Oxford Brookes University

MSc by Research Thesis

Modification of a plant-based expression system to produce Lysosomal Acid Lipase (LAL) in plant cells using genetic modification of endogenous plant N-glycosylation

---

Author

Alastair McGinness

Supervisors

Dr. Verena Kriechbaumer

Prof. David Evans

Dr. David Gervais

Thesis submitted in partial fulfilment of the requirements of the award of MSc by Research

Department of Biological and Medical Sciences,

Faculty of Health and Life Sciences,

Oxford Brookes University

In collaboration with Porton Biopharma Ltd

August 2020

Acknowledgements

Throughout this project I have been fortunate enough to have constant discussion, support and insight from both of my supervisors at Oxford Brookes, Dr Verena Kriechbaumer and Prof. David Evans, alongside the excellent community that is found in the Sinclair Annex at Oxford Brookes. In particular I would like to thank both Dr. Kriechbaumer and Professor Evans for their endless patience during the writing process of this final thesis, as well as their expert guidance during the management of the entire project, that has ultimately led to my continuing the pursuit of research.

I would also like to thank all my colleagues within the plant endomembrane group at Oxford Brookes for their support, insight and near inexhaustible supply of biscuits that made this lab such a joy to be part of. Their camaraderie and insight into the world of research has been everything a new postgraduate student could ask for. I would like to direct particular thanks to Charlotte Pain and Tattiana Spatola Rossi for bearing the brunt of my endless questions within the lab itself. Likewise, thanks go out to our ever-patient lab technician Alice Wingfield, both for her professional capacity and as part of the Annex community.

In addition, I would like to thank all my friends and family for their unflinching support, and continuing willingness to listen to me talk about plants. Special credit goes to Stefan Wojcik, without whom I would never have discovered this excellent research group, and who somehow both lived and worked with me for a year without serious complaint.

Thanks go to Dr. David Gervais for both his consultation on the course of the project, alongside the associated support of Porton Biopharma Ltd that made this project possible.

Abstract

As plant cells are capable of complex N-glycosylation they are a valid expression system to produce safe and correctly glycosylated human proteins. To modify plant glycosylation patterns towards fully human patterns, the plant glycosylation machinery requires substitution of a series of plant enzymes with human ones.

Here we show that the mammalian glycosylation enzymes B4GALT1, ST6GAL, GNTIV and GNTV can be expressed in plant cells, and localise to specific cisternae of the Golgi body. Only ST6GAL and GNTV localise in the *medial/trans*-Golgi but fusion of a sialyltransferase-derived N-terminal signal peptide is capable of targeting B4GALT1 to the *medial/trans*-Golgi.

As a proof of concept for human therapeutic enzyme production in plant cells, Lysosomal Acid Lipase (LAL) was expressed in tobacco leaf cells. Here LAL was shown to localise to *medial/trans*-Golgi cisternae. Protein expression of human LAL was enhanced by removal of an endogenous signal peptide associated with mammalian intracellular targeting that possessed no functional role in the therapeutic efficacy of LAL.

1. Introduction

1.1 Plant-based production systems for therapeutics

Vector selection for the production of biologics, especially for human therapeutics, presents challenges both scientific and economic in nature (Sack et al., 2015). The biological parameters of any system must be capable of supporting the intended application of the end product. Large quantities of simple proteins for industrial use may be an ideal product for a bacterial cell suspension system, but the same system would not be capable of producing complex human proteins, or be economically viable to do so (Chen et al., 2015). The innate biological properties of existing systems, such as endogenous contaminants, capability for post-translational modification and glycomodification are examples of inflexible determinants of a system's suitability. Economic factors of a production system may be more generalized, such as the ability to produce a uniform and stable product in a manner that is cost-effective, and affordably scalable enough to support the abundance required, but are equally vital to the viability of a production system's use (Chen et al., 2016).

Plant systems designed for the production of recombinant protein have been a fast-growing development, and with good reason. Major benefits of plant-based systems are low cost production, fast and sustainable scalability, reliable stable or transient expression, and flexibility of genetic modification (Chen et al., 2016). Innate tolerance in mammals of plant proteins, through constant contact and diet, makes plant-produced therapeutics highly desirable for their lack of a substantial immunogenic profile, when compared to more traditional hosts such as bacteria (Nagels et al., 2011). Yeast and bacterial systems can generate high risk immunogenic contaminants such as bacterial lipopolysaccharides (LPS) and terminally mannosylated glycan structures, which are a significant risk factor for therapeutic applications as they can trigger detrimental immune responses in patients (Gomes et al., 2016).

In comparison to the majority of existing drug production systems that require costly equipment, facilities and consumables, plant systems are relatively low-cost (Tekoah et al., 2015). The convenient ability to quickly generate small yields of protein from transiently-transformed mature plants facilitates niche applications, such as a rapid vaccine production in response to an outbreak, which combined with low initial investment costs makes them an attractive proposition in the developing world. Plant systems have a greater potential compared to many existing

systems to shift closer to a carbon-neutral production platform in an ever ecologically conscious age (Buyel, 2019). Plant systems cannot easily be contaminated with mammalian pathogens or support their growth as some bacterial, yeast and mammalian cell systems can. This is an important advantage to create products safe for human therapeutic applications (Tekoah et al., 2015).

Recent history has seen plant-based platforms for the production of recombinant proteins begin to challenge existing, well-developed and trusted mammalian and bacterial systems, as industry begins to see the economic and biological advantages of biopharming (Moon et al., 2019). The therapeutic potential for plant-manufactured proteins has already seen real-world demonstrations of effectiveness, from the pandemic response capability of tobacco-manufactured vaccines for the 2014 Ebola outbreak (Sack et al., 2015), to the first plant-manufactured enzyme replacement therapy of recombinant Taliglucerase Alfa therapy, utilized for Gaucher's disease and produced in a suspended carrot cell culture (Gregory et al., 2014; Rup et al., 2017). But in general, plant systems have yet to come into their own as a serious competitor for existing expression systems which dominate the current production market, which are chiefly yeast, bacterial and mammalian cell lines (Gomes et al., 2016).

Use of Agrobacterium-mediated transformation of *Nicotiana tabacum* was chosen due to the technical demands of the project regarding glycomodification, and practical considerations of the research. There is a strong basis in the literature for the use of tobacco plants for complex glycomodification (Strasser et al., 2008), with ability to perform N-linked glycosylation similar to mammalian cells being an essential requirement for the purposes of this study. Use of well-documented crop species as a vector is useful for the production of recombinant protein for therapeutic uses, due to known growth conditions and the molecular toolkit available (Moon et al., 2019). Agrobacterium-mediated transformation of mature plant tissue was useful as a high-throughput method to rapidly test expression of multiple constructs and fluorophores within a relatively short time frame via confocal analysis, while typically promising relatively high expression levels sufficient for gram yields of product (Moon et al., 2019). This advantage of utilizing *Nicotiana tabacum* in various forms allows the possibility to investigate production of recombinant proteins from transiently-transformed plants, such as in this work, but also in stably transformed plants and cell lines, including the possibility of manipulating this system in future to better suit industrially scalable production methods (Xu et al 2012).

1.2 Lysosomal Acid Lipase: Characteristics and deficiency

LAL is an essential human lysosomal lipid metabolism enzyme responsible for the hydrolysis of cholesterol esters (CEs) and triglycerides (TGs) into fatty acids (FAs) within lysosomes (Du et al., 2008, Thelwall et al., 2013). The deficiency of endogenous LAL supply will lead to dysfunction of lipid metabolism and a spectrum of disease phenotypes resulting from chronic lipid accumulation. LAL deficiency (LAL-D) is considered an rare autosomal recessive condition caused by alterations in the *LIPA* gene encoding LAL (Kim et al, 2016) and occurring in a rate of less than 1/100,000 newborns (Aguisanda et al., 2017). The spectrum of LAL-D pathology is multisystem and largely dependent upon endogenous LAL activity. Absolute at-birth deficiency is known as Wolman's disease, and is typically neonatal lethal, with a life expectancy of approximately 6 months to one year, whilst milder forms with some residual activity are broadly termed Cholesterol Ester Storage Disease (CESD) (Balwani et al., 2013). LAL dysfunction is considered a life-limiting condition with a severely impaired at-birth life expectancy (Strebinger et al., 2019). At time of writing, the only effective treatment option for LAL dysfunction is enzyme replacement therapy (ERT), consisting of typically bi-weekly infusions of Sebelipase Alfa© (produced by Alexion Pharmaceuticals), which is a recombinant human LAL analogue produced in genetically modified chicken eggs (Su et al., 2016).

ERT requires a final product that is capable of meeting nessecary demands: sufficient quantity, with correct structure and activity, pure and as close to the human analogue as possible. This is to ensure a threshold of replacement enzyme activity is reached, but with minimal off target effects occurring during repeated exposure, as has been noted in existing ERT options (Kizhner et al., 2015). Existing mammalian, bacterial and yeast systems could each match these criteria to some extent, but not as completely as desired, especially where off target effects are concerned. Current ERTs utilized for lysosomal storage diseases have recorded cases of allergic reactions which could prevent patients from receiving effective treatment options (Zimran et al., 2011), and the current enzyme replacement Sebelipase alfa has been identified as having more commonplace adverse reactions in infants in particular (Su et al., 2016).

A major issue with a plant production system producing enzyme for ERT in humans is that the end product should be as analogous to the human enzyme as possible, including post-translational modifications, which may not be possible in a wild-type plant. Of specific concern to this project was the arrangement of N-linked glycans on the synthesized LAL, which are attached via N-linked glycosylation enzymes residing in the last step of protein synthesis within the Golgi apparatus. N-Linked glycosylation refers to the stepwise enzymatic attachment of sugar moieties onto specific asparagine residues residing on synthesized protein (Bosch et al., 2013), and have a strong impact on protein localisation and function, being the major posttranslational modification in eukaryotic cells (Strasser et al., 2014). In plant Golgi bodies, this glycosylation machinery differs from the mammalian equivalent, and so could not generate a “complete” mammalian glycosylation profile (Schoberer et al., 2018). In order to overcome this, mammalian post-translational glycosylation enzymes would need to be expressed in plants, within the appropriate Golgi cisternae, so as to attach mammalian-type glycans onto the synthesized protein and generate a mammalian glycosylation profile on the functional plant-derived LAL.

The differences between plant and mammalian N-glycosylation machinery could hinder the expression of viable mammalian proteins. For example, absence of terminal sialic acid significantly reduces plasma half-life in subject testing, and is essential in many drugs for optimal therapeutic potency, as absence would significantly diminish therapeutic effectiveness of treatment (Bosch et al., 2013). These issues can be mitigated by the targeted knock-in of mammalian glycosylation enzyme, as has been observed in previous studies, to add essential human N-glycans to LAL, and so improve therapeutic viability (Strasser, 2014). Effectiveness of enzyme knock-in strategies will be dependent on correct localisation within the plant Golgi stack, and the ability of induced genes to compete with endogenous plant enzymes for production.

1.3 The plant Golgi body and N-linked glycosylation

The plant Golgi apparatus is a significant site of both protein synthesis and glycomodification within the plant cell, and so a focal point for production of recombinant proteins (Gomord et al., 2010; Schoberer et al., 2018). In contrast to the immobile, perinuclear Golgi body found in mammalian cells, Golgi bodies in higher plants are discrete stacks that can number several hundred per cell, and are extremely mobile throughout the cytoplasm (for review see Robinson et al., 2020). The basic structure of the individual Golgi body (Figure 1) is a stack of disc-like cisternae characterized by ordering within the stack, the *cis*-cisternae as the entry point for cargo exiting the ER, the *medial*-cisternae and the *trans*-cisternae. From the *trans*-cisternae product is delivered to the Trans Golgi Network (TGN) mediating intracellular delivery of ER products (for review see Robinson et al., 2020).

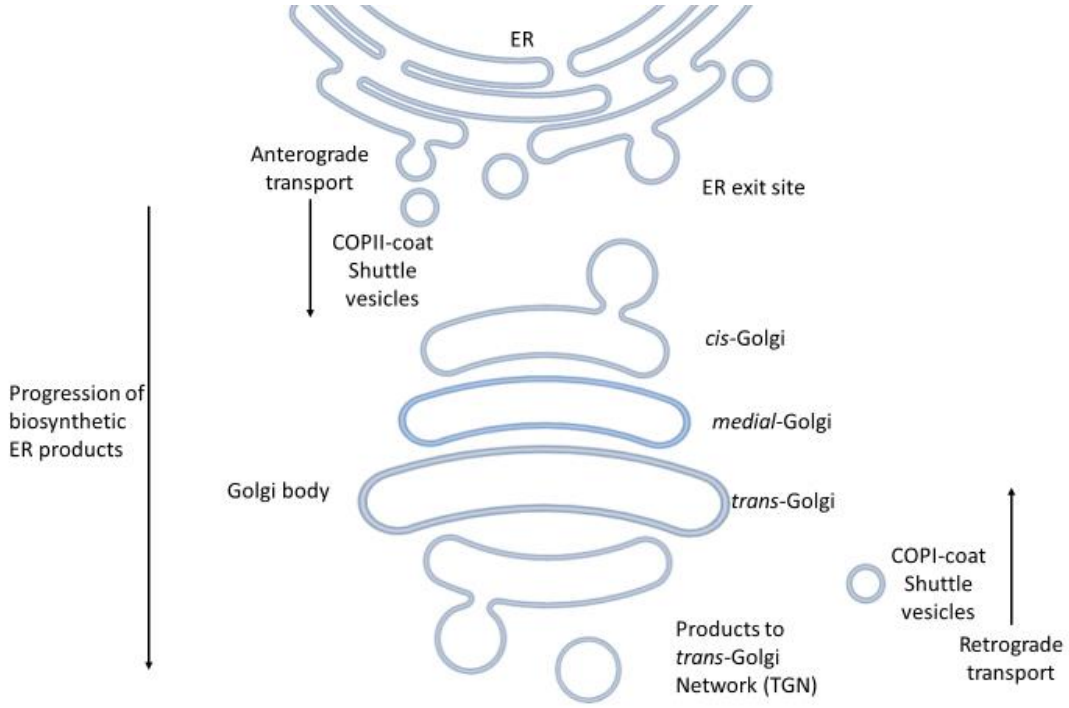


Figure 1: Organization of the plant Golgi structure and transport for protein biosynthesis (modified from Pigino, 2012).

The composition of glycosylation enzymes in the medial/trans-Golgi differs between plant and mammalian systems (Figure 2, 3). Adaptation of the plant machinery to produce human glycan pattern requires the plant enzymes XylT (β1,2-xylosyltransferase), FUT11 and 12 (α1,3-fucosyltransferases) to be replaced or outcompeted with the mammalian enzymes GNTIV and GNTV (N-acetylglucosaminyltransferase IV and V), B4GALT1 (Beta-1,4-galactosyltransferase), and ST (α2,6-sialyltransferase) (Table 1). The plant enzymes to be replaced or outcompeted are located in the *medial/trans*-Golgi, hence GNTIV, GNTV, ST6GAL and B4GALT1 were also to be targeted to the *medial*/*trans*-Golgi for effective enzymatic function on N-glycan intermediates (Schoberer et al., 2018). N-acetylglucosaminyltransferases (GnTs; GnTI, GnTII, GnTIV and GnTV) use uDP-*N*-acetylglucosamine (UDP-GlcNAc) in order to transfer N-acetylglucosamine onto various points of the developing glycan structure, the former steps (GnTI and GnTII) are conserved between plants and mammalian cells (Figure 3), whilst the latter and subsequent steps responsible for the intended end-terminal sialylation (GNTIV, GNTV, B4GALT1 and ST6GAL) are not present in plants and must be introduced for sialic acid capping to occur (Macdonald et al., 2014).

Such adaption may avoid the production of, for example, core fucose, xylose and Lewis-a bodies which are all a feature of plant glycosylation (Figure 3) and are conserved in most of the higher plants being screened for use in therapeutic production. These motifs have raised concerns for their potential immunogenic profile when applied to use for human therapeutics (Bardor et al., 2003; Ree et al., 2000), and in the circumstances of lifelong drug regimens such as enzyme replacement therapy, could significantly limit the long-term efficacy of treatment with increased risk of hypersensitivity reactions. Therefore, a drug production system with a lack of host organism endogenous contaminants, alongside improved general production parameters offered by plant systems, would be of great scientific and commercial significance.

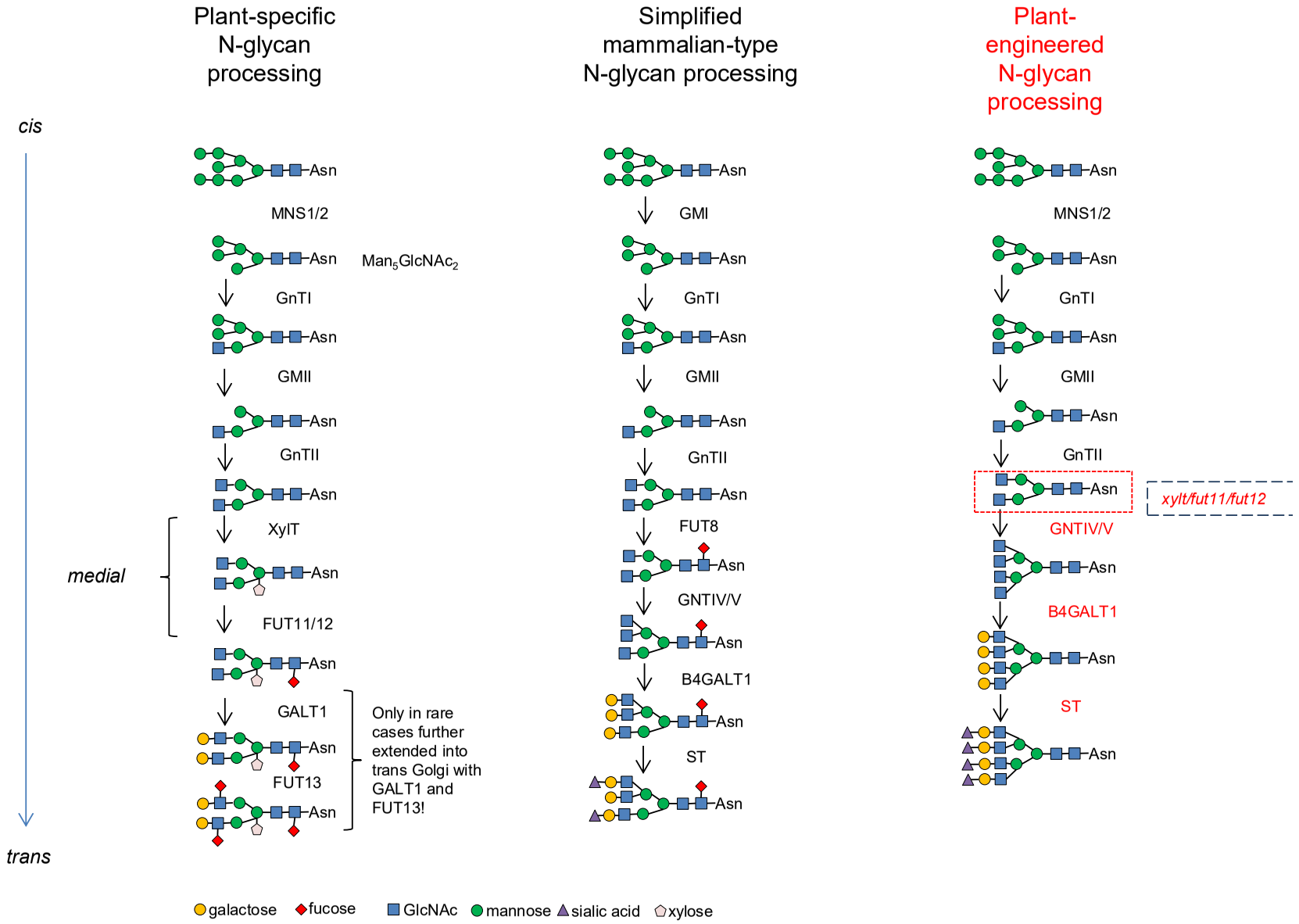

Plant-specific N-glycan processing
Simplified mammalian-type N-glycan processing
Plant-engineered N-glycan processing
cis
trans
medial
Asn
MNS1/2
Man5GlcNAc2
GnTI
GMII
GnTII
XylT
FUT11/12
GALT1
FUT13
Only in rare cases further extended into trans Golgi with GALT1 and FUT13!
GMI
FUT8
GNTIV/V
B4GALT1
ST
xylt/fut11/fut12
galactose
fucose
GlcNAc
mannose
sialic acid
xylose

Figure 2: Schematic representation of plant and mammalian glycosylation pathways and a summary of steps that need to be changed/replaced/added in the plant systems for humanization of plant glycosylation. MNS: α-mannosidase, GnT: N-acetylglucosaminyltransferase, XylT: β1,2-xylosyltransferase, FUT11/12: α1,3-fucosyltransferase, GALT1: β1,3-galactosyltransferase, FUT13: α1,4-fucosyltransferase, B4GALT1: Beta- 1,4-galactosyltransferase, ST: sialyl transferase. Figure modified from Schoberer et al. (2018).

Table 1: Mammalian Glycosylation enzymes and their function

Each of the major glycosylation enzymes investigated within this project and their relevant activity towards a humanized LAL end-product.

Glycosylation enzymes	Enzymatic function	Relation to product
GNTIV & GNTV	Introduce bi- and tri- antennary branched glycans into original chain, N-acetylglucosaminyltransferase enzymes, utilize UDP- *N*-acetylglucosamine to transfer N-acetylglucosamine onto the structure at several branch points (McDonald et al, 2014)	Glycan structure not found in plants but is an essential foundation for downstream mammalian glycomodification
B4GALT1	Addition of β1,4-galactose intermediates	Addition of subsequent and intermediary glycans onto branched chains for the precursor to sialic acid addition
ST6GAL	Addition of terminal sialic acid residues	Addition of terminal sialic acid residue essential to mimicking mammalian glycosylation profile, massively increases half-life and stability by reducing elimination from circulation

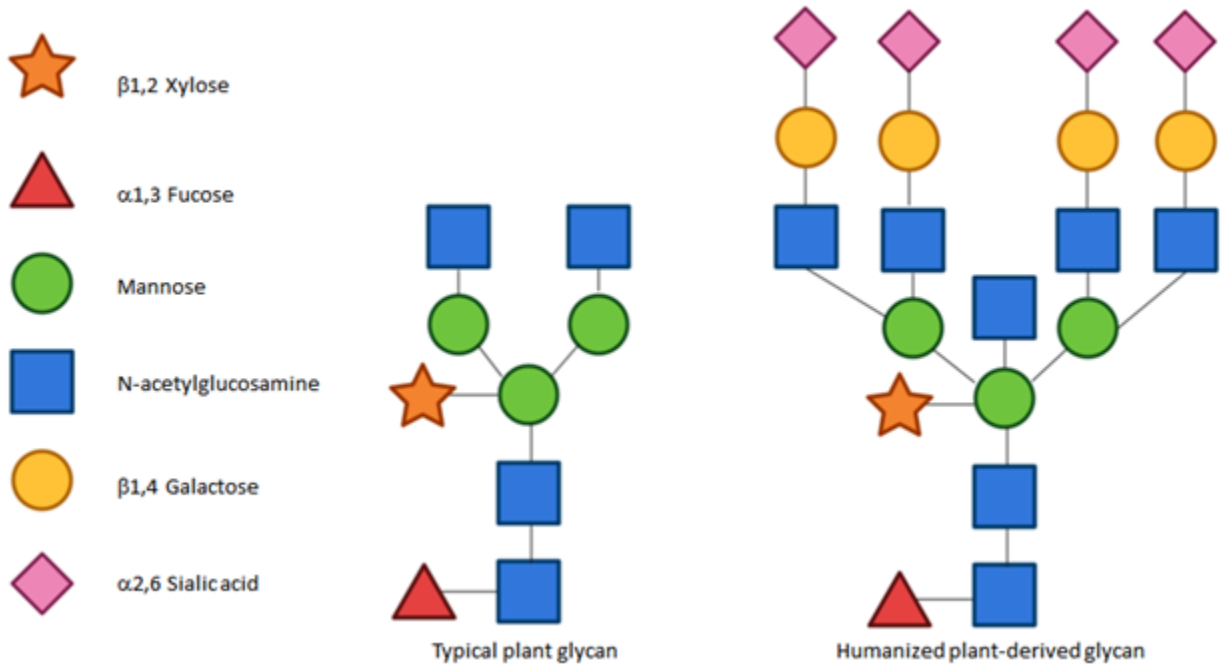


Figure 3: Potential effects on plant N-glycan structure by addition of mammalian post-translational glycosylation enzymes, demonstrating addition of further mammalian glycosylation motifs onto a typical plant glycan to create a humanized hybrid structure. Engineering of glycosylation pathway to include mammalian enzymes causes enzymatic activity majorly responsible for creation of branching structure and attachment of intermediary linkers of N-acetylglucosamine (GNTIV, GNTV) in addition to β1,4 Galactose intermediate glycans (B4GALT1) that facilitate subsequent attachment of sialic acid residues (ST6GAL) resulting in a humanized plant glycan hybrid structure (Macdonald et al., 2014, Schoberer et al., 2018).

Aims and Objectives

The aim of this project was to work towards an efficient plant-based protein production system using LAL as a proof-of-concept production target.

This is addressed by adapting the plant glycosylation patterns in tobacco plants towards human patterns by expressing the mammalian glycosylation enzymes B4GALT1, GNTIV/V, and ST6GAL transiently as fluorescent protein fusions in tobacco leaf cells. Correct localisation in the plant Golgi apparatus was assessed by comparison with known marker proteins using high-resolution confocal microscopy.

We further determined the subcellular localisation of LAL in tobacco leaf cells and investigated possibilities to enhance LAL protein production in plant cells.

The research described in this thesis has been undertaken to work towards the establishment of an efficient plant production system for the synthesis of human therapeutics. As a proof of concept, Lysosomal Acid Lipase (LAL) protein is expressed in *Nicotiana tabacum* plants.

2. Material and Methods

2.1 Selection of LAL and human glycosylation enzyme sequences

The initial sequence for LAL was provided by Porton Biopharma Ltd (Salisbury, England, https://www.portonbiopharma.com/). Selection of candidate human glycosylation enzymes was intended to generate a human posttranslational N-linked glycosylation pathway in a plant, by insertion or replacement of the final steps in endogenous plant glycan synthesis necessary for a humanized glycosylation profile, which would not otherwise be present in *N. tabacum.* A short literature review was conducted to generate a list of candidate human genes necessary for glycosylation, including B4GALT1, GNTIV, GNTV and ST6GAL. The constitutive human sequences were taken from the NCBI database. LAL and the mammalian glycosylation enzymes were codon-optimised for tobacco and synthesized by TWIST Biosciences (San Francisco, United States, https://www.twistbioscience.com/).

2.2 Vector preparation

Standard Gateway® cloning protocols were followed. All primers were ordered from Eurofins Genomics (Ebersberg, Germany, https://www.eurofinsgenomics.eu/) and are listed in Table 2.

Table 2: Primers generated for molecular cloning

Primers were designed to function with New England Biolabs OneTaq and Q5 Polymerase molecular cloning kits. Annealing temperatures were calculated utilizing the online Tm Calculator toolkit. Gene sequences are shown in capital letters, Gateway overhangs are added in small letters.

Primer name	Annealing temperature (℃)	Sequence
ST-SPfor (B4GALT1SPgatefor)	58	ggggacaagtttgtacaaaaaagcaggctctATGATCCACACAAATCTTAAG
LALfor	48	ATGAAAATGCGGTTTTTGGG
LALrev	48	CTGATACTTGCGCATAAGG
LALgatefor	48	ggggacaagtttgtacaaaaaagcaggctctATGAAAATGCGGTTTTTGG
LALgateforsp	60	ggggacaagtttgtacaaaaaagcaggctctATGTCAGGTGGAAAGTTG
LALgaterev	60	ggggaccactttgtacaagaaagctgggtcCTGATACTTGCGCATAAGG

B4GALT1for	49	ATGAGGCTGAGGGAAC
B4GALT1rev	49	CCATAGCAGTCTCCTAATTTTTTC
B4GALT1gatefor	60	ggggacaagtttgtacaaaaaagcaggctctATGAGGCTGAGGGAA
B4GALT1gaterev	60	ggggaccactttgtacaagaaagctgggtcCCATAGCAGTCTCCTAATTTTT
B4GALT1signalfor	56	ATGATCCACACAAATCTTAAGAAAAAATTTTCTTGTTGTGTTCTAGTCTTTTTACTTTTTGCAGTTATTTGTGTATGG AGGCTGAGGGAAC
B4GALT1gaterevlow	58	ggggaccactttgtacaagaaagctgggtcCCATAGCAGTCTCCTAATTT
ST6GAL1for	48	ATGATCCACACAAATCTTAAGAAAAAATT
ST6GAL1rev	48	GCAGTGGATTGTTCGG
ST6GAL1gatefor	60	ggggacaagtttgtacaaaaaagcaggctctATGATCCACACAAATCTTAAGAAA
ST6GAL1gaterev	60	ggggaccactttgtacaagaaagctgggtcGCAGTGGATTGTTCGG
GNTIVfor	49	ATGAGATTACGTAATGGGACTG
GNTIVrev	49	ATTTGTAGCCTTTTTGATATGAATTTCG
GNTIVgatefor	60	ggggacaagtttgtacaaaaaagcaggctctATGAGATTACGTAATGGGAC
GNTIVgaterev	60	ggggaccactttgtacaagaaagctgggtcATTTGTAGCCTTTTTGATATGAATTTC
GNTIVsignalfor	58	ATGATCCACACAAATCTTAAGAAAAAATTTTCTTGTTGTGTTCTAGTCTTTTTACTTTTTGCAGTTATTTGTGTATGGAGATTACGTAATGGGACTG
GNTVfor	48	ATGGCGCTGTTTACTCCATG
GNTVrev	48	AAGGCAATCTTTGCACAAG
GNTVgatefor	59	ggggacaagtttgtacaaaaaagcaggctctATGGCGCTGTTTACTC
GNTVgaterev	59	GGGGACCACTTTGTACAAGAAAGCTGGGTCAAGGCAATCTTTGCACAA

Genes were first inserted into a pDONR/221 Gateway Entry vector via BP clonase II Enzyme Mix (ThermoFisher Scientific) according to the manufacturer's instructions, using half volumes to conserve reagents. Resulting plasmids were then transformed into DH 5-alpha competent high efficiency *E. coli* cells (New England Biolabs) according to manufacturer guidelines, however using 25µl of cells and 200µl SOC recovery media. Selection was undertaken on agar plates with selection agent Kanamycin 50µg mL^{-1}. After 24hr incubation at 37℃, isolated single colonies were tested for the plasmid of interest via colony PCR. Once the presence of the plasmid of interest was confirmed, selected colonies were transferred into 3ml liquid culture of LB and appropriate antibiotics, and were kept overnight at 37℃, with shaking at 200rpm. Plasmid DNA extraction was performed using the NEB Monarch Plasmid Miniprep Kit (New England Biolabs) in accordance with manufacturer instructions. Gene sequences in plasmid DNA were confirmed by sequencing (Eurofins Genomics). For correct pDONR221 constructs an LR reaction was performed using an LR Clonase II Enzyme Mix (ThermoFisher) following manufacturer guidelines, but with half the suggested volume of reagents, to generate the desired plant expression vector (pB7WGC2 or pB7FWG2 for RFP or GFP-fusions, respectively; Karimi et al., 2005). A second transformation into competent *E. coli* was then undertaken, selected with Spectinomycin 50µg mL^{-1}, constructs confirmed via colony PCR and plasmid extracted (Table 3).

Table 3: List of constructs utilized for investigation

Constructs generated and marker proteins used for confocal analysis.

Construct	Plasmid backbone	Source	Citation
RFP-LAL	pB7WGR2	Alastair McGinness	Unpublished
LAL-RFP	pB7RWG2	Alastair McGinness	Unpublished
RFP- LALΔSP	pB7WGR2	Alastair McGinness	Unpublished
LALΔSP-RFP	pB7RWG2	Alastair McGinness	Unpublished
RFP-B4GALT1	pB7WGR2	Alastair McGinness	Unpublished
B4GALT1-RFP	pB7RWG2	Alastair McGinness	Unpublished
RFP-B4GALT1SP	pB7WGR2	Alastair McGinness	Unpublished
B4GALT1SP-RFP	pB7RWG2	Alastair McGinness	Unpublished
RFP-ST6GAL	pB7WGR2	Alastair McGinness	Unpublished
ST6GAL-RFP	pB7RWG2	Alastair McGinness	Unpublished
RFP-GNTIV	pB7WGR2	Alastair McGinness	Unpublished
GNTIV-RFP	pB7RWG2	Alastair McGinness	Unpublished
RFP-GNTV	pB7WGR2	Alastair McGinness	Unpublished
GNTV-RFP	pB7RWG2	Alastair McGinness	Unpublished
GFP-LAL	pB7WGF2	Alastair McGinness	Unpublished
LAL-GFP	pB7FWG2	Alastair McGinness	Unpublished

GFP-B4GALT1	pB7WGF2	Alastair McGinness	Unpublished
B4GALT1-GFP	pB7FWG2	Alastair McGinness	Unpublished
GFP-ST6GAL	pB7WGF2	Alastair McGinness	Unpublished
ST6GAL-GFP	pB7FWG2	Alastair McGinness	Unpublished
GNTIV-GFP	pB7FWG2	Alastair McGinness	Unpublished
GNTV-GFP	pB7FWG2	Alastair McGinness	Unpublished
LAL-CLVR	pB7FWG2	Alastair McGinness	Unpublished
B4GALT1-CLVR	pB7FWG2	Alastair McGinness	Unpublished
ST6GAL-CLVR	pB7FWG2	Alastair McGinness	Unpublished
GNTIV-CLVR	pB7FWG2	Alastair McGinness	Unpublished
GNTV-CLVR	pB7FWG2	Alastair McGinness	Unpublished
ST-RFP	pVKH18En6	Federica Brandizzi	Renna et al. (2005)
ST-CFP	pVKH18En6	Federica Brandizzi	Renna et al. (2005)
ST-GFP	pVKH18En6	Federica Brandizzi	Renna et al. (2005)
MNS1-RFP	p31 (Hüttner et al., 2012)	Jennifer Schoberer	Schoberer et al. (2013)
MNS1-GFP	p20 (Schoberer et al., 2009)	Jennifer Schoberer	Schoberer et al. (2013)

2.3 Production of competent agrobacteria

The *Agrobacterium tumefaciens* strain GV3101 was prepared as a large initial volume of approximately 200ml in LB (Lysogeny Broth media, with 25µg mL^{-1} rifampicin) from a 28℃-overnight culture. The initial volume of agrobacterium was then pelleted by centrifugation at 3500rpm for 30 minutes, at 4℃. The pellet was resuspended in cold 1M $CaCl_2$ and then spun again at 3500rpm for 10 mins at 4℃. The pellet was re-suspended in 15ml 1M $CaCl_2$ and the bacterial solution was split into 400µl aliquots snap-frozen in liquid nitrogen and stored at -80℃.

2.4 Transformation of agrobacteria

100µl of competent agrobacteria (section 2.3) were combined with 400-600ng of the expression plasmid of interest, and incubated on ice for 5 mins. The mixture was then flash-frozen either by liquid nitrogen immersion or by placement on metal shelving in the -80℃ for 15 mins. Agrobacteria were incubated in a 37℃ water bath for 4 mins. The mixture was rapidly transferred to a 15ml Falcon tube with 1ml LB, and was then allowed to recover in a 28℃ incubator with shaking in 1ml fresh LB for 2-3 hours. Agrobacteria were then spread onto an agar plate (25µg mL^{-1} rifampicin, 50µg mL^{-1} of gentamycin and either 50µg mL^{-1} of kanamycin or spectinomycin dependent on the vector used), the plates were incubated at 28℃ for 48h, at which point colonies of agrobacterium could be observed. Transformed agrobacterium colonies were then transferred into liquid culture for use in agrobacterium-mediated plant transformation (section 2.5), or for long term storage they were placed at -80℃ in 50% sterile glycerol.

2.5 Transient expression in tobacco leaf epidermal cells of fluorescent proteins

Transformed agrobacterium liquid cultures were pelleted by centrifugation at 2200g at room temperature for 5 mins. Infiltration buffer (5mg mL^{-1} glucose, 50mM MES, 2mM $Na_3PO_4 \cdot 12H_2O$ and 0.1mM acetosyringone) was made fresh from stocks and used to wash the pellet once, re-pellet by a repeat centrifugation step and then to resuspend the agrobacterium into a final solution. The bacterial suspension was diluted in the infiltration buffer to an OD_{600} 0.1. The infiltration medium was injected into the abaxial leaf surface of 2-week-old mature tobacco plants (grown under a 22hr light cycle) with a blunt 1ml syringe, after first making a small puncture using a pipette tip. A permanent marker on the leaf surface was used to indicate the spread of the medium within the leaf tissue to ensure only infiltrated leaf surface is later used. Infiltrated plants were incubated at 23℃ for 72h before imaging (section 2.6).

2.6 Confocal microscopy

For the purposes of imaging tobacco leaf epidermal cells, small leaf sections ($25mm^2$) were cut from infiltrated tobacco leaves and the abaxial surface images. Sections were mounted in water for confocal microscopy. All confocal microscopy was performed using a Zeiss LSM 880 with Airyscan detector. Images were collected with either a Zeiss PlanApo 100x/1.46 NA oil immersion or a Zeiss PlanApo 63x/1.4 NA oil immersion objective. 512 x 512 images were collected in 8-bit with 2-line averaging and excitation at 488nm (GFP) and 561 (RFP), and emission at 495-550nm and 570-615nm, respectively.

2.7 Co-localisation analysis methods

Side-profile images of Golgi bodies were collected for the purposes of co-localisation analysis. All co-localisation analysis was performed on the Fiji package for ImageJ image processing software. The plugin JACoP (Just another Co-localization Plugin) was used for overlap coefficient analysis (Figure 5) and generation of cytofluorograms (Bolte et al., 2006), alongside the RGB profiler plugin (Figure 4) for generation of color line intensity profiles (Laummonerie et al., 2004).

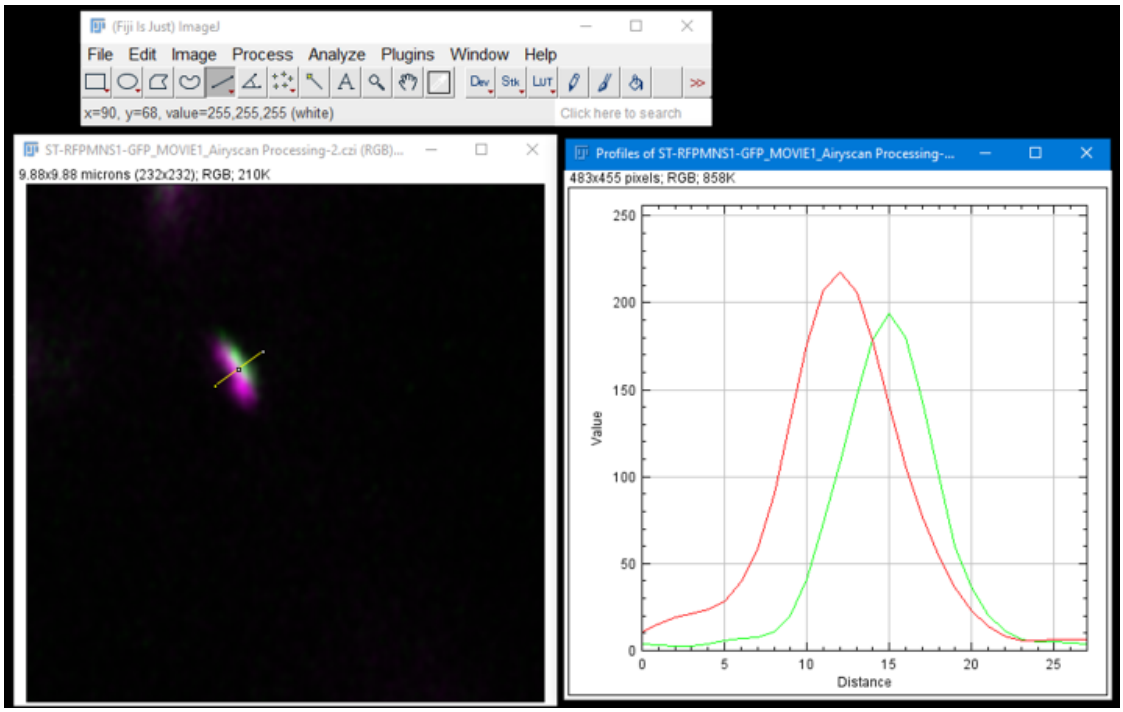


Figure 4: Workflow process image for use of RGB profiler plugin for Fiji. ST-RFP (shown in magenta/red) and MNS1-GFP (shown in green) marker image used as an example image.

For the generation of a line profile (Figure 4), a confocal image of a side-view Golgi body was selected, and the line tool used to directly bisect the fluorescent regions. Using the “Plugins” menu within the Fiji interface to open RGB profiler generated a line profile analysis graph output based on the fluorescence intensity value along line selected.

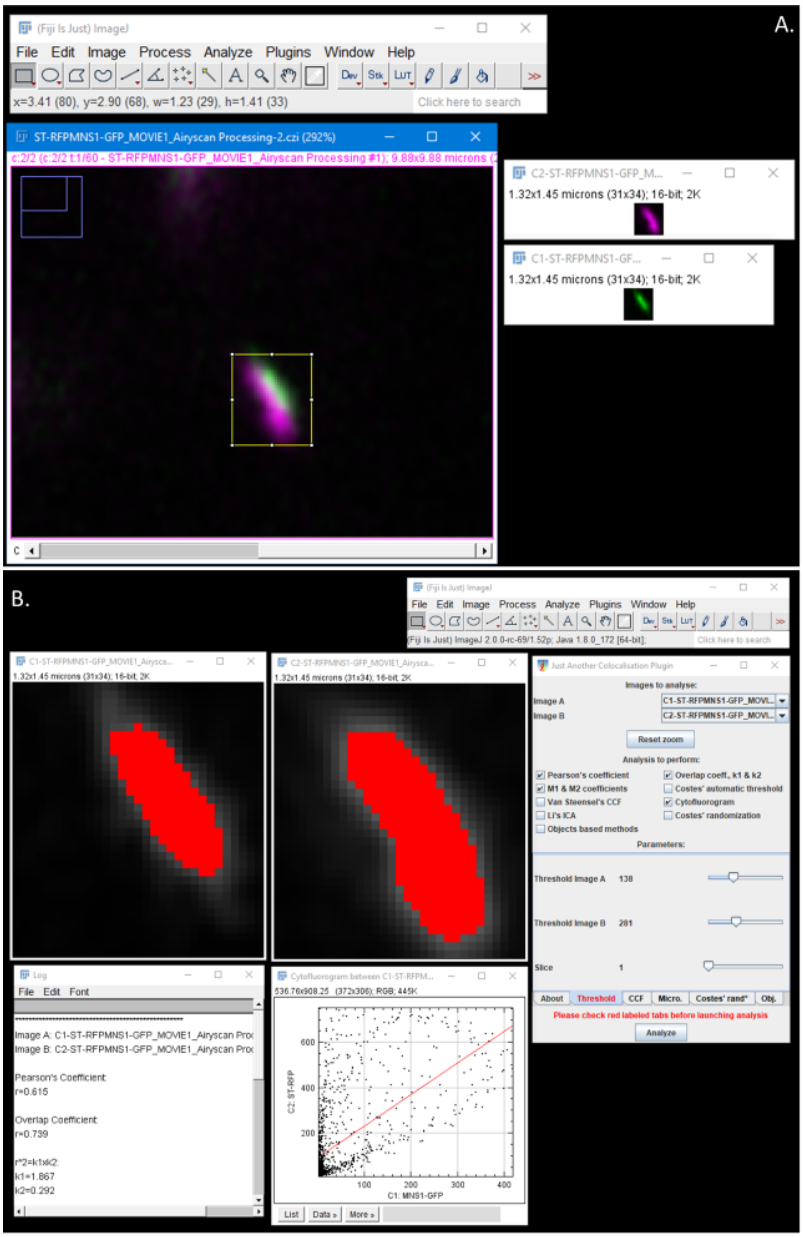


Figure 5: Workflow image process using the JACoB plugin in Fiji for processing a side-view image of a Golgi body. ST-RFP (shown in magenta) and MNS1-GFP (shown in green) marker image used as an example.

Generation of cytofluorograms and Pearson's coefficient (Figure 5) required a rectangular selection area to be made of the Golgi body, then duplicated into a separate image, and split into separate channels. Selection of JACoB from "Plugins" menu opened user interface for the tool, where channels for analysis were selected, and intensity threshold for analysis set. Running analysis generated an output text log of coefficient analysis values and a cytofluorogram graph.

3. Results

3.1 Subcellular localisation of mammalian glycosylation enzymes in tobacco cells

Comparison of glycosylation pathways in mammalian and plant cells revealed that the following mammalian enzymes would have to be added to the plant glycosylation machinery: B4GALT1, GNTIV, GNTV and ST6GAL. Sequences for the glycosylation enzymes were codon-optimised for expression in tobacco cells and sent for synthesis to TWIST biosciences. Subcellular localisation analysis was investigated via fluorescent constructs, generation of which is described in subsection 2.2 of the methods section, and combinations used can be seen in table 3.

Using an Agrobacterium-mediated transformation protocol (section 2.5), all glycosylation enzymes fused to fluorescent proteins were transiently expressed in tobacco leaf epidermal cells. Golgi body localisation was investigated by co-expression of all constructs with the known *medial/trans*-Golgi marker protein sialyltransferase (ST) (Figures 6, 7). Proteins fused to a GFP fluorophore localised mainly to the ER (Figure 6) or distinctly separated from the Golgi body (appendices 1).

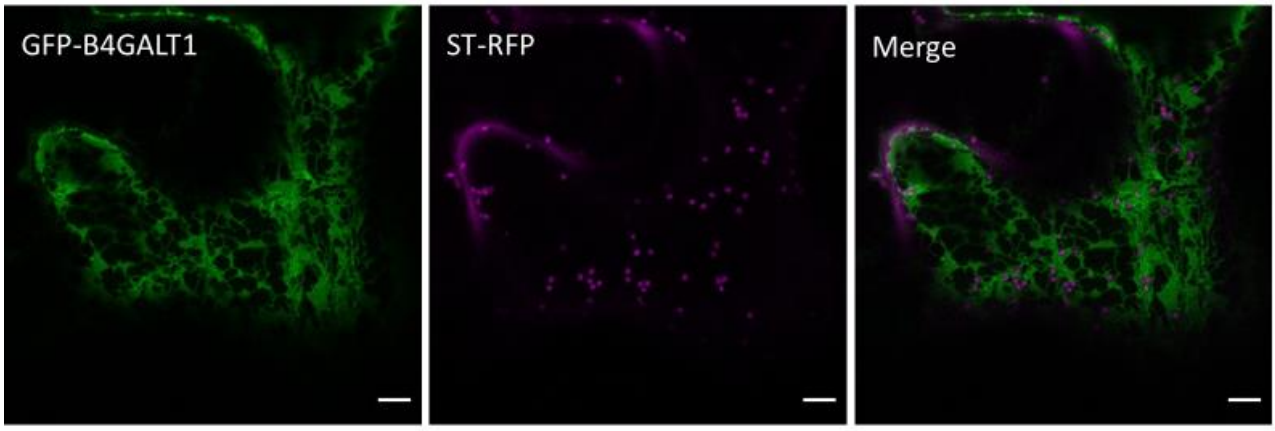


Figure 6: Confocal data for expression of GFP-B4GALT1 with the Golgi marker ST-RFP.
Co-expression of GFP-B4GALT1 (shown in green) and ST-RFP (shown in magenta) does not show co-localisation between the two constructs. GFP-B4GALT1 shows ER localisation. Scale bar = 5 µm. A representative image is shown; n=3 with 5 technical replicas each.

RFP-tagged proteins, however, localised to Golgi bodies and showed co-localisation with ST-GFP (Figure 7).

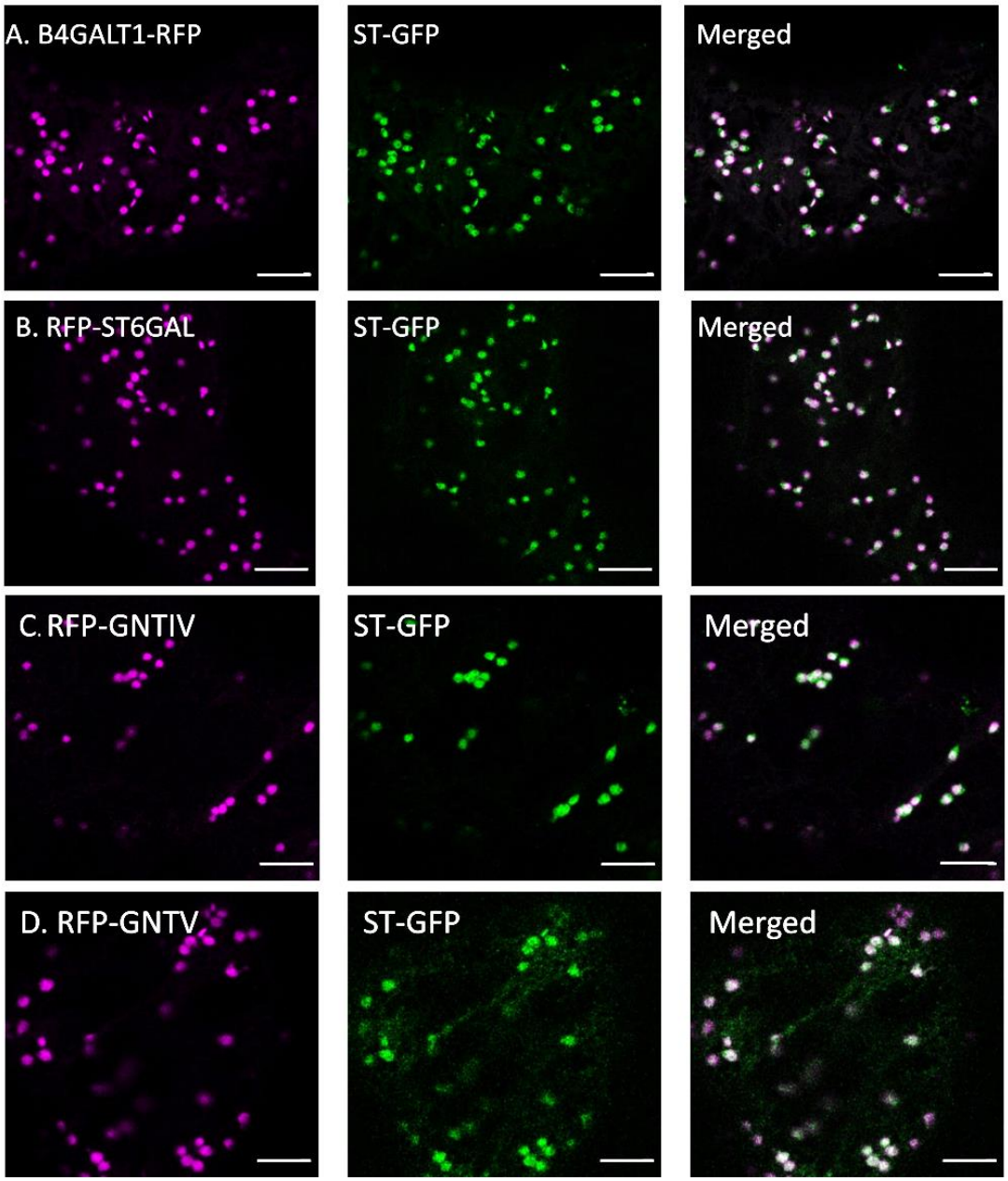
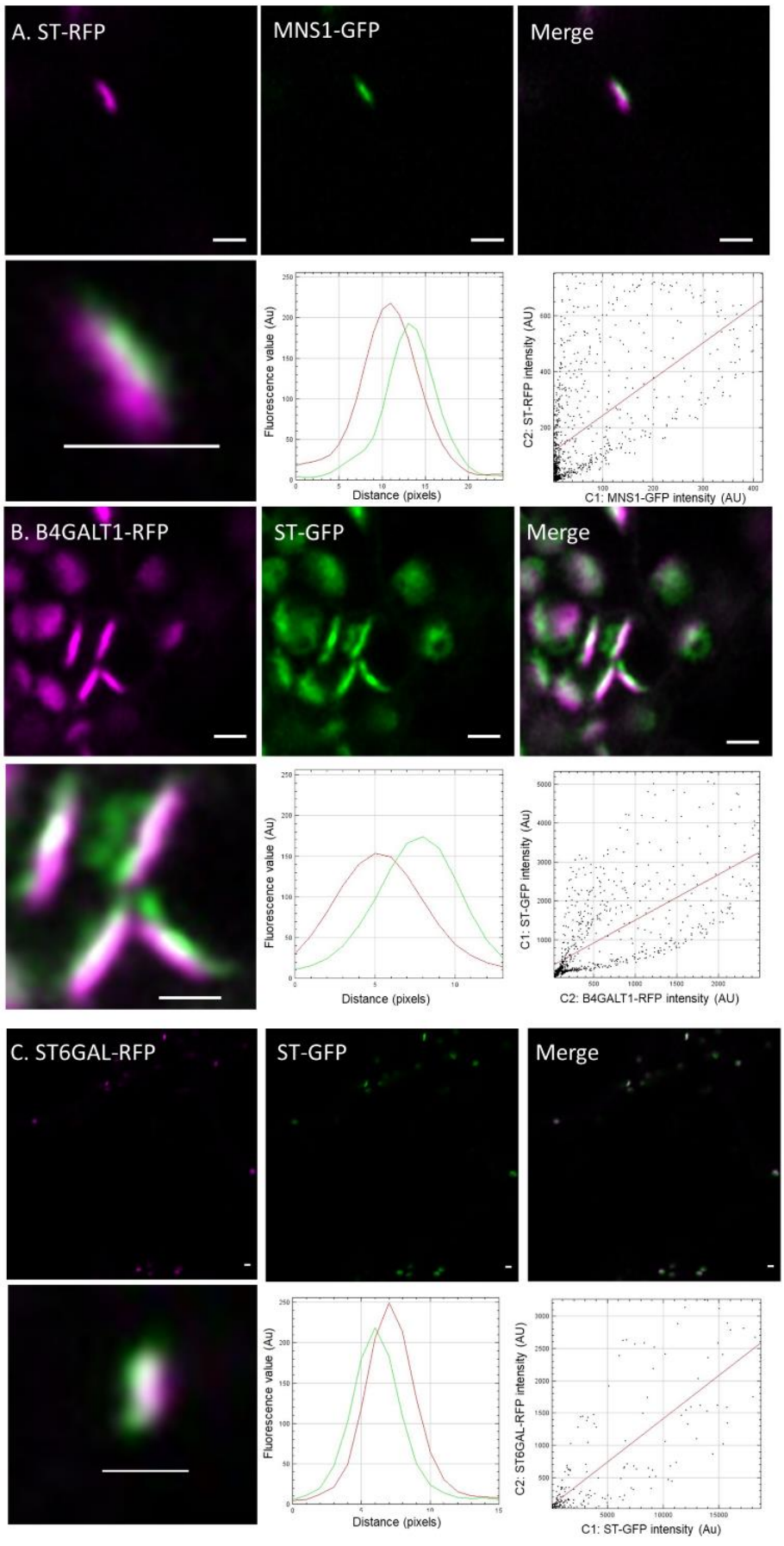


Figure 7: Co-localisation of the mammalian glycosylation enzymes with the *medial/trans*-Golgi marker ST-GFP. Subcellular view for localisation of glycosylation enzymes, expressed alongside ST-GFP (shown in green) to indicate localisation to within the Golgi body: A) B4GALT1-RFP, B) RFP-ST6GAL, C) RFP-GNTIV, D) RFP-GNTV (all shown in magenta). Size bars = 10 µm. Representative images are shown; n=3 with 5 technical replicas each.

3.2 Suborganellar localisation of mammalian glycosylation enzymes in tobacco cells

For the mammalian glycosylation enzymes to function properly and to produce the correct glycosylation pattern, they must be expressed in the *medial/trans*-Golgi (Figures 1, 2). To investigate this, all enzymes were tested for sub-organellar expression within the Golgi stacks by confocal microscopy. To determine sub-organellar localisation the enzymes of interest were co-expressed with the *medial/trans*-Golgi marker ST (Ito et al., 2012; Figure 2).

The quantitative cytofluorogram output for the individual construct combinations (Figure 8) is given by Pearson's coefficient (r), a value for measuring linear correlation between two variables. A value of r = 1 indicates direct overlap of the fluorescence intensity of the marker ST-GFP/RFP and the fluorescent protein fusion of a human enzyme within a Golgi body. For the purposes of clarifying localisation within relatively close suborganellar compartments, in order to be considered as a negative control for co-localization (Figure 8, A.), a known *medial/trans*-Golgi marker ST-RFP was expressed with the *cis*-Golgi marker MNSI (Liebminger et al., 2009; Schoberer et al., 2019), as these are known to localise in different Golgi cisternae. The resulting Pearson's Coefficient r = 0.59 indicates that they do not co-localise, and any degree of variable overlap of localisation is likely due to unspecific fluorophore overlap and resolution restrictions. The relatively low Pearson's coefficient of ST-GFP with B4GALT1 (Figure 8 B, r = 0.66) and GNTIV (Figure 8 D, r = 0.79) indicates that they did not co-localise directly with ST-GFP and so were not expressing within the *medial/trans*-Golgi, and are likely confined to the *cis*-Golgi. Expression of ST6GAL and GNTV indicate a strong degree of co-localisation with ST-GFP (Figure 8 C ST6GAL r = 0.87, 8 E GNTV, r = 0.93) suggesting they are expressing within the *medial/trans*-Golgi cisternae. Co-localisation with ST-GFP indicates that only GNTV and ST6GAL co-localise with ST-GFP, and so are located in the *medial/trans*-Golgi (Figure 8).


A. ST-RFP
MNS1-GFP
Merge
Fluorescence value (Au)
Distance (pixels)
C2: ST-RFP intensity (AU)
C1: MNS1-GFP intensity (AU)
B. B4GALT1-RFP
ST-GFP
Merge
Fluorescence value (Au)
Distance (pixels)
C1: ST-GFP intensity (Au)
C2: B4GALT1-RFP intensity (AU)
C. ST6GAL-RFP
ST-GFP
Merge
Fluorescence value (Au)
Distance (pixels)
C2: ST6GAL-RFP intensity (AU)
C1: ST-GFP intensity (Au)

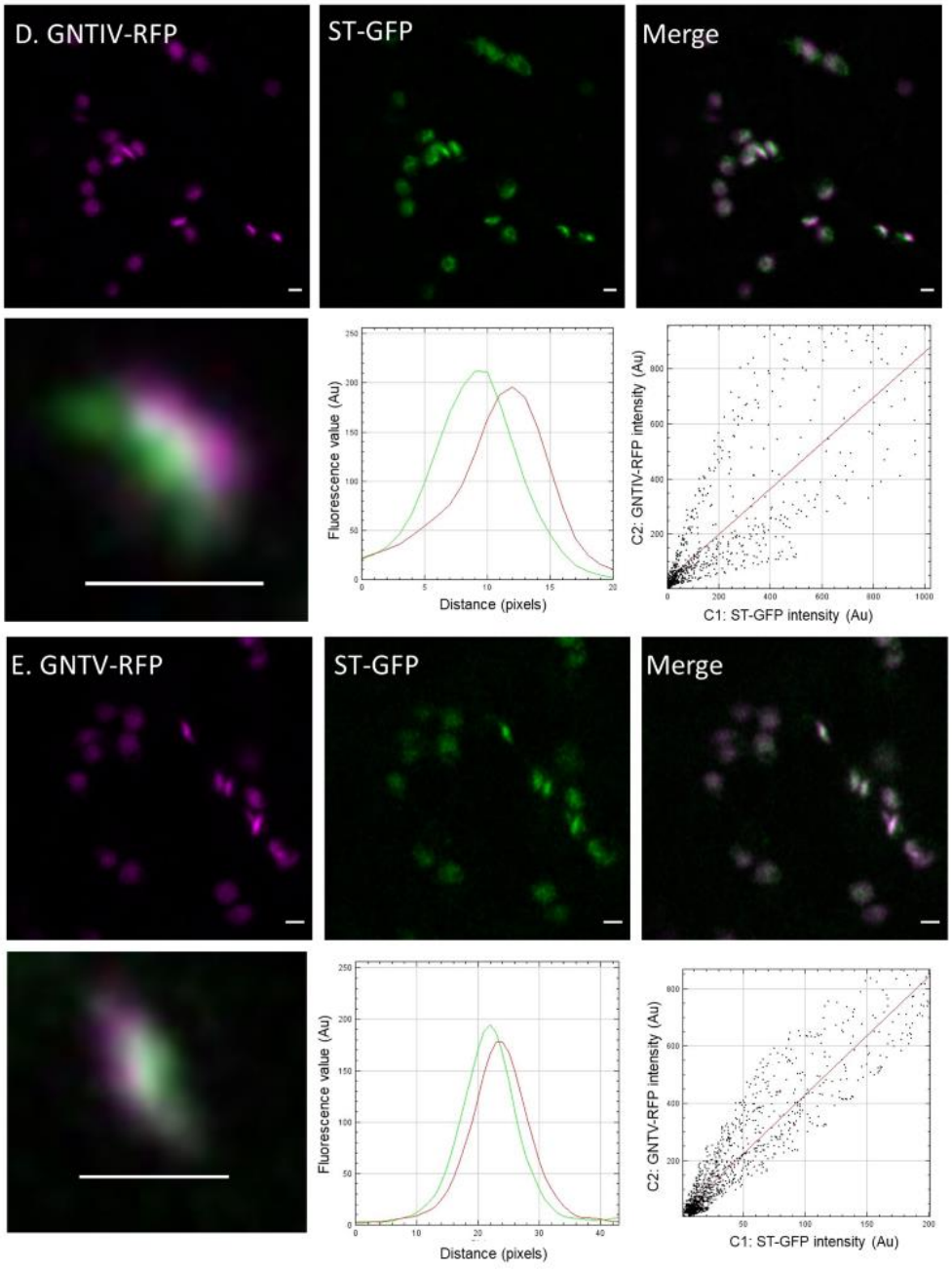


Figure 8: Co-localisation of mammalian glycosylation enzymes with *cis*- and *medial/trans*-Golgi markers. A) Negative marker control for *cis*-(MNS1-GFP, shown in green) and *medial/trans*-Golgi (ST-RFP, shown in magenta) localisation. The *medial/trans*-Golgi marker ST-GFP (shown in green) is co-expressed with the mammalian glycosylation enzymes in tobacco leaf epidermal cells: B) B4GALT1-RFP, C) RFP-ST6GAL, D) GNTIV-RFP, E) GNTV-RFP (all shown in magenta). Size bars = 1 µm. R value n=1 as generated from individual Golgi body shown in inset image, minimum 3 representative Golgi body images samples displaying comparable localisation. Insets show enlarged structures for co-localisation analysis, line profiles generated via RGB profiler and cytofluorogram generated via JACoB. A) Negative control markers, Pearson's Coefficient r = 0.59 B) B4GALT1 r = 0.66 C) ST6GAL r = 0.87 D) GNTIV r = 0.79, E) GNTV r = 0.93. Representative images are shown; n=3 with 5 technical replicas each.

3.3 Engineering of B4GALT1 to target the enzyme to the *medial/trans*-Golgi

In order for the selected glycosylation enzymes to elicit any functional effect on post-translational N-linked glycosylation, they must express in an active state within the *medial/trans*-Golgi cisternae. The initial testing of the glycosylation enzymes indicated they do not all co-localise with ST-GFP (Figure 8), and only ST6GAL and GNTV were localised to the *medial/trans*-Golgi, and so the proposed glycomodification activity would be out of sequence to facilitate intended synthesis of mammalian glycans.

An approach to this issue involved a sialyltransferase signal peptide from ST6GAL known to target proteins to *trans*-Golgi cisternae (Schoberer et al., 2018; Figure 8 C) which was fused to B4GALT1 (B4GALT1SP-RFP, Table 2). B4GALT1 was selected as the most reliably expressing candidate among the glycosylation enzymes in addition to requiring localisation to the *medial/trans*-Golgi for optimal efficacy (Bosch et al., 2013). To test the suborganellar localisation of B4GALT1SP-RFP, the construct was co-expressed with the *medial/trans*-Golgi marker ST-GFP (Figure 9). B4GALT1SP-RFP and ST-GFP show co-localisation (Figure 9 A) and an overlapping line profile (Figure 9 B). Pearson's Coefficient analysis shows a high degree of correlation indicating co-localisation (Figure 9 C; r= 0.99 with values = 1 indicating full co-localisation).

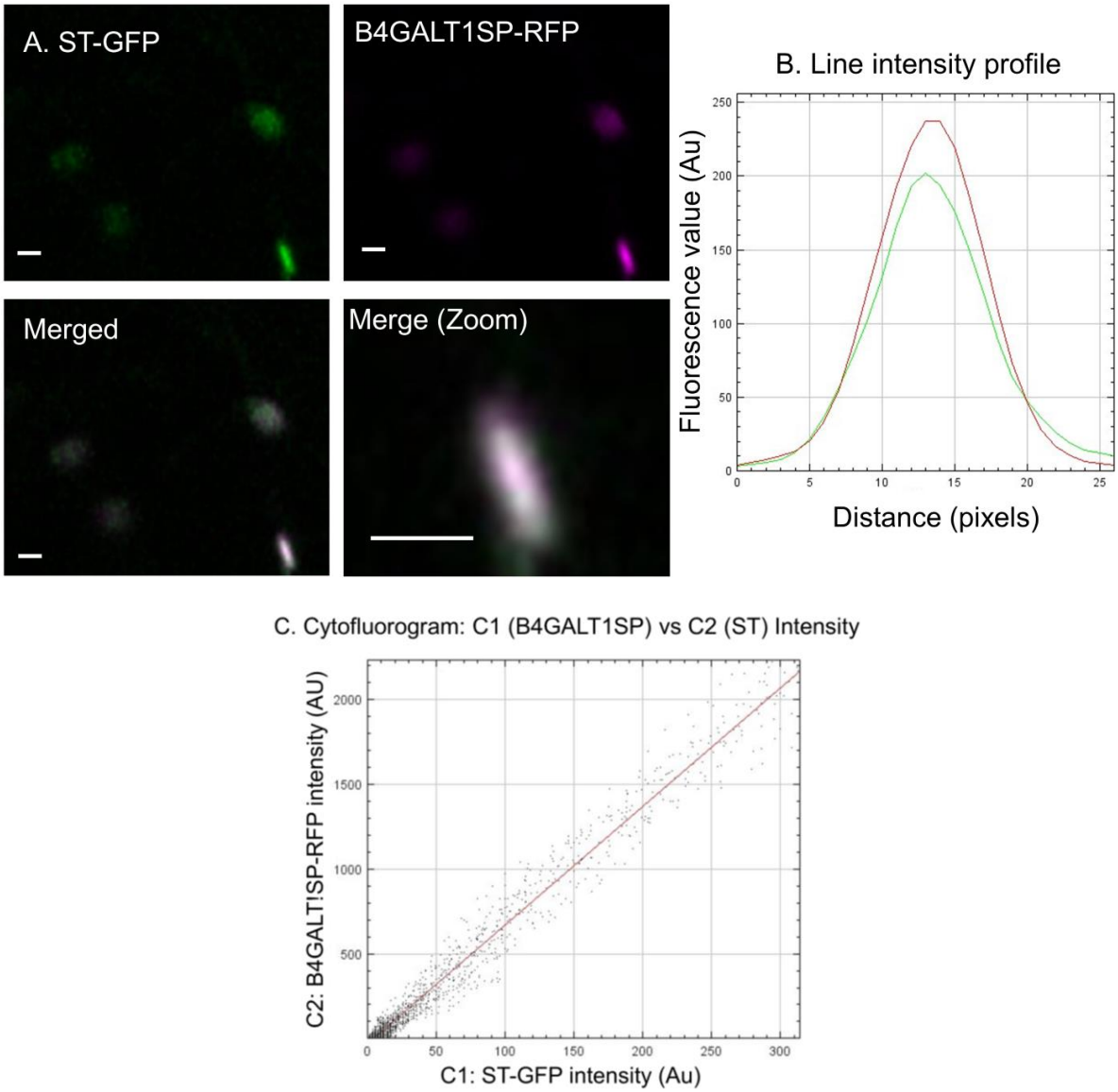


Figure 9: Co-localisation analysis for B4GALT1SP-RFP and ST-GFP *medial/trans*-Golgi marker.

A) Co-expression of B4GALT1SP-RFP (magenta) and ST-GFP (green) shows co-localisation of both constructs. Scale bars= 1 µm. R value n=1 as generated from individual Golgi body shown in inset image, minimum 3 representative images samples displaying similar localisation. Inset shows enlarged structures for co-localisation analysis B) Signal analysis by RGB Profiler plugin for Fiji indicates direct overlap of fluorescence signals, and so direct co-localization. C) Cytofluorogram, Pearson's Coefficient r = 0.99, indicating high degree of co-localisation. Representative images are shown; n=3 with 6 technical replicas each.

3.4 Expression and subcellular localisation of LAL in tobacco leaves

An RFP fluorescent tag was used to assess expression of the human glycosylation enzymes (section 3.2), and the same approach was applied to LAL protein expression. LAL was expressed in tobacco epidermal leaf cells in a transient manner, and localised to the plant Golgi body. A suborganellar view shows co-localisation with the *medial/trans*-Golgi marker ST-GFP (Figure 10 B.). LAL-RFP and ST-GFP show visual co-localisation (Figure 10 A-C) and an overlapping line profile (Figure 10 D). Pearson's Coefficient analysis shows good correlation (Figure 10 E; Pearson's r = 0.88 with values = 1 indicating full co-localisation) indicating direct localisation to the *medial/trans*-Golgi.

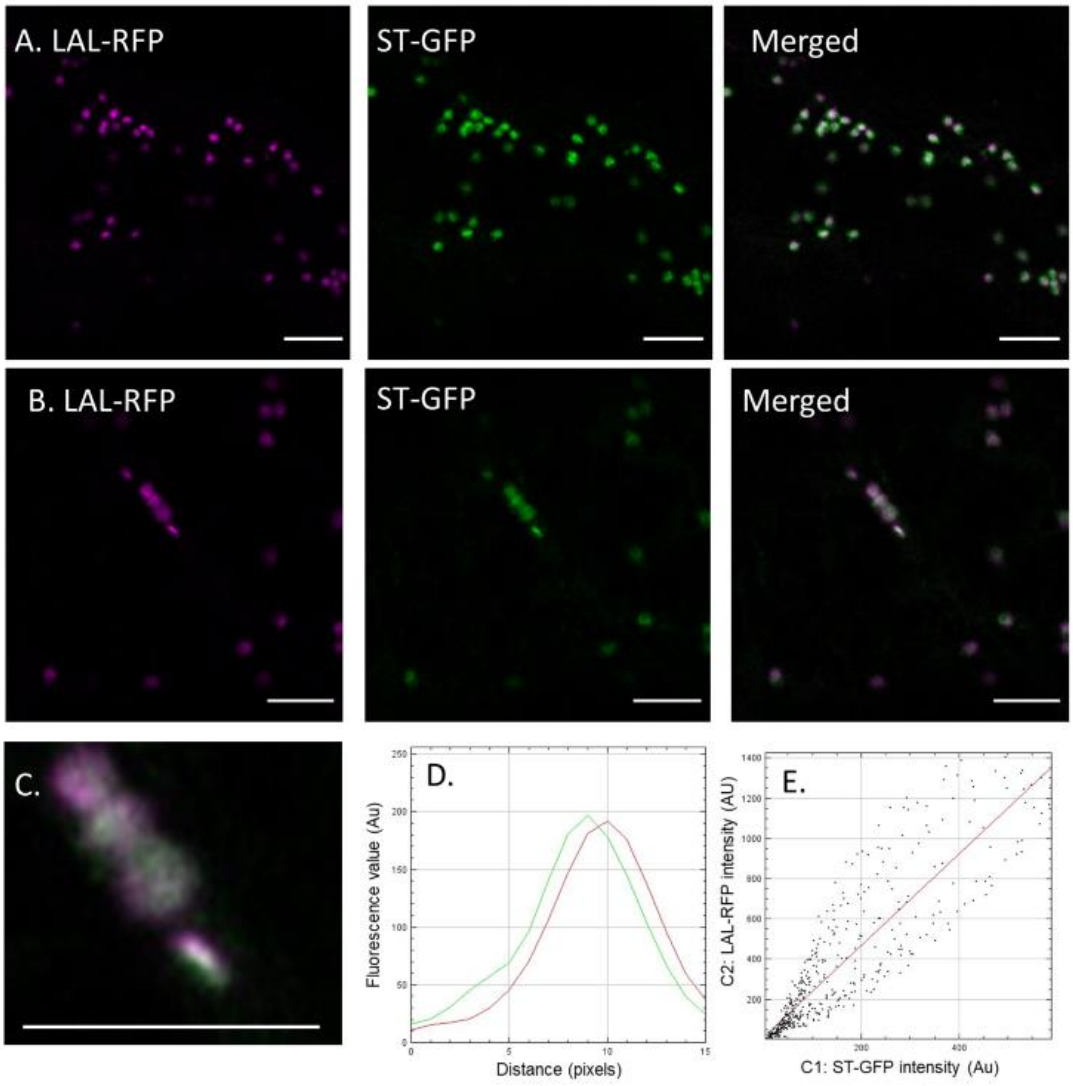


Figure 10: Co-localisation of LAL-RFP with the Golgi marker ST-GFP. LAL-RFP (shown in magenta) is co-expressed with the *medial/trans*-Golgi marker ST-GFP (shown in green). A) Displays a subcellular view of LAL-RFP with ST-GFP to indicate distinct localisation to the Golgi body and B) High resolution image with the Golgi body in side-profile showing co-localisation of LAL-RFP and ST-GFP. C) Inset shows enlarged structures for co-localisation analysis. R value n=1 as generated from individual Golgi body shown in inset image, minimum 3 representative Golgi body images samples displaying comparable localisation. D. Line profile intensity analysis on inset image for B) (analysis by RGB profiler plugin for Fiji) is shown. E) Further Cytofluorogram analysis indicates that signals converge and so co-localize, Pearson's coefficient r = 0.88. Size bars = 10 µm. Representative images are shown; n=3 with 5 technical replicas each.

3.4 Bioengineering of the LAL protein structure

Problematic for biotechnological approaches is that full length human LAL as an RFP fluorophore fusion was found to express in an extremely low proportion of tobacco cells tested, with a minimal number of cells strongly expressing the construct. This made it especially important to increase cellular transformation efficiency and therefore protein expression levels in order to improve efficiency as a potential production system.

In an attempt to identify any potential solutions within the original LAL protein sequence, SignalP analysis was run on the LAL aa sequence to identify potential explanations of low expression yields in the sequence (Figure 11). The original human LAL sequence was predicted to feature an endogenous signal peptide at aa 1-22. As this signal peptide would likely be cleaved as part of targeted localisation within the mammalian Golgi during synthesis it is possible that the plant system is lacking the correct mechanism for this process due to differences in host architecture, and this could lead to the poor expression identified.

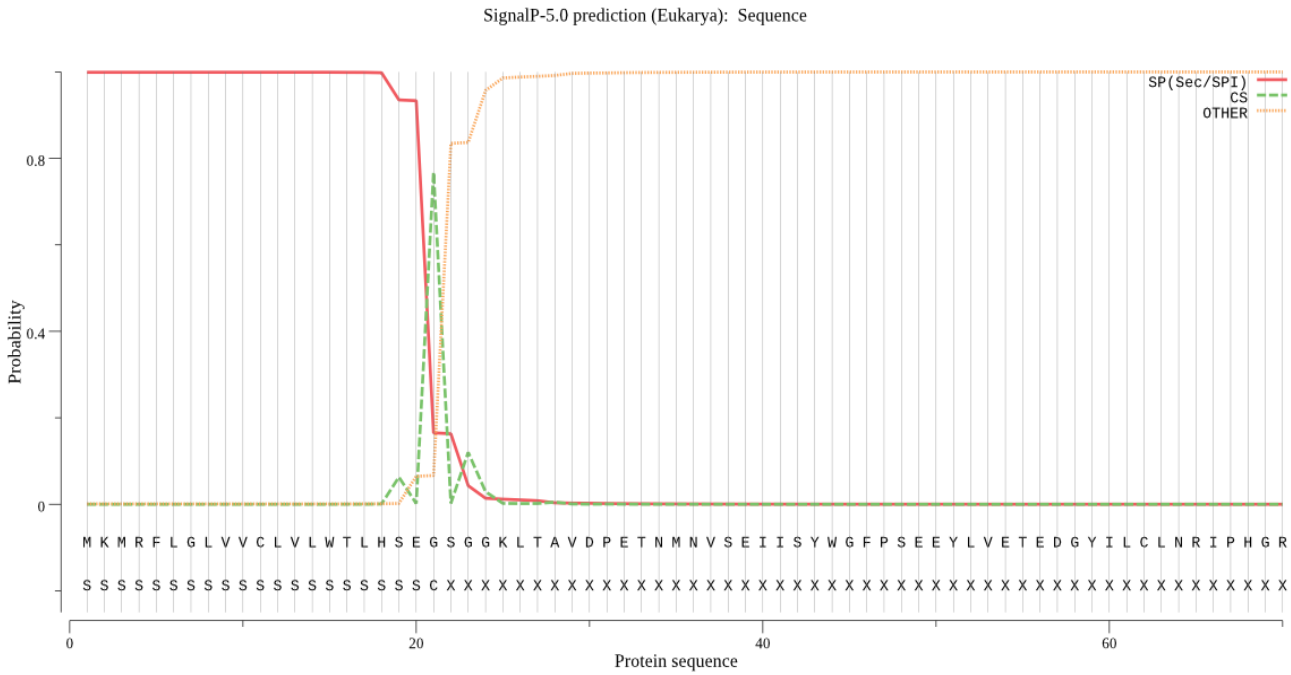


Figure 11: Sequence analysis of LAL using SignalP.

SignalP output of LAL sequence indicates a predicted signal peptide site position in the sequence between 1-22aa with a probability of 0.7682.

To test this, a modified LAL construct lacking the endogenous signal peptide was created and termed LALΔSP. LALΔSP was fused to a C-terminal RFP and co-expressed with the *medial/trans*-Golgi marker ST-GFP (Figure 12). LALΔSP-RFP showed the same localisation as LAL-RFP (Figure 10) and co-localised with the *medial/trans*-Golgi marker ST-GFP (Figure 12 A). Cytofluorogram analysis resulted in a high Pearson's coefficient r= 0.87 indicating co-localisation. The removal of the signal peptide does not seem to impact on localisation.

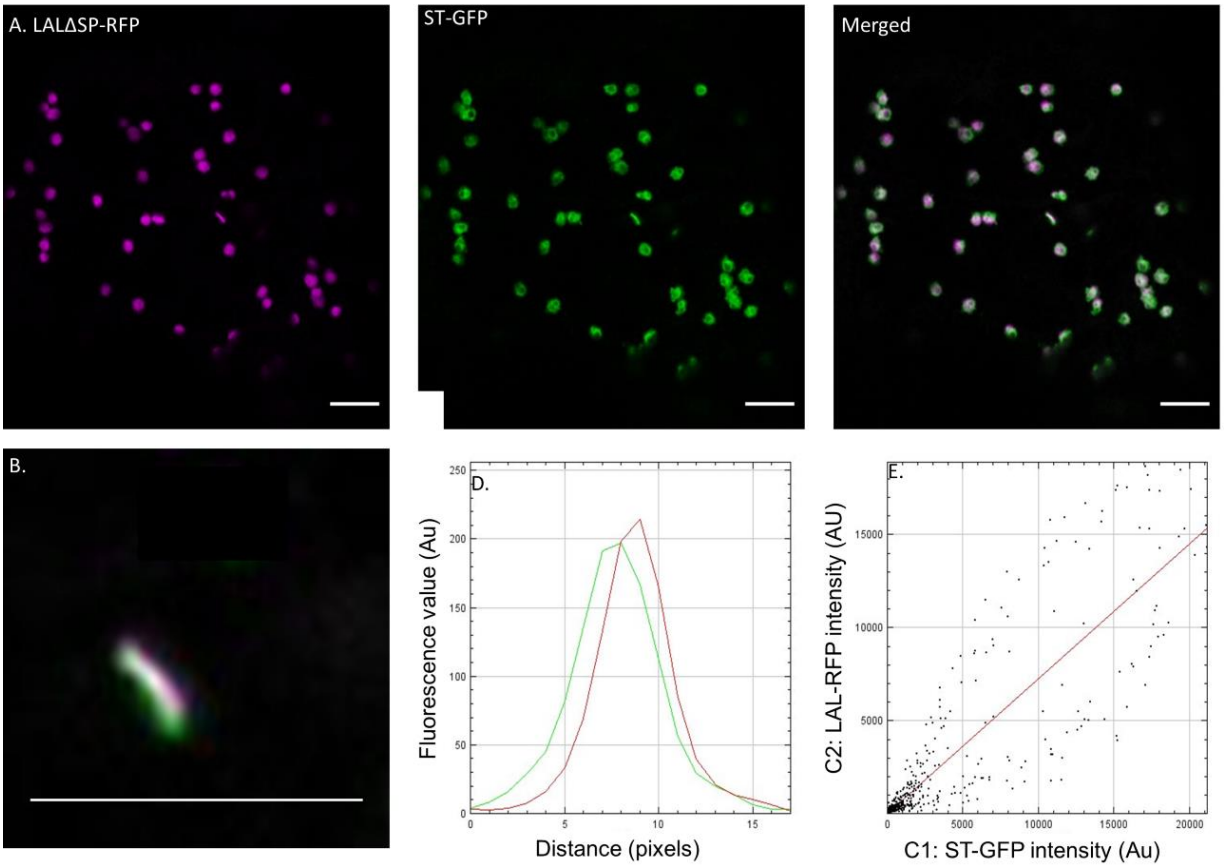


Figure 12: Co-localisation analysis of A) original sequence human LAL and B) LALΔSP. Scale bars= 5 µm. A) Co-expression of LAL-RFP and ST-GFP indicates localisation of the two constructs with overlap. B) Inset shows enlarged structures for co-localisation analysis, R value n=1 as generated from individual Golgi body shown in inset image, minimum 3 representative Golgi body images samples displaying comparable localisation. D) Signal analysis with the RGB Profiler plugin for Fiji, indicating curves for signal co-localisation. E) Cytofluorogram for indicates signals convergence and so constructs do co-express, Pearson's coefficient r = 0.87. Representative images are shown; n=3, with 5 technical repeats each.

In order to quantitatively assess improvements in expression, cell counts were performed using a 10x objective lens to assess the proportion of cells expressing LALΔSP (Figure 13). Visual cell counts were undertaken of cells clearly expressing the LALΔSP construct (Figure 13, magenta bordered) and of cells presenting the marker protein ST-GFP (Figure 13, Green bordered) within the given field of view, for a sample number of 34 images from three separate biological samples. As the original sequence LAL construct expressed exceptionally poorly, surface cell counts could not be accurately conducted to enable comparison within the duration of the project, but should be included in any future work (section 4.4). Of 568 cells expressing infiltrated constructs, 447 expressed only ST-GFP and 121 expressed also LALΔSP, an average of approximately 27% of cells actively expressing LALΔSP. The removal of the signal peptide therefore results in increased transformation efficiency and protein expression levels.

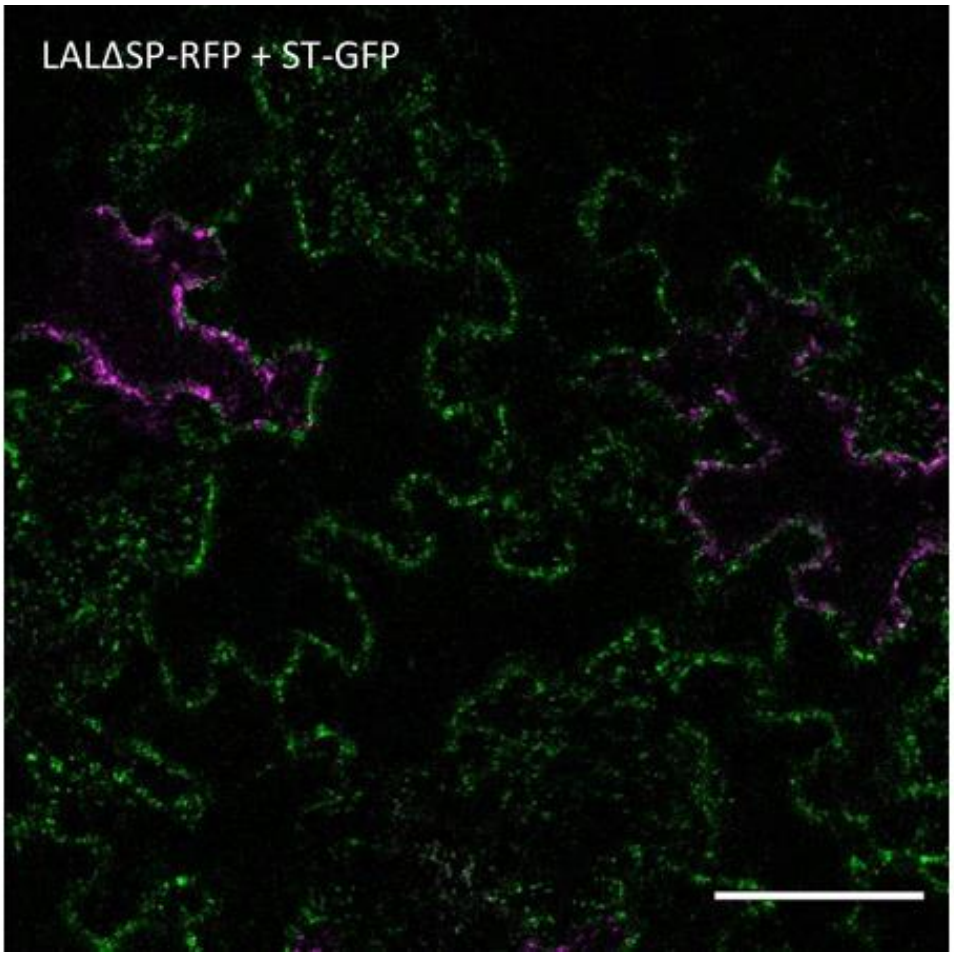


Figure 13: Representative image for leaf surface under 10x objective zoom, merged image of cells expressing LALΔSP-RFP (magenta) construct alongside a ST-GFP (green) *medial/trans*-Golgi marker. Scale bar = 100 µm. n=3 with 10 technical replicas each.

4. Discussion

The aim of this project was to work towards a proof-of-concept plant-based system for the production of humanized protein, utilizing LAL as an example protein. Several objectives directed towards this have been achieved within the scope of this research. Firstly the transient expression of mammalian glycosylation enzymes in tobacco leaf cells has been achieved, and a *medial/trans*-Golgi cisternae specific targeting strategy implemented for B4GALT1. Secondly the expression and analysis of the subcellular localisation of LAL in tobacco leaf cells has been carried out and a strategy implemented to enhance LAL production.

4.1 Protein expression with fluorophore fusions

To study the subcellular and suborganellar localisation of LAL and the mammalian glycosylation enzymes the proteins were fused to fluorophores to allow visualization. For the glycosylation enzymes, both N- and C-terminal GFP/ RFP fluorophore fusions were tested.. The fusion terminus did not affect localisation significantly, but whereas RFP fusions localised to the Golgi body (Figure 7, 8), GFP-fusions were less successful, staying in the ER or localising otherwise outside of the Golgi body (Figure 6 and appendices 1). This could potentially be due to the GFP constructs having higher expression levels and the produced protein not being transported fast enough to the Golgi body. However, as imaging was performed 3 days after infiltration and the marker ST is reaching the Golgi body, this is unlikely. An additional possibility is that while GFP is prone to forming noncovalent dimers, RFP (here mRFP) is a monomer and this difference could impact on protein conformation and targeting mechanisms (Campbell et al., 2002).

In an established cell expression system expression of LAL without fluorophores would be preferential without fluorescent tags as these would have to be removed before LAL could be used for ERT resulting in additional cost and potential differences between batches. All mammalian glycosylation enzymes within this project were preferentially given C-terminal fluorophore tags to asses localisation, as they are all type II membrane proteins, and natively possess a N-terminal signal-anchor domain that directs localisation within the mammalian Golgi and acts as a membrane anchor for the embedded enzyme (Saint-jore-Dupas et al., 2006). Use of a C-terminal fluorescent tag would therefore limit interference a N-terminal fluorescent tag could have on the original targeting domain. Enzymes were expressed as fluorescent protein fusions for the purposes of localisation, any assays on enzyme actively would be performed on untagged constructs.

4.2 Signal peptide targeting of expression

Initial confocal investigation into sub-organellar targeting of the mammalian glycosylation enzymes indicated expression distinct from the intended sites of localisation (GNTIV & GNTV to the *medial*-Golgi, ST6GAL and B4GALT1 to the *medial/trans*-Golgi). Literature described the implementation of several targeting options to alter the localisation of B4GALT1(Vézina et al., 2009; Schoberer et al., 2011). Firstly, a fusion of B4GALT1 with a xylosyltransferase transmembrane domain derived from *Arabidopsis thaliana* for targeting to the *medial*-Golgi was suggested (Vézina et al., 2009), and secondly a rat α2,6 sialyltransferase signal peptide for the purposes of targeting exogenous enzymes to the *medial/trans*-Golgi (Schoberer et al., 2011). As research suggested that *medial/trans*-Golgi expression of B4GALT1 was most effective, a sialyltransferase signal peptide was adapted from the N-terminus of ST6GAL, as this had already demonstrated effective localisation to the *medial/trans*-Golgi (Figure 8). Results indicated this strategy was effective for targeting expression of B4GALT1SP to the *medial/trans*-Golgi (Figure 9).

An additional Golgi-targeting system may increase efficacy for the enzymes GNTIV and GNTV. Literature (Bosch et al., 2013; Castilho et al., 2013) suggests the targeting of these enzymes to the *medial*-Golgi generates a greater yield of bi- and tri-antennary glycan structures (Figure 2 & table 3) than expression in the *trans*-Golgi. Expression of GNTIV and GNTV has to be analysed with a *medial*-Golgi marker such as FUT11. If the localisation requires changing to the *medial*-Golgi an existing strategy for this is fusing the GNTV catalytic domain to the cytoplasmic tail of FUT11 or XylT (Castilho et al., 2011).

4.3 LAL protein expression in tobacco cells

Initial infiltrations of human LAL as a fluorescent protein fusion had shown successful expression and localisation to Golgi bodies (Figure 7); but low transformation efficiency of the construct posed an issue to achieving efficient production of significant yields. This was problematic for a proof of concept or the downstream conversion of this system to a successful production vector as, in order for production to be commercially viable, transformants need to be expressing LAL in a majority of cells at high levels. A low transformation efficiency would increase costs of purification and processing of plant tissue to obtain product, decreasing commercial viability of this system. A potential intervention to resolve this issue was identified through aa sequence analysis of the original human LAL structure, which identified a signal peptide site between aa 1-22.

Such a signal peptide is a common feature on precursor polypeptides of soluble lysosomal proteins, and is responsible in part for translocation of LAL to the lumen of the mammalian ER in non-deficient individuals (Li et al., 2019). Within mammalian cells, LAL undergoes distinct signal peptide cleavage steps and a series of post-translational modifications in order to generate a mature lipase capable of translocation from the ER to the lysosomal lumen (Li et al., 2019). The translated LAL protein is a 399 aa sequence, with an initial 27 aa signal peptide being cleaved as part of ER translocation and subsequent signal peptidase cleavage. The resulting 372 aa pro-protein undergoes co/post-translational glycosylation within the ER before transport to the Golgi for addition of mannose 6-phosphate (M6P) residues, which facilitate lysosomal targeting of the mature LAL. Specific proteolytic cleavage of the first 49 aa from the N-terminus releases the 323 aa enzymatically active mature lipase, the site of action for this is unknown (Li et al., 2019). However, the generation of exogenous LAL for use in human intravenous enzyme replacement therapy only requires presence of the M6P residues, which are attached during glycosylation. Presence of M6P facilitate binding of exogenous LAL to the M6P receptor and subsequent localisation to the lysosome, where the M6P receptor dissociates from LAL to generate a mature lipase (Gomaraschi et al., 2019). Presence of signal peptides do not play a functional role on the therapeutic targeting of exogenous LAL delivered as ERT.

Within a plant system there is a possibility correctly translated LAL may not undergo complete cleavage of the initial 27 aa signal peptide or proteolytic processing of the 49 aa sequence necessary to generate mature enzymatically active LAL. As the described translocation and cleavage activity is specific to mammalian cells, this was unlikely to have any functional role within the plant host. This was supported by the findings of this thesis as removal of the signal peptide from LAL resulted in an increased expression efficiency without disrupting localisation (Figure 10, 12). Low expression of the original human LAL construct may be attributed to unnecessary clearance or inactivation caused by presence of the mammalian signal peptide. Differences in the ER environment or localisation of original LAL could lead to nonfunctional protein or poor levels of expression. Post-translational glycomodification within the plant Golgi would be essential for generating a therapeutically effective recombinant human LAL capable of uptake by the human lysosome, which would be achieved by co-expression of the proposed mammalian glycosylation enzymes to generate essential M6P structures.

4.4 Future work

Following the signal protein targeting strategy for B4GALT1SP, targeting strategies for the glycosylation enzymes will be expanded to include targeting GNTIV and GNTV to the *medial*-Golgi. This will provide proof of an effective strategy for expressing each of the exogenous mammalian enzymes in the desired Golgi cisternae, and would further validate the efficacy of this system. An addendum to this work would be a dual-marker quantitative analysis of localisation for all constructs expressing the glycosylation enzymes together with both *medial*/*trans*- (ST) or *medial*- (FUT11) and *cis*- (MNS1) markers for suborganellar localisation within the Golgi body, and later attempting to express multiple glycosylation enzymes together.

Future work on LAL would include investigating expression levels for industrial production, including strategies for protein quantification in order to investigate potential yields of LAL, and the enzymatic activity of synthesized product. The first step would be conducting quantification assays such as Western blots or quantitative ELISA for protein abundance to ensure a detectable level of LAL synthesis, with the downstream goal of producing sufficient protein for enzyme activity assays. This could be achieved through a combination of embedded enzyme assays to measure objective catalytic activity on cholesterol esters, and structural analysis via mass spectrometry (Vézina et al., 2009). In tandem with this would be strategies for the extraction of LAL from plant tissue without losing catalytic activity, and strategies for improving this plant system itself to facilitate this. It is worth mention that any continuation of work on this aspect of the product should include the same quantitative assessment of the expression of unmodified LAL as was performed for LALΔSP in section 3.2 (Figure 13) to enable direct comparison between the two, as this was not completed for full length human LAL during the initial project.

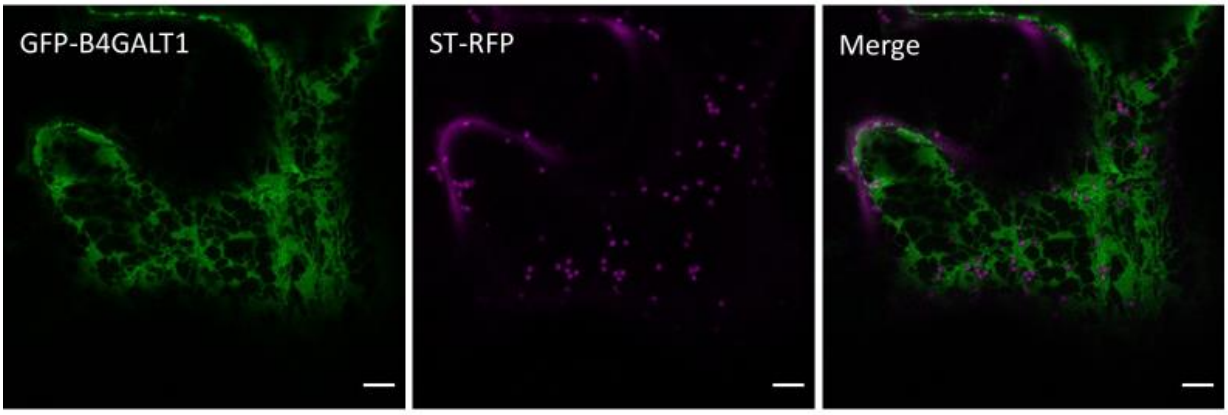
Further investigations of the glycosylation state and activity of synthesized LAL will be an essential step in validating this platform as viable for ERT application. Determinants of therapeutic efficacy will include specific N-linked glycan structures. Addition of M6P as part of N-linked glycosylation within the modified Golgi will be essential for end-target therapeutic effects, as this structure is essential for receptor-mediated endocytosis of extracellular LAL into the lysosome and subsequent generation of the mature lipase (Li et al., 2019; Gomaraschi et al., 2019)). Specific targeting of exogenous LAL to the lysosome via the macrophage mannose receptor on reticuloendothelial cells alongside the M6P receptor is an essential feature of the current LAL ERT Sebelipase Alfa© for LAL-D (Su et al., 2016), and any recombinant LAL would need to replicate this effect.

Several alternative strategies of targeted drug delivery to the lysosome for lysosomal storage disorders have been discussed in previous research, and may warrant further investigation to improve efficacy of LAL delivery and ERT for lysosomal storage disorders in general (Desnick et al., 2012). Use of octadecyl-rhodamine B modified liposomes have been suggested as a viable pharmaceutical carrier for the delivery of ERT to the lysosome for the treatment of lysosomal storage disorders, with the potential to alleviate low stability and poor delivery of therapeutic exogenous enzymes to resistant tissue sites (Koshkaryev et al., 2011). Alternative research has been undertaken into the use of a chimeric protein of the insulin-like growth factor II (IGF-II) fused to ERT cargo for delivery to the corresponding binding site on the bifunctional IGF-II/M6P receptor, and a second strategy making use of the 11aa human immunodeficiency Tat domain fused to ERT cargo for delivery via absorptive endocytosis (Grubb et al., 2010). Alternative targeting strategies warrant further investigation for application in treatment-resistant patients and specific tissues (i.e. impermeability of blood brain barrier) but also towards wider applicability of this system.

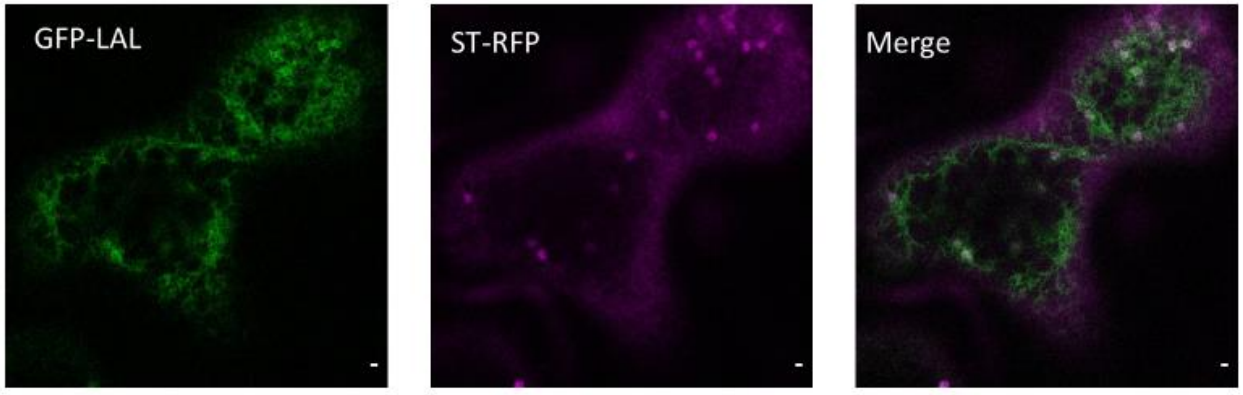
Effective production of terminal sialylated proteins in this system likely will require co-expression of the Golgi CMP-Sialic acid transporter (CST) and the necessary precursor feed for CMP-sialic acid biosynthesis required (Schoberer et al., 2018), and would additionally require further investigation into inducible expression of multiple recombinant enzymes within the *medial/trans*-Golgi cisternae.

Strasser et al. (2008) discusses the endogenous activity of β1,2-xylotransferase (XylT) and α1,3-fucosyltransferase (FucT) in generation of potentially immunogenic glycans and the downregulation of their activity via RNAi interference, which may be a viable strategy for long term risk reduction within this system. The acceptance of plant production systems has in the past been hampered by concerns on immunogenic hypersensitivity reactions to these glycan structures, especially in the context of therapeutic recombinant proteins. Additional concerns over the overall therapeutic efficacy of therapeutics bearing these glycans are warranted, as antibodies specific to these glycans may lead to rapid immune clearance of plant-made proteins bearing these glycosylation motifs, limiting therapeutic effect of product synthesized in this system (Vézina et al., 2009). However, it may be of theoretical interest to assess the glycosylation motifs present on LAL synthesized within a wild-type system to assess the structural impacts these impart.

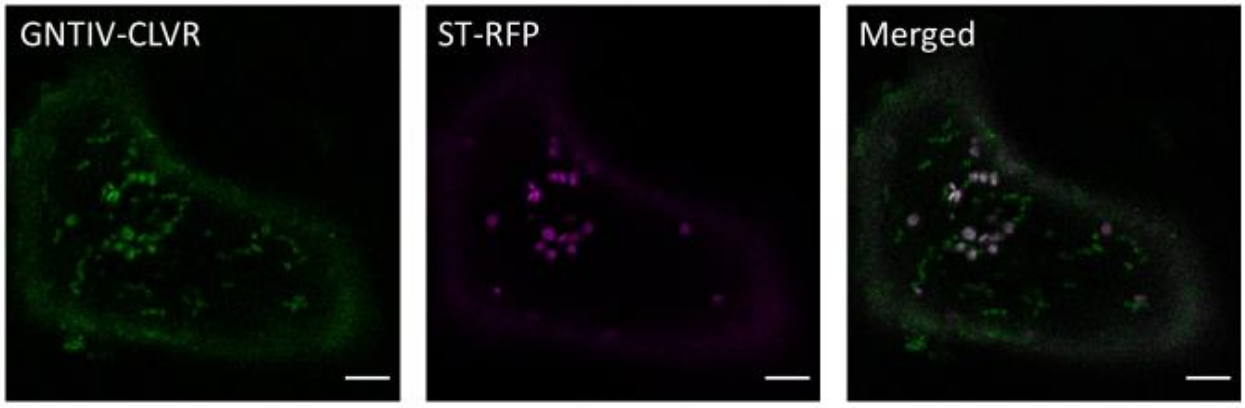
Appendices: Off target expression of fluorophores



Appendices figure 1: Identical to figure 6. scale bar = 5 μm. Expressing *Medial/Trans* Golgi *marker* ST-RFP (magenta) construct alongside a GFP-B4GALT1 (green)



Appendices figure 2: Identical to figure 6. scale bar = 1 μm. Expressing *Medial/Trans* Golgi marker ST-RFP (magenta) construct alongside a GFP-LAL (green)



Appendices figure 3: Identical to figure 6. scale bar = 5 μm. Expressing *Medial/Trans* Golgi marker ST-RFP (magenta) construct alongside a GNTIV-CLVR (green)

Glossary

Abbreviation	Term	Definition
LAL	Lysosomal Acid Lipase	Essential human lysosomal enzyme for metabolism of dietary fats, LAL for enzyme replacement therapy (ERT) serves as a proof of concept for plant systems
FA	Fatty acids	Product of LAL hydrolysis on dietary lipids
TG	Triglyceride	Dietary lipid, substrate for LAL
CE	Cholesterol esters	Dietary lipid, substrate for LAL
LAL-D	Lysosomal Acid Lipase Deficiency	Dysfunction of *LIPA* gene leading to lack of endogenous LAL leading to disease, either Wolman's or CESD
CESD	Cholesterol Ester Storage Disorder	Metabolic dysfunction causes by partial inactivity of *LIPA* gene/lack of endogenous LAL activity leading to cellular lipid accumulation, spectrum disease with varying severity
B4GALT1	Beta-1,4- Galactosyltransferase 1	Human post-translational N-linked glycosylation enzyme. Type II membrane bound protein, transfer of GlcNAc onto several acceptor substrates

ST6GAL	Alpha 2,6-sialyltransferase	Human post-translational N-linked glycosylation enzyme. Type II membrane bound protein, catalyses transfer of sialic acid from cmp-sialic precursors onto galactose containing substrate sites
GNTIV	N-Acetylglucosaminyltransferase-IV	Human post-translational N-linked glycosylation enzyme. Type II membrane bound protein, responsible for production of tetra-antennary sugar chains and subsequent addition of galactose residues for later sialic acid residue addition.
GNTV	N-Acetylglucosaminyltransferase-V	Human post-translational N-linked glycosylation enzyme. Type II membrane bound protein, responsible for production of tetra-antennary sugar chains and subsequent addition of galactose residues for later sialic acid residue addition.
ST	Sialyltransferase	enzymes that transfer sialic acid to nascent oligosaccharide
MNS1	Golgi-α-mannosidase I	key enzyme in N-glycan processing
PTM	Post Translational Modification	the covalent and generally enzymatic modification of proteins following protein biosynthesis
GFP	Green Fluorescent protein	*Protein* composed of 238 amino acid residues (26.9 kDa) that exhibits bright *green fluorescence* when exposed to light in the blue to ultraviolet range.
RFP	Red Fluorescent protein	fluorophore that fluoresces red-orange when excited

TGN	Trans Golgi Network	Series of interconnected tubules and vesicles at the *trans* face of the Golgi stack. The *medial/trans*-Golgi network functions in the processing and sorting of glycoproteins and glycolipids at the interface of the biosynthetic and endosomal pathways.
ER	Endoplasmic reticulum	Organelle comprised by a network of membrane tubules continuous with the nuclear membrane and site of ribosome attachment, major site of protein and lipid synthesis in Eukaryotic cells
XylT	β1,2-xylosyltransferase	Endogenous plant specific glycosylation enzymes involved in generation of β1,2 xylose residues
FUT11	α1,3-fucosyltransferase	Endogenous plant specific glycosylation enzyme involved in generation of core α1,3-fucose residues
UDP-GlcNAc	uDP-*N*-acetylglucosamine	Substrate form of N-Acetylglucosamine for attachment via the GnT family of enzymes
COPI	Coat Protein Complex I	Protein complex that coats vesicles transporting proteins from the cis end of the Golgi complex back to the endoplasmic reticulum.
NEB	New England Biolabs	Provider from some materials

References

Aguisanda F, Thorne N, Zheng W, Targeting Wolman Disease and Cholesteryl Ester Storage Disease: Disease Pathogenesis and Therapeutic Development, *Current Chemical Genomics and Translational Medicine,* 2017, 30:1-18

Balwani M, Breen C, Enns GM, Deegan PB, Honzík T, Jones S, Kane JP, Malinova V, Sharma R, Stock EO, Valayannopoulos V, Wraith JE, Burg J, Eckert S, Schneider E, Quinn AG, Clinical effect and safety profile of recombinant human lysosomal acid lipase in patients with cholesteryl ester storage disease, *Hepatology*, 2013, 58:950-957

Bardor M, Faveeuw C, Fitchette AC, Gilbert D, Galas L, Trottein F, Faye L, Lerouge P, Immunoreactivity in mammals of two typical plant glyco-epitopes, core α(1,3)-fucose and core xylose. *Glycobiology*, 2003, 13:427-434

Bolte S, Cordelières FP, A guided tour into subcellular co-localization analysis in light microscopy, *Journal of Microscopy*, 2006, 224:213-232

Bosch D, Castilho A, Loos A, Schots A, Steinkellner H, N-glycosylation of plant-produced recombinant proteins, *Current Pharmaceutical Design*, 2013, 19:5503-5512

Buyel J, Plant Molecular Farming - Integration and Exploitation of Side Streams to Achieve Sustainable Biomanufacturing, *Frontiers in Plant Science*, 2019, 18:1893

Campbell R, Tour O, Palmer A, Steinbach P, Baird G, Zacharias D, Tsien R, A monomeric red fluorescent protein, *Proceedings of the National Academy of Sciences*, 2002, 99:7877-7882

Castilho A, Gattinger P, Grass J, Jez J, Pabst M, Altmann F, Gorfer M, Strasser R, Steinkellner H, N-Glycosylation engineering of plants for the biosynthesis of glycoproteins with bisected and branched complex *N*-glycans, *Glycobiology*, 2011, 21:813–823

Castilho A, Neumann L, Gattinger P, Strasser R, Vorauer-Uhl K, Sterovsky T, Altmann F, Steinkellner H, Generation of biologically active multi-sialylated recombinant human EPOFc in plants, *PLOS ONE*, 2013, 8:e54836

Chen Q, Santi L, Zhang C, Plant-made biologics, *BioMed Research International*, 2014, PMC:418064

Chen Q, Davis KR, The potential of plants as a system for the development and production of human biologics, *F1000Research,* 2016, 5:912

Desnick R, Schuchman E, Enzyme Replacement Therapy for Lysosomal Diseases: Lessons from 20 Years of Experience and Remaining Challenges, *Annual Review of Genomics and Human Genetic*s ,2012, 13:307-335

Du H, Cameron TL, Garger SJ, Pogue GP, Hamm LA, White E, Hanley KM, Grabowski GA, Wolman disease/cholesteryl ester storage disease: efficacy of plant-produced human lysosomal acid lipase in mice, *Journal of Lipid Research*, 2008, 49:1646-1657

Geuze H, Slot J, Strous G, Hasilik A, von Figura K, Possible pathways for lysosomal enzyme delivery, *Journal of Cell Biology,* 1985, 101:2253-62

Gomes A, Byregowda S, Veeregowda B, Vinayagamurthy B, An overview of heterologous expression host systems for the production of recombinant proteins. *Advances in Animal and Veterinary Sciences*, 2016, 4:346-356

Gomord V, Fitchette A.-C, Menu-Bouaouiche L, Saint-Jore-Dupas C, Plasson C, Michaud D, Faye L, Plant-specific glycosylation patterns in the context of therapeutic protein production*, Plant Biotechnology Journal,* 2010, 8:564-587

Gomaraschi M, Bonacina F, Norata G, Lysosomal Acid Lipase: From Cellular Lipid Handler to Immunometabolic Target, *Trends in Pharmacological Sciences*, 2019, 40: 104-11

Grabowski G, Golembo M, Shaaltiel Y, Taliglucerase alfa: An enzyme replacement therapy using plant cell expression technology. *Molecular Genetics and Metabolism*, 2014, 112:1-8

Grubb J, Vogler C, Sly W, New strategies for enzyme replacement therapy for lysosomal storage diseases, *Rejuvenation Research* 2010, 13:229-36

Hamilton B, Wilson J, Shumakovitch M, Fisher A, Brooks J, Pontes A, Naran R, Heiss C, Gao C, Kardish R, Heimburg-Molinaro J, Azadi P, Cummings R, Merritt J, Delisa M, A library of chemically defined human N-glycans synthesized from microbial oligosaccharide precursors. *Scientific Reports*, 2017, 21:15907

Hüttner S, Veit C, Schoberer J, Grass J, Strasser R, Unraveling the function of Arabidopsis thaliana OS9 in the endoplasmic reticulum-associated degradation of glycoproteins, *Plant Molecular Biology,* 2012, 79:21-33

Ito Y, Uemura T, Shoda K, Fujimoto M, Ueda T, Nakano A, *cis*-Golgi proteins accumulate near the ER exit sites and act as the scaffold for Golgi regeneration after brefeldin A treatment in tobacco BY-2 cells. *Molecular Biology of the Cell*, 2012, 23:3203–3214

Kim HS, Jeon JH, Lee KJ, Ko K, N-glycosylation modification of plant-derived virus-like particles: an application in vaccines, *BioMed Research International*, 2014, 2014:249519

Kizhner T, Azulay Y, Hainrichson M, Tekoah Y, Arvatz G, Shulman A, Ruderfer I, Aviezer D, Shaaltiel Y, Characterization of a chemically modified plant cell culture expressed human α-Galactosidase-A enzyme for treatment of Fabry disease, *Molecular Genetics and Metabolism*, 2015, 114:259-267

Koshkaryev A, Thekkedath R, Pagano C, Meerovich I, & Torchilin V P, Targeting of lysosomes by liposomes modified with octadecyl-rhodamine B, *Journal of drug targeting*, 2011, 19: 606–614

Laummonerie C, Mutterer J, RGB profiler, *IBMP-CNRS Strasbourg,* 2004

Li F, Zhang H, Lysosomal Acid Lipase in Lipid Metabolism and Beyond, *Arteriosclerosis, Thrombosis, and Vascular Biology,* 2019, 39:850-856

Liebminger E, Hüttner S, Vavra U, Fischl R, Schoberer J, Grass J, Blaukopf C, Seifert GJ, Altmann F, Mach L, Strasser R, Class I alpha-mannosidases are required for N-glycan processing and root development in *Arabidopsis thaliana*. *Plant Cell*, 2009, 21:3850–3867

McDonald A, Hayes J, Bezak T, Głuchowska S, Cosgrave E, Struwe W, Stroop C, Kok H, van de Laar T, Rudd P, Tipton K, Davey G, Galactosyltransferase 4 is a major control point for glycan branching in N-linked glycosylation. *Journal of Cell Sci*ence, 2014, 127:5014–5026

Moon KB, Park JS, Park YI, Song IJ, Lee HJ, Cho HS, Jeon JH, Kim HS. Development of Systems for the Production of Plant-Derived Biopharmaceuticals, *Plants*, 2019, 9:30

Nagels B, Van Damme EJ, Pabst M, Callewaert N, Weterings K, Production of complex multiantennary N-glycans in *Nicotiana benthamiana* plants. *Plant Physiology*, 2011, 155:1103–1112

Pigino G, Morfini G, Brady S, Chapter 7 - Intracellular Trafficking, Basic Neurochemistry (Eighth Edition). *Academic Press*, 2012, 119-145

Piron R, Santens F, De Paepe A, Depicker A, Callewaart N, Using GlycoDelete to produce proteins lacking plant-specific *N*-glycan modification in seeds. *Nature Biotechnology, 2015,* 33:1135–1137

Ree R, Cabanes-Macheteau M, Akkerdaas J, Milazzo J, Loutelier-Bourhis C, Rayon C, Villalba M, Koppelman S, Aalberse R, Rodriguez R, Faye L, Lerouge P, β(1,2)-Xylose and α(1,3)-Fucose residues have a strong contribution in IgE binding to plant glycoallergens. *Journal of Biological Chemistry,* 2000, 275:11451-11458

Rhee SW, Starr T, Forsten-Williams K, Storrie B, The steady-state distribution of glycosyltransferases between the Golgi apparatus and the endoplasmic reticulum is approximately 90:10. *Traffic*, 2005, 6:978-990

Rup B, Alon S, Amit-Cohen BC, Brill Almon E, Chertkoff R, Immunogenicity of glycans on biotherapeutic drugs produced in plant expression systems—The taliglucerase alfa story. *PLOS ONE,* 2017, 12: e0186211

Robinson DG, Plant Golgi ultrastructure, *Journal of Microscopy*, 2020, doi: 10.1111/jmi.12899

Saint-Jore-Dupas C, Nebenführ A, Boulaflous A, Follet-Gueye M, Plasson C, Hawes C, Driouich A, Faye L, Gomord V, Plant N-Glycan Processing Enzymes Employ Different Targeting Mechanisms for Their Spatial Arrangement along the Secretory Pathway, *The Plant Cell*, 2006, 18:3182-3200

Sack M, Hofbauer A, Fischer R, Stoger E, The increasing value of plant-made proteins. *Current Opinion Biotechnol*ogy, 2015, 32:163–170

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

Schoberer J, Strasser R, Sub-Compartmental Organization of Golgi-Resident N-Glycan Processing Enzymes in Plants, *Molecular Plant,* 2011, 4:220-228.

Schoberer J, Strasser R, Plant glyco-biotechnology, *Seminars in Cell and Developmental Biology*, 2018, 80:133-141

Schoberer J, Liebminger E, Botchway SW, Strasser R, Hawes C, Time-resolved fluorescence imaging reveals differential interactions of N-glycan processing enzymes across the Golgi stack in planta, *Plant Physiology,* 2013, 161:1737-1754

Schoberer J, König J, Veit C, Vavra U, Liebminger E, Botchway SW, Altmann F, Kriechbaumer V, Hawes C, Strasser R, A signal motif retains Arabidopsis ER-α-mannosidase I in the *cis*-Golgi and prevents enhanced glycoprotein ERAD. *Nature Communications,* 2019, 10:3701

Shaaltiel Y, Tekoah Y, Plant specific *N*-glycans do not have proven adverse effects in humans, *Nature Biotechnology,* 2016, 34:706–708

Strasser R, Stadlmann J, Schähs M, Stiegler G, Quendler H, Mach L, Glössl J, Weterings K, Pabst M, Steinkellner H, Generation of glyco-engineered *Nicotiana benthamiana* for the production of monoclonal antibodies with a homogeneous human-like *N*-glycan structure, *Plant Biotechnology Journal*, 2008, 6:392-402

Strasser R, Biological significance of complex *N*-glycans in plants and their impact on plant physiology, *Frontiers in Plant Science,* 2014, 5:363

Strebinger G, Müller E, Feldman A, Aigner E, Lysosomal acid lipase deficiency - early diagnosis is the key, *Hepatic Medicine,* 2019, 11:79-88

Su K, Donaldson E, Sharma R, Novel treatment options for lysosomal acid lipase deficiency: critical appraisal of sebelipase alfa, *The Application of Clinical Genetics*, 2016, 9:157-167

Thelwall P, Smith F, Leavitt M, Canty D, Hu W, Hollingsworth K, Thoma C, Trenell M, Taylor R, Rutkowski J, Blamire A, Quinn A, Hepatic cholesteryl ester accumulation in lysosomal acid lipase deficiency: non-invasive identification and treatment monitoring by magnetic resonance, *Journal of Hepatology*, 2013, 59:543-549

Vézina LP, Faye L, Lerouge P, D'Aoust MA, Marquet-Blouin E, Burel C, Lavoie PO, Bardor M, Gomord V, Transient co-expression for fast and high-yield production of antibodies with human-like N-glycans in plants, *Plant Biotechnology Journal*, 2009, 7:442-455

Ward B, Landry N, Trépanier S, Mercier G, Dargis M, Couture M, D'Aoust MA, Vézina LP, Human antibody response to N-glycans present on plant-made influenza virus-like particle (VLP) vaccines. *Vaccine*, 2014, 32:6098-6106

Xu J, Dolan M, Medrano G, Cramer C, Weathers P, Green factory: Plants as bioproduction platforms for recombinant proteins, *Biotechnology Advances*, 2012, 30:1171-1184

Zimran A, Brill-Almon E, Chertkoff R, Petakov M, Blanco-Favela F, Muñoz ET, Solorio-Meza SE, Amato D, Duran G, Giona F, Heitner R, Rosenbaum H, Giraldo P, Mehta A, Park G, Phillips M, Elstein D, Altarescu G, Szleifer M, Hashmueli S, Aviezer D, Pivotal trial with plant cell-expressed recombinant glucocerebrosidase, taliglucerase alfa, a novel enzyme replacement therapy for Gaucher disease, *Blood*, 2011 24:5767-5773

Zimran A, Durán G, Mehta A, Giraldo P, Rosenbaum H, Giona F, Amato DJ, Petakov M, Muñoz ET, Solorio-Meza SE, Cooper PA, Varughese S, Chertkoff R, Brill-Almon E, Long-term efficacy

and safety results of taliglucerase alfa up to 36 months in adult treatment-naïve patients with Gaucher disease, *Americal Journal of Hematology*, 2016, 91:656-660