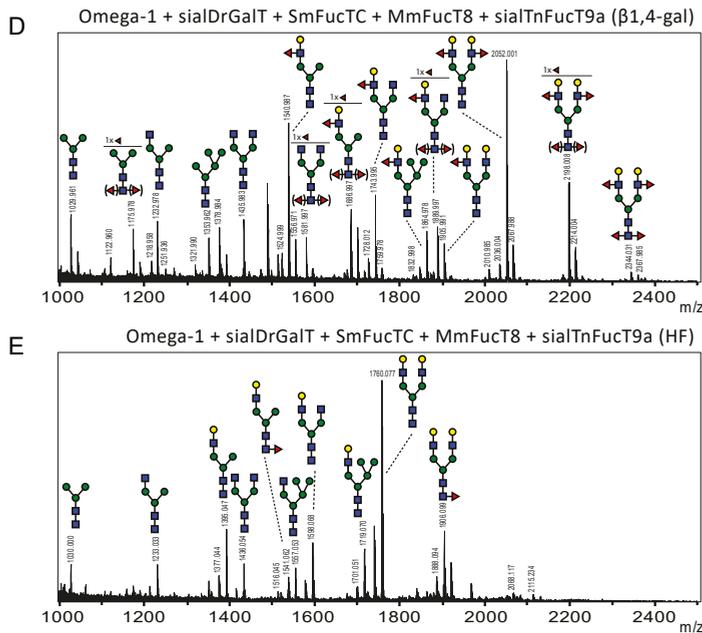
5

FIGURE 5 | Synthesis of double core fucosylated diantennary LeX carrying N-glycans on omega-1. Continued
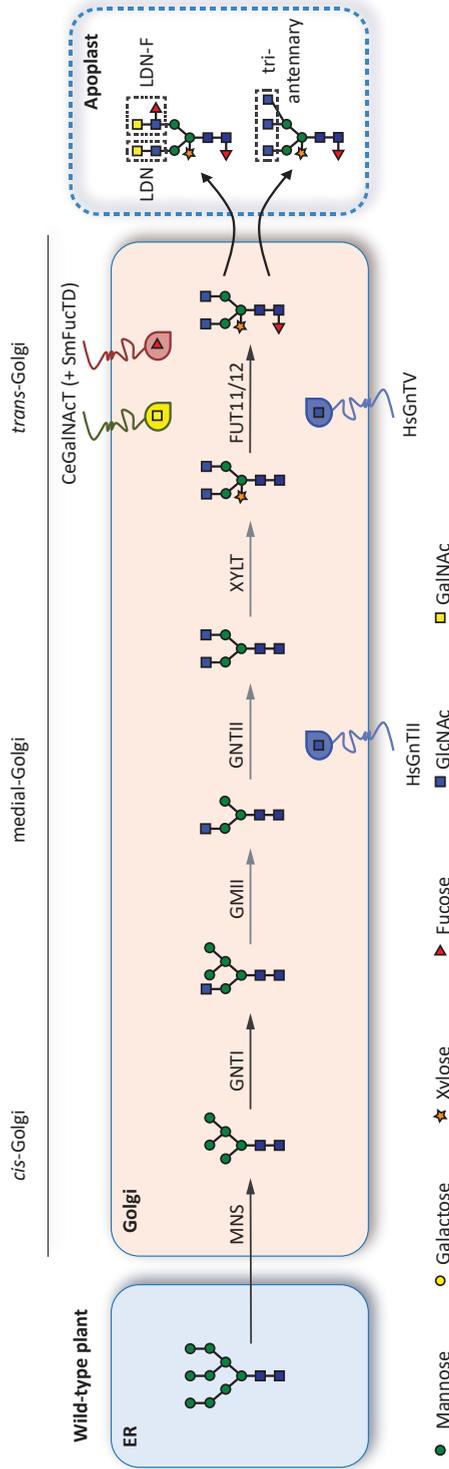
FIGURE 6 | Glyco-engineering for the production of helminth glycoprotein omega-1 with tailored N-glycans. A schematic overview of the N-glycan modifying steps in the plant Golgi-system of ΔXT/FT *Nicotiana benthamiana* plants. The plant N-glycosylation machinery was engineered by introducing (hybrid) glycosyltransferases that allow core fucosylation (MmFucT8 and SmFucTC) and synthesis of diantennary LeX carrying N-glycans (HsGnTII, sialDrGalT and sialTnFucT9a) upon agro co-infiltration with omega-1. Since, omega-1 is mainly produced with one core fucose, depending on the co-expressed core-FucT, both core fucoses are indicated in brackets. MNS: Class I mannosidases MNS1, 2 and 3; GnTI: N-acetyl-glucosaminyltransferase I; GMII: Golgi-α-mannosidase II; GnTII: N-acetyl-glucosaminyltransferase II; XYLT: β1,2-xylosyltransferase; FUT11/12: core α1,3-fucosyltransferase. The red cross indicates the two *N. benthamiana* glycosyltransferases that are down regulated in ΔXT/FT *N. benthamiana* plants. Arrows are indicated in grey, when alternative N-glycan processing routes have been postulated.

Discussion

Omega-1 is one of the most promising helminth glycoproteins for use as modulator of inflammatory conditions, including type 2 diabetes. Omega-1 can be extracted from *S. mansoni* SEA, however extraction is laborious and yields are relatively low. In this study, we investigated the production of helminth-like omega-1 in plants as an alternative and show that one plant can produce >0.5 mg omega-1. Moreover, we were able to produce omega-1 with native diantennary LeX carrying N-glycans with a fucosylated glycan core by introduction of HsGnTII, sialDrGalT, SmFucTC, MmFucT8 and sialTnFucT9a (Figure 6) [14].

Production and efficient isolation of recombinant omega-1 was achieved by codon optimisation of the open reading frame, addition of a chitinase signal peptide for secretion and transcription via a dual 35S promoter from the plant expression vector pHYG [20], [21]. Due to secretion 90% plant produced omega-1 could be recovered from the apoplast. This is remarkable efficient in comparison to for instance plant produced IL-22 of which only 35% could be retrieved from the apoplast under the same conditions [22]. The efficient secretion enabled one-step purification of >0.5 mg of omega-1 per plant (3-4 gram fresh leaf material), due to low abundance of endogenous plant proteins in the apoplast. N-glycan analysis of omega-1 revealed that omega-1 mainly displays terminal mannose residues on typical plant paucimannosidic N-glycans with core α1,3-fucose and β1,2-xylose residues. For the addition of core α1,3-fucose and β1,2-xylose FucT and XylT require preceding GnTI activity. This indicates that the terminal GlcNAc residues were probably cleaved later in the secretory pathway or in the apoplast by β-hexosaminidases (HEXO) [23], [24]. Two GlcNAc cutting HEXOs have been characterised in *N. benthamiana*, NbHEXO1 and NbHEXO3. NbHEXO1 is localised in the vacuole and NbHEXO3 is localised at the plasma membrane and/or the apoplast. As both NbHEXOs carry Golgi-processed N-glycans, NbHEXO1 and NbHEXO3 probably pass the secretory pathway similar to omega-1 [24]. In the acidic Golgi compartments or in the apoplast NbHEXOs are able to cleave the accessible terminal GlcNAc residues from omega-1 N-glycans, leading to paucimannosidic N-glycans on omega-1.

We have shown efficient omega-1 production and purification from plants with typical plant N-glycans. To mimic production of native omega-1 carrying diantennary LeX N-glycans in plants we transiently co-expressed omega-1 with glycosyltransferases that allow synthesis of LeX. Overexpression or medial-Golgi localisation of β1,4-GalT interferes with endogenous glycan modifying enzymes such as Golgi α-mannosidase (GM) II, GnTII, XylT and FucT, resulting in hybrid LeX-type N-glycans [16], [25], [26]. Therefore, we targeted DrGalT to the *trans*-Golgi with the CTS domain of *Rattus norvegicus* α2,6-sialyltransferase (sialDrGalT). Expression of sialDrGalT under control of the strong dual 35S promoter resulted in hybrid LeX-type N-glycans without core fucose and xylose as also reported

by Rouwendal and colleagues (2009). This indicates that the localisation and activity of sialDrGalT is not restricted to the *trans*-Golgi when the enzyme is expressed under the control of the strong dual 35S promoter. Therefore, we refined sialDrGalT expression by introduction of a weak constitutive promoter from the plant resistance gene Gpa2, driven from two different expression vectors pBIN and pHYG.

Whereas expression from pBIN resulted in a significant proportion of typical plant paucimannosidic N-glycans, expression from pHYG resulted in relatively homogeneous N-glycans carrying monoantennary LeX. The use of both vectors resulted in N-glycans carrying monoantennary and diantennary LeX motifs containing a core β1,2-xylose. Moreover, no hybrid LeX-type N-glycans were observed. So, not only targeting β1,4-GalT to the proper Golgi position is important, but also control of sialDrGalT expression to prevent overflow to other sub-Golgi compartments. Our findings demonstrate that controlled expression of a *trans*-Golgi targeted sialDrGalT is required to prevent inhibition of endogenous GMII and XylT in the medial-Golgi. These findings point at another layer of control in the Golgi, as both a proper CTS domain and controlled expression of glycosyltransferases are required for correct localisation of their biological activity in the Golgi [27].

Controlled expression of *Gpa2*:sialDrGalT from pHYG resulted in relative high amounts of monoantennary LeX carrying N-glycans on omega-1. However, the addition of galactose on the α1,3-mannose branch still appeared to interfere with the activity of *N. benthamiana* GnTII (NbGnTII). Co-expression of AtGnTII and HsGnTII resulted in an increased fraction of diantennary LeX carrying N-glycans on omega-1. The inability of CeGnTII to yield diantennary LeX carrying N-glycans can be because of a difference in localisation and/or glycan acceptor in comparison to NbGnTII. In *N. benthamiana*, NbGnTII is probably localised after NbXylT, since expression of sialDrGalT from pHYG does interfere with NbGnTII activity, whereas it does not interfere with NbXylT activity. So, because of NbXylT localisation before GnTII, the presence of a core β1,2-xylose on the glycan acceptor can interfere with the activity of exogenous GnTIIs, when originally core β1,2-xylose is absent on the exogenous N-glycans. This may explain the inactivity of CeGnTII, although HsGnTII does not seem to be inhibited by NbXylT activity.

Next to, interference with NbGnTII, expression of *Gpa2*:sialDrGalT from pHYG also interferes with core fucosylation. Therefore, one of the two earlier characterised core α1,3-FucTs (SmFucTC and NbFucTA1, Chapter 3 and 4, respectively) or core α1,6-FucT MmFucT8 was introduced for introduction of core fucosylation. All core FucTs were able to introduce a core fucose upon co-expression of omega-1 with *Gpa2*:sialDrGalT in pHYG and HsGnTII. However, introduction of double core fucosylation by co-expression of the α1,6-FucT MmFucT8 and SmFucTC or NbFucTA1 revealed an incompatibility between MmFucT8 and NbFucTA1 or SmFucTC. MmFucT8 carries the CTS domain of the Arabidopsis core α1,3-

FucT AtFUT11 and therefore probably localises at the same Golgi position as the core α1,3-FucTs SmFucTC and NbFucTA1. This suggests that MmFucT8 and the core α1,3-FucTs may compete for the same glycan acceptor. When sialTnFucT9a is co-expressed in this mixture of transferases, core α1,3-fucosylation almost seems to be restored and increased core α1,6-fucosylation was observed.

Remarkable in the experiments with diantennary LN glycan synthesising enzymes is the absence of LN on the majority of the omega-1 N-glycans. Since, addition of sialTnFucT9a resulted in diantennary LeX carrying N-glycans on omega-1, it seems that the terminal galactoses are cleaved off in the absence of sialTnFucT9a. Goulet and colleagues [28] have shown that expression of galactosidases is upregulated upon *Agrobacterium* infection, which could explain cleavage of terminal galactose. Striking is that in an earlier paper published by Strasser and colleagues [18] terminal galactoses are not cleaved off, when anti-HIV monoclonal antibody (mAb) 4E10 was produced in HsGalT expressing ΔXT/FT plants. However, for mAbs is shown that addition of a core α1,3-fucose results in a conformational change of the N-glycan that influences glycan accessibility [29]. Since, expression of mAb 4E10 in wild type (wt) plants seems to result in galactose cleavage, the absence of a core α1,3-fucose on the N-glycans of mAb 4E10 produced in ΔXT/FT plants could prevent glycan accessibility and as a result prevent terminal galactose cleavage.

Co-expression of omega-1 with HsGnTII, *Gpa2*:sialDrGalT in pHYG, sialTnFucT9a, MmFucT8 and SmFucTC resulted in omega-1 carrying diantennary LeX N-glycans with core fucose. With this glycan structure the effect of N-glycan core fucosylation on the activity of omega-1 can be assessed. In the publication corresponding to this chapter we show that the presence of one antennae with terminal LeX enhances Th2 polarisation in mice [30]. IL-10 mRNA levels in DCs were enhanced by binding to DC SIGN. This suggests an additional DC-SIGN-mediated Th2-related signal to DCs, next to the MR-mediated effect observed after internalisation and subsequent RNase activity of non-glycan engineered plant produced omega-1. The question remains if N-glycan core fucosylation and/or diantennary LeX further modulate the activity of omega-1.

In this study, we introduce plants as a production platform for helminth secreted glycoproteins. Thereby introducing an alternative for the use of live parasites, their secretions or recombinant proteins with inappropriate glycosylation to investigate the immunomodulatory properties of helminths. The ability to mimic the helminth glycome in plants fulfils the increasing demand for helminth glycoproteins by academic and industrial research communities to investigate how helminths are able to dampen host immune responses. Moreover, with this platform the effect of protein glycosylation on protein function can be studied, by controlled changes in protein glycosylation. Ultimately, this research could lead to the development of a new class of biopharmaceuticals for the treatment of autoimmune diseases and other chronic inflammatory disorders, such as type 2 diabetes.

Materials and Methods

Construction of expression vectors

The complete sequence encoding the *S. mansoni* protein omega-1 was codon optimised in-house [31]. The protein sequence was preceded by a signal peptide from the *A. thaliana* chitinase gene (cSP) and a 6x histidine-FLAG tag (H6F) was included at the N-terminus for detection and purification purposes (Supplemental Figure 5). The optimised gene was synthetically constructed at GeneArt. The full sequence was cloned into a pHYG expression vector [20]. Hybrid α1,3-FucT IXa from *Tetraodon nigroviridis* (TnFucT9a) and hybrid β1,4-GalT from *Danio rerio* (DrGalT) driven from the pBIN-PLUS expression vector were used to synthesise LeX structures [16], [19], [32]. The N-terminal CTS domain of *Rattus norvegicus* α2,6-sialyltransferase was used to enable *trans*-Golgi targeting of both transferases, from now on referred to as sialTnFucT9a and sialDrGalT, respectively. For controlled expression of sialDrGalT we replaced YFP in the pRAP-pGpa2::Gpa2-3′ UTR-YFP vector with sialDrGalT, thereby placing the expression of sialDrGalT under the control of the Gpa2 gene promoter region [33]. The entire Gpa2:sialDrGalT expression cassette was transferred to pBIN and pHYG. To introduce a second antennae on omega-1 N-glycans upon co-expression of sialDrGalT three exogenous GnTIIs were co-expressed, from *A. thaliana* (AtGnTII), *C. elegans* (CeGnTII) and *H. sapiens* (HsGnTII). The construct for overexpression of AtGnTII was kindly provided by Prof. H. Steinkellner (University of Natural Resources and Life Sciences, Vienna) [34]. The sequence of CeGnTII was synthesised by GeneArt with flanking BspHI and KpnI restriction sites and subsequently cloned into pHYG via the flanking restriction sites (Supplemental Figure 5). HsGnTII was amplified using illustra PuReTaq Ready-To-Go PCR Beads (GE Healthcare) from human MGAT2 gene cDNA clone (Sino Biological) flanking BspHI and BsrGI restriction sites and cloned into the pHYG plant expression vector (primers indicated in Supplemental Table 1). To introduce core fucosylation on omega-1 N-glycans, two core α1,3-FucTs and one core α1,6-FucT were co-expressed. The two core α1,3-FucTs were earlier characterised as core α1,3-FucT from *S. mansoni* and *N. benthamiana*, SmFucTC (Chapter 2 and 3) and NbFucTA1 (Chapter 4) respectively. As core α1,6-FucT, Fut8 of *Mus musculus* fused with the CTS domain of *A. thaliana* FUT11 was co-expressed (MmFucT8) [22]. In all experiments the silencing suppressor P19 from tomato bushy stunt virus in pBIN61 was co-infiltrated to enhance expression [35]. All constructs were transformed into *Agrobacterium tumefaciens* strain MOG101 for plant expression.

Agroinfiltration

A. tumefaciens clones were cultured for 16 hrs at 28 °C/250 rpm in LB medium (10 g/L pepton140, 5 g/L yeast extract, 10 g/L NaCl, pH 7.0) containing 50 µg/mL kanamycin and 20 µM acetosyringone. The bacteria were suspended in MMA infiltration medium (20 g/L sucrose, 5 g/L MS-salts, 1.95 g/L MES, pH 5.6) containing 200 µM acetosyringone. For

5

co-infiltration experiments *Agrobacterium* cultures were mixed while maintaining a final OD of 0.5 per culture, except MmFucT8 which was infiltrated with a final OD of 0.1. The two youngest fully expanded leaves of four to six week old wt or ΔXT/FT *N. benthamiana* plants were infiltrated completely by injecting the *Agrobacterium* suspension into a *N. benthamiana* leaf at the abaxial side using a 1 mL needles syringe [15]. *N. benthamiana* plants were maintained in a controlled greenhouse compartment (UNIFARM, Wageningen) and infiltrated leaves were harvested at five to six days post infiltration.

Protein extraction

For isolation of total apoplast proteins, harvested leaves were submerged in ice-cold extraction buffer (50 mM phosphate-buffered saline, pH 8.0, containing 100 mM NaCl and 0.1% v/v Tween-20) after which vacuum was applied for 10 min and subsequently slowly released to ensure infiltration of the apoplast. To collect the apoplast fluid leaves were placed in 10 mL syringes and centrifuged for 10 min/2000 ×g. Apoplast fluids were clarified by centrifugation for 5 min/16.000 ×g at 4 °C. If required, remaining intracellular proteins were isolated from the leaves by homogenisation in liquid nitrogen. Homogenised plant material was ground in ice-cold extraction buffer (50 mM phosphate-buffered saline, pH 8.0, containing 100 mM NaCl, 0.1% v/v Tween-20 and 2% w/v immobilized polyvinylpolypyrrolidone (PVPP)) using 2 mL/g fresh weight. Crude extracts were clarified by centrifugation for 5 min/16.000 ×g at 4 °C. Total protein content was analysed by a Pierce Bicinchoninic Acid Protein Assay (BCA, Fisher Scientific). Total soluble plant proteins were separated under reducing conditions by SDS-PAGE on a 12% Bis-Tris gel (Invitrogen) and subsequently stained with Coomassie brilliant blue staining.

Purification from the apoplast fluid

Plant produced omega-1 was purified from the apoplast fluid on Pierce Strong Cation Exchange Mini Spin Columns (Fisher Scientific). In brief, apoplast fluids were transferred over G25 Sephadex columns to exchange the extraction buffer for CEX binding buffer (10 mM Sørensen's phosphate buffer, pH 6.0, containing 100 mM NaCl). Column loading, washing and elution was done by centrifugation for 5 min/2000 x*g*. Bound omega-1 was eluted with elution buffer (Tris-HCl buffer, pH 9.0 containing 2 M NaCl). After elution samples were dialysed against PBS and protein concentration was determined with BCA (Fisher Scientific). Aliquots of different purification steps and eluted omega-1 were separated under reducing conditions by SDS-PAGE on a 12% Bis-Tris gel (Invitrogen) and subsequently stained with Coomassie brilliant blue staining.

RNase activity assay

Total RNA was isolated from 5.0*10^6 bone marrow-derived DCs with the RNeasy Mini Kit (Qiagen) according to the suppliers protocol. 1 μg of total RNA was incubated with 500 ng

purified omega-1 in PBS for 1 hr/37 °C. PBS only was used as negative control and 500 ng RNase A was used as positive control. RNA breakdown was analysed on a 1.5% agarose gel and visualised with UV light.

Lectin binding assay

Core fucosylation of omega-1 N-glycans was analysed by a lectin binding assay with biotinylated *Aleuria aurantia* lectin (AAL, Bio-Connet). *Pholiota squarrosa* lectin (PhoSL) was used to screening for core α1,6-fucosylation. PhoSL was kindly provided by Dr. A. Varrot (Université Grenoble Alpes, Grenoble, France) and was biotinylated in house following manufactures protocol (Pierce) [36]. For the lectin binding assay purified omega-1 in PBS was coated overnight at 4 °C on Microtiter plates at a protein concentration of 1 µg/mL for binding by AAL and 10 µg/mL for binding by PhoSL. Plates were blocked with carbohydrate-free blocking buffer (Vector Laboratories). All subsequent steps were performed in blocking buffer. Plates were then incubated for 1 hr/RT with biotinylated lectin at a concentration of 2 µg/mL AAL or 1.8 µg/mL PhoSL. Subsequently, plates were incubated with avidin-HRP (eBioscience) for 30 min/RT. After every incubation step the Microtiter plate was washed 5 times with PBST (PBS containing 0.05 % v/v Tween-20). Lectin binding was visualised by TMB substrate (Fisher Scientific) and absorbance was measured at a wavelength of 450 nm while using 655 nm as reference filter.

SDS-PAGE and western blot

For screening of core α1,3-fucosylation, samples were denatured and subsequent incubated with Peptide:N-glycosidase (PNGase) F for 1 hr/37 °C . After PNGase F treatment samples were analysed on a 12 % Bis-Tris SDS-PAGE gel (Fisher Scientific) followed by transfer to a PVDF membrane by wet blotting. For glycan motif detection on western blot, proteins were also run on a 12 % Bis-Tris gel under reducing conditions and subsequently transferred to a PVDF membrane by wet blotting. After blotting the membrane was blocked for 1 hr at RT or o/n at 4 °C with 5 % w/v BSA or elk powder in PBST (PBS containing 0.1 % v/v Tween-20) for anti-glycan or anti-FLAG western blots, respectively. Subsequent the membrane was incubated with antibodies: 1:5000 diluted HRP labelled anti-FLAG for FLAG-tag detection; 1:1000 diluted rat IgM 5750 for LeX detection [37] followed by 0.75 µg/mL HRP labelled donkey anti-mouse IgM (Jackson ImmunoResearch laboratories). After every incubation step the membrane was washed 5 times with PBST. The anti-HRP labelled antibodies were detected with a 1:1 SuperSignal West Femto:Dura substrate (Fisher Scientific) in the G:BOX Chemi System (Syngene).

Matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry

The N-glycan composition was analysed with matrix-assisted laser desorption/ionisation time-of-flight (MALDI-TOF) mass spectrometry (MS). Thereto, 1-2 µg purified omega-1

5

was denatured by incubation with 1.3 % w/v SDS and 0.1 % v/v β-mercaptoethanol for 10 min/95 °C. SDS was neutralised by adding 1.3 % v/v NP-40, after which proteins were digested with trypsin (Sigma-Aldrich) linked to NHS-activated Sepharose (GE Healthcare) o/n at 37 °C. Trypsin beads were removed by centrifugation for 3 min/400 rpm and the supernatant was transferred and dried under vacuum. Dried samples were dissolved in 1 M sodium acetate pH 4.5 and sonicated for 5 min. Subsequently the N-glycans were released by incubation with PNGase A (Roche Diagnostics) 24-48 hrs/37 °C. The released N-glycans were separated from peptides using C18 Bakerbond™ SPE cartridges (JT Baker) and subsequent binding of the N-glycans to Extract Clean™ Carbo SPE columns. Eluates were dried o/n under vacuum and reconstituted in MQ. N-glycans were then labelled with anthranilic acid (Sigma-Aldrich), by incubation for 2 hrs/65 °C with labelling mix (DMSO:Acetic acid (10:3) with 48 mg/mL 2-aminobenzoic acid (Sigma-Aldrich) and 107.23 mg/mL 2-picoline-borane complex (Sigma-Aldrich)). Samples were cooled to RT and subsequently desalted by hydrophilic interaction chromatography on Biogel P10 (BioRad). If necessary, samples were cleaned using C18 ZipTip (Millipore) and eluted by 2 μL of matrix solution (20 mg/mL 2,5-dihydroxybenzoic acid in 50% acetonitrile, 0.1% v/v TFA) and subsequently pipetted on a polished steel target plate. Otherwise, samples in 75% acetonitrile were mixed with 2 μL of matrix solution and were dried under a stream of warm air. MALDI-TOF mass spectra were obtained using an Ultraflex II mass spectrometer (Bruker Daltonics) in negative-ion reflection mode as previously described [38].

To confirm the presence of specific glycan motifs, prior to ZipTip C18 clean-up, N-glycans were treated with enzymes or hydrofluoric acid (Sigma-Aldrich), which hydrolyses labile glycosidic bonds (Supplemental Table 2). The following glycosidases were used according to the suppliers protocols: α1,3/4-fucosidase from *Xanthomonas* sp. (Sigma-Aldrich), α1,2/4/6-fucosidase O from *Omnitrophica* (New England Biolabs) and/or β1,4/6-galactosidase 5012 from Jack bean (Prozyme).

Acknowledgements

References

[1] K. de Ruiter *et al.*, "Helminths, hygiene hypothesis and type 2 diabetes," *Parasite Immunol.*, vol. 39, no. 5, pp. 1–11, 2017.

[2] Y. Chen *et al.*, "Association of previous schistosome infection with diabetes and metabolic syndrome: A cross-sectional study in rural China," *J. Clin. Endocrinol. Metab.*, vol. 98, no. 2, pp. 283–287, 2013.

[3] S. W. Shen *et al.*, "The potential long-term effect of previous schistosome infection reduces the risk of metabolic syndrome among Chinese men," *Parasite Immunol.*, vol. 37, no. 7. pp. 333–339, 2015.

[4] A. E. Wiria *et al.*, "Infection with soil-transmitted helminths is associated with increased insulin sensitivity," *PLoS One*, vol. 10, no. 6, 2015.

[5] V. Aravindhan *et al.*, "Decreased prevalence of lymphatic filariasis among diabetic subjects associated with a diminished pro-inflammatory cytokine response (CURES 83)," *PLoS Negl. Trop. Dis.*, vol. 4, no. 6, p. e707, 2010.

[6] D. Wu *et al.*, "Eosinophils sustain adipose alternatively activated macrophages associated with glucose homeostasis," *Science*, vol. 332, no. 6026, pp. 243–247, 2011.

[7] L. Hussaarts *et al.*, "Chronic helminth infection and helminth-derived egg antigens promote adipose tissue M2 macrophages and improve insulin sensitivity in obese mice," *FASEB J.*, vol. 29, no. 7, pp. 3027–3039, 2015.

[8] P. Bhargava *et al.*, "Immunomodulatory glycan LNFPIII alleviates hepatosteatosis and insulin resistance through direct and indirect control of metabolic pathways," *Nat. Med.*, vol. 18, no. 11, pp. 1665–1672, 2012.

[9] C. H. Smit *et al.*, "Glycomic analysis of life stages of the human parasite Schistosoma mansoni reveals developmental expression profiles of functional and antigenic glycan motifs.," *Mol. Cell. Proteomics*, vol. 14, no. 7, pp. 1750–1769, 2015.

[10] B. Everts *et al.*, "Schistosome-derived omega-1 drives Th2 polarization by suppressing protein synthesis following internalization by the mannose receptor," *J. Exp. Med.*, vol. 209, no. 10, pp. 1753–1767, 2012.

[11] B. Everts *et al.*, "Omega-1, a glycoprotein secreted by Schistosoma mansoni eggs, drives Th2 responses," *J. Exp. Med.*, vol. 206, no. 8, pp. 1673–1680, 2009.

[12] E. Hams *et al.*, "The helminth T2 RNase ω1 promotes metabolic homeostasis in an IL-33- and group 2 innate lymphoid cell-dependent mechanism," *FASEB Journal*, vol. 30, no. 2. pp. 824–835, 2016.

[13] S. Steinfelder *et al.*, "The major component in schistosome eggs responsible for conditioning dendritic cells for Th2 polarization is a T2 ribonuclease (omega-1)," *J. Exp. Med.*, vol. 206, no. 8, pp. 1681–1690, 2009.

[14] M. H. J. Meevissen *et al.*, "Structural characterization of glycans on omega-1, a major schistosoma mansoni egg glycoprotein that drives Th2 responses," *J. Proteome Res.*, vol. 9, no. 5, pp. 2630–2642, 2010.

[15] R. Strasser *et al.*, "Generation of glyco-engineered Nicotiana benthamiana for the production of monoclonal antibodies with a homogeneous human-like N-glycan structure," *Plant Biotechnology J.*, vol. 6, no. 4. pp. 392–402, 2008.

[16] G. J. A. Rouwendal, D. E. A. Florack, T. Hesselink, J. H. Cordewener, J. P. F. G. Helsper, and D. Bosch, "Synthesis of Lewis X epitopes on plant N-glycans," *Carbohydr. Res.*, vol. 344, no. 12, pp. 1487–1493, 2009.

[17] H. Bakker *et al.*, "An antibody produced in tobacco expressing a hybrid β-1,4- galactosyltransferase is essentially devoid of plant carbohydrate epitopes," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 103, no. 20, pp. 7577–7582, 2006.

[18] R. Strasser *et al.*, "Improved virus neutralization by plant-produced anti-HIV antibodies with a homogeneous β1,4-galactosylated N-glycan profile," *Journal of Biological Chemistry*, vol. 284, no. 31. pp. 20479–20485, 2009.

[19] T. Hesselink, G. J. A. Rouwendal, M. G. L. Henquet, D. E. A. Florack, J. P. F. G. Helsper, and D. Bosch, "Expression of natural human β1,4-GalT1 variants and of non-mammalian homologues in plants leads to differences in galactosylation of N-glycans," *Transgenic Research*, vol. 23, no. 5. pp. 717–728, 2014.

[20] L. B. Westerhof *et al.*, "3D Domain Swapping Causes Extensive Multimerisation of Human Interleukin-10 When Expressed In Planta," *PLoS One*, vol. 7, no. 10, pp. 1–10, 2012.

5

[21] L. B. Westerhof *et al.*, "Transient expression of secretory IgA in Planta is optimal using a multi-gene vector and may be further enhanced by improving joining chain incorporation," *Front. Plant Sci.*, vol. 6, p. 1200, 2016.

[22] R. H. P. Wilbers *et al.*, "The N-glycan on Asn54 affects the atypical N-glycan composition of plant-produced interleukin-22, but does not influence its activity," *Plant Biotechnol. J.*, vol. 14, no. 2, pp. 670–681, 2016.

[23] P. Bencúr *et al.*, "Arabidopsis thaliana β1,2-xylosyltransferase: An unusual glycosyltransferase with the potential to act at multiple stages of the plant N-glycosylation pathway," *Biochemical Journal*, vol. 388, no. 2. pp. 515–525, 2005.

[24] Y. J. Shin *et al.*, "Reduced paucimannosidic N-glycan formation by suppression of a specific β-hexosaminidase from Nicotiana benthamiana," *Plant Biotechnology J.*, vol. 15, no. 2. pp. 197–206, 2017.

[25] N. Q. Palacpac *et al.*, "Stable expression of human β1,4-galactosyltransferase in plant cells modifies N-linked glycosylation patterns," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 96, no. 8. pp. 4692–4697, 1999.

[26] H. Bakker *et al.*, "Galactose-extended glycans of antibodies produced by transgenic plants," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 98, no. 5, pp. 2899–2904, 2001.

[27] S. Kallolimath, C. Gruber, H. Steinkellner, and A. Castilho, "Promoter Choice Impacts the Efficiency of Plant Glyco-Engineering," *Biotechnol. J.*, vol. 13, no. 1, pp. 1–7, 2018.

[28] C. Goulet, C. Goulet, M. C. Goulet, and D. Michaud, "2-DE proteome maps for the leaf apoplast of Nicotiana benthamiana," *Proteomics*, vol. 10, no. 13. pp. 2536–2544, 2010.

[29] A. Castilho *et al.*, "Processing of complex N-glycans in IgG Fc-region is affected by core fucosylation," *MAbs*, vol. 7, no. 5, pp. 863–870, 2015.

[30] R. H. P. Wilbers *et al.*, "Production and glyco-engineering of immunomodulatory helminth glycoproteins in plants," *Sci. Rep.*, vol. 7, p. 45910, 2017.

[31] L. B. Westerhof, "Plant as a production platform for high-value proteins," 2014.

[32] W. J. van Engelen, F. A., Molthoff, J. W., Conner, A. J., Nap, J. P., Pereira, A., & Stiekema, "pBINPLUS/ARS: An improved plant transformation vector based on pBIN19," *Transgenic Res.*, vol. 4, no. 4, pp. 288–290, 1995.

[33] M. A. Sacco *et al.*, "The cyst nematode SPRYSEC protein RBP-1 elicits Gpa2- and RanGAP2-dependent plant cell death," *PLoS Pathog.*, vol. 5, no. 8, p. e1000564, 2009.

[34] J. D. Schneider *et al.*, "Expression of human butyrylcholinesterase with an engineered glycosylation profile resembling the plasma-derived orthologue," *Biotechnology Journal*, vol. 9, no. 4. pp. 501–510, 2014.

[35] O. Voinnet, S. Rivas, P. Mestre, and D. Baulcombe, "An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus," *Plant J.*, vol. 33, no. 5, pp. 949–956, 2003.

[36] A. Cabanettes, L. Perkams, C. Spies, C. Unverzagt, and A. Varrot, "Recognition of Complex Core-Fucosylated N-Glycans by a Mini Lectin," *Angew. Chemie - Int. Ed.*, vol. 57, no. 32, pp. 10178–10181, 2018.

[37] E. Hennen, T. Czopka, and A. Faissner, "Structurally distinct LewisX glycans distinguish subpopulations of neural stem/progenitor cells," *J. Biol. Chem.*, vol. 286, no. 18, pp. 16321–16331, 2011.

[38] M. H. J. Meevissen *et al.*, "Targeted glycoproteomic analysis reveals that kappa-5 is a major, uniquely glycosylated component of schistosoma mansoni egg antigens," *Mol. Cell. Proteomics*, vol. 10, no. 5, pp. M110-005710, 2011.

5

Supplemental Figures and Tables

SUPPLEMENTAL TABLE 1 | Primers. Primers used for amplification of HsGnTII. The part of the sequences that corresponds to the open reading frame is written in capital letters.

Primer name	Sequence 5' to 3'
HsGnTII - *Forward*	acggccgccagtgtgctgtcATGAGGTTCCGCATCTACAAACG
HsGnTII - *Reverse*	cgccagtgtgatggatatctgcatgtacaTCACTGCAGTCTTCTATAACTTTTACAGAG

SUPPLEMENTAL TABLE 2 | Glycosidases. This table depicts glycosidases used to screen for specific N-glycan motifs prior to MALDI-TOF analysis. The for this chapter relevant glycan specificity is indicated and the arrow indicates the site of cleavage.

Name of the glycosidase	Abbreviation used in this thesis	Used in screening for	Specificity
β1,4/6-galactosidase	β1,4-gal	LeX vs LeA	Galβ1-4GlcNAcβ-R
α1,3/4-fucosidase	fuc	core vs LeX/LeA	Galβ1-4(Fucα1-3)GlcNAcβ-R Galβ1-3(Fucα1-4)GlcNAcβ-R
α1,2/4/6-fucosidase O	fuc-o	LeX/LeA vs core	R-GlcNAcβ1-4(Fucα1-3)GlcNAcβ-AA
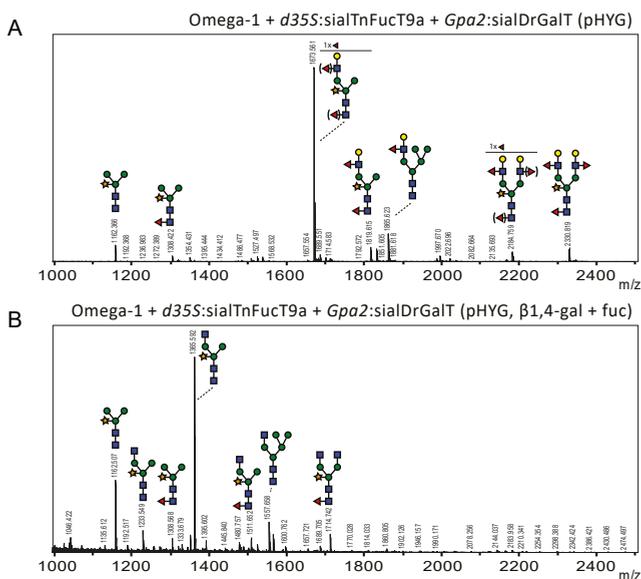
SUPPLEMENTAL FIGURE 1 | Confirmation of the presence of LeX glycan motifs. (A) N-glycan profile for purified omega-1 produced in *Nicotiana benthamiana* plants upon co-expression of *d35S*:sialTnFucT9a and *Gpa2*:sialDrGalT, the latter being expressed using the pHYG vector. (B) Profile of the same N-glycans upon treatment with α1,3/4-fucosidase from *Xanthomonas sp.* (fuc) and β1,4/6-galactosidase from Jack bean (β1,4-gal). This profile confirms the presences of LeX glycan motifs.
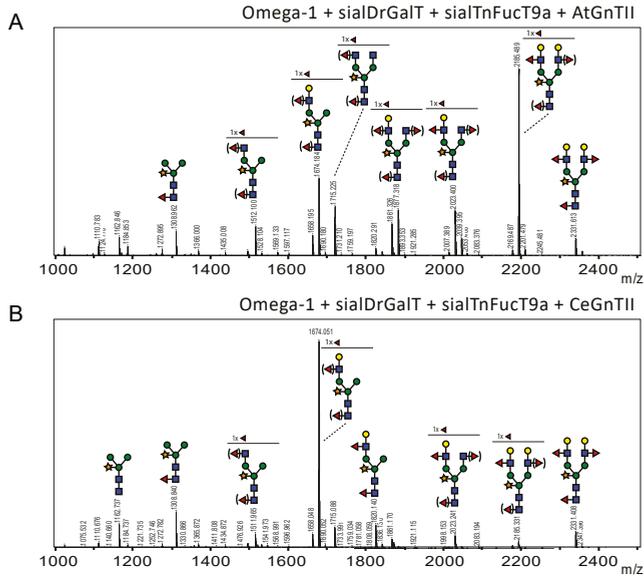
SUPPLEMENTAL FIGURE 2 | Co-expression of *Arabidopsis thaliana* and *Caenorhabditis elegans* GnTIIs for synthesis of diantennary LeX carrying N-glycans on omega-1 in plants. Engineering of diantennary LeX carrying N-glycans on omega-1 was attempted by transient co-expression of *Gpa2*:sialDrGalT in pHYG, sialTnFucT9a and exogenous AtGnTII or CeGnTII. (A-B) N-glycan profile of purified omega-1 upon co-expression of *Gpa2*:sialDrGalT in pHYG, sialTnFucT9a and AtGnTII (A) or CeGnTII (B*)*. When a peak represents multiple N-glycan structures of identical mass, the number of residues of which the position on the N-glycan is not clear is indicated above the glycan structure. The possible positions of these sugar residues on the N-glycan are indicated between brackets.
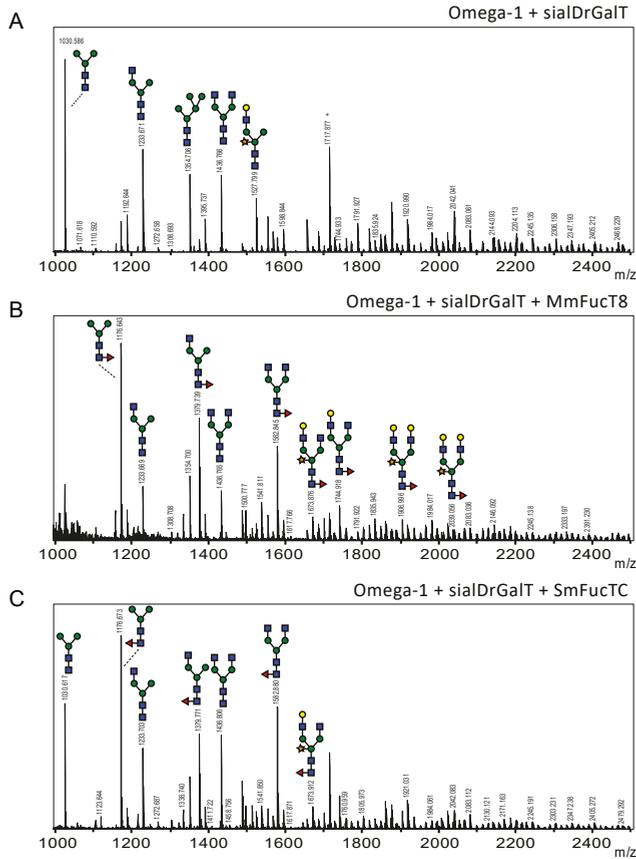
SUPPLEMENTAL FIGURE 3 | N-glycan profiles of single and double core fucosylated diantennary N-glycans on omega-1. (A-E) MALDI-TOF MS N-glycan profiles of omega-1 upon co-expression of HsGnTII, *Gpa2*:sialDrGalT in pHYG (A) and the core FucTs MmFucT8 (B), NbFucTA1 (C), SmFucTC (D) or MmFucT8 and NbFucTA1 (E) in ΔXT/ FT *Nicotiana benthamiana* plants. When a peak represents multiple N-glycan structures of identical mass, the number of residues of which the position on the N-glycan is not clear, is indicated above the glycan structure. The possible positions of these sugar residues on the N-glycan are indicated between brackets. * indicates that under the mentioned conditions no N-glycan structure could be annotated for this peak.
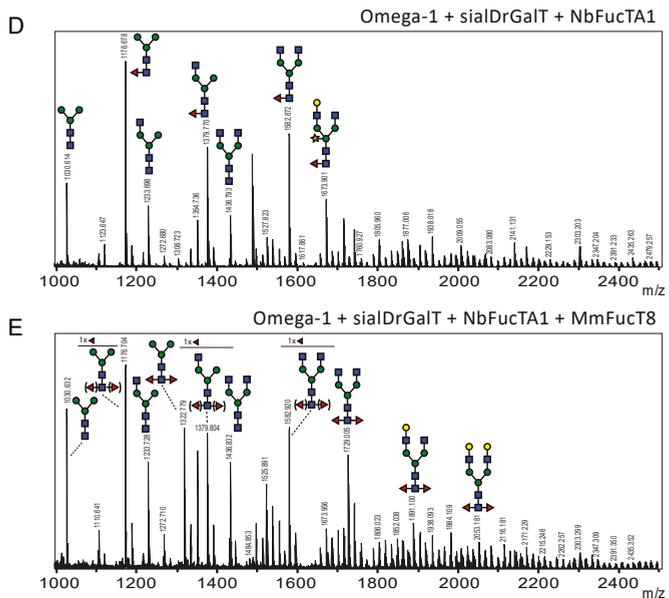
SUPPLEMENTAL FIGURE 3 | N-glycan profiles of single and double core fucosylated diantennary N-glycans on omega-1. Continued
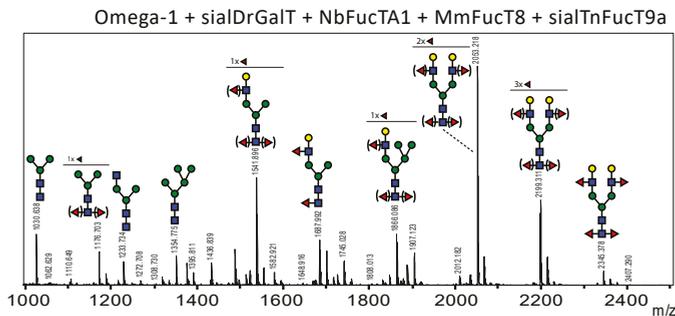
SUPPLEMENTAL FIGURE 4 | MALDI-TOF MS N-glycan profiles of double core fucosylated diantennary LeX carrying N-glycans on omega-1. MALDI-TOF MS N-glycan profile of purified omega-1 upon co-expression of HsGnTII, *Gpa2*:sialDrGalT in pHYG, MmFucT8 and NbFucTA1. When a peak represents multiple N-glycan structures of identical mass, the number of residues of which the position on the N-glycan is not clear is indicated above the glycan structure. The possible positions of these sugar residues on the N-glycan are indicated between brackets.

Omega-1

atggccaagaccaacctcttcctcttcctcatcttctccctcctgctctccctctcctccgccgtcgaggacgCTAGTCAGAACTCCTGGGACTAC
TACGTCTTCTCCGTCACCTGGCCCCCCACCTACTGCGAGTCCATCCAGTGCCGTCTCCCCCGTGGTCTCCGTGACTTCAC
CATCCACGGTCTCTGGCCCACCATCTTCCCCAACCGTCAGCCCAACTGCACCGGTTCCCTCCGTTTCGACATCCGTCGTCTC
CAGGGTATCCGTAACGAGCTTGACCTCATGTGGCCCCACCTCAAGAACTACCGTGAGTCCCCCTCCTTCTGGAAGCAC
GAGTTCGAGAAGCACGGTCTCTGCGCCGTCGAGGACCCCCAGGTCTTCAACCAGTACGGTTACTTCAAGTTCGGTATC
CAGCTCATGCAGAAGCTCAACCTCCTCAAGACCCTCATGAAGTACAGGATCTCCCCCCACGACTCCCGTCAGTACGA
CACCATCAACCTCATGAACGTCCTCGAGCGTGAGTTCGGTTACAACGGTTCCGCCAACTGCATCCGTAAGCCCGGTC
GTCGTGGTATGTACCACCTCGAGGAGGTCCACGTCTGCCTCAACCGTAAGCACGAGTTCATGAACTGCCCCTTCCTCGG
TAACTGCCCCAAGAAGTTCATCTTCCCCACCCTTCCAGGAGCTCgactacaaggacgatgacgacaagcaccatcaccaccatcactga

ORF Sequence CeGnTII

ATGATGGTCTATCGACGGATGCACCGTTTGGCAAATGCTGTTATAGCGTGTTGTTTATTCGGGTTCATCGTCATTTTTCTGAAAGCACCTG
GAGAAGATCAACGGCTACGAGACGGTGTTCCTGTATTATCACAAAATCCTGTGTGGTCAGTTAATGACTGGAATGGTTTGAATA
AAGAAGTCGTAGATTTATTGAAAAATGAATCGAATCGTCTGTCGATTAACGAAAAGCCCGAATTGAGTGGATGGAATTTTAAAAC
TAAAGAAGACATGTCACTGAAAGGATCGGAAATCGTCGAATCTGTTAGTTTTCTTAATGAAAATTTTGACATTTTGAACGCTGC
GAAATTTGGAGATTTGTCCACTGTGAAGACAATTCTAGTGATTCAGGTACACGATCGGCCGGTCTATTTGCAATACCTCATTGAGT
CAATGCGAAATACAAAAGGAATAGAAGACACATTACTAGTTTTCTCGCATGACATCAATGTCGGAATAATCAATGAAATGATCCG
TAACATCACATTCGCTCGTGTCTACCAAATCTTCTATCCATACAATTTGCAACTGTTTCCGACAGTTTTTCCAGGACAGTCTCCTTC
CGATTGTCCGGAGAAAATGAAGAGAGATAAAGCACAAGAGACGAATTGTTCAAATTGGAGTAGCCCGGATAAATATGGAAACTATC
GAGTAGCTCAATTGACACAGATAAAACATCATTGGTGGTGGAAGATGAACTTTGTGTTTGATGGAATTGTTGAGAAGTATTCGAT
GAAAGATCCATGGGTTCTACTACTGGAAGAAGATCACATGCTGGCTCCAGATGCATTACATGTTCTTGATATTATTGTATCAAATCGTC
CAAAATATTGTGAAAACTGCGAAATAATATCTCTGGGATTTTATTTGAAATCCACTAACAAATACGGTCAAGACATAGCTCATCTCG
GAGTTCACCCATGGTACAGTAGCAAGCATAATATGGGAATGGCTCTTCAGAAGAACACGTCGCAGAAGATCAAAGGATGCTCC
GAGATGTTTTGCAAATGGGATGACTATAACTGGGATTGGTCGTTGATGCAGATTTCAGCAAAATGTCTACCTCAAAGGTTTCGGGT
CATATTCACAAAAAGTCCACGAGTTATACATATAGGAGACTGTGGAGTCCATACGCACAGATGTGAAGCACATAAAGCTCTCCAATC
GACACAAGAACTATTCCGTCAGCATAAGGATCTTCTGTTCCCGACAAGTTTATCTGTGACAGACACTTCGAGAAGATCTTTGAAGC
CGTCGAAAGAAAATGGTGGATGGGGTGACATTAGAGACCGTCAACTATGCGAAATCAACAAATCTCCATTGGTTAGAGTTTCTTCT
CAATCTGCATCTGTACTCCATAAATTGCTCAACTCAAAAATCCAGTTCTCGTCATCCAACAAAACAATCACATCTACAACTTCTTAA

SUPPLEMENTAL FIGURE 5 | Sequences. Codon optimised sequence of omega-1 (incl. signal peptide and C-terminal FLAG/HIS-tag in lowercase) and the sequence of CeGnTII.

5

Chapter 6

Production and glyco-engineering of *Schistosoma mansoni* kappa-5 in plants

Kim van Noort [1], Linh Nguyen [2], Jorge Armero Gimenez [1],
Lotte B. Westerhof [1], Cornelis H. Hokke [2], Ruud H.P. Wilbers [1] and Arjen Schots [1]

[1] Laboratory of Nematology, Wageningen University and Research, Wageningen, the Netherlands;
[2] Department of Parasitology, Leiden University Medical Centre, Leiden, The Netherlands.

Abstract

Worldwide, many people are infected with Digenean trematodes of the genus *Schistosoma*. During host infection *Schistosoma mansoni* secretes proteins to modulate the host immune system. In many cases, glycans on secreted proteins have shown to be involved in immunomodulation. Among secreted immunomodulatory proteins are the three most abundant glycosylated soluble egg antigens alpha-1/IPSE, omega-1 and kappa-5. Alpha-1/IPSE induces secretion of IL-4 by basophils and thereby alternatively activates macrophages. Omega-1 modulates dendritic cells to induce a T-helper cell type 2 response. The function of the third most abundant soluble egg antigen, kappa-5, is still unknown. It is known, however, that the LDN and LDN-F glycan motifs on kappa-5 are involved in granuloma formation or immunomodulation, respectively. Kappa-5 is the major glycoprotein in soluble egg antigens that carries LDN and LDN-F N-glycans. This makes kappa-5 an interesting protein to study. At the moment kappa-5 can only be extracted from *S. mansoni* soluble egg antigens in a time consuming and inefficient manner. However, high amounts of kappa-5 are required to study its biological function. To support fundamental studies to the function of kappa-5, a recombinant production platform is required that can produce a natural mimic of kappa-5 with helminth-like N-glycans. Here we show production of kappa-5 with N-glycans carrying LDN and LDN-F glycan motifs in *Nicotiana benthamiana*. Production of helminth glycoproteins, such as kappa-5, enables research on their function, immunomodulatory properties and possible future use as biopharmaceuticals.

Introduction

Worldwide, 252 million people are infected with Digenean trematodes of the genus *Schistosoma*, causing the disease schistosomiasis [1]. Before trematodes of the species *S. mansoni* can infect humans, miracidia infect freshwater snails of the species *Biomphalaria glabrata* as an intermediate host wherein asexual reproduction takes place. After four to six weeks cercariae leave the snails in search for a human host. When humans come in contact with contaminated fresh water, the free swimming cercariae penetrate the human skin. After skin penetration, schistosomula migrate into the bloodstream and ultimately enter veins around the liver via the arteries of the heart and lungs. In the portal vein the worms mature, mate and reproduce sexually. Thousands of eggs produced by the female schistosome migrate to the colon to be secreted with the faeces. The migration process of the eggs is inefficient, and many eggs are trapped in organs, such as the liver or the intestine, where they cause granuloma formation. When faeces containing schistosome eggs comes in contact with fresh water the eggs hatch and snail infection by miracidia will continue the life cycle.

Parasitic helminths, such as schistosomes, achieve long-term infections and survive by actively modulating the host immune system [2], [3]. The immunomodulatory properties of helminths are for a large part mediated by their secretory glycoproteins, which induce T helper 2 (Th2) immunity and activate regulatory networks via glycan-dependent mechanisms [4], [5]. For instance, *S. mansoni* deposited eggs produce glycoproteins that influence the human immune system. Three of the most abundant soluble egg antigens (SEA) are the glycoproteins IPSE/alpha-1, omega-1 and kappa-5. IPSE/Alpha-1 induces secretion of IL-4 by basophils, alternatively activates macrophages and downregulates granuloma formation [6]–[9]. Omega-1 programs dendritic cells (DCs) to induce a Th2 response via its ribonuclease activity [4], [10]. In contrast to ISPE/alpha-1 and omega-1 the host reaction to kappa-5 is still unknown.

S. mansoni glycans are highly diverse, although they do display typical characteristics such as lack of sialylation, a high proportion of fucosylation, the possible occurrence of β1,2-xylose and α1,3-fucose core modifications, and a high abundance of antennary LDN(-F) or LeX. The core-modified N-glycans, as well as N-glycans that carry LeX have been identified as functional elements, such as for the immunomodulatory helminth glycoprotein omega-1 [11], [12]. Both, IPSE/alpha-1 and omega-1 carry N-glycans with diantennary LeX glycan motifs with a double fucosylated core [12], [13]. The N-glycans of omega-1 are important for its uptake by DCs and Th2 induction. In contrast to IPSE/alpha-1 and omega-1, the third most abundant protein in SEA, kappa-5, is glycosylated in a different manner and mainly carries triantennary N-glycans with terminal glycan motifs LDN and LDN-F and a double fucosylated and xylosylated core [14], [15]. The glycan motifs on kappa-5 N-glycans, LDN and LDN-F, are implicated in granuloma formation or immunomodulation, respectively

6

[15], [16]. Moreover, the human IgE response against kappa-5 is directed against its N-glycans [15]. Kappa-5 is the primary component in SEA that carries LDN and LDN-F glycans. This makes kappa-5 with terminal LDN and LDN-F glycan motifs an interesting glycoprotein for host-parasite interaction studies.

Since, kappa-5 cannot be isolated in sufficient quantities from SEA, to support fundamental studies on the function of kappa-5, a recombinant production platform is required. For such a platform the controlled production of recombinant kappa-5 with native helminth N-glycans is essential, since the N-glycans can influence protein functionality. Plants have emerged as a versatile expression platform for the production of recombinant glycoproteins, as plants are highly compliant with engineering of their glycosylation pathway [17], [18]. Also, in contrast to other production platforms, such as yeast and mammalian cells, plant-produced glycoproteins have a remarkable homogeneous N-glycan profile [17], [19]. The glycosylation profile in for instance cultured mammalian cells can differ through a slight change in pH, cell status or nutrient composition [20]. Moreover, some of the typical characteristics of the limited plant glycome match those of *S. mansoni*, including the α1,3-fucose and β1,2-xylose core modifications found on kappa-5 N-glycans. Furthermore, tools are available for introduction of non-plant N-glycan structures, such as highly branched N-glycans or GalNAc. Branched N-glycans have been introduced for the production of human erythropoietin (EPO) and human transferrin (TF) with mammalian-type N-glycans, carrying bisected, tri- and tetra-antennary N-glycans [21]. GalNAc-carrying glycans have been introduced in plants for the synthesis of mammalian mucin-type O-glycans and this required the co-expression of human GalNAc transferase 2 in combination with a uridine diphosphate (UDP) N-acetylglucosamine C4-epimerase and the UDP-GalNAc transporter nstp-419 [22]. More recent reports showed that it is possible to synthesise GalNAc containing O-glycans on several recombinant proteins in the absence of a transporter and/or epimerase [23], [24].

The above mentioned characteristics inspired us to investigate the production of kappa-5 carrying branched N-glycans with LDN or LDN-F glycan motifs *in planta*. This resulted in a platform for the transient production of large amounts of recombinant kappa-5 in leaves of *Nicotiana benthamiana* plants. Furthermore, we show engineering of the plant glycosylation pathway for the synthesis of triantennary N-glycans and introduction of LDN and LDN-F glycan motifs. Production of recombinant kappa-5 will enable fundamental studies to the function of kappa-5 and its N-glycans.

Results

Schistosoma mansoni glycoprotein kappa-5 is efficiently produced in *Nicotiana benthamiana*

For production of recombinant kappa-5 in *N. benthamiana*, codon optimised kappa-5 was expressed by leaf agroinfiltration. Expression of kappa-5 was analysed with and without P19 co-expression in crude plant extracts 3, 5 and 7 days post infiltration (dpi) on western blot using anti-FLAG antibodies (Figure 1A). Non-reduced kappa-5, corresponding to the kappa-5 dimer, was detected between 75 and 100 kDa. The monomer of kappa-5 was detected between 37 and 50 kDa upon reduction. Kappa-5 accumulation is highest at 5 dpi upon co-expression of the silencing suppressor P19. Accumulation of kappa-5 in crude extracts and apoplast fluids was analysed by SDS-PAGE and Coomassie staining (Figure 1B). The majority of kappa-5 (~90%) was recovered from the apoplast, which indicates that kappa-5 is secreted with remarkable efficiency. Only few endogenous plant proteins were present in the apoplast fluid, which enabled single-step purification of 1 mg kappa-5 per plant (3–4 grams of fresh leaf material) by Cation Exchange (CEX) chromatography (Figure 1C). The N-glycan composition of purified kappa-5 was assessed by Endo H / PNGase F digestion and MALDI-TOF MS analysis of released N-glycans by PNGase A. The predominant N-glycans on kappa-5 carry typical plant core α1,3-fucose and β1,2-xylose with terminal GlcNAc residues (Figure 1D). Taken together, our data show that glycosylated kappa-5 can be produced in plants, which is easily purified from the apoplast due to efficient secretion.

Synthesis of LDN and LDN-F glycan motifs on kappa-5

To achieve LDN synthesis on kappa-5 N-glycans in plants, we tested co-expression of kappa-5 with different combinations of *Caenorhabditis elegans* β1,4-*N*-acetyl-galactosaminyltransferase (CeGalNAcT) with UDP-GalNAc transporters (sqv-7, nstp-4 and nstp-5) and/or UDP-GlcNAc C4-epimerase. Soybean agglutinin (SBA) lectin binding assays were performed to screen for the presence of LDN carrying N-glycans on total apoplast proteins. This assay showed that co-expression of only CeGalNAcT is sufficient for LDN synthesis on the N-glycans of kappa-5 in plants (Figure 2A). Co-expression of the C4 epimerase increased the binding of SBA to total apoplast proteins, however increased binding to purified kappa-5 was not observed (data not shown). MALDI-TOF MS analysis confirmed the presence of LDN motifs on the N-glycans released from purified kappa-5 (Figure 2B). Approximately 40% of the isolated N-glycans from kappa-5 carried a single LDN motif as was shown by enzymatic digestion with β-N-acetylhexosaminidase from *Streptomyces plicatus* and β-N-acetylglucosaminidase from *Xanthomonas manihotis* and subsequent MALDI-TOF MS analysis (Supplemental Figure 1).

6
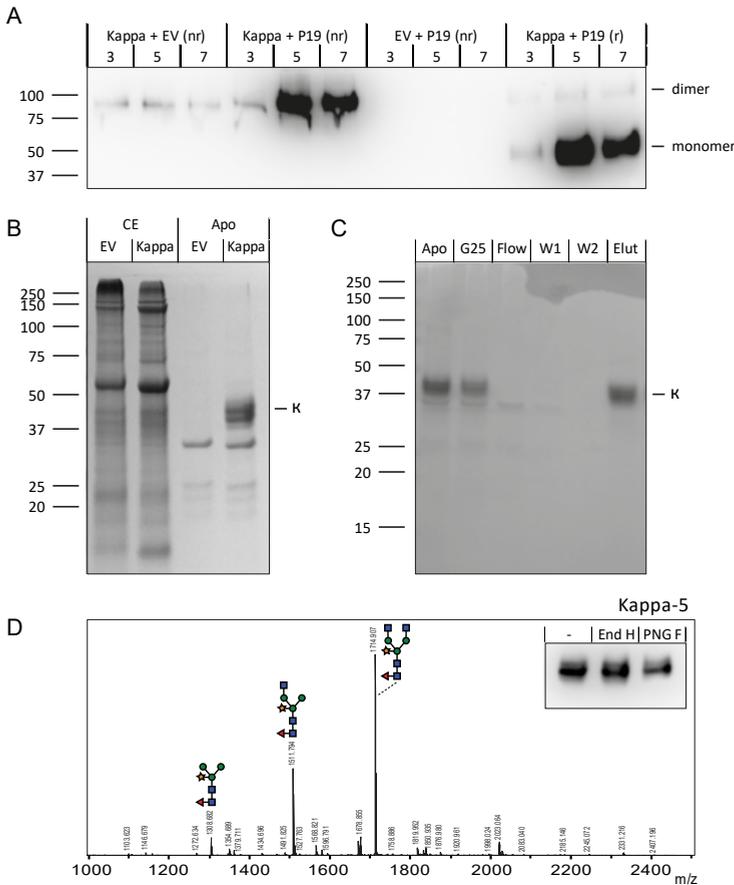
FIGURE 1 | Plant-based production of kappa-5. Kappa-5 was expressed in *Nicotiana benthamiana* plants. (A) Western blot treated with anti-FLAG antibodies containing apoplast samples of kappa-5 or empty vector (EV) infiltrated *N. benthamiana* plants harvested 3, 5 or 7 days post infiltration (dpi). At 5 dpi highest expression of kappa-5 upon P19 co-expression is observed. Samples were non-reduced (nr) or reduced (r). (B) Coomassie stained SDS-PAGE gel with crude extracts (CE) and apoplast fluids (AF) of kappa-5 or EV infiltrated *N. benthamiana* plants reveals efficient secretion of kappa-5. (C) Coomassie stained SDS-PAGE gel of different purification steps (apoplast wash (Apo) ; Sephadex G25 (G25) ; flow through (Flow) ; wash 1 (W1) ; wash 2 (W2) ; elution (Elut)). Single-step cation exchange chromatography purification of kappa-5 from leaf apoplast fluid was remarkably efficient. (D) N-glycan composition of purified kappa-5 was analysed by MALDI-TOF MS and glycan releasing enzymes Endo H (End H) and PNGase F (PNG F). The predominant N-glycans on kappa-5 carry typical plant core α1,3-fucose and β1,2-xylose with terminal GlcNAc residues.
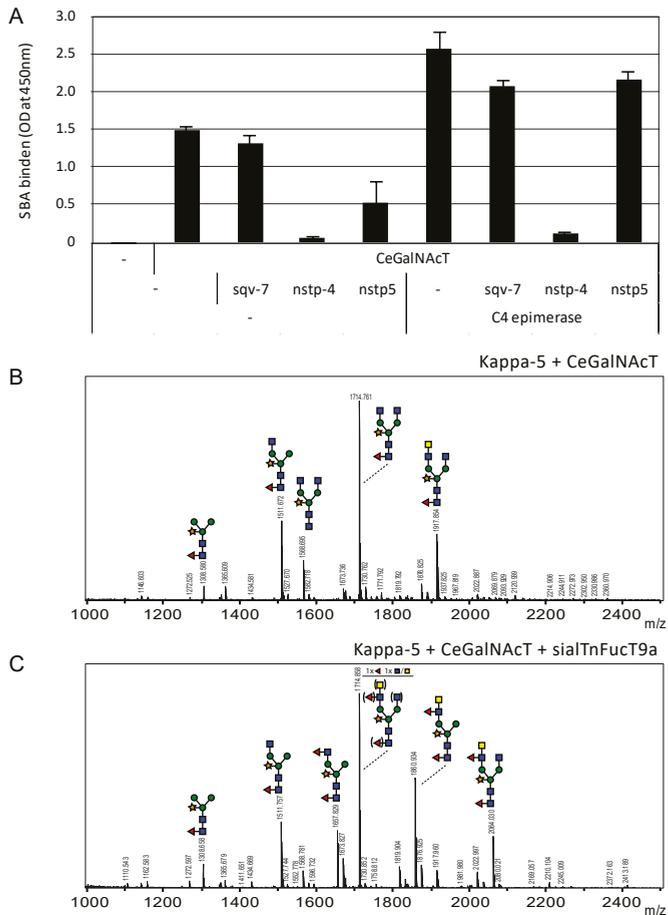
FIGURE 2 | Engineering of LDN carrying N-glycans in plants. Different combinations of CeGalNAcT, UDP-GalNAc transporters (sqv-7, nstp-4 and nstp-5) and/or C4 epimerase were co-expressed to determine which of these genes are required for *in planta* engineering of LDN carrying N-glycans. (A) Soybean agglutinin (SBA) binding assay on total soluble proteins from apoplast fluids reveals that expression of CeGalNAcT is sufficient for the synthesis of LDN carrying N-glycans. (B) MALDI-TOF MS N-glycan profile for purified kappa-5 upon co-expression of CeGalNAcT reveals the synthesis of LDN motifs. (C) MALDI-TOF MS N-glycan profile for purified kappa-5 upon co-expression of *trans*-Golgi-targeted sialTnFucT9a and CeGalNAcT reveals the synthesis of LDN-F motifs. When a peak represents multiple N-glycan structures of identical mass, the peak indicates the major N-glycan present based on enzymatic digestions (for details see Supplemental Figure 1 and 2). When the sugar position remains unclear, the number of residues of which the position on the N-glycan is not clear is indicated above the N-glycan structure. The possible positions of these sugar residues on the N-glycan are indicated between brackets.

Next, successful synthesis of LDN-F was achieved upon co-expression of kappa-5, CeGalNAcT and sialTnFucT9a (Figure 2C). Digestions with β-N-acetylhexosaminidase from *Streptomyces plicatus* and β-N-acetylglucosaminidase from *Xanthomonas manihotis* confirmed that ~50% of the N-glycans carried a single LDN-F motif (Supplemental Figure 2). Fucosylation of the LDN motif enhances the accumulation of GalNAc-containing N-glycans. In conclusion, LDN and LDN-F glycan motifs can be synthesised on the N-glycans of plant produced kappa-5, although CeGalNAcT introduction seems to hamper endogenous GnTII activity as diantennary LDN or LDN-F containing N-glycans were not observed.

Synthesis of kappa-5 carrying diantennary N-glycans

To restore GnTII activity, exogenous GnTIIs from *Arabidopsis thaliana* (AtGnTII), *C. elegans* (CeGnTII) and *Homo sapiens* (HsGnTII) were co-expressed with kappa-5, CeGalNAcT and FucTD from *S. mansoni* (SmFucTD) (Chapter 2 and 3). In this experiment SmFucTD was used instead of sialTnFucT9a, because earlier experiments (Chapter 3) showed that LDN-F synthesis by SmFucTD is more efficient then LDN-F synthesis by sialTnFucT9a. Furthermore, sialTnFucT9a can fucosylate terminal GlcNAc without the presence of GalNAc, which makes result interpretation more difficult (Supplemental Figure 3). Total apoplast proteins were analysed on western blot using anti-LDN-F antibodies. This western blot revealed increased binding to kappa-5 upon co-expression of AtGnTII and HsGnTII (Figure 3A). This observation suggested synthesis of N-glycans carrying diantennary LDN-F. MALDI-TOF MS analysis on the N-glycans released from purified kappa-5 confirmed the presence of diantennary N-glycans upon co-expression of HsGnTII or AtGnTII (Figure 3C and Supplemental Figure 4A). Although, a second GlcNAc was introduced upon co-expression of HsGnTII or AtGnTII, N-glycans carrying diantennary LDN-F were not observed after enzymatic digestion with α1,2/4/6-fucosidase O from *Omnitrophica* and subsequent MALDI-TOF MS analysis (Supplemental Figure 5E and H). In conclusion, upon co-expression of AtGnTII and HsGnTII diantennary N-glycans can be synthesised on kappa-5, although diantennary LDN or LDN-F carrying N-glycans were not observed.

Synthesis of kappa-5 carrying triantennary N-glycans

To investigate whether instead of two antennae also multiple antennae could be introduced on kappa-5 N-glycans, kappa-5, CeGalNAcT and HsGnTII were co-expressed with *H. sapiens* GnTIV and/or GnTV (respectively, HsGnTIV and HsGnTV). Kappa-5 N-glycans were released by PNGase A and analysed by MALDI-TOF MS. Only upon co-expression of HsGnTV the main peak shifted (from 1714 to 1918 m/z), indicating that HsGnTV may add a third antenna to the N-glycans of kappa-5 (Figure 4). To check if HsGnTIV and/or HsGnTV activity was influenced by plant core β1,2-xylose, the same experiment was repeated in ΔXT/FT *N. benthamiana* plants. This resulted in similar N-glycan profiles as in wild type

plants, suggesting that an extra antenna was only synthesised upon co-expression of HsGnTV (Supplemental Figure 6).

Since, GlcNAc and GalNAc cannot be distinguished based on mass, an enzymatic digestion was performed on the released N-glycans to detect the amount of LDN carrying N-glycans on plant produced kappa-5. N-glycan profiles of enzymatic digestion with β-N-acetyl-glucosaminidase from *Streptococcus pneumoniae* revealed that the majority of branches is cleaved (Supplemental Figure 7). This indicates that the majority of N-glycans carries GlcNAc branches and only a low amount of LDN carrying N-glycans is present. Thus, the shift (from 1714 to 1918 m/z) upon co-expression of HsGnTV indicates the synthesis of triantennary N-glycans on kappa-5. In conclusion, triantennary N-glycans can be synthesised on the N-glycans of plant produced kappa-5, but lack LDN.
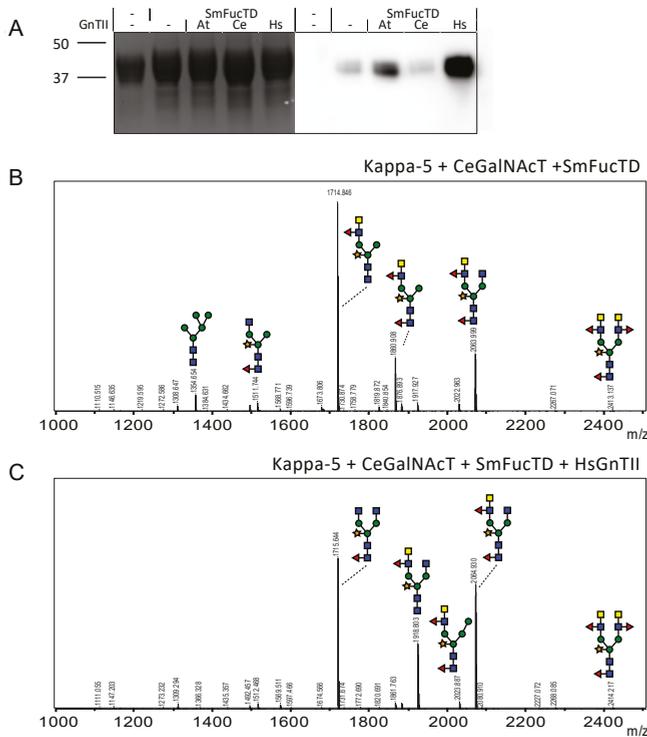
FIGURE 3 | Co-expression of HsGnTII for synthesis of diantennary LDN-F carrying N-glycans on plant produced kappa-5. Engineering of diantennary LDN-F carrying N-glycans on kappa-5 was attempted by transient co-expression of CeGalNAcT, SmFucTD and HsGnTII. (A) The N-glycan composition on purified kappa-5 was analysed by Coomassie stained SDS-PAGE gel (left) and a western blot treated with anti-LDN-F antibodies (right). The western blot reveals variable band intensity upon co-expression of AtGnTII and HsGnTII. (B-C) The N-glycan composition of purified kappa-5 was analysed by MALDI-TOF MS upon co-expression of CeGalNAcT with SmFucTD (B) and HsGnTII (C). When a peak represents multiple N-glycan structures of identical mass, the peak indicates the major N-glycan present based on enzymatic digestions (for details see Supplemental Figure 5).
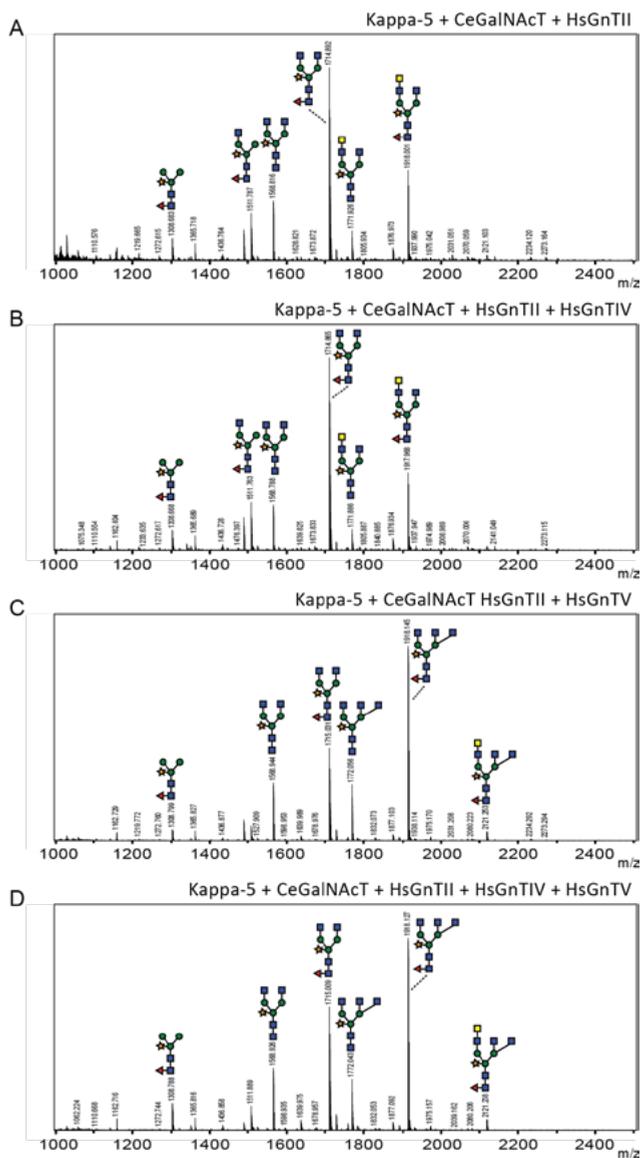
FIGURE 4 | Engineering triantennary N-glycans on plant produced kappa-5. The N-glycan composition of purified kappa-5 was analysed by MALDI-TOF MS upon co-expression of CeGalNAcT and HsGnTII (A) with either HsGnTIV (B), HsGnTV (C) or HsGnTIV and HsGnTV (D). When a peak represents multiple N-glycan structures of identical mass, the peak indicates the major N-glycan present based on enzymatic digestions (for details see Supplemental Figure 7).
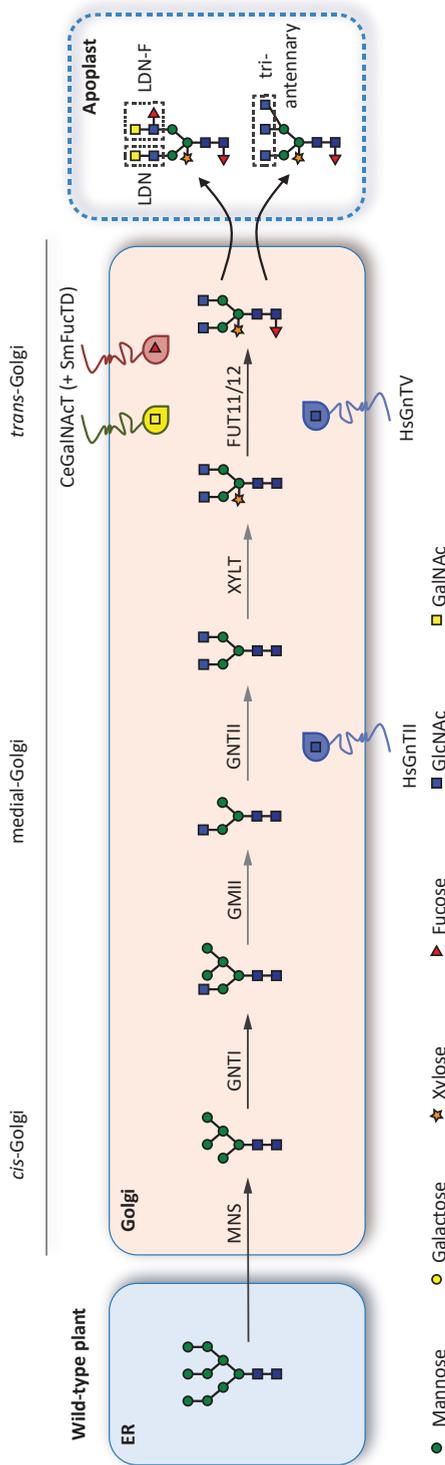
FIGURE 5 | Engineering of the plant glycosylation pathway for the production of helminth glycoprotein kappa-5 with tailored N-glycans. A schematic overview of the N-glycan modifying steps in the plant Golgi-system of wild type *Nicotiana benthamiana* plants. The plant N-glycosylation machinery was engineered by introducing (hybrid) glycosyltransferases that allow synthesis of three antennae (HsGnTII and HsGnTV) carrying N-glycans and LDN or LDN-F (CeGalNAcT or CeGalNAcT and SmFucTD) carrying N-glycans upon agro co-infiltration with kappa-5. MNS: Class I mannosidases MNS1, 2 and 3; GnTI: N-acetyl-glucosaminyltransferase I; GMII: Golgi-α-mannosidase II; GnTII: N-acetyl-glucosaminyltransferase II; XYLT: β1,2-xylosyltransferase; FUT11/12: core α1,3-fucosyltransferase. Arrows are indicated in grey, when alternative N-glycan processing routes have been postulated.

Discussion

During human infection the helminth *S. mansoni* produces immunomodulatory molecules, such as those present in SEA during the egg stage. One of the major SEA, kappa-5, carries N-glycans containing the glycan motifs LDN and LDN-F. Although the function of kappa-5 itself is still unknown, the glycan motifs on its N-glycans, LDN and LDN-F, are implicated in granuloma formation or immunomodulation, respectively [15], [16]. The fact that kappa-5 is one of the major SEA that carries LDN and LDN-F glycan motifs, makes kappa-5 an interesting protein for functional studies. However, functional studies are limited by the amounts of kappa-5 isolated from laboriously acquired SEA. Therefore, we introduced *N. benthamiana* as production platform to produce large amounts of kappa-5 with native N-glycan motifs LDN and LDN-F or triantennary branches (Figure 5).

In this chapter, we showed that *N. benthamiana* is a promising production platform for kappa-5, as over 1 mg kappa-5 per plant (3-4 gram fresh leaf material) could be purified from apoplast fluids. High expression levels and remarkable efficient secretion allowed single-step purification. Purified kappa-5 carried typical complex plant N-glycans with terminal GlcNAc residues. The presence of terminal GlcNAc residues indicates that kappa-5 N-glycans are not sensitive towards GlcNAc cleavage by *N. benthamiana* β-N-acetylhexosaminidases (NbHEXOs) present in the plant apoplast, as was observed for omega-1 (Chapter 5 and [25]).

We have shown efficient production and purification of kappa-5 with complex N-glycans in plants. However, native kappa-5 N-glycans carry LDN and LDN-F glycan motifs, whereas plants naturally do not synthesise these glycan motifs [15]. Therefore, the plant glycosylation pathway was adjusted. Engineering of GalNAc-carrying glycans in plants has to our knowledge only been attempted for the synthesis of mammalian mucin-type O-glycans [22]–[24]. These O-glycans were previously synthesised with and without co-expression of UDP-GlcNAc C4-epimerase and UDP-GalNAc transporter nstp-4. Our data show that the C4-epimerase is not required for the introduction of GalNAc on kappa-5 N-glycans, which indicates that the substrate UDP-GalNAc is present in plants. Also, the addition of UDP-GalNAc transporters sqv-7, nstp-4 and nstp-5 was not required and does not enhance GalNAc introduction on kappa-5 N-glycans. The SBA binding assay showed that nstp-4 even blocks the synthesis of LDN on kappa-5 N-glycans, which could indicate that nstp-4 re-directs UDP-GalNAc to another sub-Golgi compartment where LDN synthesis on N-glycans does not occur. Expression of CeGalNAcT is sufficient to synthesis LDN on N-glycans in plants, although for synthesis of GalNAc on O-glycans and glycolipids co-expression of epimerases and/or transporters may still be required.

Next, successful synthesis of LDN-F was achieved upon co-expression of kappa-5, CeGalNAcT and sialTnFucT9a. Addition of the α1,3-fucose to the GlcNAc in LDN enhanced the amount of GalNAc carrying N-glycans. Furthermore, a significant proportion of

N-glycans on kappa-5 contained a fucosylated terminal GlcNAc. This suggests that fucosylated LDN motifs are probably more, but not completely, resistant to GalNAc cleavage. On the other hand, we have shown that sialTnFucT9a is able to fucosylate a terminal GlcNAc without presence of GalNAc, as was also observed for XylTTnFucT9a in absence of galactose by Rouwendal and colleagues [26]. This would indicate that fucosylation of the GlcNAc in LDN protects the GalNAc from cleavage. *In vitro* HEXO activity tests with *A. thaliana* HEXOs showed next to GlcNAc cleavage also activity towards pNP-GalNAc [27]. NbHEXO3 localises in the apoplast and may therefore cleave the terminal GalNAc residues on glyco-engineered kappa-5 N-glycans in the apoplast [25]. However, the presence of terminal GlcNAc residues on non-glyco-engineered kappa-5 N-glycans suggest that the N-glycans on kappa-5 are not accessible to NbHEXOs. Still, the more outward locating GalNAc residues in LDN on glyco-engineered kappa-5 N-glycans could be accessible to NbHEXOs. Therefore, generation of NbHEXO knockout plants, using the strategy as described in Chapter 4, would be useful for synthesis of diantennary LDN or LDN-F carrying N-glycans on kappa-5.

Production of kappa-5 carrying LDN-F N-glycans showed that LDN-F synthesis on kappa-5 N-glycans is restricted to one antenna. Furthermore, upon co-expression of CeGalNAcT the amount of diantennary N-glycans is reduced. The second antenna was reintroduced upon co-expression of AtGnTII or HsGnTII. However, this second antenna, did not carry the glycan motifs LDN or LDN-F. *In vitro*, CeGalNAcT showed activity towards branched N-glycans, however its soluble form showed much less activity towards diantennary N-glycans, in comparison to human β4-GalT [28]. Furthermore, a three times higher activity is observed towards the O-glycan acceptor GlcNAcβ1–6Gal in comparison to the N-glycan acceptors. This suggests that CeGalNAcT is probably also less active to branched N-glycans *in vivo*. Therefore, introduction of a N-glycan synthesising GalNAcT of *S. mansoni* may be more suitable for generation of diantennary LDN or LDN-F carrying N-glycans on kappa-5. However, *S. mansoni* GalNAcTs first need to be characterised following the strategies described in Chapter 3 and 4.

6

As plant N-glycans carry only two antennae and the majority of native kappa-5 N-glycans carry three and sometimes even four antennae, additional antennae had to be introduced for the production of kappa-5 with native helminth N-glycans [15]. N-glycans with three or four antennae have been introduced in *N. benthamiana* plants to produce EPO and TF with humanised N-glycans [21]. Co-expression of kappa-5 with CeGalNAcT, HsGnTII and HsGnTV did result in triantennary N-glycans on kappa-5, whereas co-expression of GnTIV instead of HsGnTV did not. Combined co-expression of HsGnTIV and HsGnTV also resulted in triantennary N-glycans on kappa-5. The lack of the GlcNAcβ1-4Manα1-3R antennae upon co-expression of GnTIV could be due to the protein environment. Co-expression of HsGnTIV and HsGnTV in combination with different glycopeptides of EPO and TF showed that HsGnTIV and HsGnTV activity is influenced by the protein environment [29]. Only a

small proportion of native kappa-5 N-glycans carries four antennae. This could explain why four antennae on kappa-5 N-glycans upon co-expression of CeGalNAcT, HsGnTII, HsGnTIV and HsGnTV were not observed. Furthermore, since we only observed activity of HsGnTV and not of HsGnTIV this may indicate that the α1,3 mannose antenna is not accessible to HsGnTIV. At the moment it is unknown which antennae are present on native kappa-5 N-glycans, therefore activity of HsGnTV and inactivity of HsGnTIV may suggest that the majority of native kappa-5 N-glycans carry the antennae: GlcNAcβ1-2Manα1-3R, GlcNAcβ1-2Manα1-6R and GlcNAcβ1-6Manα1-6R. Although, GnTV and GnTIV of *S. mansoni* could show differences in activity in comparison to HsGnTIV and HsGnTV.

In this chapter, we showed production of kappa-5 containing plant N-glycans that can be helminthisised. The ability to efficiently isolate and purify plant produced kappa-5 in high amounts fulfils the demand for fundamental studies to the function of kappa-5. Furthermore, introduction of LDN or LDN-F carrying N-glycans on kappa-5 enables studies to the function of the LDN and LDN-F glycan motifs in granuloma formation or immunomodulation, respectively. The introduction of a third antennae holds the promise to produce full mimics of *S. mansoni* kappa-5 in plants in the near future.

Materials and Methods

Construction of expression vectors

The complete sequence encoding *S. mansoni* glycoprotein kappa-5 was codon optimised in-house. The protein sequence was preceded by the signal peptide from the *A. thaliana* chitinase gene (cSP) for secretion and a 6x histidine-FLAG tag (H6F) was included at the C-terminus for detection and purification purposes (Supplemental Figure 8). The optimised gene was synthetically constructed at GeneArt and subsequently cloned into the pHYG expression vector [30]. To synthesise terminal GalNAc residues the native β 1,4-N-acetylgalactosaminyltransferase (bre-4, referred to as CeGalNAcT) and the UDP-GalNAc transporters nstp-4, nstp-5 and sqv-7 were amplified from *Caenorhabditis elegans* cDNA and cloned into the pBIN-PLUS expression vector [31]. Furthermore, a codon optimised C4 epimerase from *Pseudomonas aeruginosa* (WbpP) was synthetically constructed at GeneArt and cloned into the pBIN-PLUS expression vector [24], [31]. Hybrid α1,3-fucosyltransferase IXa from *Tetraodon nigriviridus* (TnFucT9a) with the N-terminal CTS domain (C, cytoplasmic tail; T, transmembrane domain; S, stem region) of *Rattus norvegicus* α2,6-sialyltransferase for *trans*-Golgi targeting driven from the pBIN vector was used to enable LDN-F synthesis (from now on referred to as sialTnFucT9a) [26], [31]. To introduce a second antenna on N-glycans of plant produced kappa-5 upon co-expression of CeGalNAcT, three exogenous GnTIIs were co-expressed, AtGnTII, CeGnTII and HsGnTII. The construct for overexpression of AtGnTII was kindly provided by Prof. H. Steinkellner (University of Natural Resources and Life Sciences, Vienna) [32]. The sequence

of CeGnTII was synthesised by GeneArt with flanking BspHI and KpnI restriction sites and subsequently cloned into pHYG (Supplemental Figure 8). HsGnTII was amplified from a human MGAT2 gene cDNA clone (Sino Biological) with flanking BspHI and BsrGI restriction sites using illustra PuReTaq Ready-To-Go PCR Beads (GE Healthcare) and cloned into pHYG. To introduce three or four antennae on plant produced kappa-5 N-glycans *H. sapiens* GnTIV (HsGnTIV) and GnTV (HsGnTV) were co-expressed (constructs were kindly provided by Prof. H. Steinkellner University of Natural Resources and Life Sciences, Vienna) [21]. To enhance expression, the P19 silencing suppressor from tomato bushy stunt virus pBIN61 was co-infiltrated in all experiments, unless stated differently [33]. For plant expression, all constructs were transformed into *Agrobacterium tumefaciens* strain MOG101 except HsGnTV, which was transformed into GV3101.

Agroinfiltration

A. tumefaciens clones were cultured for 16 hrs at 28 °C/250 rpm in LB medium (10 g/L pepton140, 5 g/L yeast extract, 10 g/L NaCl, pH 7.0) with 20 µM acetosyringone and 50 µg/mL kanamycin for pHYG or 100 µg/mL spectinomycin and 1.25 µg/mL tetracycline for GV3101. The bacteria were suspended in MMA infiltration medium (20 g/L sucrose, 5 g/L MS-salts, 1.95 g/L MES, pH 5.6) containing 200 µM acetosyringone. For co-infiltration experiments *Agrobacterium* cultures were mixed while maintaining a final optical density (OD) of 0.5 per culture, except for introduction of a third and fourth antenna experiments in which HsGnTII, HsGnTIV and HsGnTV were infiltrated with a final OD of 0.1. The two youngest fully expanded leaves of four to six week old wild type or ΔXT/FT *N. benthamiana* plants were infiltrated completely by injecting the *Agrobacterium* suspension into a *N. benthamiana* leaf at the abaxial side using a 1 mL needles syringe [34]. *N. benthamiana* plants were maintained in a controlled greenhouse compartment (UNIFARM, Wageningen) and infiltrated leaves were harvested at three to six days post infiltration.

Protein extraction

For isolation of total apoplast proteins, harvested leaves were submerged in ice-cold extraction buffer (50 mM phosphate-buffered saline, pH 8.0, containing 100 mM NaCl and 0.1% v/v Tween-20) after which vacuum was applied for 10 min. Vacuum was released slowly to ensure infiltration of the apoplast. Leaves were carefully rolled up and placed in 10 mL syringes and centrifuged for 10 min/2000 ×g. Apoplast fluids were clarified by centrifugation for 5 min/16.000 ×g at 4 °C. If required, remaining intracellular proteins were isolated from the leaves by homogenisation in liquid nitrogen. Homogenised plant material was ground in ice-cold extraction buffer (50 mM phosphate-buffered saline, pH 8.0, containing 100 mM NaCl, 0.1% v/v Tween-20 and 2% w/v immobilized polyvinylpolypyrrolidone (PVPP)) using 2 mL/g fresh weight. Crude extracts were clarified by centrifugation for 5 min/16.000 ×g at 4 °C. Total protein content was analysed by a

6

Pierce Bicinchoninic Acid Protein Assay (BCA, Fisher Scientific). Total soluble plant proteins were separated under reducing conditions by SDS-PAGE on a 12% Bis-Tris gel (Invitrogen) and subsequently stained with Coomassie brilliant blue staining.

Purification from the apoplast fluid

Plant produced kappa-5 was purified from the apoplast fluid on Pierce Strong Cation Exchange Mini Spin Columns (Fisher Scientific). In brief, apoplast fluids were transferred over G25 Sephadex columns to exchange the extraction buffer for CEX binding buffer (10 mM Sørensen's phosphate buffer, pH 6.0, containing 100 mM NaCl). Column loading, washing and elution was done by centrifugation for 5 min/2000 xg. Bound kappa-5 was eluted with CEX binding buffer containing 2 M NaCl. After elution samples were dialysed against PBS and protein concentration was determined with BCA (Fisher Scientific). Aliquots of different purification steps and eluted kappa-5 were separated under reducing conditions by SDS-PAGE on a 12% Bis-Tris gel (Invitrogen) and subsequently stained with Coomassie brilliant blue staining.

Lectin binding assay

Soybean agglutinin (SBA, Bio-Connect) was used to determine fucosylation of LDN, since SBA binding to the GalNAc residue in LDN is inhibited by fucosylation of LDN. For this purpose, 10 µg/mL of total apoplast proteins or purified kappa-5 in PBS were coated overnight at 4 °C on microtiter plates. Plates were blocked with carbohydrate-free blocking buffer (Vector Laboratories) for 1 hr/room temperature (RT). Plates were then incubated for 1 hr/RT with biotinylated lectin at a concentration of 5 µg/mL SBA. Subsequently, plates were incubated with avidin-HRP (eBioscience) for 30 min/RT. After every incubation step the microtiter plate was washed 5 times with PBST (PBS containing 0.05 % v/v Tween-20). Lectin binding was visualised by TMB substrate (Fisher Scientific) and absorbance was measured at a wavelength of 450 nm while using 655 nm as reference filter.

Western blot

For glycan detection on western blot, purified proteins were run on a 12 % Bis-Tris gel under reducing conditions and subsequently transferred to a PVDF membrane by wet blotting. After blotting the membrane was blocked with 5 % w/v BSA in PBST (PBS containing 0.1 % v/v Tween-20). Subsequently the membrane was incubated with 1:500 diluted primary antibody mouse IgM 290-2E6 for LDN-F detection [35]. Subsequently, the membrane was incubated with 0.75 µg/mL secondary antibody, HRP labelled donkey anti-mouse IgM (Jackson ImmunoResearch laboratories). The anti-HRP labelled antibodies were detected with a 1:1 SuperSignal West Femto:Dura substrate (Fisher Scientific) in the G:BOX Chemi System (Syngene).

Matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry

The glycan composition was analysed with matrix-assisted laser desorption/ionisation time-of-flight (MALDI-TOF) mass spectrometry (MS). Thereto,1-2 µg purified kappa-5 was denatured by incubation with 1.3 % w/v SDS and 0.1 % v/v β-mercaptoethanol for 10 min/95 °C. SDS was neutralised by adding 1.3 % v/v NP-40, after which proteins were digested with trypsin (Sigma-Aldrich) linked to NHS-activated Sepharose (GE Healthcare) o/n at 37 °C. Trypsin beads were removed by centrifugation for 3 min/400 rpm and the supernatant was transferred and dried under vacuum. Dried samples were dissolved in 1 M sodium acetate pH 4.5 and sonicate for 5 min. Subsequently the N-glycans were released by incubation with PNGase A (Roche Diagnostics) 24-48 hrs/37 °C. The released N-glycans were separated from peptides using C18 Bakerbond™ SPE cartridges (JT Baker) and subsequent binding of the N-glycans to Extract Clean™ Carbo SPE columns. Eluates were dried o/n under vacuum and reconstituted in MQ. N-glycans were then labelled with anthranilic acid (Sigma-Aldrich), by incubation for 2 hrs/65 °C with labelling mix (DMSO:Acetic acid (10:3) with 48 mg/mL 2-aminobenzoic acid (Sigma-Aldrich) and 107.23 mg/mL 2-picoline-borane complex (Sigma-Aldrich)). Samples were cooled to RT and subsequently desalted by hydrophilic interaction chromatography on Biogel P10 (BioRad). If necessary, samples were cleaned using C18 ZipTip (Millipore) and eluted by 2 µL of matrix solution (20 mg/mL 2,5-dihydroxybenzoic acid in 50% acetonitrile, 0.1% v/v TFA) and subsequently pipetted on a polished steel target plate. Otherwise, samples in 75% acetonitrile were mixed with 2 µL of matrix solution and were dried under a stream of warm air. MALDI-TOF mass spectra were obtained using an Ultraflex II mass spectrometer (Bruker Daltonics) in negative-ion reflection mode as previously described [15].

To confirm the presence of specific glycan motifs, N-glycans were treated with enzymes, prior to ZipTip C18 clean-up (Supplemental Table 2). The following glycosidases were used according to the suppliers protocols: β-N-acetyl-hexosaminidase from *Streptomyces plicatus* (New England Biolabs), β-N-acetyl-glucosaminidase from *Xanthomonas manihotis* or *Streptococcus pneumoniae* (New England Biolabs) or/and α1,2/4/6-fucosidase O from *Omnitrophica* (New England Biolabs).

6

Acknowledgements

References

[1]   P. J. Hotez *et al.*, "The Global Burden of Disease Study 2010: Interpretation and Implications for the Neglected Tropical Diseases," *PLoS Negl. Trop. Dis.*, vol. 8, no. 7, p. e2865, 2014.

[2]   N. M. Girgis, U. M. Gundra, and P. Loke, "Immune Regulation during Helminth Infections," *PLoS Pathog.*, vol. 9, no. 4, p. e1003250, 2013.

[3]   H. J. McSorley and R. M. Maizels, "Helminth infections and host immune regulation," *Clin. Microbiol. Rev.*, vol. 25, no. 4, pp. 585–608, 2012.

[4]   B. Everts *et al.*, "Schistosome-derived omega-1 drives Th2 polarization by suppressing protein synthesis following internalization by the mannose receptor," *J. Exp. Med.*, vol. 209, no. 10, pp. 1753–1767, 2012.

[5]   E. J. Klaver, L. M. Kuijk, T. K. Lindhorst, R. D. Cummings, and I. Van Die, "Schistosoma mansoni soluble egg antigens induce expression of the negative regulators SOCS1 and SHP1 in human dendritic cells via interaction with the mannose receptor," *PLoS One*, vol. 10, no. 4, pp. 1–18, 2015.

[6]   G. Schramm *et al.*, " Cutting Edge: IPSE/alpha-1, a Glycoprotein from Schistosoma mansoni Eggs, Induces IgE-Dependent, Antigen-Independent IL-4 Production by Murine Basophils In Vivo ," *J. Immunol.*, vol. 178, no. 10, pp. 6023–6027, 2007.

[7]   G. Schramm *et al.*, "Molecular characterization of an interleukin-4-inducing factor from Schistosoma mansoni eggs," *J. Biol. Chem.*, vol. 278, no. 20, pp. 18384–18392, 2003.

[8]   P. Smith *et al.*, "Schistosoma mansoni secretes a chemokine binding protein with antiinflammatory activity," *J. Exp. Med.*, vol. 202, no. 10, pp. 1319–1325, 2005.

[9]   G. Schramm *et al.*, "IPSE/alpha-1: A major immunogenic component secreted from Schistosoma mansoni eggs," *Mol. Biochem. Parasitol.*, vol. 147, no. 1, pp. 9–19, 2006.

[10]   B. Everts *et al.*, "Omega-1, a glycoprotein secreted by Schistosoma mansoni eggs, drives Th2 responses," *J. Exp. Med.*, vol. 206, no. 8, pp. 1673–1680, 2009.

[11]   N. S. Prasanphanich, M. L. Mickum, J. Heimburg-Molinaro, and R. D. Cummings, "Glycoconjugates in host-helminth interactions," *Front. Immunol.*, vol. 4, no. 240, pp. 1–22, 2013.

[12]   M. H. J. Meevissen *et al.*, "Structural characterization of glycans on omega-1, a major schistosoma mansoni egg glycoprotein that drives Th2 responses," *J. Proteome Res.*, vol. 9, no. 5, pp. 2630–2642, 2010.

[13]   M. Wuhrer *et al.*, "IPSE/alpha-1, a major secretory glycoprotein antigen from schistosome eggs, expresses the Lewis X motif on core-difucosylated N-glycans," *FEBS J.*, vol. 273, no. 10, pp. 2276–2292, 2006.

[14]   G. Schramm *et al.*, "Molecular characterisation of kappa-5, a major antigenic glycoprotein from Schistosoma mansoni eggs," *Mol. Biochem. Parasitol.*, vol. 166, no. 1, pp. 4–14, 2009.

[15]   M. H. J. Meevissen *et al.*, "Targeted glycoproteomic analysis reveals that kappa-5 is a major, uniquely glycosylated component of schistosoma mansoni egg antigens," *Mol. Cell. Proteomics*, vol. 10, no. 5, pp. M110-005710, 2011.

[16]   K. K. Van de Vijver, A. M. Deelder, W. Jacobs, E. A. Van Marck, and C. H. Hokke, "LacdiNAc- and LacNAc-containing glycans induce granulomas in an in vivo model for schistosome egg-induced hepatic granuloma formation," *Glycobiology*, vol. 16, no. 3, pp. 237–243, 2006.

[17]   D. Bosch, A. Castilho, A. Loos, A. Schots, and H. Steinkellner, "N-Glycosylation of Plant-produced Recombinant Proteins," *Curr. Pharm. Des.*, vol. 19, no. 31, pp. 5503–5512, 2013.

[18]   E. Stoger, R. Fischer, M. Moloney, and J. K.-C. Ma, "Plant Molecular Pharming for the Treatment of Chronic and Infectious Diseases," *Annu. Rev. Plant Biol.*, vol. 65, no. 1, pp. 743–768, 2014.

[19]   R. H. P. Wilbers *et al.*, "The N-glycan on Asn54 affects the atypical N-glycan composition of plant-produced interleukin-22, but does not influence its activity," *Plant Biotechnol. J.*, vol. 14, no. 2, pp. 670–681, 2016.

[20]   S. A. Brooks, "Strategies for analysis of the glycosylation of proteins: Current status and future perspectives," *Mol. Biotechnol.*, vol. 43, no. 1, pp. 76–88, 2009.

[21]   A. Castilho *et al.*, "N-Glycosylation engineering of plants for the biosynthesis of glycoproteins with bisected and branched complex N-glycans," *Glycobiology*, vol. 21, no. 6, pp. 813–823, 2011.

[22]   S. M. Daskalova *et al.*, "Engineering of N. benthamiana L. plants for production of N-acetylgalactosamine-glycosylated proteins - towards development of a plant-based platform for production of protein therapeutics with mucin type O-glycosylation," *BMC Biotechnol.*, vol. 10, no. 62, pp. 1–18, 2010.

[23] A. Castilho *et al.*, "Engineering of sialylated mucin-type O-glycosylation in plants," *J. Biol. Chem.*, vol. 287, no. 43, pp. 36518–36526, 2012.

[24] Z. Yang *et al.*, "Engineering mammalian mucin-type O-glycosylation in plants," *J. Biol. Chem.*, vol. 287, no. 15, pp. 11911–11923, 2012.

[25] Y. J. Shin *et al.*, "Reduced paucimannosidic N-glycan formation by suppression of a specific β-hexosaminidase from Nicotiana benthamiana," *Plant Biotechnology J.*, vol. 15, no. 2. pp. 197–206, 2017.

[26] G. J. A. Rouwendal, D. E. A. Florack, T. Hesselink, J. H. Cordewener, J. P. F. G. Helsper, and D. Bosch, "Synthesis of Lewis X epitopes on plant N-glycans," *Carbohydr. Res.*, vol. 344, no. 12, pp. 1487–1493, 2009.

[27] R. Strasser *et al.*, "Enzymatic properties and subcellular localization of arabidopsis β-N-acetylhexosaminidases," *Plant Physiol.*, vol. 145, no. 1, pp. 5–16, 2007.

[28] Z. S. Kawar, I. Van Die, and R. D. Cummings, "Molecular cloning and enzymatic characterization of a UDP-GalNAc:GlcNAcβ-R β1,4-N-acetylgalactosaminyltransferase from Caenorhabditis elegans," *Journal of Biological Chemistry*, vol. 277, no. 38. pp. 34924–34932, 2002.

[29] B. Nagels *et al.*, "Biologically active, magnICON®-expressed EPO-Fc from stably transformed Nicotiana benthamiana plants presenting tetra-antennary N-glycan structures," *J. Biotechnol.*, vol. 160, no. 3–4, pp. 242–250, 2012.

[30] L. B. Westerhof *et al.*, "3D Domain Swapping Causes Extensive Multimerisation of Human Interleukin-10 When Expressed In Planta," *PLoS One*, vol. 7, no. 10, pp. 1–10, 2012.

[31] W. J. van Engelen, F. A., Molthoff, J. W., Conner, A. J., Nap, J. P., Pereira, A., & Stiekema, "pBINPLUS/ARS: An improved plant transformation vector based on pBIN19," *Transgenic Res.*, vol. 4, no. 4, pp. 288–290, 1995.

[32] J. D. Schneider *et al.*, "Expression of human butyrylcholinesterase with an engineered glycosylation profile resembling the plasma-derived orthologue," *Biotechnology Journal*, vol. 9, no. 4. pp. 501–510, 2014.

[33] O. Voinnet, S. Rivas, P. Mestre, and D. Baulcombe, "An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus," *Plant J.*, vol. 33, no. 5, pp. 949–956, 2003.

[34] R. Strasser *et al.*, "Generation of glyco-engineered Nicotiana benthamiana for the production of monoclonal antibodies with a homogeneous human-like N-glycan structure," *Plant Biotechnology J.*, vol. 6, no. 4. pp. 392–402, 2008.

[35] A. v. Remoortere, C. H.Hokke, G. J. v. Dam, I. v. Die, A. M.Deelder, and D. H. v. d. Eijnden, "Various stages of Schistosoma express Lewisx, LacdiNAc, GalNAc 1-4 (Fuc 1-3)GlcNAc and GalNAc 1-4(Fuc 1-2Fuc 1-3)GlcNAc carbohydrate epitopes: detection with monoclonal antibodies that are characterized by enzymatically synthesized neoglycoproteins," *Glycobiology*, vol. 10, no. 6, pp. 601–609, 2000.

6

Supplemental Figures and Tables

SUPPLEMENTAL TABLE 1 | Primers. This table depicts primers used for amplification of HsGnTII. The part of the sequences that corresponds to the open reading frame is written in capital letters.

Primer name	Sequence 5' to 3'
HsGnTII - Forward	acggccgccagtgtgctgtcATGAGGTTCCGCATCTACAAACG
HsGnTII - Reverse	cgccagtgtgatggatatctgcatgtacaTCACTGCAGTCTTCTATAACTTTTACAGAG

SUPPLEMENTAL TABLE 2 | Glycosidases. This table depicts glycosidases used to screen for specific N-glycan motifs prior to MALDI-TOF analysis. The for this chapter relevant glycan specificity is indicated and the arrow indicates the site of cleavage.

Name of the glycosidase	Abbreviation used in this thesis	Used in screening for	Specificity
α1,2/4/6-fucosidase O	fuc-o	LDN-F vs core	R-GlcNAcβ1-4(Fucα1-3)GlcNAcβ-AA
β-N-acetyl-hexosaminidase	hexo	presence of LDN or fucosylated LDN	GalNAcβ1-4GlcNAcβ-R GlcNAcβ1-4,6Manα1-3,6-R
β-N-acetyl-glucosaminidase	gluc	presence of LDN or fucosylated LDN	GlcNAcβ1-2,4,6Manα1-3,6-R
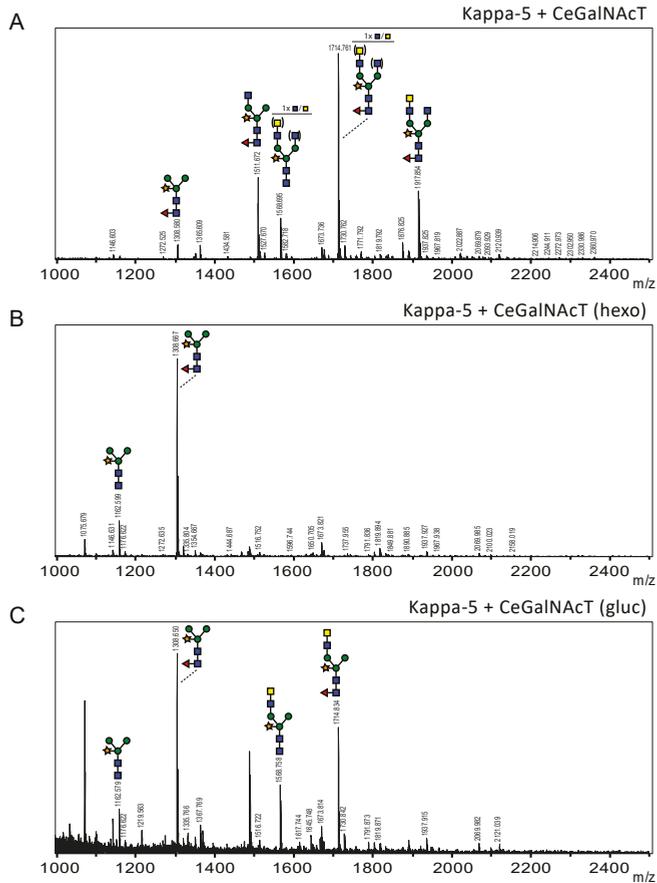
SUPPLEMENTAL FIGURE 1 | Confirmation of the presence of LDN glycan motifs. To confirm the presence of LDN structures, PNGase A released N-glycans from purified kappa-5 co-expressed with CeGalNAcT were analysed by MALDI-TOF-MS. (A) N-glycan profile for kappa-5 from wild type *Nicotiana benthamiana* plants upon co-expression of CeGalNAcT. (B) Profile of the same N-glycans upon treatment with β-N-acetyl-hexosaminidase (hexo) from *Streptomyces plicatus* (hexo). (C) Profile of the same N-glycans upon treatment with β-N-acetyl-glucosaminidase from *Xanthomonas manihotis* (gluc), which reveals successful synthesis of LDN. Sugar residues for which the position is not clear prior to enzymatic digestion are placed between brackets. The number of sugar residues of which the position on the N-glycan is not clear, is indicated above the glycan structure.
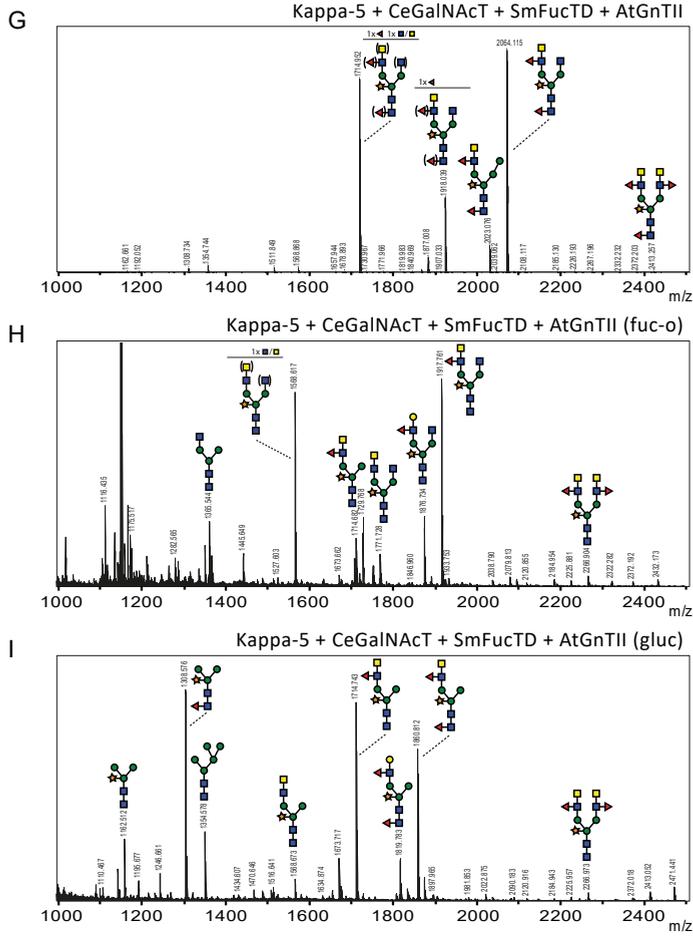
6
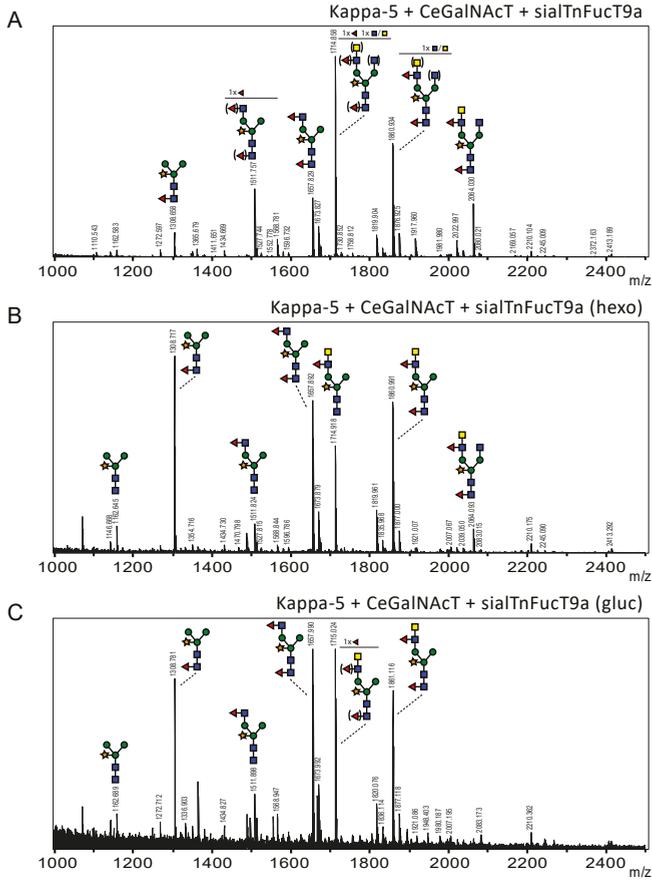
SUPPLEMENTAL FIGURE 2 | Confirmation of the presence of LDN-F glycan motifs. To confirm the presence of LDN-F structures, PNGase A released N-glycans from purified kappa-5 co-expressed with CeGalNAcT and sialTnFucT9a were analysed by MALDI-TOF-MS. (A) N-glycan profile for kappa-5 from wild-type *Nicotiana benthamiana* plants upon co-expression of CeGalNAcT and sialTnFucT9a. (B) Profile of the same N-glycans upon treatment with β-N-acetyl-hexosaminidase from *Streptomyces plicatus* (hexo), which reveals successful synthesis of LDN-F. (C) Profile of the same N-glycans upon treatment with β-N-acetyl-glucosaminidase from *Xanthomonas manihotis* (gluc), which confirms successful synthesis of LDN-F. Sugar residues for which the position is not clear prior to enzymatic digestion are placed between brackets. The number of sugar residues of which the position on the N-glycan is not clear, is indicated above the glycan structure.
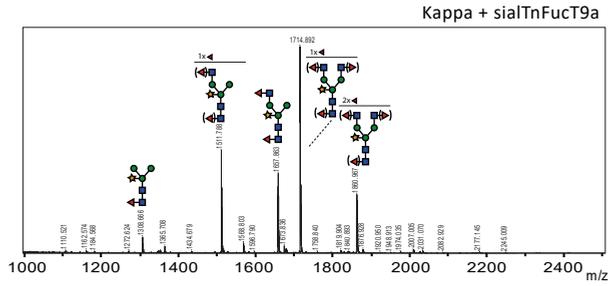
SUPPLEMENTAL FIGURE 3 | GlcNAc fucosylation by sialTnFucT9a in the absence of GalNAc. MALDI-TOF MS N-glycan profile of purified kappa-5 upon co-expression of sialTnFucT9a reveals that sialTnFucT9a is able to fucosylate the terminal GlcNAc. When a peak represents multiple possible N-glycan structures of identical mass, the number of sugar residues of which the position on the N-glycan is unclear, is indicated above the glycan structure. The possible positions of these sugar residues on the N-glycan are indicated between brackets.

6
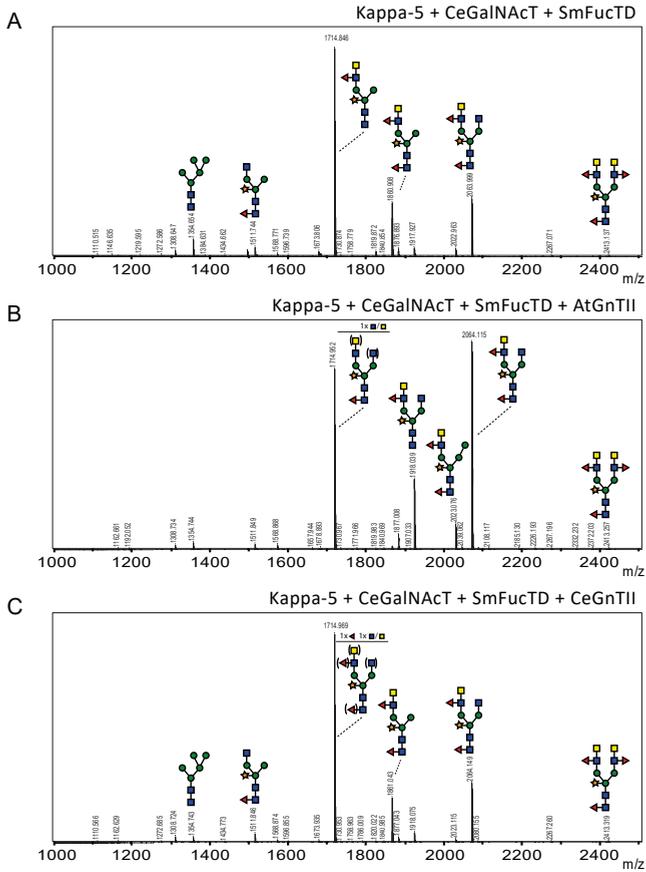
SUPPLEMENTAL FIGURE 4 | Engineering diantennary LDN-F carrying N-glycans on plant produced kappa-5 by CeGnTII and AtGnTII. The N-glycan composition of purified kappa-5 was analysed by MALDI-TOF MS upon co-expression of CeGalNAcT and SmFucTD (A), with either AtGnTII (B) or CeGnTII (C). When a peak represents multiple possible N-glycan structures of identical mass, the peak indicates the major N-glycan present based on enzymatic digestions (Supplemental Figure 5). When no enzymatic digestions were performed or when the sugar position remains unclear the possible sugar residues on the N-glycan are indicated between brackets. The number of sugar residues of which the position on the N-glycan is not clear is indicated above the glycan structure.

6
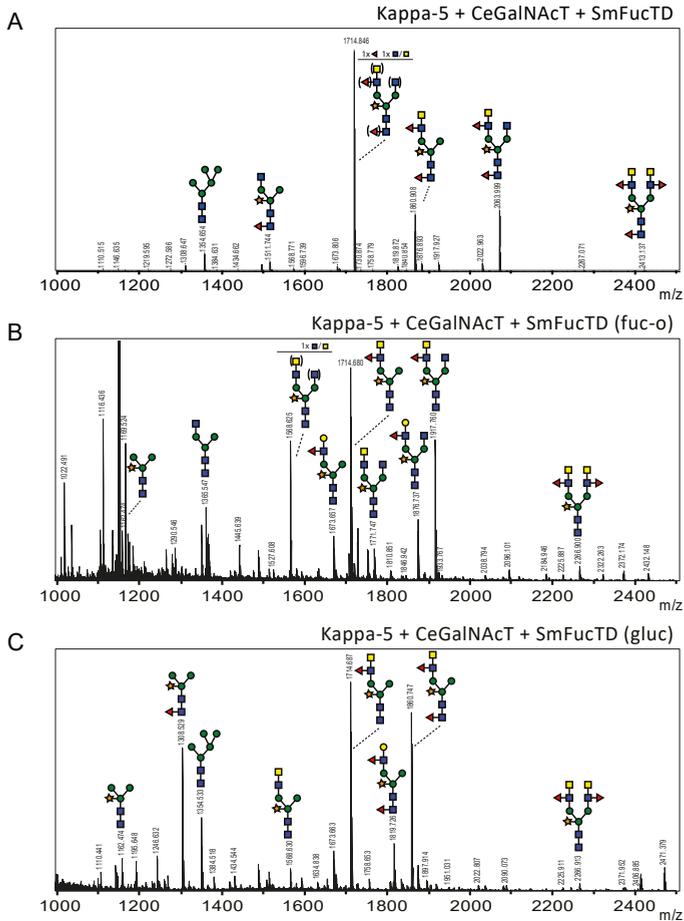
SUPPLEMENTAL FIGURE 5 | Confirmation of LDN-F glycan motifs on diantennary N-glycans of kappa-5.
To confirm the presence of LDN-F glycan motifs on diantennary N-glycans of kappa-5, PNGase A released N-glycans from purified kappa-5 were analysed by MALDI-TOF-MS. N-glycan profiles of purified kappa-5 upon co-expression of CeGalNAcT and SmFucTD (A-C) and with either HsGnTII (D-F) or AtGnTII (G-I). Released N-glycans were treated with the exoglycosidases α1,2/4/6-fucosidase O from *Omnitrophica* (fuc-o) or β-N-acetyl-glucosaminidase from *Streptococcus pneumoniae* (gluc). Sugar residues for which the position is not clear (prior to enzymatic digestion) are placed between brackets. The number of sugar residues of which the position on the N-glycan is not clear is indicated above the glycan structure.

6
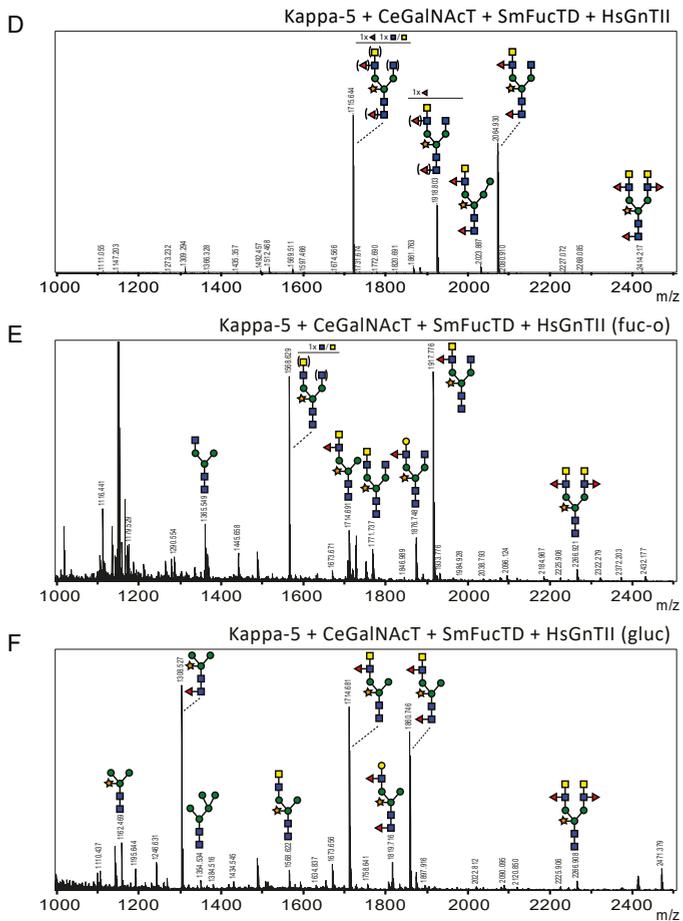
SUPPLEMENTAL FIGURE 5 | Confirmation of LDN-F glycan motifs on diantennary N-glycans of kappa-5.
Continued

SUPPLEMENTAL FIGURE 5 | Confirmation of LDN-F glycan motifs on diantennary N-glycans of kappa-5.
Continued
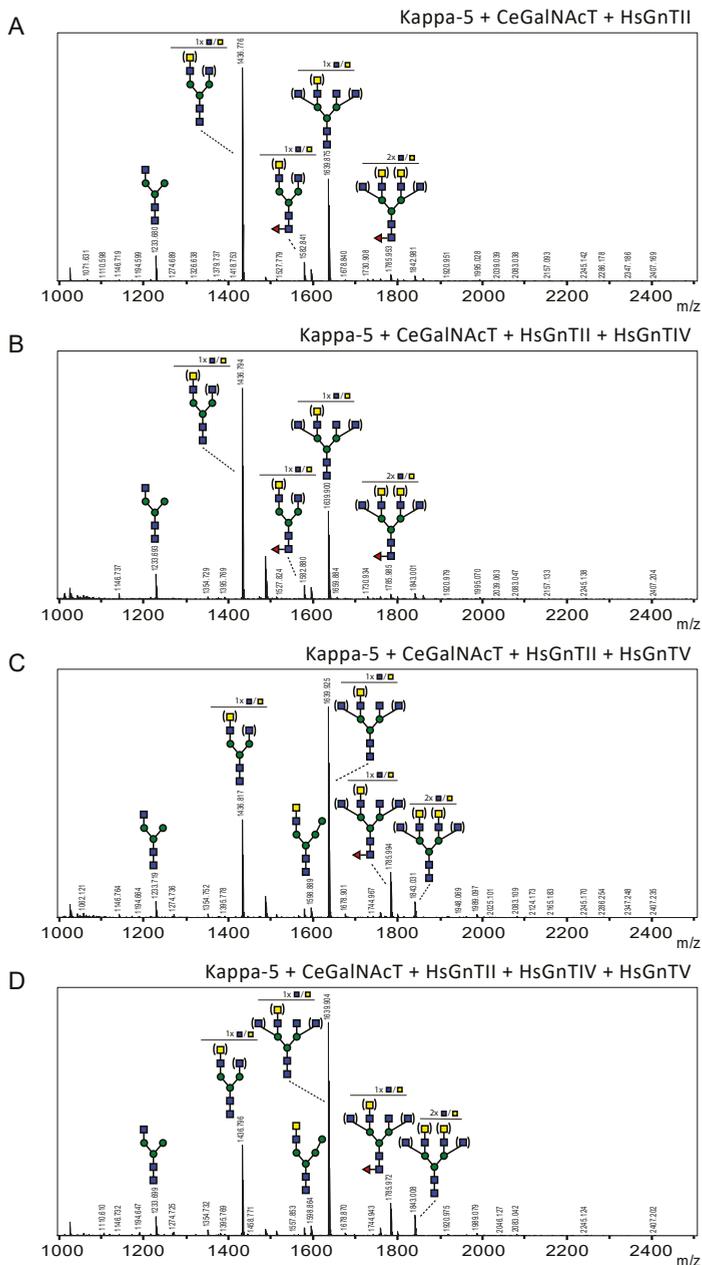
SUPPLEMENTAL FIGURE 6 | Engineering triantennary N-glycans on kappa-5 produced in ΔXT/FT *Nicotiana benthamiana* plants. The N-glycan composition of purified kappa-5 was analysed by MALDI-TOF MS upon co-expression of CeGalNAcT and HsGnTII (A) with either HsGnTIV (B), HsGnTV (C) or HsGnTIV and HsGnTV (D) in ΔXT/FT *Nicotiana benthamiana* plants. When a peak represents multiple N-glycan structures of identical mass the number of sugar residues of which the position on the N-glycan is not clear, is indicated above the glycan structure. The possible positions of these sugar residues on the N-glycan are indicated between brackets.
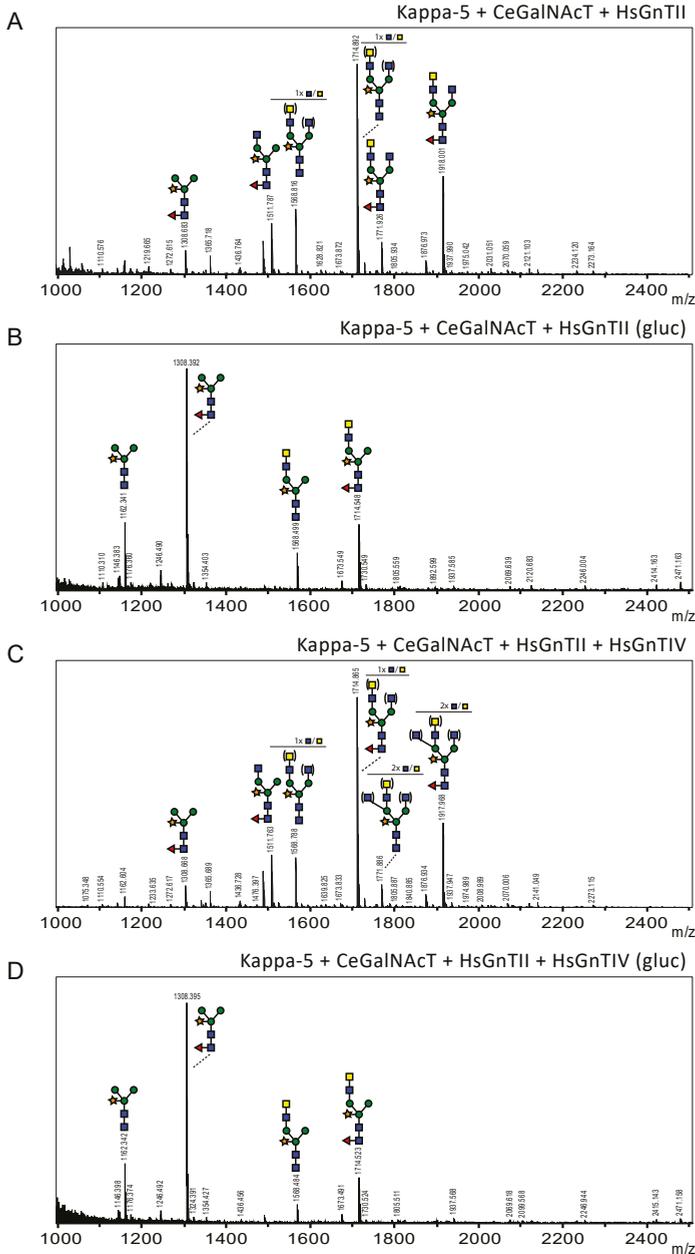
SUPPLEMENTAL FIGURE 7 | Triantennary N-glycans on kappa-5 in wild type *Nicotiana benthamiana* plants. N-glycan profiles of purified kappa-5 upon co-expression of CeGalNAcT, SmFucTD and HsGnTII (A-B) with either HsGnTIV(C-D) HsGnTV (E-F) or HsGnTIV and HsGnTV (G-H) in wild type *Nicotiana benthamiana* plants. Released N-glycans were treated with the exoglycosidases β-N-acetyl-glucosaminidase from *Streptococcus pneumoniae* (gluc). Sugar residues for which the position is not clear (prior to enzymatic digestion) are placed between brackets. The number of sugar residues of which the position on the N-glycan is not clear, is indicated above the glycan structure.

6
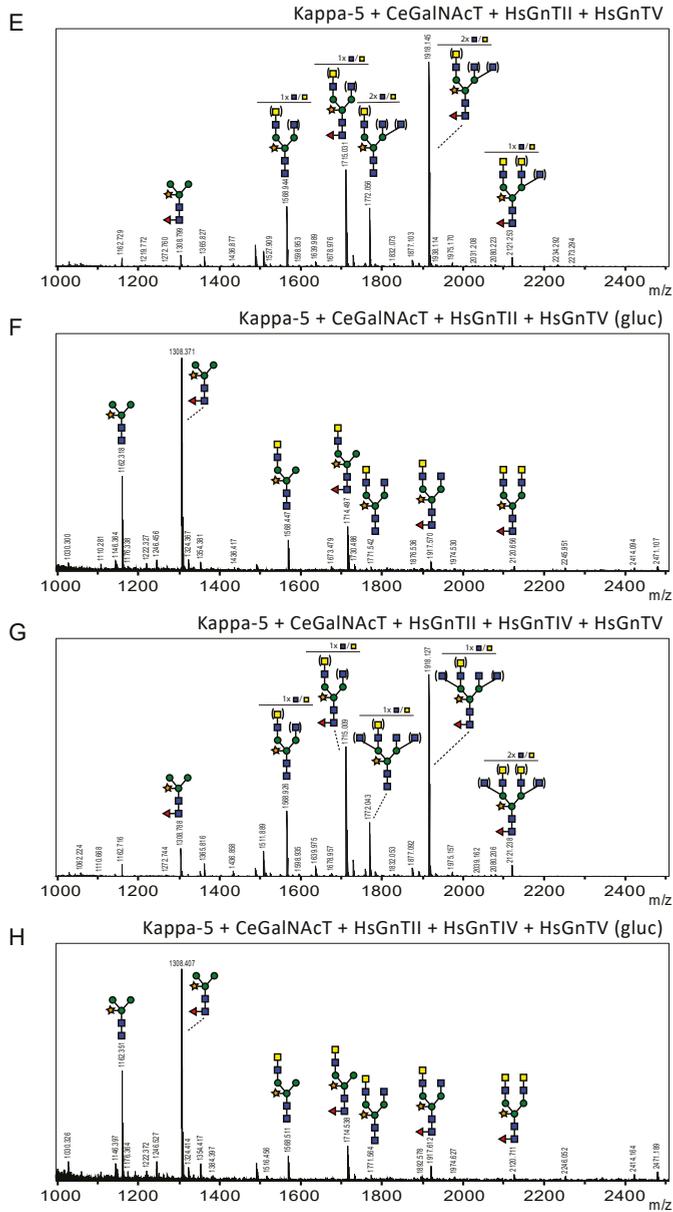
SUPPLEMENTAL FIGURE 7 | Triantennary N-glycans on kappa-5 in wild type *Nicotiana benthamiana* plants. Continued

Kappa-5

atggccaagaccaacctcttcctcttcctcatcttctccctcctgctctccctctcctccgccgtcgaggacgctagccaccatcaccaccatcacgactacaaggac
gatgacgacaagactagtCAGTCCCCCACCAACGACGAGATGCACGCCACCATCTCCGAGTACGGTCGTCTCTACATCACCAACATC
CACATCCACTACCGTCTCCTCATCATGGCCCTCTCCCCCAACATGAAGTTCACCCCCGGTGAGGCCGACAACATCCTCCACAAGTC
CGAGGAGGAGCACCAGGTCAAGTGGGCCCTCAACTACCTCAACGCCGCCCGTTCCACCTGGAAGCTCGAGAACGAGGACAT
GTTCAAGAAGACCATGTCCTCCTACGTCGGTGTCAACAAGACCATCGTCCCCAACTTCTGCAAGACCATGCTCCAGCAGTCCTC
CGCCCGTAACTGGGACACCAACATCCAGAAGAACATCACCTACGGTTGCGAGGTCCTCAAGAAGTACTCCGAGCTtAAGTGGGGT
GCCCGTAAGAAGCTCGACCTCATGACCATCCGTTGGCTCAACGGTTCCGACGAGAACGGTCAGTCCCAGCACATCTCCGA
CAAGGGTTTCAACTACCACTCCAAGAAGGAGTACCTCGAGTGCGGCCCAGTCCGTCATGAAGATCCACCGTACCAAGGCCGAGGTC
GACTGCCGTTCCACCGTCGGTGAGTTCCTCAAGCTCCAGCAGGTCAAGGACACCCCCTCCTCCAAGACCCAGGCCATCCAGTTC
GACAAGATCAACGAGAACTTCAACAAGTCCCTCGTCGAGCTtAAGCGTCGTGTCGAGAACCTCGAGATCAACCGTATCAACTACCT
CAAGTTCATGAACCCCATGGAGCGTGTCATCTCCGTCATCGACGCCATGGAGGAGGAGGTCGACCACAAGTACGGTGTCTACATGTAA

CeGnTII

ATGATGGTCTATCGACGGATGCACCGTTTGGCAAATGCTGTTATAGCGTGTTGTTTATTCGGGTTCATCGTCATTTTTCTGAAAGCACCTG
GAGAAGATCAACGGCTACGAGACGGTGTTCCTGTATTATCACAAAATCCTGTGTGGTCAGTTAATGACTGGAATGGTTTGAATA
AAGAAGTCGTAGATTTATTGAAAAATGAATCGAATCGTCTGTCGATTAACGAAAAGCCCGAATTGAGTGGATGGAATTTTAAAAC
TAAAGAAGACATGTCACTGAAAGGATCGGAAATCGTCGAATCTGTTAGTTTTCTTAATGAAAATTTTGACATTTTGAACGCTGC
GAAATTTGGAGATTTGTCCACTGTGAAGACAATTCTAGTGATTCAGGTACACGATCGGCCGGTCTATTTGCAATACCTCATTGAGT
CAATGCGAAATACAAAAGGAATAGAAGACACATTACTAGTTTTCTCGCATGACATCAATGTCGGAATAATCAATGAAATGATCCG
TAACATCACATTCGCTCGTGTCTACCAAATCTTCTATCCATACAATTTGCAACTGTTTCCGACAGTTTTTCCAGGACAGTCTCCTTC
CGATTGTCCGGAGAAAATGAAGAGAGATAAAGCACAAGAGACGAATTGTTCAAATTGGAGTAGCCCGGATAAATATGGAAACTATC
GAGTAGCTCAATTGACACAGATAAAACATCATTGGTGGTGGAAGATGAACTTTGTGTTTGATGGAATTGTTGAGAAGTATTCGAT
GAAAGATCCATGGGTTCTACTACTGGAAGAAGATCACATGCTGGCTCCAGATGCATTACATGTTCTTGATATTATTGTATCAAATCGTC
CAAAATATTGTGAAAACTGCGAAATAATATCTCTGGGATTTTATTTGAAATCCACTAACAAATACGGTCAAGACATAGCTCATCTCG
GAGTTCACCCATGGTACAGTAGCAAGCATAATATGGGAATGGCTCTTCAGAAGAACACGTCGCAGAAGATCAAAGGATGCTCC
GAGATGTTTTGCAAATGGGATGACTATAACTGGGATTGGTCGTTGATGCAGATTTCAGCAAAATGTCTACCTCAAAGGTTTCGGGT
CATATTCACAAAAAGTCCACGAGTTATACATATAGGAGACTGTGGAGTCCATACGCACAGATGTGAAGCACATAAAGCTCTCCAATC
GACACAAGAACTATTCCGTCAGCATAAGGATCTTCTGTTCCCGACAAGTTTATCTGTGACAGACACTTCGAGAAGATCTTTGAAGC
CGTCGAAAGAAAATGGTGGATGGGGTGACATTAGAGACCGTCAACTATGCGAAATCAACAAATCTCCATTGGTTAGAGTTTCTTCT
CAATCTGCATCTGTACTCCATAAATTGCTCAACTCAAAAATCCAGTTCTCGTCATCCAACAAAACAATCACATCTACAACTTCTTAA

SUPPLEMENTAL FIGURE 8 | Protein sequences. Codon optimised sequence of kappa-5 (incl. signal peptide and N-terminal HIS/FLAG-tag in lowercase) and the sequence of CeGnTI

6

Chapter 7

General discussion

Kim van Noort

Laboratory of Nematology, Wageningen University and Research, Wageningen, the Netherlands.

Plants have shown to be a promising platform for the production of biopharmaceuticals. The first plant-derived recombinant protein is on the market and even more are enrolled in clinical trials. The achieved "humanisation" of the plant glycosylation pathway shows the possibility of glyco-engineering in plants. However, the main focus of the plant molecular pharming community so far was on "humanisation" of the plant glycosylation pathway, whereas plants are an interesting research tool for production of a wide variety of glycosylated proteins. Thanks to their ability to allow glyco-engineering, and thereby enable production of glycoproteins with (near to) native N-glycans. Therefore, plants could also enable production of immunomodulatory helminth glycoproteins, which could be used as vaccine target or to treat autoimmune or other inflammatory disorders. Isolation of these helminth proteins from natural sources is laborious and yields only low amounts of proteins. Therefore, production of these proteins relies on heterologous expression systems. However, research with recombinant proteins carrying non-native glycans may not be representative, because glycosylation has shown to be important for protein activity and vaccine development. This thesis describes engineering of the N-glycosylation pathway in *Nicotiana benthamiana* for the production of helminth glycoproteins. Such a production platform enables further research to the biology of helminth glycoproteins and their development as potential new biopharmaceuticals or vaccines. In this chapter, I will elaborate on the function of the plant Golgi, where an important part of glycosylation takes place. Furthermore, I will discuss our findings in glyco-engineering of the plant glycosylation pathway.

Plant Golgi functioning

In the Golgi removal and addition of specific sugar residues takes place in a spatiotemporal manner. Hence, it is crucial to have proper understanding of Golgi functioning with regard to the localisation of glycosyltransferases in the Golgi as well as the timing of their activity. In contrast to the human Golgi that consists of a single stacked continuous compartment linked via tubules, plant cells contain numerous small Golgi bodies [1]. These Golgi bodies consist of multiple stacked, flattened and polarised cisternae. Golgi bodies are polarised from *cis*- to medial- to *trans*-Golgi cisternae, which are distinguished based on structural differences and luminal width [2]. Protein cargo from the endoplasmic reticulum (ER) enters the Golgi at the Golgi *cis*-cisternae and leaves the Golgi at the *trans*-Golgi continuing into the trans Golgi network (TGN) to the point of exit. The plant Golgi bodies are highly mobile and can move up to 2.2 µm/sec in a cytoplasmic stream, although differences in motility were observed in for instance cotyledons and the hypocotyl [3], [4]. To enable this high mobility, the Golgi cisternae are held together by the Golgi matrix. The Golgi matrix consists of structural proteins that surround the Golgi or are present in between Golgi cisternae. These structural proteins are involved in Golgi biogenesis and tethering

events. Golgi bodies move together with the ER, although the exact mechanism behind this ER-associated movement is still unknown, Golgi body movement seems to depend on myosin. Sparkes and colleagues [5] showed that Golgi body movement was inhibited by a truncated form of myosin. A physiological association between the ER and the Golgi was shown by movement of a laser trapped Golgi body and subsequent ER movement [6]. This physiological association was confirmed by a weakened ER-Golgi interaction after over-expression of a truncated *Arabidopsis thaliana* Golgin AtCASP [7]. Recently, a model is proposed in which myosin-driven actin filament sliding drives Golgi movement by tethering of the Golgi to the ER and simultaneous tethering of the ER to actin [3].

The tethering connection between the ER and the Golgi does not necessarily imply that cargo between the ER and the Golgi is transported over a permanent luminal connection between the two [8]. Coat protein (COP) II coated vesicles are present in plants, which are proposed to move protein cargo from ER exit sites (ERES) to the Golgi. However, the formation and movement of ~65 nm COPII vesicles over the small distance between the ER and Golgi (200-500 nm) is questioned and a tubular continuity for cargo transfer is suggested [9]–[11]. During confocal analysis of GFP tagged *Schistosoma mansoni* fucosyltransferases (SmFucTs), we sometimes observed tubular like structures, which suggest a tubular connection between the ER and the Golgi (Chapter 2, Supplemental Figure 3). However, these tubular-like structures could also consist of a series of COPII vesicles. Nowadays, the consensus in the field is that cargo from the ER is transported largely by COPII vesicles, although this may not be the only way of transport from the ER [12]. A way to fuse Golgi stacks with the ER is by the use of Brefeldin A (BFA), a fungal metabolite that inhibits formation of COPI coated vesicles [13], [14]. BFA disruption of the Golgi and the ER-Golgi interface showed that two vacuole localised protein pumps move to the tonoplast independently from the ER-Golgi interface or post-Golgi-trafficking [15]. Thereby showing another form of transport from the ER than COPII vesicle transport to the Golgi.

Two types of COPI coated vesicles were identified, enabling retrograde transport from the *cis*-Golgi to the ER by COPIa coated vesicles and retrograde transport from the medial/*trans*-Golgi to the *cis*/medial-Golgi by COPIb coated vesicles [16], [17]. Upon BFA treatment the Golgi stacks fuse with the ER in a *trans* to *cis*-Golgi manner, including the *trans*- and *cis*-Golgi localised matrix proteins [14]. However, proteins solely active at the ER/Golgi interface, such as ER-α-mannosidase I (MNS3), localise in punctuated structures and are suggested to function as starting point for Golgi biogenesis [18]. Upon washout of BFA the Golgi was build up again from the ERES in a *cis*- to *trans*-Golgi fashion, starting with the *cis*-matrix followed by the *cis*/medial-Golgi proteins and subsequently the *trans*-matrix and *trans*-Golgi proteins [14], [19]. This suggests that Golgi matrix proteins are important for proper Golgi assembly. The retrograde transport by COPI vesicles and the order of Golgi assembly and disassembly suggest that intra-Golgi trafficking follows the cisternal maturation model.

7

The cisternal maturation model is one of the multiple Golgi trafficking models proposed over the years. This model and the vesicle transport model are the two main models for plant Golgi trafficking [20]. In the cisternal maturation model, proteins are transported from the ER to the Golgi in COPII vesicles. A few of these vesicles are fused together to form the first cisternae [11] (Figure 1). From the *cis*- to the *trans*-side of the Golgi the cisternae mature, by budding and fusion of COPI vesicles containing for instance glycosyltransferases. Glycosyltransferases are constantly shuttled back to earlier cisternae by retrograde transport via COPI vesicles. In the vesicle transport model, not the glycosyltransferases are shuttled by COPI vesicle transport, but the protein cargo. In this model the different Golgi cisternae with corresponding glycosyltransferases are stationary and the protein cargo is transported to the next Golgi compartment via COPI vesicles. Retrograde transport is used to recycle trafficking components to the ER via COPI vesicles with a diameter of ~45nm [11].

Following the cisternal maturation model, development of a new Golgi stack via fission starts with assembly of two half size *cis*-cisternae from the same ERESs [17], [19] (Figure 1). These new *cis*-cisternae mature to later Golgi compartments, after which the ERESs is divided in two and the cisternae increase towards two normal sized Golgi stacks. Golgi biogenesis via this "fission model" is based on the secretory model unit, which predicts that the Golgi is continuously associated with the ERESs. Another model for Golgi biogenesis, based on the secretory model unit, is *de novo*-construction from non-Golgi associated ERESs [12]. This Golgi biogenesis model is based on the observation of non-Golgi associated ERESs. During co-localisation analysis with confocal microscopy donut/ring structures were observed. The donut/ring structures probably correspond to the outside of the cisternae where the glycosyltransferases are localised in the membrane or incorporated in COP vesicles for transport. Sometimes multiple donut/ring structures were observed in close association, which could represent development of new Golgi stacks via fission (Chapter 2, Supplemental Figure 3). However, the start of Golgi fission is observed at the *cis*-Golgi, whereas our observations showed these double donut/ring structures only for the more *trans*-Golgi localised α1,3-SmFucTs and not for the co-expressed markers [17] (Figure 1). Another explanation for the observed double donut/ring structures could be that the glycosyltransferase is present in two closely related cisternae in one Golgi-body.
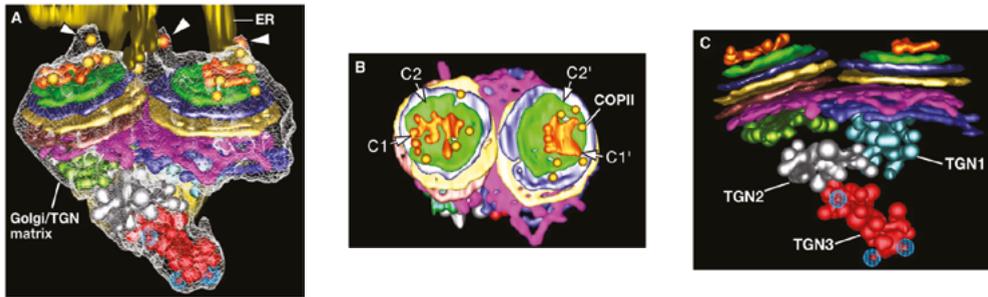
7

FIGURE 1 | 3D tomographic model of dividing Golgi stack. This 3D tomographic model, of a dividing Golgi stack docked to an ER export site in an *Arabidopsis* root meristem cell, illustrates the Golgi composition and its relation to the ER. Furthermore, it illustrates development of a new Golgi stack via fission and *de novo* assembly of cis-Golgi cisternae from COPII vesicles as postulated by the cisternal maturation model. (A) The two sets of *cis-* and medial-cisternae are held together by a larger *trans-*Golgi cisterna (pink). The Golgi stack is connected to budding COPII vesicles via interactions between the COPII scaffolds and the *cis-*Golgi matrix (arrowheads). (B) Face-on view of the dividing Golgi stack. The C1 (orange) and C2 (green) cisternae of the two *cis-*side stacks (C1, C2, C1', and C2' marked with arrows) display *de novo* assembly of cis-Golgi cisternae from COPII vesicle (yellow spheres). (C) Side view of the dividing Golgi stack. On the *trans-*side, three trans Golgi network cisternae (TGN1–TGN3) are indicated at different stages of maturation. This picture is copyrighted by the American Society of Plant Biologists and is reprinted from the publication of Staehelin and Kang [17] with permission.

Glyco-engineering

Glycosyltransferase localisation and function

In order to produce helminth glycoproteins with native N-glycosylation, the glycosylation pathway of plants must be adjusted. This adjustment is done by transient co-expression of glycosyltransferases via agro-infiltration. Since, formation of specific N-glycans takes place in the Golgi, these glycosyltransferases need to be localised in the plant Golgi. Correct Golgi sub-localisation of glycosyltransferases is important for availability of the correct donor and acceptor. The function of a glycosyltransferase is determined by the glycan motif it synthesises, which is enabled by the donor and acceptor. Therefore, the function of a glycosyltransferase is related to its sub-Golgi localisation. For the SmFucTs we observed subtle differences in sub-Golgi localisation, corresponding to the function found in our characterisation studies (Chapters 2 and 3). Glycans of *S. mansoni* are fucosylated in a specific order, starting with core α1,6-fucosylation followed by core α1,3-fucosylation and subsequently fucosylation of terminal glycan motifs, such as LeX [21]. Therefore, the core α1,6-SmFucT SmFucTH was expected to localise before the core α1,3-SmFucT SmFucTC. Terminal glycan motif fucosylating enzymes SmFucTA, D, E and F were expected to localise after SmFucTH and SmFucTC. Corresponding to their functions, SmFucTC localises less to the *trans-*Golgi than SmFucTA, D, E and F (Figure 2 and Chapter 2). Also, the plant core α1,3-FucTs, FUT11 and FUT12, localise earlier in the Golgi than the plant terminal LeA synthesising FUT13

[22], [23]. This shows that although plants cannot naturally synthesise the terminal motifs of *S. mansoni*, the order of fucosylation in the glycosylation pathway is similar. This suggests that the same mechanisms are in place for correct Golgi localisation of core α1,3- and terminal fucosylating FucTs. Since, core α1,6-fucosylation of *S. mansoni* glycans was shown to happen before other types of fucosylation, localisation of SmFucTH was expected in early parts of the Golgi. However, SmFucTH showed co-localisation with all three Golgi markers, corresponding to *cis*/medial-, medial- and *trans*-Golgi localisation (Figure 2 and Chapter 2). Although, this localisation was unexpected other mechanisms or environmental conditions can influence glycosyltransferase activity [24]. For instance, the pH in the secretory pathway becomes more acidic towards the end of the secretory pathway and can influence protein activity rates through the glycosylation pathway. Therefore, localisation of SmFucTH throughout the Golgi does not necessarily equal activity in the whole Golgi. On the other hand, since plants lack core α1,6-fucosylation, SmFucTH localisation could be aberrant in plants.
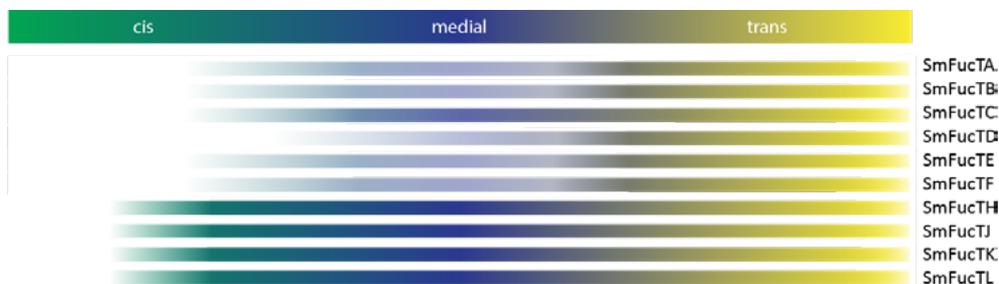
FIGURE 2 | Schematic representation of *Schistosoma mansoni* fucosyltransferase (SmFucT) Golgi localisation in plants. Based on analysis of SmFucT sub-Golgi localisation performed in Chapter 2 a schematic representation was created of SmFucT Golgi localisation in *Nicotiana benthamiana*. The three colours, green, blue and yellow indicate the polarisation from *cis*-, to medial- to *trans*-Golgi.

Aberrant localisation

Throughout the course of our studies to engineer the plant glycosylation pathway, we have made several observations of aberrant localisation. One of these observations was the localisation of SmFucTH. Following the cisternal maturation model, SmFucTH is probably shuttled back towards the *cis*/medial Golgi via COPI vesicles. However, since plants lack core α1,6-fucosylation, core α1,6-SmFucT transport signals may be absent in plants. This can result in retrograde transport of SmFucTH at the end of the Golgi instead of the *medial*-Golgi and subsequent localisation through the whole Golgi.

The importance of retrograde transport via COPI vesicles was shown for the sub-Golgi localisation of β1,2-N-acetylglucosaminyltransferase (GnT) I. In the absence of COPI vesicle transport, GnTI accumulated in the *trans*-Golgi and was subsequently transported

to the apoplast or vacuole [25]. During our co-localisation studies with the fluorescently tagged CTS domain of *Rattus norvegicus* α2,6-sialyltransferase (ST-mRFP) we observed that ST-mRFP expression needs to be regulated to prevent aberrant localisation in other cell compartments, such as the apoplast (personal communication C. Hawes, unpublished data). This Suggests that high amounts of non-plant glycosyltransferases may overload the plant Golgi mechanisms for correct retrograde transport via COPI vesicles, which may explain the aberrant localisation observed in the apoplast or vacuole. In another confocal experiment with GFP-tagged *Homo sapiens* and *A. thaliana* GnTII (HsGnTII and AtGnTII, respectively), AtGnTII was only seen in the Golgi, whereas HsGnTII seemed to localise also in the ER (van Noort, unpublished data). Upon P19 co-expression GFP-tagged HsGnTII localisation in the ER became stronger. This suggests overload of the plant Golgi mechanisms for correct ER to Golgi transport via COPII vesicles, which may explain the aberrant localisation in the ER. So, not only COPI vesicle transport is important for glycosyltransferases Golgi localisation, but also COPII vesicle transport from the ER to the Golgi. Nevertheless, both AtGnTII and HsGnTII were able to synthesise diantennary N-glycans on omega-1 and kappa-5 (Chapter 5 and 6).

Diantennary N-glycans on omega-1 were synthesised by HsGnTII instead of *N. benthamiana* GnTII (NbGnTII), because upon co-expression of *Danio rerio* and *H. sapiens* β1,4-galactosyltransferase (β1,4-DrGalT and β1,4-HsGalT, respectively) NbGnTII is inhibited. Too high expression of both β1,4-GalTs leads even to inhibition of more *N. benthamiana* glycosylation related enzymes, such as Golgi α-mannosidase II (GMII) and xylosyltransferase (XylT), suggesting that β1,4-DrGalT and β1,4-HsGalT localise earlier in the Golgi and thereby inhibit the activity of GMII and XylT [26]. To prevent aberrant Golgi localisation of β1,4-DrGalT or β1,4-HsGalT in *N. benthamiana* expression needs to be controlled tightly [26], [27]. In mammals β1,4-GalT shows homodimers, whereas in plants no homodimers are formed for plant β1,3-GalT [28]–[30]. Furthermore, in plants β1,3-GalT forms heterodimers with GMII, whereas β1,4-GalT does not in mammals. So, differences are observed between GalT dimerisation in plants and mammals.

Following the oligomerisation model, dimerisation can prevent incorporation in vesicles for further transport. Furthermore, a double mutation in the transmembrane domain (TMD) of β1,4-HsGalT resulted in blocked homodimerisation and subsequently reduced Golgi retention [31]. This suggests that homodimerisation could be a form of controlled Golgi localisation for β1,4-GalT. However, upon reduced expression of β1,4-GalT in plants no aberrant glycosylation was observed. This suggest that only too high expression levels result in aberrant localisation and may overload the plant Golgi mechanisms for correct Golgi localisation, retention or retrograde transport [26], [27].

7

Protein affinity and expression

Next to overload of the plant Golgi localisation mechanisms by high glycosyltransferase levels, high glycosyltransferase or nucleotide sugar levels can make low affinity binding of sugar residues to glycosyltransferases and thereby low affinity enzymatic reactions more likely. For instance, the core α1,6-FucT of *Mus musculus* and the core α1,3-FucT of *A. thaliana* can core galactosylate a N-glycan *in vitro* at low fucose and high galactose concentrations [32]. This indicates that regulation of the nucleotide sugar concentration and localisation can influence the N-glycan composition. This is for instance also seen in hepatocellular carcinoma in which a correlation was found between the expression of a GDP-fucose transporter and the level of cellular fucosylation [33].

Our characterisation studies suggest that SmFucTF can fucosylate LN and LDN to synthesise LeX and F-LDN, respectively. Depending on the available glycan acceptor LeX or F-LDN is synthesised by SmFucTF, although F-LDN synthesis seems far less efficient (Chapter 3). The improved fucosylation of the GalNAc in LDN by SmFucTF upon co-expression of LDN-F synthesising enzymes SmFucTD or SmFucTE, suggests that SmFucTF has a higher affinity for LDN-F in comparison to LDN. On the other hand, SmFucTF could require the presence or dimerisation of SmFucTD or E to create successive fucosylation of LDN, as is observed for many other successive acting glycosyltransferases in plants and mammals [28]–[30]. Dimerisation can enhance glycosyltransferase activity by enzyme stabilisation and facilitation of cooperate substrate binding and catalysis [34].

Another option for the lower efficiency in F-LDN synthesis by SmFucTF is LDN cleavage by β-hexosaminidases (HEXOs) in the Golgi before fucosylation by SmFucTF can take place (Wilbers *et al.*, unpublished data). In insect cells a HEXO is found that processes N-glycans inside the Golgi [35], [36]. *N. benthamiana* HEXOs (NbHEXOs) are assumed to localise in the vacuole and apoplast [37]. However, NbHEXOs carry Endo H digestion resistant N-glycans that are processed in the Golgi, indicating NbHEXOs are transported to the vacuole and apoplast via the Golgi. The pH optimum for *A. thaliana* HEXO activity is around 5.0 [38]. Through the secretory pathway the pH goes down, from 7.1 in the ER to 6.3 in the TGN [39]. The pH in the apoplast is variable and can for instance change as a consequence of salt stress [40]. Normal pH in the apoplast of *N. tabacum* seems to be more alkaline than the pH in the *trans*-Golgi [41]. This indicates that HEXOs may be activated by the more acidic pH at the end of the Golgi and the TGN. Experiments with HEXO overexpression in combination with LDN-F synthesising enzymes on kappa-5 resulted in a reduction of LDN-F glycan motifs, whereas experiments with F-LDN synthesising enzymes on kappa-5 and HEXO RNAi silencing constructs seemed to result in more F-LDN glycan motifs (van Noort, unpublished data). This suggests that HEXOs can indeed be active in the Golgi and hamper F-LDN synthesis. The HEXO overexpression experiment shows that the level of HEXO expression can influence formation of N-glycans in the Golgi.

These different examples show that N-glycosylation can be adjusted by regulation of for instance glycosyltransferase, glycosidase and transporter expression levels and thereby the availability of glycan acceptors or nucleotide sugars. Indicating that plants can regulate the N-glycan composition on their glycoproteins by regulation of protein expression levels.

Differences in N-glycosylation of plant-produced glycoproteins

Next to the Golgi micro-environment, also protein intrinsic properties play a role by protein N-glycosylation. Upon production of glycoproteins in plants we observed differences in N-glycosylation. Plant-produced kappa-5 shows high proportions of N-glycans with terminal GlcNAcs, whereas omega-1 carries paucimannosidic N-glycans when isolated from the apoplast of *N. benthamiana* (Chapter 5 and 6). Differences in glycan composition on plant produced kappa-5 and omega-1 are probably introduced by HEXO activity, which can cleave terminal GlcNAc from protein N-glycans [37], [38]. Since, kappa-5 N-glycans carry terminal GlcNAcs, these GlcNAc residues may be protected against HEXO binding by protein intrinsic properties. Similarly, AAL lectin binding assays with kappa-5 showed that AAL could not bind to the core α1,3-fucose present on kappa-5 N-glycans, suggesting glycan protection (Chapter 4, van Noort, unpublished data). On the other hand, upon HEXO overexpression the GlcNAc residues on kappa-5 N-glycans were cleaved, showing that HEXOs are able to cleave the terminal GlcNAcs of kappa-5 N-glycans when enough HEXOs are present. Furthermore, kappa-5 N-glycans can be extended by addition of terminal GalNAc and fucosylation of GlcNAc and/or GalNAc. This suggest that the N-glycans of kappa-5 are probably only partly protected and HEXO and AAL binding may be influenced by protein intrinsic properties.

Next to omega-1 and kappa-5, we made use of the glycoprotein IL-22 in our studies. The N-glycans on wild type plant produced IL-22 carry non α1,3-fucosylated N-glycans with terminal LeA [42]. Therefore, it was remarkable that upon co-expression of IL-22 with NbFucTA1, D5 or F11 in ΔXT/FT plants, core α1,3-fucosylation was observed on the N-glycans of IL-22 (Chapter 4) [43]. The presence of the terminal LeA glycan motif found on plant-produced IL-22 probably depends on protein intrinsic properties. The presence of LeA is highly reduced upon introduction of the N54S mutation, in one of the glycosylation sites [42]. Moreover, introduction of this mutation resulted in a majority of core α1,3-fucose carrying N-glycans on IL-22. This suggests that the N-glycan on this glycosylation site can influence the introduction of core α1,3-fucose on other glycosylation sites. Also, upon production of IL-22 in ΔXT/FT plants the presence of LeA is reduced. Thus, the absence of LeA in ΔXT/FT plants, may be the reason that core α1,3-fucose could be introduced on IL-22 N-glycans upon co-expression of active NbFucT proteins. These observations show that protein intrinsic properties affect N-glycosylation. Another option could be that similar to GlcNAc cleavage by HEXO overexpression, IL-22 is core α1,3-fucosylated when enough FucTs are present.

7

While producing helminth glycoproteins omega-1 and kappa-5 *in planta*, we have tried to increase the protein production by agroinfiltration with the *Agrobacterium tumefaciens* strain AGL-1. Production of kappa-5 was higher upon agroinfiltration with the *A. tumefaciens* strain AGL-1 in comparison to the MOG101 strain that was used commonly (van Noort, unpublished data). However, processing of kappa-5 N-glycans was incomplete, as was shown by the high mannose carrying N-glycans on purified kappa-5 (van Noort, unpublished data). This suggests that upon too high expression of kappa-5 the glycosylation pathway seems to become overloaded, resulting in non-Golgi processed N-glycans. Another possibility is that kappa-5 was transported from the ER to the apoplast bypassing the Golgi via unconventional protein secretion [44]. Expression of omega-1 using AGL-1 did result in paucimannosidic plant N-glycans. Production of omega-1 upon agroinfiltration in MOG101 is lower than production of kappa-5 under the same conditions. Probably, the amount of omega-1 produced upon agroinfiltration with the AGL-1 strain is still manageable by the Golgi. This shows that next to the protein intrinsic properties the amount of carrier protein produced by the plant can influence N-glycosylation processing.

Conclusion

Different models surround the working of the Golgi, from its interaction with the ER to the transport between different cisternae and Golgi biogenesis. This makes glyco-engineering challenging and aberrant results difficult to interpret. It is clear that the localisation of the glycosyltransferases is important for their function and that the right expression level is important to regulate their localisation. On the other hand, glycan synthesis can be influenced by regulation of other glycosyltransferases, glycosidases and transporter expression levels.

Although, we do not know exactly how the Golgi works, we show in this thesis that characterisation of glycosyltransferases and synthesis of various glycan motifs in plants is feasible. With this knowledge we do have the tools in hand to produce proteins with helminth-like N-glycans, to study protein function and to develop vaccines and proteins relevant for use in diagnostics. The production of glycoproteins with tailored N-glycans enables studies to the effect of N-glycans on protein function. Furthermore, characterisation of glycosyltransferases enlarges our glyco-engineering toolbox and enables research to the function of glycosyltransferase in for instance parasite biology or development. At last, screening the synthesis of specific glycan motifs with carrier glycoproteins enables studies on glycosyltransferase functioning under different micro-environmental conditions in the Golgi.

References

[1]   M. Pavelka and D. G. Robinson, "The Golgi Apparatus in Mammalian and Higher Plant Cells: A Comparison," in *Annual Plant Reviews online*, vol. 9, 2018, pp. 17–38.

[2]   L. A. Staehelin, T. H. Giddings, J. Z. Kiss, and F. D. Sack, "Macromolecular differentiation of Golgi stacks in root tips of Arabidopsis and Nicotiana seedlings as visualized in high pressure frozen and freeze-substituted samples," *Protoplasma*, vol. 157, no. 1–3, pp. 75–91, 1990.

[3]   J. F. Mckenna, S. E. D. Webb, V. Kriechbaumer, and C. Hawes, "Is Actin Filament Sliding Responsible for Endoplasmic Reticulum and Golgi Movement?," *Biorxiv*, p. 347187, 2018.

[4]   P. Boevink, K. Oparka, S. S. Cruz, B. Martin, A. Betteridge, and C. Hawes, "Stacks on tracks: The plant Golgi apparatus traffics on an actin/ER network," *Plant J.*, vol. 15, no. 3, pp. 441–447, 1998.

[5]   I. A. Sparkes, N. A. Teanby, and C. Hawes, "Truncated myosin XI tail fusions inhibit peroxisome, Golgi, and mitochondrial movement in tobacco leaf epidermal cells: A genetic tool for the next generation," *J. Exp. Bot.*, vol. 59, no. 9, pp. 2499–2512, 2008.

[6]   I. A. Sparkes, T. Ketelaar, N. C. A. De Ruijter, and C. Hawes, "Grab a golgi: Laser trapping of golgi bodies reveals in vivo interactions with the endoplasmic reticulum," *Traffic*, vol. 10, no. 5, pp. 567–571, 2009.

[7]   A. Osterrieder *et al.*, "Stacks off tracks: a role for the golgin AtCASP in plant endoplasmic reticulum-Golgi apparatus tethering," *J. Exp. Bot.*, vol. 68, no. 13, pp. 3339–3350, 2017.

[8]   C. Hawes, P. Kiviniemi, and V. Kriechbaumer, "The endoplasmic reticulum: A dynamic and well-connected organelle," *J. Integr. Plant Biol.*, vol. 57, no. 1, pp. 50–62, 2015.

[9]   D. G. Robinson, F. Brandizzi, C. Hawes, and A. Nakano, "Vesicles versus tubes: Is endoplasmic reticulum-golgi transport in plants fundamentally different from other eukaryotes?," *Plant Physiol.*, vol. 168, no. 2, pp. 393–406, 2015.

[10]   C. Hawes, "The ER/Golgi interface - Is there anything in-between?," *Front. Plant Sci.*, vol. 3, p. 73, 2012.

[11]   B. S. Donohoe *et al.*, "Cis-Golgi Cisternal Assembly and Biosynthetic Activation Occur Sequentially in Plants and Algae," *Traffic*, vol. 14, no. 5, pp. 551–567, 2013.

[12]   F. Brandizzi, "Transport from the endoplasmic reticulum to the Golgi in plants: Where are we now?," *Semin. Cell Dev. Biol.*, vol. 80, pp. 94–105, 2018.

[13]   M. K. Singh and G. Jürgens, "Specificity of plant membrane trafficking – ARFs, regulators and coat proteins," *Semin. Cell Dev. Biol.*, vol. 80, pp. 85–93, 2018.

[14]   J. Schoberer, J. Runions, H. Steinkellner, R. Strasser, C. Hawes, and A. Osterrieder, "Sequential depletion and acquisition of proteins during Golgi stack disassembly and reformation," *Traffic*, vol. 11, no. 11, pp. 1429–1444, 2010.

[15]   C. Viotti *et al.*, "The endoplasmic reticulum is the main membrane source for biogenesis of the lytic vacuole in Arabidopsis," *Plant Cell*, vol. 25, no. 9, pp. 3434–3449, 2013.

[16]   B. S. Donohoe, B. H. Kang, and L. A. Staehelin, "Identification and of characterization of COPIa- and COPIb-type vesicle classes associated with plant and algal Golgi," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 104, no. 1, pp. 163–168, 2007.

[17]   L. A. Staehelin and B. H. Kang, "Nanoscale architecture of endoplasmic reticulum export sites and of Golgi membranes as determined by electron tomography," *Plant Physiol.*, vol. 147, no. 4, pp. 1454–1468, 2008.

[18]   J. Schoberer *et al.*, "A signal motif retains Arabidopsis ER-α-mannosidase I in the cis-Golgi and prevents enhanced glycoprotein ERAD," *Nat. Commun.*, vol. 10, no. 1, p. 3701, 2019.

[19]   C. Hawes, J. Schoberer, E. Hummel, and A. Osterrieder, "Biogenesis of the plant Golgi apparatus," *Biochem. Soc. Trans.*, vol. 38, no. 3, pp. 761–767, 2010.

[20]   B. S. Glick and A. Luini, "Models for Golgi traffic: A critical assessment," *Cold Spring Harb. Perspect. Biol.*, vol. 3, no. 11, pp. 1–16, 2011.

[21]   K. Paschinger, E. Staudacher, U. Stemmer, G. Fabini, and I. B. H. Wilson, "Fucosyltransferase substrate specificity and the order of fucosylation in invertebrates," *Glycobiology*, vol. 15, no. 5, pp. 463–474, 2005.

[22]   R. Strasser, "Localization of plant N-glycan processing enzymes along the secretory pathway," *Plant Biosyst.*, vol. 143, no. 3, pp. 636–642, 2009.

7

[23] R. Léonard *et al.*, "The presence of Lewis a epitopes in Arabidopsis thaliana glycoconjugates depends on an active α4-fucosyltransferase gene," *Glycobiology*, vol. 12, no. 5, pp. 299–306, 2002.

[24] S. Kellokumpu, "Golgi pH, Ion and Redox Homeostasis: How Much Do They Really Matter?," *Front. Cell Dev. Biol.*, vol. 7, p. 93, 2019.

[25] J. Schoberer *et al.*, "The golgi localization of gnti requires a polar amino acid residue within its transmembrane domain," *Plant Physiol.*, vol. 180, no. 2, pp. 859–873, 2019.

[26] R. H. P. Wilbers *et al.*, "Production and glyco-engineering of immunomodulatory helminth glycoproteins in plants," *Sci. Rep.*, vol. 7, p. 45910, 2017.

[27] S. Kallolimath, C. Gruber, H. Steinkellner, and A. Castilho, "Promoter Choice Impacts the Efficiency of Plant Glyco-Engineering," *Biotechnol. J.*, vol. 13, no. 1, pp. 1–7, 2018.

[28] A. Hassinen *et al.*, "Functional organization of Golgi N- and O-glycosylation pathways involves pH-dependent complex formation that is impaired in cancer cells," *J. Biol. Chem.*, vol. 286, no. 44, pp. 38329–38340, 2011.

[29] S. Kellokumpu, A. Hassinen, and T. Glumoff, "Glycosyltransferase complexes in eukaryotes: Long-known, prevalent but still unrecognized," *Cell. Mol. Life Sci.*, vol. 73, no. 2, pp. 305–325, 2016.

[30] J. Schoberer, E. Liebminger, S. W. Botchway, R. Strasser, and C. Hawes, "Time-resolved fluorescence imaging reveals differential interactions of N-glycan processing enzymes across the golgi stack in planta," *Plant Physiol.*, vol. 161, no. 4, pp. 1737–1754, 2013.

[31] N. Yamaguchi and M. N. Fukuda, "Golgi retention mechanism of β-1,4-galactosyltransferase: Membrane-spanning domain-dependent homodimerization and association with α- and β-tubulins," *J. Biol. Chem.*, vol. 270, no. 20. pp. 12170–12176, 1995.

[32] H. Ohashi, T. Ohashi, H. Kajiura, R. Misaki, S. Kitamura, and K. Fujiyama, "Fucosyltransferases produce N-glycans containing core L-galactose," *Biochem. Biophys. Res. Commun.*, vol. 483, no. 1, pp. 658–663, 2017.

[33] K. Moriwaki *et al.*, "A high expression of GDP-Fucose transporter in hepatocellular carcinoma is a key factor for increases in fucosylation," *Glycobiology*, vol. 17, no. 12, pp. 1311–1320, 2007.

[34] D. Harrus, S. Kellokumpu, and T. Glumoff, "Crystal structures of eukaryote glycosyltransferases reveal biologically relevant enzyme homooligomers," *Cell. Mol. Life Sci.*, vol. 75, no. 5, pp. 833–848, 2018.

[35] F. Altmann, H. Schwihla, E. Staudacher, J. Glossl, and L. Marz, "Insect cells contain an unusual membrane-bound β-N-acetylglucosaminidase probably involved in the processing of protein N-glycans," *J. Biol. Chem.*, vol. 270, no. 29. pp. 17344–17349, 1995.

[36] R. Léonard, D. Rendić, C. Rabouille, I. B. H. Wilson, T. Préat, and F. Altmann, "The Drosophila fused lobes gene encodes an N-acetylglucosaminidase involved in N-glycan processing," *J. Biol. Chem.*, vol. 281, no. 8, pp. 4867–4875, 2006.

[37] Y. J. Shin *et al.*, "Reduced paucimannosidic N-glycan formation by suppression of a specific β-hexosaminidase from Nicotiana benthamiana," *Plant Biotechnology J.*, vol. 15, no. 2. pp. 197–206, 2017.

[38] R. Strasser *et al.*, "Enzymatic properties and subcellular localization of arabidopsis β-N-acetylhexosaminidases," *Plant Physiol.*, vol. 145, no. 1, pp. 5–16, 2007.

[39] J. Shen *et al.*, "Organelle pH in the arabidopsis endomembrane system," *Mol. Plant*, vol. 6, no. 5, pp. 1419–1437, 2013.

[40] D. Gao, M. R. Knight, A. J. Trewavas, B. Sattelmacher, and C. Plieth, "Self-reporting arabidopsis expressing pH and [Ca2+] indicators unveil ion dynamics in the cytoplasm and in the apoplast under abiotic stress," *Plant Physiol.*, vol. 134, no. 3, pp. 898–908, 2004.

[41] P. V. Jutras *et al.*, "Modulating secretory pathway pH by proton channel co-expression can increase recombinant protein stability in plants," *Biotechnol. J.*, vol. 10, no. 9, pp. 1478–1486, 2015.

[42] R. H. P. Wilbers *et al.*, "The N-glycan on Asn54 affects the atypical N-glycan composition of plant-produced interleukin-22, but does not influence its activity," *Plant Biotechnology J.*, vol. 14, no. 2, pp. 670–681, 2016.

[43] R. Strasser *et al.*, "Generation of glyco-engineered Nicotiana benthamiana for the production of monoclonal antibodies with a homogeneous human-like N-glycan structure," *Plant Biotechnology J.*, vol. 6, no. 4. pp. 392–402, 2008.

[44] D. G. Robinson, Y. Ding, and L. Jiang, "Unconventional protein secretion in plants: a critical assessment," *Protoplasma*, vol. 253, no. 1, pp. 31–43, 2016.

Addendum

Summary

Human parasitic trematodes of the genus *Schistosoma* infect 252 million people worldwide. Research to *Schistosoma mansoni*, one of the species causing schistosomiasis, has shown that *S. mansoni* secretes immunomodulatory proteins during its life cycle to influence the host immune system. These secreted proteins are often glycosylated and the glycans have shown to play an important role in immune modulation. To study the immunomodulatory properties of these proteins and the effect of their glycans, high amounts of protein are required, which cannot be isolated from the worm or its secretions. Current recombinant production systems cannot mimic the highly fucosylated N-glycans of *S. mansoni*. To mimic helminth glycosylation in the expression host, glycosyltransferases have to be co-expressed that synthesise the required glycan motifs. Although many putative *S. mansoni* fucosyltransferases (SmFucTs) are known, they are poorly characterised. Therefore, in chapter 2 and 3 we focussed on SmFucT characterisation.

Glycosylation is a process that takes place in an orderly fashion in the Golgi where sugars are detached and attached starting at the *cis*-Golgi that is closest to the endoplasmic reticulum. In sequential steps the glycan-moiety is build up in the medial- and *trans*-Golgi from where the glycoprotein is secreted or transported to another subcellular compartment. FucTs are type II membrane proteins that localise in the Golgi, where they couple fucose to glycans through interaction with GDP-fucose and specific glycan acceptors. Since, GDP-fucose and the correct glycan acceptor need to be available for synthesis of a specific glycan motif, sub-Golgi localisation of a FucT is important for its function.

In chapter 2, we investigated the sub-Golgi localisation of six α1,3-SmFucTs and four α1,6-SmFucTs by co-localisation studies with three reference Golgi markers in *Nicotiana benthamiana*. Fucosylation in *S. mansoni* has a strict order. Core α1,6-fucosylation occurs before core α1,3-fucosylation that is followed by synthesis of more complex terminally fucosylated glycan motifs, such as LeX and LDN-F. Therefore, it was not surprising to find subtle differences in Golgi localisation between the various SmFucTs. The four α1,6-SmFucTs showed localisation through the whole Golgi, whereas the α1,3-SmFucTs localised more to the *trans* side of the Golgi. All α1,3-SmFucTs, except SmFucTC, co-localised significantly more with the *trans*-Golgi marker in comparison to the *cis*- and *cis*/medial-Golgi markers. SmFucTC localised more towards the medial-Golgi in comparison to the other α1,3-SmFucTs. This suggests that SmFucTC has a core α1,3-fucosylating function whereas synthesis of more complex terminally fucosylated glycan motifs is realised by one or more of the other α1,3-SmFucTs.

In chapter 3 these suggestions were further investigated by functional characterisation in *N. benthamiana*. The SmFucTs were transiently expressed in *Nicotiana benthamiana* together with a carrier glycoprotein and other glycosyltransferases to synthesise

A

various N-glycans. A few days after infiltration, the leaves were harvested and the carrier glycoprotein was isolated from the plant leaves. Subsequently, the N-glycan composition of the carrier glycoprotein was analysed for fucose addition and/or formation of specific glycan motifs. With this method we characterised SmFucTs that add an α1,3- or α1,6-linked fucose to the N-glycan core or synthesise more complex glycan motifs, such as LeX, LDN-F or F-LDN-F. We show that *N. benthamiana* is a promising *in vivo* platform for characterisation of novel glycosyltransferases.

In chapter 4 we focussed on *N. benthamiana* as protein production platform. Plants have shown to be a promising host for the production of biopharmaceuticals. However, plant N-glycans can differ from helminth or human N-glycans. Typical plant N-glycans have an α1,3-fucosylated and β1,2-xylosylated core. The enzymes that add an α1,3-fucose to the glycan core of *N. benthamiana* N-glycans are core α1,3-FucTs (NbFucTs). So, in order to generate a plant without core α1,3-fucose on its N-glycans these NbFucTs need to be knocked out. *N. benthamiana* is an allotetraploid and therefore multiple sequence variants of one protein can be present. Studies have already focused on the production of NbFucT knockout plants by targeting all or just a few NbFucT sequence variants. However, time and effort can be saved when targeting only the NbFucT sequence variants that are active. We found ten NbFucT sequences in our *N. benthamiana* plants of which six encoded full-length enzymes. To test the activity of these six full-length sequences, they were transiently co-expressed with carrier glycoproteins in ΔXT/FT *N. benthamiana* plants (plants in which the NbFucTs and xylosyltransferases are down regulated by RNA interference). After a few days the carrier glycoproteins were isolated and their N-glycan composition was analysed for the addition of a core α1,3-fucose. This method revealed that only three of the six full-length NbFucT sequences were active. Therefore, only three of the ten sequences have to be targeted in order to generate a plant for the production of N-glycosylated proteins without core α1,3-fucose. Targeting three instead of ten sequences could save time, effort and resources.

In Chapter 5 and 6 we focussed on *N. benthamiana* as a production platform for immunomodulatory helminth glycoproteins, omega-1 and kappa-5. The capacity of helminths to establish long lasting infections can be attributed to their ability to modulate the human immune system in the direction of a modified type 2 immune response. This modified type 2 immune response is established by helminth secreted immunomodulatory components. Research to helminth infection has revealed to be inversely correlated with metabolic syndrome and type-2 diabetes. This may be explained by the fact that helminths are strong inducers of type 2 immune responses, which are possibly beneficial for metabolic homeostasis. For instance, mice on a high-fat-diet treated with soluble egg antigens (SEA) of *S. mansoni* showed improved glucose tolerance and insulin sensitivity. In order to enable *in vivo* studies to SEA components, such as omega-1 and kappa-5, high amounts of protein are required. However, isolation of *S. mansoni* SEA is laborious, costly,

A

time consuming and protein isolation from SEA results in too low amounts of protein for biological studies. Therefore, a production platform is required that produces high amounts of protein with native helminth like N-glycans.

In chapter 5 we investigated the production of omega-1 in *N. benthamiana*. Omega-1 is a ribonuclease (RNase) that is taken up by dendritic cells via its N-glycans. After uptake omega-1 induces a Th2 response via its RNase activity. Furthermore, omega-1 has shown to induce weight loss and improved metabolic homeostasis in obese mice. Since, it has been shown that the N-glycans on omega-1 are important for its function, we investigated the production of omega-1 with native N-glycans. We showed efficient production and purification of active glycosylated omega-1 in *N. benthamiana*. Plant produced omega-1 carries paucimannosidic N-glycans with core α1,3-fucose and β1,2-xylose, whereas native omega-1 carries diantennary LeX and a double fucosylated core. Synthesis of LeX and core α1,6-fucosylation does not occur in plants. Therefore, glycosyltransferases that synthesise LeX and/or add an α1,6-linked fucose to the N-glycan core were introduced in *N. benthamiana*. LeX was synthesised on omega-1 N-glycans upon co-expression of a hybrid *Danio rerio* β1,4-galactosyltransferase and a hybrid *Tetraodon nigriviridus* α1,3-FucT IXa. Both enzymes were preceded with the CTS domain of *Rattus norvegicus* α2,6-sialyltransferase (sialDrGalT and sialTnFucT9a, respectively) in order to localise the two enzymes in the *trans*-Golgi. However, sialDrGalT interfered with endogenous enzymes in the plant glycosylation pathway. Therefore, optimal sialDrGalT expression was investigated. Optimal LeX formation was achieved by expression of sialDrGalT from pHYG under the weaker constitutive *Gpa2* promoter, which reduced sialDrGalT expression. However, even in these optimal settings sialDrGalT still interferes with the plant N-acetyl-glucosaminyltransferase (GnT) II, resulting in monoantennary instead of diantennary LeX carrying N-glycans. Therefore, three exogenous GnTIIs were introduced, of which introduction of Human GnTII (HsGnTII) and Arabidopsis GnTII resulted in production of omega-1 with diantennary LeX carrying N-glycans. Next, we focussed on omega-1 core fucosylation. The native N-glycans on omega-1 are core α1,3- and α1,6-fucosylated. On plant N-glycans only core α1,3-fucose can be found, although upon co-expression of omega-1 with glycosyltransferases that synthesise diantennary LeX no core α1,3-fucose is found on its N-glycans. Therefore, two core α1,3-FucTs, SmFucTC and *N. benthamiana* FucTA1 (NbFucTA1), and one core α1,6-FucT, *Mus musculus* FucT8 (MmFucT8), were introduced. Expression of either of the three core FucTs resulted in core fucosylation of omega-1, whereas co-expression of a core α1,3- and α1,6-FucT resulted in less core fucosylation and only minor proportions of double core fucosylated N-glycans. Upon co-expression of omega-1 with sialDrGalT, sialTnFucT9a, MmFucT8 and SmFucTC, diantennary LeX carrying N-glycans with a core fucose can be synthesised on omega-1.

A

In chapter 6 we showed the production of *S. mansoni* kappa-5 in *N. benthamiana*. Kappa-5 is a protein produced in the eggs of *S. mansoni*. The function of kappa-5 is still unknown

despite the fact that it is the most abundant SEA that carries LDN and LDN-F glycan motifs on its N-glycans. LDN and LDN-F are known to be involved in granuloma formation or immunomodulation, respectively. To enable studies on the biological function of kappa-5 or its N-glycans we investigated the production of kappa-5 with native N-glycans. Also for kappa-5 high amounts of protein can be produced by and purified from *N. benthamiana*. Plant produced kappa-5 carries N-glycans with terminal GlcNAcs and an α1,3-fucosylated and β1,2-xylosylated core. Since, native kappa-5 N-glycans carry LDN and LDN-F glycan motifs that naturally do not occur in plants, the plant glycosylation pathway was engineered. We showed that LDN can be synthesised on kappa-5 N-glycans by introduction of native *Caenorhabditis elegans* β1,4-N-acetylgalactosaminyltransferase (CeGalNAcT). Subsequently, LDN-F was synthesised on kappa-5 N-glycans by introduction of sialTnFucT9a or the more effective SmFucTD. However, only monoantennary LDN or LDN-F carrying N-glycans were found on plant produced kappa-5, even upon introduction of a second antenna by HsGnTII. Native kappa-5 carries tri- or even tetra-antennary N-glycans. Therefore, production of tri and tetra-antennary carrying N-glycans on kappa-5 was investigated by co-expression with HsGnTIV and/or HsGnTV. Triantennary carrying N-glycans were synthesised on kappa-5 upon co-expression with HsGnTII, CeGalNAcT and HsGnTV, but without LDN.

Altogether, the results presented in this thesis show the suitability of plants as both a platform for characterisation of novel glycosyltransferases and the production of helminth glycoproteins carrying native N-glycan motifs. It has been shown before that plants are a promising platform for the production of human glycoproteins, but here we show new possibilities for plant molecular farming to be applied in the field of (parasite) glycobiology, parasitology and immunology.

A

Samenvatting

Wereldwijd zijn 252 miljoen mensen geïnfecteerd met humane parasitaire wormen behorende tot het geslacht *Schistosoma*. Onderzoek naar één van deze *Schistosoma* soorten, *Schistosoma mansoni*, laat zien dat tijdens infectie *Schistosoma mansoni* een scala aan stoffen uitscheidt die het immuunsysteem kunnen beïnvloeden. Dit proces wordt ook wel immunomoduleren genoemd. Veelal zijn de eiwitten in deze secreties bekleed met suikerketens, ook wel glycanen genoemd. Deze geglycosyleerde eiwitten spelen een belangrijke rol in de modulatie van het immuunsysteem. Onderzoek naar de immunomodulerende eigenschappen van deze eiwitten en hun glycanen wordt echter belemmerd door de kleine hoeveelheden eiwit die uit worm secreties geïsoleerd kunnen worden, terwijl veel grotere hoeveelheden eiwit nodig zijn. Daarom is een productiesysteem nodig dat deze eiwitten met de van nature aanwezige glycanen kan produceren, maar zo'n productiesysteem is momenteel nog niet beschikbaar. De glycanen op de eiwitten van *Schistosoma mansoni* bevatten een grote verscheidenheid aan motieven met de suiker fucose. Aangezien deze gefucosyleerde motieven van nature niet vaak voorkomen in het gebruikte productiesysteem moeten enzymen geïntroduceerd worden die de fucose op een specifieke manier koppelen, zogenaamde fucosyltransferases (FucTs). De synthese van specifiek gefucosyleerde motieven wordt echter belemmerd door de beperkte kennis over welke enzymen door de parasiet zelf worden gebruikt voor de synthese van dergelijke motieven. In hoofdstuk 2 en 3 achterhalen we welke motieven gemaakt worden door zes van de tien geselecteerde fucosyltransferases van *Schistosoma mansoni* (SmFucTs). Hiervoor maken we gebruik van de tabaksplant *Nicotiana benthamiana*, een plant die al tientallen jaren gebruikt wordt voor de productie van farmaceutische eiwitten. In hoofdstuk 4 richten we ons op de tabaksplant als productiesysteem. De glycanen op een plant-geproduceerd eiwit verschillen van de glycanen op eiwitten van een worm of mens. Een verschil is bijvoorbeeld de aanwezigheid van een gefucosyleerd motief genaamd core α1,3-fucose in planten. Om de glycanen op plant-geproduceerde eiwitten meer te laten lijken op worm of menselijke eiwitten, achterhalen we welke FucTs in de tabaksplant verantwoordelijk zijn voor het toevoegen van de core α1,3-fucose. Door vervolgens deze FucTs uit te schakelen kan een plant gemaakt worden die glycanen produceert zonder deze core α1,3-fucose. Naast het uitschakelen van FucTs is het ook mogelijk om enzymen die suikers toevoegen, zogenaamde glycosyltransferases, te introduceren om zo de plant glycanen meer te laten lijken op worm of menselijke glycanen. In hoofdstuk 5 en 6 produceren we de worm eiwitten omega-1 en kappa-5 in de tabaksplant met de van nature aanwezige motieven op hun glycanen door de introductie van verschillende glycosyltransferases. Hiermee laten we zien dat wormachtige glycanen gesynthetiseerd kunnen worden in planten.

A

Het toevoegen en afsplitsen van suikers op glycanen, het glycosylerings proces, speelt zich voornamelijk af in het Golgi complex in een cel. Hierbij zijn glycosyltransferases verantwoordelijk voor het toevoegen en glycosidases verantwoordelijk voor het afsplitsen van suikerresiduen. Dit proces gebeurt op een georganiseerde manier, omdat tegelijkertijd de juiste acceptor, glycosyltransferase of glycosidase en eventueel donor aanwezig moeten zijn. Dit betekent dat de lokalisatie van bijvoorbeeld FucTs in het Golgi complex iets kan zeggen over welk motief deze FucTs kunnen synthetiseren. Om te achterhalen welke motieven gemaakt worden door verschillende SmFucTs bekijken we in hoofdstuk 2 daarom de lokalisatie van deze enzymen in het Golgi complex. Subtiele verschillen in de Golgi lokalisatie van de SmFucTs werden geobserveerd tijdens deze studies in de tabaksplant *Nicotiana benthamiana*. Deze subtiele verschillen in Golgi lokalisatie suggereren dat bepaalde SmFucTs betrokken zijn bij de synthese van motieven die in het midden dan wel aan het einde van het Golgi complex gesynthetiseerd worden.

Deze suggesties zijn in hoofdstuk 3 verder onderzocht, door tegelijkertijd SmFucTs, andere glycosyltransferases en model eiwitten die glycanen kunnen ontvangen te introduceren in de tabaksplant. Een aantal dagen later werd de glycaan compositie op deze eiwitten geanalyseerd om zo de synthese van gefucosyleerde motieven vast te kunnen stellen. Op deze manier hebben we SmFucTs gekarakteriseerd die betrokken zijn bij core α1,3- of α1,6-fucosylering dan wel betrokken zijn bij de synthese van motieven, zoals Lewis X, LDN-F en/of F-LDN-F. Hierbij laten we zien dat de tabaksplant een veelbelovend platform is voor de karakterisatie van glycosyltransferases, zoals SmFucTs.

In hoofdstuk 4 richten we ons op de tabaksplant als productiesysteem. In de afgelopen jaren hebben planten laten zien dat ze een veelbelovend productiesysteem zijn voor farmaceutische eiwitten. Echter bevatten de glycanen op plant-geproduceerde eiwitten motieven die bijvoorbeeld niet op humane eiwitten voorkomen, zoals een core α1,3-fucose dat wordt gesynthetiseerd door specifieke FucTs in de plant. Een plant zonder core α1,3-fucose kan gegenereerd worden door het uitschakelen van deze FucTs. In onze tabaksplanten hebben we zes volledige FucT (NbFucT) genen gevonden. De activiteit van deze zes genen is daarna onderzocht door ze te introduceren met model eiwitten in transgene ΔXT/FT tabaksplanten (planten waarin de expressie van NbFucTs onderdrukt wordt door RNA interference). Deze studie liet zien dat introductie van slechts drie van de zes genen zorgde voor core α1,3-fucosylering. Deze kennis laat zien dat voor het maken van een tabaksplant zonder core a1,3-fucose, het voldoende is om enkel drie NbFucTs doelgericht uit te schakelen. Dit kan tijd, geld en moeite besparen ten opzichte van het uitschakelen van alle FucT genen.

In hoofdstuk 5 onderzoeken we of omega-1 geproduceerd kan worden in de tabaksplant met de van nature aanwezige glycanen. Omega-1 is een geglycosyleerd eiwit dat uitgescheden wordt door de eieren van *Schistosoma mansoni* en behoord tot de

A

drie meest voorkomende oplosbare ei-antigenen. Omega-1 wordt via zijn glycanen opgenomen door immuuncellen, waarin het door zijn RNA afbrekende vermogen het immuunsysteem via deze cellen kan moduleren. Hiernaast heeft men laten zien dat in een muismodel voor diabetes het toedienen van omega-1 leidt tot gewichtsverlies en een verbeterd metabolisch evenwicht. In dit hoofdstuk laten we zien dat omega-1 efficiënt in tabaksbladeren geproduceerd en gezuiverd kan worden. Plant-geproduceerd omega-1 is actief en wordt geglycosyleerd met typische plant glycanen met een core α1,3-fucose en een β1,2-xylose. De natuurlijke glycanen op omega-1 dragen het motief Lewis X, een α1,3- en α1,6-gefucosyleerde core en geen β1,2-xylose. Vandaar dat we onderzocht hebben of het glycosylerings proces aangepast kan worden, zodat omega-1 geproduceerd kan worden in de tabaksplant met de van nature aanwezige glycanen. In dit hoofdstuk laten we zien dat omega-1 momenteel in planten geproduceerd kan worden met glycanen met twee Lewis X motieven en een enkel gefucosyleerde core.

In hoofdstuk 6 laten we de productie zien van kappa-5 in de tabaksplant. Naast omega-1 behoort ook kappa-5 tot de drie meest voorkomende oplosbare ei-antigenen van *Schistosoma mansoni*. De functie van kappa-5 is echter nog onbekend, alhoewel bekend is dat de LDN en LDN-F motieven op de glycanen van kappa-5 betrokken zijn bij respectievelijk granuloom vorming of immuunmodulatie. Aangezien kappa-5 het meest voorkomende oplosbare ei-antigeen is dat deze twee motieven op de glycanen draagt, maakt dit kappa-5 een interessant eiwit om te bestuderen. Om de functie van kappa-5 en zijn glycanen te kunnen bepalen zijn grotere hoeveelheden eiwit nodig dan momenteel uit de eieren gehaald kan worden. In dit hoofdstuk laten we zien dat ook kappa-5 in tabaksbladeren geproduceerd en gezuiverd kan worden. Plant-geproduceerd kappa-5 draagt typische plant glycanen met core α1,3-fucose en β1,2-xylose. Aangezien de glycanen op kappa-5 van nature LDN en LDN-F motieven dragen die van nature niet in planten voorkomen zijn deze motieven in de plant gesynthetiseerd door introductie van verschillende glycosyltransferases. Hiermee laten we zien dat de motieven LDN en LDN-F succesvol geïntroduceerd kunnen worden in planten. Echter is de natuurlijke glycaan op kappa-5 nog complexer. Kort samengevat, kappa-5 kan in planten geproduceerd worden met verschillende van nature aanwezig motieven, maar voor het produceren van kappa-5 met zijn volledig natuurlijke glycaan structuur is meer onderzoek nodig.

De gepresenteerde resultaten in dit proefschrift tonen aan dat planten gebruikt kunnen worden voor de karakterisatie van glycosyltransferases. Hiernaast wordt duidelijk dat planten geschikt zijn als productiesysteem voor geglycosyleerde eiwitten van de humane parasitaire worm *Schistosoma mansoni* met de van nature aanwezige motieven op hun glycanen. Voorheen heeft men laten zien dat planten geschikt zijn voor de productie van geglycosyleerde menselijke eiwitten, echter laat dit proefschrift zien dat eiwitproductie in planten veel breder inzetbaar is en ook toegepast kan worden in diverse onderzoeksvelden, zoals (parasitaire) glycobiologie, parasitologie en immunologie.

A

Acknowledgements

I love to read or listen to stories, but I don't like to write or tell them. Therefore, these acknowledgements will not be a story, but a letter partly in English and partly in Dutch to my family to express my gratitude for their support, chats and the nice time I had.

To my family,

This letter I would like to start by thanking my **Nema**-family. Figuratively speaking that's what Nema is to me a family with many different members and many different opinions, but above all a supporting group. Thank you all for the nice work environment and atmosphere at Nema, with social gatherings like the yearly bbq, Sinterklaas, WE-day and Christmas lunch or the organisation of events such as laser gaming and dinner with sushi or hot pot. But also, of course the many celebrations with drinks and too much cake. Personally, I will never forget how people were sprinting through Radix during a game of Monopoly. Thank you also for the positive vibe, with smiles, "good morning" / "goeiemorgen" or enthusiastic waves in the morning and the nice breaks during the day.

Within Nema, I would like to thank the two people who made this thesis possible, my supervisors **Arjen** and **Ruud**. **Arjen**, without you I would never had the opportunity to start my PhD at Nema. I still recall the day that I almost settled for a job at the Utrecht University and you emailed me that a PhD position would become available at Nema. I never doubted my choice and still love the research I am doing. During my PhD you gave me the freedom to develop myself and pursue my ideas. Furthermore, when I asked for your help I knew I could count on you and even in busy periods you would help me as best you could. Arjen, we had many discussions over the years and I value your opinion, but on one thing we will never agree, because when there are "bears on the way" you cannot drive or walk through them. **Ruud**, thank you for your guidance and advice. You were always there for me when I needed you on the lab or during the writing period. Your positive and optimistic attituded balanced my more careful attitude and kept me from a (big) PhD dip. I loved our discussion on all the different research lines, new ideas and student projects. I truly hope you can pursue all or at least many of them over the coming years and I am very curious to the results.

Furthermore, I would like to thank **Lotte** and **Koen** from the LMA-group. **Lotte**, when you left LMA and eventually Nema this was a shock for me. In the time you were there you learned me a lot and guided me as a starting PhD. Later in my PhD there were times that I wondered how it would have been with you next to me against "the guys". **Koen**, in the last part of my PhD you joined the LMA work discussions and office. I have to say before I started my PhD I had my doubts about us working together. However, at the end of my PhD I am happy you were there. Thank you for the not necessarily work related

A

conversations that distracted me in a good way, your cheerful attitude and engagement in my well-being. As well thank you for the introduction to Dominion, good books and bouldering, although I will never share your love for the latter.

Next, I would like to thank **Jaap** and the PI's **Aska**, **Geert**, **Hans** and **Jan** for managing the Nematology family with all it's different individuals, supporting one another when needed and interested in all research done within Nematology. Thank you for your support and open mind by setting up for instance the PoP meeting, PhD meeting and the writing retreat and thank you for thinking along on how to improve the "PhD-track" in general. Next, to the PI's, I would like to thank the Nema Postdocs present during my time as PhD. **Erik**, **Jose**, **Mark** and **Martijn**, you are each specialist in your field of study and were always open to questions whenever our research lines crossed, thanks for your help.

A chair group cannot run smoothly without proper organisation, support and administration. Therefore, I would like to thank **Christel**, **Lisette** and **Manouk** for all their help in front and behind the scenes. I am well aware that you do much more than we know and I am grateful for all the work you take from our hands. Furthermore, thank you **Christel** for your time as desk neighbour and thank you **Manouk** for not wearing the same clothes as me to Nema, although I know you have them in your closet.

Also inside the lab proper organisation, support and administration is required for smoothly running of a chair group. Therefore, I would like to thank our lab managers and technicians **Casper**, **Debbie**, **Hein**, **Joost**, **Rikus** and **Sven**. Thanks to your excellent work, I could focus on my experiments without worrying about the equipment or supplies. **Debbie** and **Rikus**, both of you I would like to thank specifically, because you are experts on the techniques I used and were never too busy to help or answer my questions. Also, I would like to thank **Hein** for taking over these tasks after you both left. Next to our lab managers, I would like to thank **Bert**, **Bertus**, **Henk** and **Erik** from Unifarm, who take care of the plants that are essential for our experiments and make sure there is a new batch of plants every week.

Then there is the organisation of among others the thesis rings, open days and various courses by **Jet** and **Liesbeth**, important for the chair group, students and their supervisors. Jet and Liesbeth thanks for your work and effort, which influenced my work indirectly. However, above all I would like to thank you as colleagues sharing my enthusiasm for Japan or tennis and reminding me of subscribing for tennis events so I would not miss the deadline.

Of course I also would like to thank my fellow PhDs. **Amalia**, **Casper Q**, **Mark** and **Sonja** you were already more senior PhDs when I started (at least in my eyes). Thank you for all the valuable information you shared with me as a younger PhD on how to write a paper, how to work with R, how to make figures and the discussions we had on what

A

to do after finishing a PhD. **Ava**, **Katharina**, **Koen**, **Lisa**, **Octavina**, **Paula** and **Yiru** we started our PhD around the same time and were soon followed by **Jaap-Jan**, **Matthijs**, **Qi** and **Yuqing**. Together with **Amalia** and **Sonja** and the recent addition of **Joris**, **Sara** and **Nina** this diverse group of PhD students formed a nice little family within Nema. I valued our PhD meetings and dinners, where I got to know you a little better, could learn from your mistakes, struggles and found support to conquer my own issues. Thanks for your help and enthusiasm, I wish you all the best and hope you will continue organising the "awesome and efficient writing retreat".

A special word to the PhDs with whom I spend the most time. Thanks to **Ava**, **Koen**, **Lisa** and **Yiru**, who made the PhD trip to Canterbury a success. I am happy our trip was not cancelled, because of the fallen snowflakes or my old lady back. **Ava**, thank you for the joy you bring in conversations, talking with you lift my spirits. **Lisa**, thank you for your ever-present enthusiasm and drive to make things better. **Octavina**, for a long time we were the only PhDs at the LMA-lab. I value our talks in the lab about our experiments, which later on switch more to office or corridor talks about the writing process. Also, thank you for sharing my enthusiasm for confocal imaging. **Amalia**, **Paula** and **Yiru** it was nice to have you and your booklets as an example and source of information for all the small things related to finishing a PhD, thanks for your help.

During the year (except maybe during the holidays) half of Nema exists of students. Therefore, I also would like to thank the Nema students, who helped to create the enthusiastic and lively environment at Nema. Specifically, I would like to thank the LMA students, with whom I had the pleasure to work more closely. The students involved in my projects **Kasper**, **Douwe**, **Ika**, **Martijn**, **Billy**, **Ilse**, **Lonneke**, **Annet**, **Ben**, **Dayoung**, **Stijn** and **Jorge**, thank you for tagging along on my PhD, helping me with my projects, giving me an extra discussion partner and broaden my view on life. Also, a thanks to the various student assistants, **Lizeth**, **Martijn**, **Nicolo**, **Annet**, **Ben**, **Imie**, **Myrna**, **Tatiana**, **Elena** and **Jorge**, for their assistance on the lab or during the course Plants and Health. It was nice to work together and see you grow.

Next to the people of Nematology I would like to thank a few people outside Nematology. From the LUMC **Bruno**, **Linh**, **Maria**, **Patrick** and **Ron**, who made the Help-T2D project possible and showed me the impact of our work on human health. **Ron**, unfortunately I was not allowed to put your name as my external supervisor at the beginning of this thesis. Therefore, here I would like thank you for your valuable feedback on my thesis, your answers to many of my questions about glycan analysis and the discussion we had on glycobiology. I know you have a busy schedule and therefore I am even more grateful for the time you reserved for me. Thanks to **Linh** for your help with the MALDI-TOF MS analysis, your answers to my questions and taking me along during lunch breaks at the LUMC. Thanks to **Patrick** for the valuable information on SEA extraction and isolation of omega-1, kappa-5 and ipse/alpha-1.

A

en te genieten van rust, regelmaat en familie. **Hanneke** en **Ton**, bedankt voor jullie steun, interesse en de vele vragen om mijn onderzoek beter te kunnen begrijpen. Natuurlijk ook alle andere familieleden, die ik hier niet allemaal bij naam noem, bedankt voor jullie steun, vertrouwen en interesse, vooral vlak voor de deadline hielpen jullie appjes, kaartjes en telefoontjes mij er doorheen. Als allerlaatste wil ik **Jurriaan** bedanken. Bedankt voor je luisterend oor, je liefde en steun door de jaren heen. Vlak voor de deadline was ik zonder jou waarschijnlijk omgekomen van de honger, dan wel had ik moeten overleven op noedels of droge rijst. Jur, zonder jou had ik dit niet kunnen doen, je bent mijn steun en toeverlaat.

All the best,

Kim

A

About the author

Kim van Noort was born on the 29th of January 1990 in Zuidland, Bernisse, Zuid-Holland, the Netherlands. She graduated from the high school Penta college CSG Angelus Merula in Spijkenisse in 2008. Because of her interest in biology, chemistry and physics, she started in 2008 with the BSc program Molecular Life Science at Wageningen University (Wageningen, the Netherlands), in which these fields are studied on molecular level. During her BSc she specialised herself in Plant Biotechnology. After finishing her BSc thesis in the group of Dr D. Weijers on plant development, she stopped her study for a year to take a seat in the board of SWU Thymos and organised sport events for students. Next, she continued her studies at Wageningen University with the master Plant Biotechnology. During the master she developed an interest in the production of pharmaceutical proteins in plants and performed her MSc thesis in the group of Dr A. Schots. Thanks to her MSc thesis she was introduced to the fascinating world of glycosylation by her supervisors Dr R.H.P. Wilbers and Dr L.B. Westerhof. As internship she decided to work four months at the UMC Utrecht (Utrecht, the Netherlands) to work on allergens in the groups of H.G. Otten and Prof. A.C. Knulst. In 2014 she graduated as Master of Science in Plant Biotechnology with a specialisation in "plants for human and animal health".

In July 2015 she was appointed as a PhD student and joined the group of Dr A. Schots at the Laboratory of Nematology. She continued to do research to plant-based production and focussed on fucosyltransferase activity for the production of helminth glycoproteins in plants. For this research she visited the group of Prof. C. Hawes at the Oxford Brookes University (Oxford, United Kingdom) for confocal studies to fucosyltransferase Golgi localisation and the group of Prof. C.H. Hokke at the LUMC (Leiden, the Netherlands) for N-glycan analysis with MALDI-TOF MS. During her PhD she was a member of the EPS PhD council and the Wageningen PhD council and organised among others the EPS Get2Gether of 2017 and 2018. Furthermore, she supervised a lot of BSc and MSc students during their thesis and co-organised/supervised the practical or group work of several courses such as Plants and Health, Immunotechnology and Introduction of Environmental Sciences. On the 27th of March 2020 she will be given the opportunity to defend her PhD thesis. Until July 2020, she will continue her research as Postdoc in the group of Dr A. Schots.

A

Education Statement of the Graduate School Experimental Plant Sciences

Issued to: **Kim van Noort**
Date: **27 March 2020**
Group: **Laboratory of Nematology**
University: **Wageningen University**

1) Start-Up Phase	*date*	*cp*
► **First presentation of your project**		
Characterisation of *Schistosoma mansoni* fucosyltransferases for glyco-engineering in planta	14 Jan 2016	1.5
► **Writing or rewriting a project proposal**		
► **Writing a review or book chapter**		
► **MSc courses**		
Subtotal Start-Up Phase		1.5

2) Scientific Exposure	*date*	*cp*
► **EPS PhD student days**		
Symposium & Workshops: EPS Get2Gether 2016, Soest, the Netherlands	28-29 Jan 2016	0.6
Symposium & Workshops: EPS Get2Gether 2017, Soest, the Netherlands	09-10 Feb 2017	0.6
Symposium & Workshops: EPS Get2Gether 2018, Soest, the Netherlands	15-16 Feb 2018	0.6
► **EPS theme symposia**		
Symposium: Theme 4 Genome Biology, Wageningen, the Netherlands	16 Dec 2016	0.3
Symposium: Theme 2 Interactions between Plants and Biotic Agents, Wageningen, the Netherlands	23 Jan 2017	0.3
► **Lunteren Days and other national platforms**		
Symposium: 27th Joint Glyobiology Meeting , Nijmegen, the Netherlands	17-18 Oct 2016	0.5
Symposium: 28th Joint Glycobiology Meeting, Aachen, Germany	17-19 Sep 2017	0.5
Consortium meeting: 2nd biannual ZonMW TOP grant consortium meeting, Leiden, the Netherlands	15 Dec 2015	0.3
Consortium meeting: 3rd biannual ZonMW TOP grant consortium meeting, Wageningen, the Netherlands	24 May 2016	0.2
Consortium meeting: 4th biannual ZonMW TOP grant consortium meeting, Wageningen, the Netherlands	29 Nov 2016	0.3
Consortium meeting: 5th biannual ZonMW TOP grant consortium meeting, Leiden, the Netherlands	4 Jul 2017	0.3
Consortium meeting: 6th biannual ZonMW TOP grant consortium meeting, Wageningen, the Netherlands	28 Mar 2018	0.3
► **Seminars (series), workshops and symposia**		
Seminar: Prof A. MacDonald, Dendritic cells: central players in coordination of Type 2 inflammation	15 Dec 2015	0.1
Seminar: Prof. H. Thordal-Christensen, EPS Flying seminar Membrane trafficking in plant cells attacked by powdery mildew fungi	12 Dec 2016	0.1
Seminar: Prof. J.A. Doudna and Ir. E.R. Westra, Rewriting our genes?	30 Sep 2016	0.2

A

Seminar: Ir. H.J.P. van der Zande, Immune regulation of metabolic homeostasis by plant derived helminth proteins	11 May 2017	0.1
Seminar: Prof. M.J. Cann, The immune receptor Rx1 remodels chromatin and chromatin interactors in immunity	11 Jul 2017	0.1
Seminar: Dr. Ir. N. Briggs, Trichuriasis, vaccine discovery	19 Oct 2017	0.1
Symposium: Protein Purification, Wageningen, the Netherlands	09 Nov 2017	0.2
Seminar: Prof. M. Bezanilla, Cytoskeletal crosstalk impacts cell shape and development	06 Nov 2017	0.1
Symposium: Novel Developments in Fluorescence Microscopy, Wageningen, the Netherlands	20 Nov 2017	0.2
Seminar: Ir. N. van 't Wout Hofland, The origin and evolution of the vascular regulatory dimer TMO5/LHW	12 Apr 2018	0.1
Seminar: Dr. Ir. W. van de Veen, The role of B cells in the regulation of allergic immune responses	02 Nov 2018	0.1
Seminar: Dr. S. Eves-van den Akker, Plant Immunity and development-altering "toolbox"of parasitic nematodes	13 Feb 2019	0.1
Seminar: Ir. C. Pain, Identifying proteins involved in Er cisternae formation	07 Aug 2019	0.1
Seminar: Ir. B. Andov, Investigating protein-protein interactions between arabidopsis Mid-Sun proteins and the transcription factor maMYB	14 Aug 2019	0.1
▶ **Seminar plus**		
▶ **International symposia and congresses**		
Congress: 2nd ISPMF (International Society for Plant Molecular Farming), Gent, Belgium	25-27 May 2016	0.9
Congress: 5th Plant Genomics and Gene Editing Congress, Amsterdam, the Netherlands	16-17 Mar 2017	0.6
Congress: Molecular and Cellular Biology of Helminths XI, Hydra, Greece	03-08 Sep 2017	1.3
Congress: 3rd ISPMF (International Society for Plant Molecular Farming), Helsinki, Finland	11-13 Jun 2018	0.9
▶ **Presentations**		
Poster: Characterization of *Schistosoma mansoni* fucosyltransferases for glyco-engineering of 'native' helminth N-glycan structures in planta, 2nd ISPMF	25 May 2016	1.0
Poster: Characterization of *Schistosoma mansoni* fucosyltransferases for glyco-engineering of 'native' helminth N-glycan structures in planta, 14th Summer Course Glycoscience	13 Jun 2016	0.0
Poster: Characterization of *Schistosoma mansoni* fucosyltransferases for glyco-engineering of 'native' helminth N-glycan structures in planta, 27th Joint Glycobiology Meeting	17-18 Oct 2016	1.0
Poster: Characterisation of *Schistosoma mansoni* fucosyltransferases for glyco-engineering of 'native' helminth N-glycan structures in planta, 3rd ISPMF	11-14 Jun 2018	1.0
Talk: Characterization of *Schistosoma mansoni* fucosyltransferases for glyco-engineering of 'native' helminth N-glycan structures in planta, 2nd ISPMF	27 May 2016	1.0

A

	date	cp
Talk: Characterization of *Schistosoma mansoni* fucosyltransferases for glyco-engineering of 'native' helminth N-glycan structures in planta, 27th Joint Glycobiology Meeting	17-18 Oct 2016	1.0
Talk: Characterization of *Schistosoma mansoni* fucosyltransferases in plants, 4th biannual ZonMW TOP grant consortium meeting	29 Nov 2016	1.0
Talk: Characterization of *Schistosoma mansoni* fucosyltransferases in plants, 5th biannual ZonMW TOP grant consortium meeting	4 Jul 2017	1.0
Talk: Using plants to characterise *Schistosoma mansoni* fucosyltransferases, Molecular and Cellular Biology of Helminths XI	03-08 Sep 2017	1.0
Talk: Using plants to characterise *Schistosoma mansoni* fucosyltransferases , 28th Joint Glycobiology Meeting	17-19 Sep 2017	1.0
Talk: Characterization of *Schistosoma mansoni* fucosyltransferases for glyco-engineering of 'native' helminth N-glycan structures in planta, One day Symposium Canterbury	02 Mar 2018	1.0
Talk: Characterisation of *Schistosoma mansoni* fucosyltransferases for glyco-engineering of 'native' helminth N-glycan structures in planta, Group meeting Oxford Brookes University	29 Aug 2018	1.0
► **IAB interview**		
► **Excursions**		
EPS company visit KeyGene	12 Oct 2017	0.2
PhD trip for Nematology PhD students to Biomolecular Research Group, Canterbury Christ Church University, Canterbury, England	02 Mar 2018	0.3
Subtotal Scientific Exposure		**21.7**

3) In-Depth Studies

	date	*cp*
► **Advanced scientific courses & workshops**		
Course: Bioinformatics - a User's Approach, Wageningen, the Netherlands	24-28 Aug 2015	1.5
Course: 14th Summer Course Glycoscience, Groningen, the Netherlands	12-16 Jun 2016	2.0
Course: The Power of RNA-seq, Wageningen, the Netherlands	10-12 Feb 2016	0.8
► **Journal club**		
Journal club, hosted by the LMA-group, Nematology, Wageningen, the Netherlands	2015-2016	0.2
► **Individual research training**		
Confocal Microscopy at Oxford Brookes University, lab of Prof. C. Hawes	01-31 Aug 2018	2.5
MALDI-TOF Mass spectrometry at University Medical Centre Leiden, lab of Prof. C.H. Hokke	04 -10 Jan 2019	0.5
Subtotal In-Depth Studies		**7.5**

4) Personal Development

	date	*cp*
► **General skill training courses**		
Course: Scientific Writing, Wageningen, the Netherlands	14 Mar - 16 May 2017	1.8
Course: Career Assessment, Wageningen, the Netherlands	Nov 2018	0.3
Course: Competence Assesment, Wageningen, the Netherlands	05 Nov 2015	0.3

A

Course: Scientific Artwork, Vector graphics and images, Wageningen, the Netherlands	19-20 Mar 2018	0.6
Symposium: Open Access, Wageningen, the Netherlands	21 Jun 2016	0.2
Workshop: Exploring the job market outside of academia: workshop for PhD candidates, Wageningen, the Netherlands	07 Feb 2017	0.1
Workshops: WGS PhD Workshop Carousel, Wageningen, the Netherlands	07 Apr 2017	0.3
Workshop: 'R&O gesprek' (Performance and Development interview), Wageningen, the Netherlands	04 Oct 2018	0.2
Workshop: 'Big 5 for life' of Young WUR, M. Zonderland-Thomassen, Wageningen, the Netherlands	28 Mar 2019	0.1
► **Organisation of meetings, PhD courses or outreach activities**		
Organisation of the EPS "Get2Gether 2017, Soest, the Netherlands (Location)	2016-2017	1.1
Assistance at the Fascination of Plants Day 2017	20 May 2017	0.3
Organisation of the EPS "Get2Gether 2018, Soest, the Netherlands (Speakers)	2017-2018	1.1
Organisation and Hosting of Nematology student symposium, Wageningen, the Netherlands	18 Oct 2018	0.3
Organisation of PhD trip for Nematology PhD students, Canterbury, England	Feb 2018	0.2
► **Membership of EPS PhD Council**		
EPS PhD Council member & Wageningen PhD Council member (Sep 2017-Jun 2018)	Mar 2016 - Mar 2018	1.4
Subtotal Personal Development		8.3
TOTAL NUMBER OF CREDIT POINTS*		**39.0**

Herewith the Graduate School declares that the PhD candidate has complied with the educational requirements set by the Educational Committee of EPS with a minimum total of 30 ECTS credits.

* A credit represents a normative study load of 28 hours of study.

A

A PhD is like walking in the mountains
every step gives you a new perspective