Degree awarded by Oxford Brookes University

School of Biological and Medical Sciences

Oxford Interdisciplinary Bioscience BBSRC DTP

Modifying N-linked glycosylation in tobacco to produce Human-type post-translational modifications

---

This thesis is submitted in partial fulfilment
of the requirements of the award of Doctor of Philosophy

January 2025

Author:

Alastair McGinness

Supervisors:

Verena Kriechbaumer[1], Susan Brooks[1], Richard Strasser[2]

1. Oxford Brookes University, Headington, OX3 0BP
2. BOKU University of Natural Resources and Life Sciences, 1180 Wien, Austria

Author's Declaration

I, Alastair James McGinness, declare that this thesis titled "Modifying N-linked glycosylation in tobacco to produce Human-type post-translational modifications" and submitted for assessment is an original work of my own. Any works of others used within this thesis, or work presented in it are properly acknowledged. A list of references employed is included.

Abstract

Plant cells are a capable system for producing high value proteins for a variety of applications, and challenge popular host organisms such as mammalian cells, bacteria or yeasts. However, plants do not perform post translation protein modification in the same manner as mammalian cells, most critically in terms of N-linked glycosylation. This can impact both protein functionality and stability, key for human therapeutics intended for production in plants. This obstacle can be approached by creating a plant-based system capable of “humanizing” proteins of interest in terms of N-linked glycosylation, resulting in plant-produced proteins with a glycan profile as would occur in a mammalian host system, making them more suitable for therapeutic applications.

For this, four human glycosylation enzymes (HuGEs) involved in N-linked glycosylation N- acetylglucosaminyltransferase IV and V (GNTIV and GNTV), β-1,4-galactosyltransferase (B4GALT1), and α-2,6-sialyltransferase (ST6GAL) need to be expressed in plant cells. As glycosylation is a stepwise process, enzymes need to localise to either medial or *trans*-Golgi body cisternae. For this, a protein targeting strategy of replacing the mammalian N-terminal Golgi targeting domains on HuGEs (Cytoplasmic-Transmembrane-Stem (CTS) regions) with plant–specific domains, MUR3 or FUT13, were tested transiently and were successful in redirecting localisation. This was conducted via high-resolution dynamic confocal microscopy using an analysis pipeline based upon distance between peak fluorescence intensity, relative to known Golgi markers MNS1 and ST. This analysis demonstrated that the modified MUR3-GNTIV and MUR3-GNTV were successfully targeted to the medial-Golgi cisternae while FUT13-ST6GAL and FUT13-B4GALT1 were targeted to *trans*-Golgi cisternae.

Following successful expression and targeting of the four HuGEs transiently, these four enzymes were combined into two expression cassettes for stable plant transformation. These consisted of a N-terminal CTS domain, the catalytic subunit of the first enzyme, a linker domain, and the second enzyme catalytic subunit without attached fluorescent protein. This halved the insert number necessary to generate stable glycomodified lines, transformed using the modified HuGEs in wild type and plant-glycan deficient (ΔXFT) lines of *N. benthamiana.* Following growth, resulting seedlings of this line were successfully screened for presence of each of the four HuGE catalytic domains by RT-PCR against a 300BP sequence specific to each enzyme. Human proteins of interest LAL, IgE and IgG were then transiently expressed in these glycomodified lines and lectin binding assays performed against the plant extracts, either dot blot or by SDS-PAGE. The results of the lectin binding assays potentially indicate altered glycomodification in the HuGE stable lines but are ultimately inconclusive.

Acknowledgements

I cannot thank everyone around me during this PhD enough for their tireless input, encouragement, support and patience throughout this project. First thanks go to Verena Kriechbaumer, who has been the supervisor most PhDs only hope for, and so the last few years have been a delight accordingly. Thank you for your incredible academic input, dedication into my development as a researcher, and occasionally almost-too-honest feedback. I would like to thank Susan Brookes for her fresh perspective and technical insight into the project, and general unflinching positivity. I would especially like to thank Richard Strasser for his keen insight into all things plant glycosylation, and for use of the ΔXFT *N. benthamiana* line during its course. I could not be more thankful to have had such incredible support.

I would like to thank everyone in the plant and parasite labs of the Sinclair annex of the last few years for the excellent company. It's been a marvellous time. Special thanks go to Stefan Wojcik for introducing me to the group, Charlotte Pain for taking me under her wing in the lab all those years ago and being a tireless sounding board for all things science since, including lab sing-alongs. Yet more thanks to Joanna Cull as a BBSRC partner in crime, Tatiana Spatola Rossi for putting up with me, Bisa Andov as lab nonna, and Nadine Field as my lab twin. I look forward to what the next generation of the plant Endomembrane lab achieves, and hope someone can read the handwriting I left behind.

I couldn't have completed this project without my incredible network of friends. Luke-Lee Brewin who successfully crossed physics to collaborate, Leah Fitzpatrick as my daily sanity check and the other half of double act here at Brookes. Babs to no small praise, and the Surrey MMA girls; absolutely my circus, absolutely my monkeys. For my Oxford friends, Leo, Emma, Pepe and Brett; we managed, good work

Thanks go out to my family for listening patiently while I explained what a Golgi was over the last few years and not asking when I'll get a job too often. Love to my siblings, mum and Terry, thanks for being there every single time. Special dedication in memory of Rose, always on my side, who didn't get to see this finished.

I would like to thank the BBSRC for support, and particularly the staff of the oxford DTP for the excellent experience that was my PhD.

Glossary

B4GALT1	β -1,4- Galactosyltransferase 1
BFP	Blue fluorescent protein
CEs	Cholesterol esters
CFP	Cyan fluorescent protein
CLVR	CLOVER fluorescent protein
CST	CMP-Sialic acid transporter
CTS	Cytoplasmic-transmembrane-stem region
DNA	Deoxyribonucleic acid
ELISA	enzyme-linked immunosorbent assay
ER	Endoplasmic reticulum
ERT	Enzyme replacement therapy
FP	Fluorescent protein
FUT11/12	α-1,3-fucosyltransferase
FUT13	α-1,4-fucosyltransferase
gDNA	Genomic DNA
GFP	Green Fluorescent protein
GnT	N-acetylglucosaminyltransferase
GNTIV	N-Acetylglucosaminyltransferase-IV
GNTV	N-Acetylglucosaminyltransferase-V

GUI	Graphic user interface
HUGEs	Human glycosylation enzymes
IgE	Immunoglobulin E
IgG	Immunoglobulin G
JaCoP	Just another co-localisation plugin
LAL	Lysosomal Acid Lipase
LALΔ1-22	Lysosomal Acid Lipase lacking signal peptide
LB	Lysogeny Broth media
mAb	Monoclonal antibody
MES	2-(N-Morpholino) ethanesulfonic acid
MNS	α-mannosidase
MNS1	Golgi-α-mannosidase I
mRNA	Messenger ribonucleic acid
MS	Murashige and Skoog medium
MUR3	β-1,2-galactosyltransferase
N.B	Nicotiana benthamiana
NEB	New England Biolabs
OD	Optical density
P2A	Porcine 2A peptide
PCR	Polymerase chain reaction

PTM	Post Translational Modification
RCA1	Ricinus communis agglutin 1 lectin
RFP	Red Fluorescent protein
RT-PCR	Reverse transcription polymerase chain reaction
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SN	Sambucus Nigra agglutin lectin
SP	Signal peptide
ST	Sialyltransferase
ST6GAL	α-2,6-sialyltransferase
TAE	Tris Acetate-EDTA buffer
TGN	Trans Golgi Network
TGs	Triglycerides
TMD	Transmembrane domain
UDP-GlcNAc	uDP-N-acetylglucosamine
WT	Wild type
XylT	β -1,2-xylosyltransferase
ΔXFT	Xylose/Fucose deficient RNAi tobacco line

This work was supported by the BBSRC and the Oxford Interdisciplinary Bioscience Doctoral Training Partnership

Contents

List of figures

Chapter 1: Introduction

Chapter 2: Methods

Chapter 3: Imaging the plant Golgi

Chapter 4: CTS design and HuGE localisation

Chapter 5: Protein production in stable plants

List of tables

Chapter 1: introduction

Chapter 2: methods

Chapter 5: Protein production in stable plants

Chapter 1: Introduction

Contents

1.1 N-linked glycosylation in plants for functional therapeutic proteins

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 has a strong impact on protein localisation and function as one of the major posttranslational modifications in eukaryotic cells (Strasser et al., 2014). In plant and mammalian cells, protein N-glycosylation begins in the ER by co- or post-translational *en bloc* transfer of an pre-assembled oligosaccharide precursor, Glc3Man9GlcNAc2, from the lipid carrier dolichol pyrophosphate onto specific Asn residues within the N-glycosylation consensus sequence Asn–X–Ser/Thr (where X can be any amino acid except Proline) (Gomord et al., 2010; Strasser et al., 2016;Schoberer et al., 2018). The two outmost glucose residues are subsequently cleaved by α-glucosidase I and II, resulting in the generation of a monoglucosylated N-glycan. This can interact with calnexin and calreticulin in order to promote folding. The downstream release from this interaction interaction depends on the trimming of the remaining glucose by α-glucosidase II. Proteins that fail final conformation here progress to ER-associated degradation to prevent misfolded protein accumulation (Dünser et al., 2025).

The composition of glycosylation enzymes in plant and mammalian systems is comparable in the early stages of N-glycan chain generation in the *cis*-Golgi cisternae but differs in the enzymes present in the medial- and *trans*-Golgi (Schoberer et al., 2018a; Schoberer et al., 2018b); (Figure 1.1). In mammalian cells, N-acetylglucosaminyltransferases (GnTs; GnTI, GnTII, GnTIV and GnTV) use Uridine diphosphate-*N*-acetylglucosamine (UDP-GlcNAc) in order to transfer N-acetylglucosamine onto various points of the developing glycan structure, onto which further specific galactosylation is performed by B4GALT1, and final α-2,6-sialic acid capping of the glycan structure by ST6GAL (Figure 1.1;McDonald et al., 2014). The former steps (GnTI and GnTII) are conserved between plants and mammalian cells (Figure 1.1) up until the formation of biantennary GnGn structures. Whilst the latter and subsequent steps are responsible for the branching (the formation of tri- and tetra-antennary N-acetylglucosamine), β1,4-Galactosylation and end-terminal α-2,6-sialylation are not present in plants and the enzymes responsible must be introduced for these to occur (McDonald et al., 2014; Nagels et al., 2011). As such sialic acid

addition can be considered a significantly limiting step in the humanisation of heterologous proteins expressed in plants. To humanise plant glycosylation patterns to produce economically and therapeutically important proteins, insertion of the human glycosylation enzymes (HuGEs) GNTIV and GNTV, B4GALT1 and ST6GAL into the medial or *trans*-Golgi cisternae is required (Figure 1.1). Effectiveness of enzyme knock-in strategies will be dependent on correct localisation within the plant Golgi body to the medial and *trans*-Golgi cisternae respectively, which can be determined by high-resolution confocal dynamic imaging.

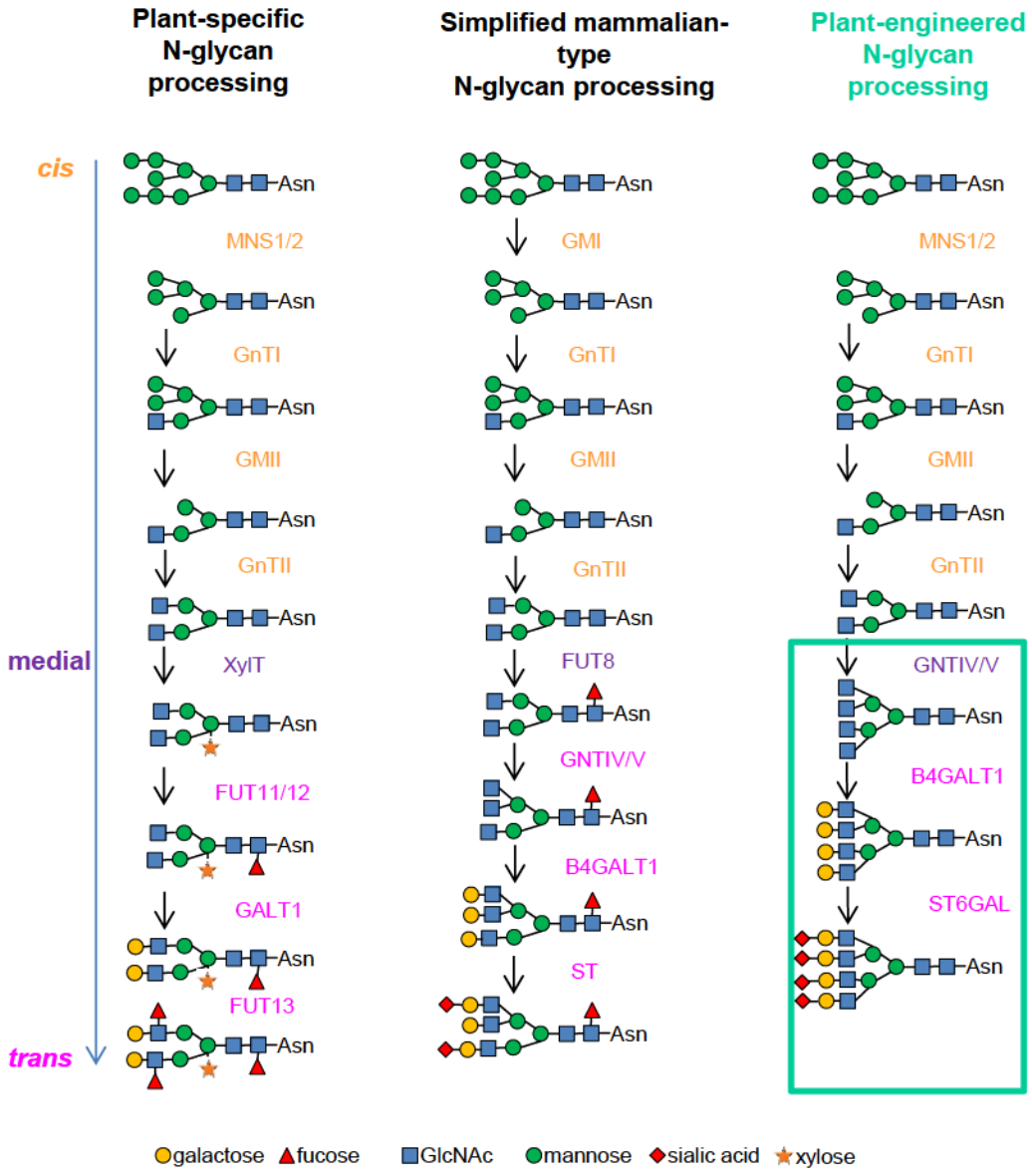


Figure 1.1: Comparison of plant and mammalian N-linked glycosylation with modifications required for humanising plant N-linked glycosylation.
Required modifications to the plant system in terms of insertion of mammalian enzymes and their resultant glycan structures are highlighted in the green box on the right hand side. Enzymes are colour-coded for localisation in *cis*-Golgi cisternae (orange), medial-Golgi cisternae (purple) and *trans*-Golgi cisternae (magenta). MNS: α-mannosidase, GnT: N-acetylglucosaminyltransferase, XylT: β-1,2-xylosyltransferase, FUT11/12: α-1,3-fucosyltransferase, GALT1: β-1,3-galactosyltransferase, FUT13: α-1,4-fucosyltransferase. Modified from Schoberer et al, 2018(b). Symbol nomenclature for graphical representation of glycans is based on Varki et al, 2015.

Glycosylation can significantly enhance the physicochemical stability of glycoproteins, N-linked glycans specifically facilitate protein folding through their role in the calnexin/calreticulin chaperone-mediated folding cycle (Margolin et al., 2020) with many functional proteins dependant on glycosylation for correct folding and so activity. Native plants enzymes can functionally N-glycosylate expressed heterologous human proteins, but they do not possess the medial- and *trans*-Golgi enzymes required for the generation of a multi-antennary N-glycan chain with end terminal sialic acid capping, which is the end structure of mammalian N-glycan modification. Adaptation of the plant machinery to produce human glycan patterns requires the insertion of key human glycosylation enzymes (Table 1.1). These are GNTIV and GNTV (N-acetylglucosaminyltransferase IV and V), B4GALT1 (β-1,4-galactosyltransferase), and ST6GAL (α-2,6-sialyltransferase). To optimise the system for therapeutic production would also require downregulation of the plant enzymes XylT (β-1,2-xylosyltransferase), FUT11 and FUT12 (α-1,3-fucosyltransferases) to avoid interference with plant glycosylation patterns or attachment of immunogenic plant-specific glycans. The plant enzymes to be replaced or outcompeted are localised to the medial and *trans*-Golgi, hence GNTIV, GNTV, B4GALT1 and ST6GAL should also be targeted to medial or *trans*-Golgi cisternae for effective enzymatic function on N-glycan intermediates (Schoberer et al., 2018b) as their activity is sequential, following the activity of the enzymes in the cis-Golgi cisternae, and sub-Golgi localisation directly effects resulting glycan structures (Bosch et al., 2013; Castilho et al., 2011). Several explanations have been made for the non-uniform distribution of glycosylation enzymes in the Golgi, from kin-recognition based on enzymes forming protein-protein interactions causing glycosylation enzymes to aggregate, or membrane thickness

models dictating distribution based on membrane size relating to the nature of glycosylation enzymes as type II membrane proteins (Nilsson et al., 2009

Table 1.1: Introduced mammalian Glycosylation enzymes and their function. Each of the major glycosylation enzymes investigated within this project and their relevant activity towards a humanized glycan profile.

Glycosylation enzymes	Enzymatic function	Relation to protein
GNTIV & GNTV α1,3-mannosyl-β1,4-*N*-acetylglucosaminyltransferase (GnT-IV isozymes a) and α1,6-mannosyl-β1,6-*N*-acetylglucosaminyltransferase (GnT-V)	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 β-1,4-galactosyltransferase	Addition of β-1,4-galactose intermediates (Schnieder et al., 2015; Bakker et al., 2001)	Addition of subsequent and intermediary glycans onto branched chains for the precursor to sialic acid addition
ST6GAL α-2,6- sialyltransferase	Addition of terminal α-2,6-sialic acid residues (Wee et al., 1998)	Addition of terminal sialic acid residue essential to mimick mammalian glycosylation profile, increases half-life and stability by reducing elimination from circulation

Plant Golgi bodies glycosylation machinery is not equivalent to the mammalian, and so is incapable of producing specific mammalian glycomotifs and a mammalian-type glycosylation profile (Figure 1.1 (Strasser et al., 2014; Schoberer et al., 2018a; Schoberer et al., 2018b). The differences between plant and mammalian N-glycosylation machinery are of importance for the expression of mammalian proteins as many economically important recombinant proteins are glycosylated (Göritzer et al., 2021). 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 significantly diminishing therapeutic effectiveness of treatment (Bosch et al., 2013), with efficacy of antibody therapy correlating with serum half life (Stelter et al., 2019) For example, the absence of complex N-glycans in an anti-rabies monoclonal antibody produced in plant systems reduces its efficacy by diminishing plasma half-life (Ko et al., 2014). The introduction or amplification of key mammalian glycans can significantly improve plasma half-life for therapeutic proteins which has been shown for a modified erythropoietin with enhanced end terminal sialyation (Su et al., 2010).

Overall, glyco-engineering in plants has been shown to be possible. Production of mammalian glycan motives in plants can be achieved by the targeted heterologous expression of mammalian glycosylation enzymes (Figure 1.2; Strasser et al., 2014a; Strasser et al., 2014b). Earlier studies demonstrating that a mammalian glycosyltransferase can convert N-glycans in plants in a similar way as in human cells (Palacpac et al., 1999; Bakker et al., 2001) paved the way for reconstructing mammalian glycosylation pathways in plant systems. Previous research on the simultaneous overexpression of six mammalian glycosylation proteins in *N. benthamiana* resulted in the synthesis of the activated sialic acid and *in planta* protein sialylation (Castilho et al., 2010).

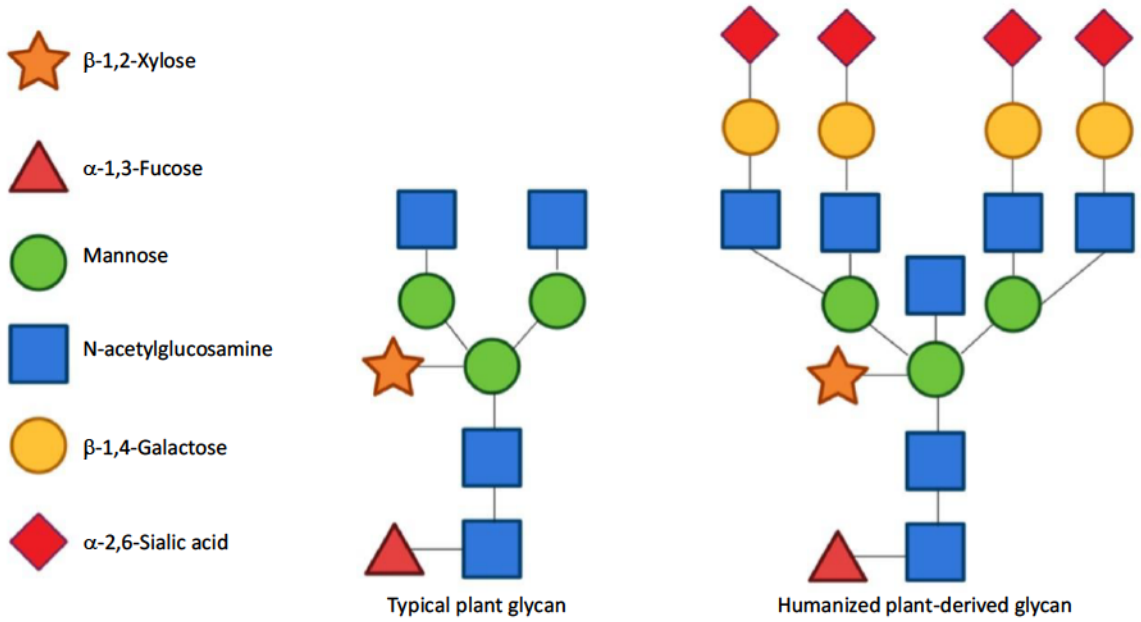


Figure 1.2: Potential effects on plant N-glycan structure by addition of mammalian post-translational glycosylation enzymes, demonstrating addition of further mammalian glycosylation motifs onto a typical plant glycan to create a humanized hybrid structure. Engineering of glycosylation pathway to include 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 β-1,4-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). Symbol nomenclature for graphical representation of glycans is based on Varki et al (2015).

The plant Golgi apparatus is a significant site of protein synthesis and glycomodification within the plant cell, and so a focal point for recombinant protein production (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 numbering several hundred per cell and are extremely mobile throughout the cytoplasm. The basic structure of the individual Golgi body (Figure 1.3) is a stack of disc-like cisternae characterized by order 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 (Robinson et al., 2020).

Transport to, from and within the plant Golgi is debated, earlier models proposing biosynthetic cargo was shuttled to the Golgi from the ER by functionally conserved proteins and COPII vesicles produced at the ER exit site (ERES), and transport between cisternae occurred in the same manner (Nilsson et al., 2009). More recent studies suggest the role of the ERES as a pre-*cis*-Golgi compartment, and cargo delivery occurs through microtubules connecting the ER and Golgi apparatus (Mcginness et al., 2002)

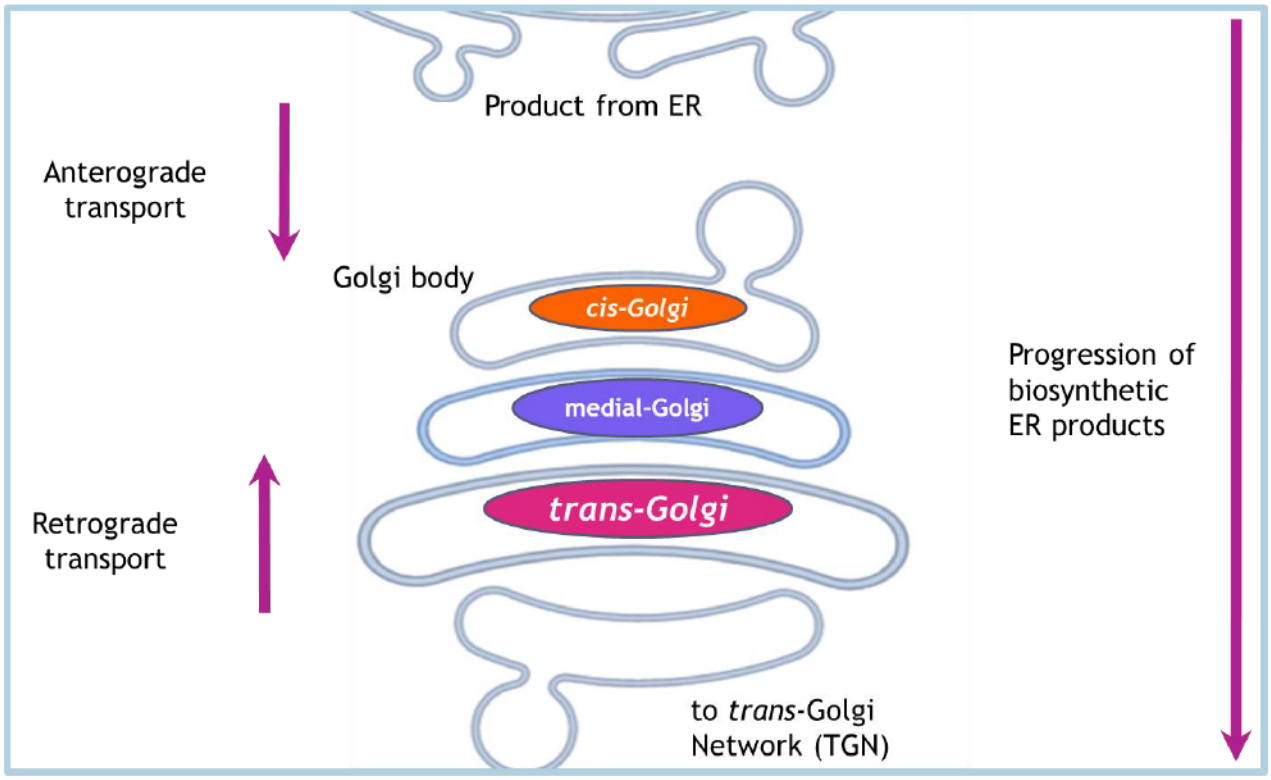


Figure 1.3: Organization of the plant Golgi structure and transport for protein biosynthesis (modified from Pigino, 2012).

1.2 Plant molecular farming for recombinant protein

More than 50% of Eukaryotic proteins are identified as glycoproteins (Apweiler et al., 1999), alongside a third of approved biopharmaceuticals (Walsh et al., 2006; Gomord et al., 2010) according to the Swiss-Prot database. Glycosylation can affect size, charge, solubility, targeting and clearance rates of therapeutic proteins, act as ligands

for lectin receptors and carry effector roles in IgG antibodies, occupying a huge role in relation to therapeutic proteins (Varki et al., 2015). Unlike many popular host systems such as bacteria, mammalian and yeasts, plants can perform complex N-glycosylation, making them an attractive host for production of recombinant proteins (Kogelmann et al., 2023).

Biopharming refers to the concept of plant molecular farming, describing the production of recombinant proteins using plants as a host organism (Fischer et al., 2004). Vector selection to produce biologics, especially 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).
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 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 most 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 transient transformation of mature plants facilitates niche applications, such as a rapid vaccine production in response to an outbreak, 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). An important advantage in products for human therapeutic applications is that plants cannot easily be contaminated with mammalian pathogens or support their growth as some bacterial, yeast and mammalian cell systems can (Tekoah et al., 2015). These factors in mind make plants an ideal system in which to develop a low-immunogen expression system for human therapeutic protein with a modified mammalian glycosylation profile.

1.3 Investigating human proteins of interest

The plant expression platform intended would be able to successfully produce, and generate mammalian-type N-glycosylation on, a variety of human proteins of interest. Proteins of interest would need to be correctly synthesized and have the correct post-translational modifications, and so three test therapeutic proteins were selected to act as reporters for this system. These were Lysosomal acid lipase (LAL), the heavy chains of immunoglobulin G (IgG) and immunoglobulin E (IgE), selected for their applications, complexity and multiple N-glycosylation sites (sequences in appendices 6).

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, Besler et al., 2022). LAL deficiency (LAL-D) is considered a 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). Endogenous deficiency of LAL leads to a spectrum of disease phenotypes resulting from chronic lipid accumulation, with LAL dysfunction is

considered a life-limiting condition with few alternative treatment options and 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), of Sebelipase Alfa© a recombinant human LAL analogue produced in genetically modified chicken eggs (Su et al., 2016). LAL is a good candidate for production in this platform, as ERTs in general require are demanding in terms of quantity, quality and minimization of sensitisation and off-target effects (Kizhner et al., 2015). Existing treatment options for LAL deficiency are known to have issues with allergic reactions that can prevent effective treatment (Zimran et al., 2011, Su et al., 2016). 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). This makes LAL an interesting candidate protein for this system, as it requires generation of a functional human enzyme with mammalian N-glycosylation, in as low immunogen system as is possible, key targets for this project.

ERT requires a final product that is capable of meeting necessary 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.

Ig's production in plants is not a new phenomenon, but their effectiveness for therapeutic applications is often limited (Beihammer et al., 2023). This is in part by the resultant N-glycosylation profile being typically under glycosylated affecting stability and binding affinity (Castilho et al., 2018), which this research may in part mitigate. IgE is particularly heavily glycosylated with seven N-Glycosylation sites (Montero-Morales et al., 2019), and with site occupation with end-terminal sialylation directly improving plasma half-life when compared to its under-glycosylated state (Chia et al.,

2023). Correct protein structure can depend upon glycosylation, which additionally directly impacts binding functionality in Igs. Relative binding affinity of synthesized Igs has previously been used as a valid metric for the measure of successful production and glyscosylation (Kogelmann et al., 2023). Relatively simple commercial tests are available to assess their functionality, streamlining a primary analysis (Göritzer et al., 2021). Whereas, to test functionality of LAL itself may not be possible due to need of an animal model, but each protein of interest selected represents a complex, glycosylated therapeutic protein structure to test the plant production system.

Chapter 2: Materials and methods

Contents

2.1 Plant material

2.1.1 Wild type material and standard growth conditions

Unless otherwise stated, all experiments were performed on wild type *Nicotiana benthamiana,* hereafter referred to as tobacco, grown under greenhouse conditions. Tobacco seeds were placed on sterile ½ MS plates (2.4 gL-1 Murashige and Skoog medium, 0.8% phytoagar in deionised water, pH 5.7), sealed with double-layered micropore tape for gas exchange, and placed in a plant growth chamber under 16hr light/dark cycle at 22°C. After 10 days seedlings were transferred to soil and grown in greenhouse conditions. Plants were used for agroinfiltration at 5-6 weeks old.

2.1.2 ΔXFT Nicotiana benthamiana

The ΔXFT *Nicotiana benthamiana* β1,2-xylotransferase (XylT) and α1,3-fucosyltransferase (FucT) RNAi knock down line was developed and provided by Richard Strasser (Strasser et al., 2008) as seeds. Use of the ΔXFT line for the generation of stable plant lineages was granted. These plants were treated with standard growth conditions described and in compliance with GM protocol.

2.2 Microscopy

2.2.1 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 imaged. Plant sections were mounted in water. All confocal microscopy was performed at Oxford Brookes University with 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. All imaging was performed at room temperature. Excitation and emission wavelengths for fluorophores used in this project are shown below (table 1).

Table 2.1: Excitation and emission wavelengths for common fluorophores

Fluorophore	Excitation wavelength (nm)	Emission spectra (nm)
Red fluorescent protein (mRFP)	561	570-615
Green fluorescent protein (eGFP)	488	495-550
Clover (CLVR)	488	493-598
Blue fluorescent protein (mTagBFP2)	400	420-540

2.2.2 Identifying Golgi localized expression with line profile analysis

Co-localisation analysis for constructs and markers in the plant Golgi body was performed using a line profile utility in Zeiss ZEN blue software (Version 3.5.093.00000; McGinness et al., 2022). Golgi in a suitable side-profile configuration (Figure 2.1, A) are obtained, usually by selection from a Airymovie time series to obtain the best configuration for reducing overlap and observing separation based on Golgi rotation, maximising proportion of difference due to actual localisation of expression. Use of airy movie time series additionally allows tracking of individual Golgi in a sample and prevents re-sampling. Analysis-quality time series are collected at 512x512 pixel density, averaging of 2 at 15x zoom on a zeiss 880 confocal at max speed using a 63x lens. The side profile of Golgi at this pixel density and zoom will typically span around 30 pixels. Tetraspeck™ fluorescent microspheres (Invitrogen T14792) were used to calibrate Airyscan detector and check laser alignment before sessions observing Golgi co-localisation. As was a control combination of MNS1-ST to check any anomalies in overall localisation.

With suitable side profile Golgi obtained, ZEN blue software was used to draw a line (sizing=5) perpendicular to the selected Golgi body (Figure 2.1,A) in order to generate a line intensity profile (Figure 2.1,B) of the two channels. Line profiles are drawn from suspected *cis*-face of the Golgi (identification of orientation dependant on construct expression) to begin and terminate within visible fluorescence. This identifies peaks of fluorescent intensity within the Golgi without significant background interference.

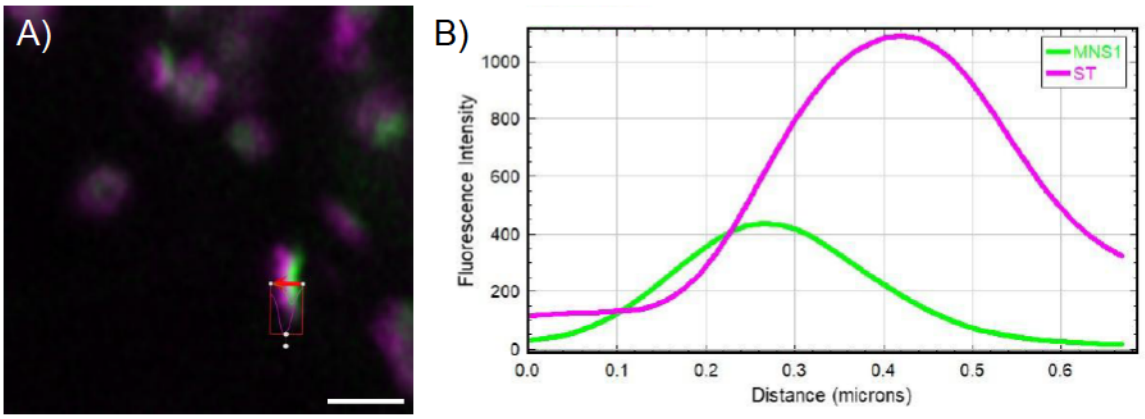
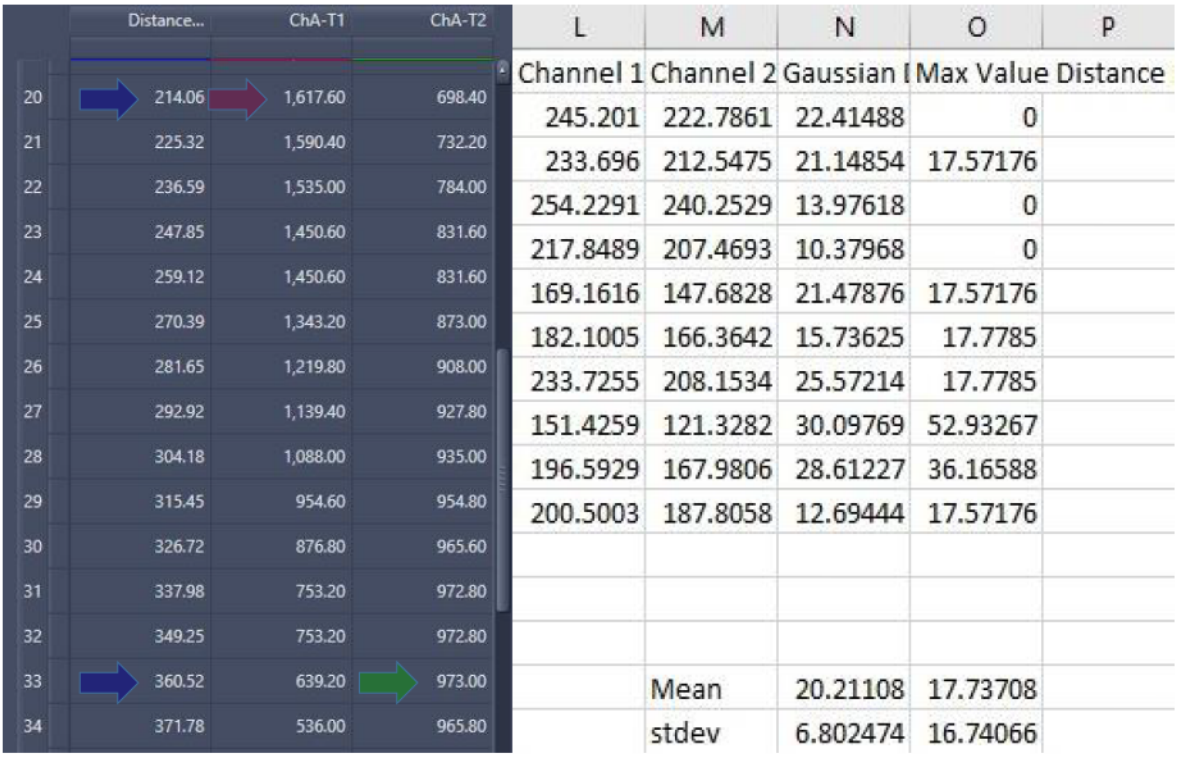


Figure 2.1. A) Example profile view confocal image and B) output of a line intensity analysis from ZEN blue for microscopy data for ST-mRFP (magenta) and MNS1-eGFP (green). Graphs were generated for each Golgi for visual representation of fluorescence distribution to identify errors.

The distance between the maximum peak intensities of each channel was collected across the line profile for each Golgi body analysed (Figure 2.2), with the data being used to determine the maximum peak intensity of each respective channel (red and green arrows), and the distance value between them in nm (blue arrows) to determine degree of co-localisation. Average and standard deviation were calculated for each protein combination. Values for distance between peaks were analysed within sessions (i.e. analysis grouped to samples taken in one microscope collection) and compared between session pools



	Distance...	ChA-T1	ChA-T2
20	214.06	1,617.60	698.40
21	225.32	1,590.40	732.20
22	236.59	1,535.00	784.00
23	247.85	1,450.60	831.60
24	259.12	1,450.60	831.60
25	270.39	1,343.20	873.00
26	281.65	1,219.80	908.00
27	292.92	1,139.40	927.80
28	304.18	1,088.00	935.00
29	315.45	954.60	954.80
30	326.72	876.80	965.60
31	337.98	753.20	972.80
32	349.25	753.20	972.80
33	360.52	639.20	973.00
34	371.78	536.00	965.80

L	M	N	O	P
Channel 1	Channel 2	Gaussian I	Max Value	Distance
245.201	222.7861	22.41488	0	
233.696	212.5475	21.14854	17.57176	
254.2291	240.2529	13.97618	0	
217.8489	207.4693	10.37968	0	
169.1616	147.6828	21.47876	17.57176	
182.1005	166.3642	15.73625	17.7785	
233.7255	208.1534	25.57214	17.7785	
151.4259	121.3282	30.09769	52.93267	
196.5929	167.9806	28.61227	36.16588	
200.5003	187.8058	12.69444	17.57176	
	Mean	20.21108	17.73708	
	stdev	6.802474	16.74066	

Figure 2.2. Data output for line profiles.

Example data output of line profile analysis in Zen is shown on the left hand side. Maximum fluorescent intensity of channel 1 (red arrow) and maximum of channel 2 (green arrow) is indicated, with the distance separating the two peaks along the line intensity profile (blue arrows). An example of a data subset is shown on the right. Each row provides a single Golgi body line profile analysis from one combination showing maximum intensities and the distance between them, processed by use of a python script (Appendix 1).

This process was automated through use of a python script (Appendix 1 with code for Python version 3.9) to iterate through multiple .csv files, obtain distance between peaks and fit to a Gaussian distribution to improve precision of results collected, as ZEN only formats distance measurements by pixel, leading to an irregular scale that lacks small increments causing data binning for close proximity combinations (Figure 2.2, right, max value column). This outputs peak intensity of each channel, as both the modified Gaussian-fitted data or the unmodified max distance data delivered by Zen, to compare for any errors, then used to produce averages and standard deviations across collected line intensity profiles.

2.3 Plant transformation and transfection

2.3.1 Regeneration for stable plant lines in tobacco

The protocol for tobacco regeneration was developed via supervisor-student correspondence. Whole plant leaves (either *N. tabacum* or *N. benthamiana*) were infiltrated according to agrobacterium-mediated expression protocol at an OD_{600} of 0.1, and incubated at 22°C with a 16hr light/dark cycle for 72 hrs. Leaves were then harvested and small (<1 cm) squares were cut from infiltrated regions. Leaf sections were sterilized in a 2.5% sodium hypochlorite solution for 10 min with shaking, then underwent two rinse cycles in sterile deionized water of 2 min with shaking. Leaf sections were then transferred to a shoot inducing solid media (For 200 ml: Mix 0.86g MS salts, 4g sucrose, 1.6g Phytoagar, 400 µl of 1 mg/ml benzylaminopurine (BAP) (for final conc. 2 mg/L) 10 µl of 1 mg/ml naphthalene acetic acid (NAA) (for final conc. 0.05 mg/L) then to final volume of 200 ml with deionized water. pH shifted to 5.7 before autoclaving. Post autoclave, when melting to pour plates add to whole bottle 200 µl of 200 mg/ml of appropriate selection antibiotic (Either Hygromycin or Kanamycin), 200 µl of 200 mg/ml timentin and 200 µl of 200 mg/ml carbenicillin. Plates were refreshed every 14 days until shoots emerged. Once shoots emerged these were transferred to a rooting media (For 200 ml: composition the same as shooting medium, but without BAP.) The same protocol was applied with rooting media until roots emerged. Plates were kept in a growth cabinet on a 16hr light cycle intended for Tobacco. For control conditions this media was used without selection agents, for constructs of interest media included 50 mg/ml timentin, 25 mg/ml carbenicillin, and then either 50 mg/ml hygromycin or 50 mg/ml kanamycin depending on vector used as a selection agent to prevent nonspecific growth. Vectors used for downstream stable plant expression were obtained from VIB-UGENT Centre for Plant Systems Biology(Karimi et al., 2002): pH7WG2 (afterwards reffered to as 388) was used to generate the MUR3-GNT45-388 expression vector and pK7WG2 (afterwards reffered to as 389) to generate the FUT13-B4ST6-389 expression vector. Once roots of sufficient size developed (~6 weeks) the plantlets were transferred to the greenhouse and potted into moist soil in covered trays to maintain high humidity for the first few weeks, then grown according to standard growth conditions for GM tobacco.

2.4 Cloning and expression plasmids

2.4.1 Gateway cloning and 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.

Genes were first inserted into a pDONR^{TM}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 25 µl of cells and 200 µl SOC recovery media. Selection was undertaken on agar plates with selection agent kanamycin 50 µg 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 3 ml liquid culture of LB and appropriate antibiotics, and were kept overnight at 37℃, with shaking at 200xg. 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 pDONR^{TM}221 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). In the vector pB7FWG2 for C-terminal eGFP fusions, the eGFP was replaced with a CLOVER green fluorescent protein (pB7FWG2: CLVR). A second transformation into competent *E. coli* was then undertaken, selected with Spectinomycin 50 µg mL^{-1}, constructs confirmed via colony PCR and plasmid extracted (Table 3).

Self-cleaving peptide constructs containing the MUR3 (fused to mTagBFP) and FUT13 CTS domains separated by the self-cleaving peptide P2A with intein sequence, were synthesised commercially (Twist) and supplied in an entry vector (for sequence see Appendix 2). An LR reaction was carried out as described above with pB7RWG2 to add a C-terminal mRFP fusions to the FUT13 CTS.

Table 2.2: Primers generated for molecular cloning. Primers were designed to function with New England Biolabs Q5 Polymerase molecular cloning kits. Annealing temperatures were calculated utilizing the online NEB Tm Calculator tool. Gene sequences are shown in capital letters, Gateway overhangs are added in small letters.

Primer name	Annealing temperature (℃)	Sequence
ST-SPfor (B4GALT1SPgatefor)	58	ggggacaagtttgtacaaaaaagcaggctctATGATCCA CACAAATCTTAAG
LALfor	48	ATGAAAATGCGGTTTTTGGG
LALrev	48	CTGATACTTGCGCATAAGG
LALgatefor	48	ggggacaagtttgtacaaaaaagcaggctctATGAAAAT GCGGTTTTTGG
LALgateforsp	60	ggggacaagtttgtacaaaaaagcaggctctATGTCAGG TGGAAAGTTG
LALgaterev	60	ggggaccactttgtacaagaaagctgggtcCTGATACTT GCGCATAAGG
B4GALT1for	49	ATGAGGCTGAGGGAAC
B4GALT1rev	49	CCATAGCAGTCTCCTAATTTTTTC
B4GALT1gatefor	60	ggggacaagtttgtacaaaaaagcaggctctATGAGGCT GAGGGAA
B4GALT1gaterev	60	ggggaccactttgtacaagaaagctgggtcCCATAGCAG TCTCCTAATTTTT
B4GALT1signalfor	56	ATGATCCACACAAATCTTAAGAAAAAATTTT CTTGTTGTGTTCTAGTCTTTTTACTTTTTGCA GTTATTTGTGTATGG AGGCTGAGGGAAC
B4GALT1gaterevlow	58	ggggaccactttgtacaagaaagctgggtcCCATAGCAG

		TCTCCTAATTT
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
MUR3-GNTIVgatefor	59	ggggacaagtttgtacaaaaaagcaggctctATGTTCCCAAGGGTG
MUR3-GNTIVgaterev	59	ggggaccactttgtacaagaaagctgggtcGATTAGTCG

		CTTTTTTGATATGG
MUR3-GNTIVfor	50	ATGTTCCCAAGGGTGTCAA
MUR3-GNTIVrev	50	AGATCCATATCAAAAAAGCGACTAATC
FUT13-B4-FOR	61	ATGCCAATGAGATATTTAAATGCTATG
FUT13-B4-REV	61	CGATATTGGAACACCTTCAC
FUT13-ST6-FOR	62	ATGCCTATGAGATACCTGAATG
FUT13-ST6-REV	62	TTGCTATTACTATCAGAAGTTCTTCC
MUR3-GN4-FOR	62	ATGTTTCCAAGAGTTAGTATGAGAC
MUR3-GN4-REV	62	CATATTAAGAAGGCCACAAACC
MUR3-GN5-FOR	62	ATGTTCCCTAGAGTCTCAATG
MUR3-GN5-REV	62	CTCTGTGCAAAGACTGTTTAC
GnT45FusFrag1FWD	62	ATGTTTCCAAGAGTTAGTATGAGAC
GnT45FusFrag1REV	62	CACAAGAACAGTTTGTGGCCTT
GnT45FusFrag2FWD	60	GGCCACAAACTGTTCTTGTGGT
GnT45FusFrag2REV	60	TAAACAGTCTTTGCACAGAG
GnT45Frag1FWDgate	60	GGGGACAAGTTTGTACAAAAAAGCAGGCTT A ATGTTTCCAAGAGTTAGTATGAGAC
GnT45Frag2REVgate	60	GGGGACCACTTTGTACAAGAAAGCTGGGTT TAAACAGTCTTTGCACAGAG
GnT5_linkerFWD	56	TGTTCTTGTGGTTCTGGTTCTAGAACGAATA GTACAAATTCTACTA
GnT5_linkerREV	56	TAAACAGTCTTTGCACA
B4_ST6Frag1FWD	61	ATGCCAATGAGATATTTAAATGCTAT
B4_ST6Frag1REV	61	CCACAAGAACATGAAGGTGTTCC

B4_ST6Frag2FWD	62	AACACCTTCATGTTCTTGTGGT
B4_ST6Frag2REV	62	GAAGAACTTCTGATAGTAATAGCAAAC
B4_ST6Frag1FWDgate	61	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAATGCCAATGAGATATTTAAATGCTAT
B4_ST6Frag2REVgate	61	GGGGACCACTTTGTACAAGAAAGCTGGGTTGAAGAACTTCTGATAGTAATAGCAAAC
ST6_linkerFWD	57	TGTTCTTGTGGTTCTGGTTCTAGAGCATCTTTCCAGGTTT
ST6_linkerREV	57	GAAGAACTTCTGATAGTAATAGC
MURG4RTFWD	50	TACAGCCCTCATATGCGA
MURG4RTREV	50	TATGTAGCGTCTCAATCAAGTATGA
MURG5RTFWD	50	TAGCTGAACCTGGC
MURG5RTREV	50	TAAACAGTCTTTGCACAGAGC
FUTB4RTFWD	50	AAATTGAGTGATAGTACTAGTCTTCCA
FUTB4RTREV	50	ATCTCCCGCCTGATTAATAACG
FUTST6RTFWD	50	GTTTACCATAGCGATATCCCAAA
FUTST6RTREV	50	GAAGAACTTCTGATAGTAATAGCAAACG

2.4.2 Protein expression plasmids

Below is a full list of plasmids used for protein expression in this work (table 3). Unless specified all constructs were C-terminally tagged as to not interfere with endogenous N-terminal localisation signals.

Table 2.3: List of constructs utilized for investigation. HuGE constructs generated and marker proteins used for confocal analysis.

Construct	**Plasmid backbone**	**Description**	**Citation**
MNS1-RFP	p31 (Hüttner et al., 2012)	*cis*-Golgi marker of C-terminal RFP fusion to plant CTS domain	Schoberer et al., (2013)
MNS1-GFP	p20 (Schoberer et al., 2009)	*cis*-Golgi marker of C-terminal GFP fusion to plant CTS domain	Schoberer et al., (2013)
MNS1-BFP-P2A-ST-RFP	pB7RWG2	Dual marker construct for *cis/trans*-Golgi localisation on one insert. Two fusions of plant CTS domains (Table 4), MNS1-BFP for *cis*-Golgi and ST-RFP for *trans*-Golgi.	McGinness et al., 2024
ST-RFP	pVKH18En6	*trans*-Golgi marker of C-terminal RFP fusion to rat CTS domain	Renna et al., (2005)
ST-CFP	pVKH18En6	*trans*-Golgi marker of C-terminal CFP fusion to rat CTS domain	Renna et al., (2005)
ST-GFP	pVKH18En6	*trans*-Golgi marker of C-terminal GFP fusion to rat CTS domain	Renna et al., (2005)
LAL-RFP	pB7RWG2	C-terminal RFP tagged LAL from original sequence supplied by Porton biopharma	Unpublished
LAL-GFP	pB7FWG2	C-terminal GFP tagged LAL from original sequence supplied by Porton biopharma	Unpublished
LAL-CLVR	pB7FWG2	C-terminal CLVR tagged LAL from original sequence supplied by	Unpublished

		Porton biopharma	
LALΔ1-22-RFP	pB7RWG2	C-terminal RFP tagged LAL with N-terminal signal peptide at positions 1-22 removed	Unpublished
IgE-CLVR	pB7FWG2	C-terminal CLVR tagged IgE heavy chain	Unpublished
IgG-CLVR	pB7FWG2	C-terminal CLVR tagged IgG heavy chain	Unpublished
GNTIV-RFP	pB7RWG2	C-terminally RFP tagged original human sequence for GNTIV from NCBI	Unpublished
GNTIV-GFP	pB7FWG2	C-terminally GFP tagged original human sequence for GNTIV from NCBI	Unpublished
GNTIV-CLVR	pB7FWG2	C-terminally CLVR tagged original human sequence for GNTIV from NCBI	Unpublished
GNTV-RFP	pB7RWG2	C-terminally RFP tagged original human sequence for GNTV from NCBI	Unpublished
GNTV-GFP	pB7FWG2	C-terminally GFP tagged original human sequence for GNTV from NCBI	Unpublished
GNTV-CLVR	pB7FWG2	C-terminally CLVR tagged original human sequence for GNTV from NCBI	Unpublished
B4GALT1-RFP	pB7RWG2	C-terminally RFP tagged original	Unpublished

		human sequence for B4GALT1 from NCBI	
B4GALT1-GFP	pB7FWG2	C-terminally GFP tagged original human sequence for B4GALT1 from NCBI	Unpublished
B4GALT1-CLVR	pB7FWG2	C-terminally CLVR tagged original human sequence for B4GALT1 from NCBI	Unpublished
ST6GAL-RFP	pB7RWG2	C-terminally RFP tagged original human sequence for ST6GAL from NCBI	Unpublished
ST6GAL-GFP	pB7FWG2	C-terminally GFP tagged original human sequence for ST6GAL from NCBI	Unpublished
ST6GAL-CLVR	pB7FWG2	C-terminally CLVR tagged original human sequence for ST6GAL from NCBI	Unpublished
MUR3-RFP	pB7RWG2	Endogenous plant enzyme CTS (table 4) domain with FP fusion to act as marker for medial-Golgi directed localisation	McGinness et al., 2024
MUR3-BFP-INT-P2A-FUT13-RFP	pB7RWG2	Dual marker construct for medial/*trans*-Golgi localisation on one insert. Two FP fusions of plant CTS domains (Table 4), MUR3-BFP for *cis*-Golgi and FUT13-RFP for *trans*-Golgi.	McGinness et al., 2024
MUR3-GNTIV-	pB7FWG2	Modified C-terminal CLVR tagged	McGinness

CLVR		GNTIV, endogenous CTS swapped for plant specific CTS targeting medial-Golgi (Table 4)	et al., 2024
MUR3-GNTV-CLVR	pB7FWG2	Modified C-terminal CLVR tagged GNTV, endogenous CTS swapped for plant specific CTS targeting medial-Golgi (Table 4)	McGinness et al., 2024
MUR3-GNT45-388	pH7WG2 (Karimi et al., 2002)	Untagged dual catalytic domain construct for stable plant transformation. Medial-Golgi targeting plant CTS domain (Table 4) followed by catalytic subunit for GNTIV and GNT5	Unpublished
FUT13-RFP	pB7RWG2	Endogenous plant enzyme CTS (table 4) domain with FP fusion to act as marker for *trans*-Golgi directed localisation.	McGinness et al., 2024
FUT13-B4GALT1-CLVR	pB7FWG2	Modified C-terminal CLVR tagged B4GALT1, endogenous CTS swapped for plant specific CTS targeting *trans*-Golgi (Table 4)	McGinness et al., 2024
FUT13-ST6GAL-CLVR	pB7FWG2	Modified C-terminal CLVR tagged ST6GAL, endogenous CTS swapped for plant specific CTS targeting *trans*-Golgi (Table 4)	McGinness et al., 2024
FUT13-B4ST6-389	pK7WG2 (Karimi et al., 2002)	Untagged dual catalytic domain construct for stable plant transformation. *Trans*-Golgi targeting plant CTS domain (Table	Unpublished

		4) followed by catalytic subunit for B4GATL1 and ST6GAL	

2.4.3 Selection of protein sequences for HuGEs and proteins of interest

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 (table 2.4) 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, as were sequences for heavy chains of IgG and IgE. Proteins of interest and the mammalian glycosylation enzymes were codon-optimised for tobacco and synthesized by TWIST Biosciences (San Francisco, United States, https://www.twistbioscience.com/).

Table 2.4: Key literature and reviews for candidate human glycosylation gene selection

Reference
Castilho, A. and Steinkellner, H. (2012), Glyco-engineering in plants to produce human-like N-glycan structures. Biotechnology Journal, 7: 1088-1098.
Schoberer, J. and R. Strasser, Plant glyco-biotechnology. Semin Cell Dev Biol, 2018(b). 80: p. 133-141.
Kallolimath, S., Castilho, A., Strasser, R., Grunwald-Gruber, C., Altmann, F., Strubl, S., et al. (2016). Engineering of complex protein sialylation in plants. *Proc. Natl. Acad. Sci. U.S.A.* 113, 9498–9503.
Montero-Morales, L., and Steinkellner, H. (2018). Advanced Plant-Based Glycan Engineering. *Front. Bioeng. Biotechnol.* 6:81.
Nagels, B., Van Damme, E. J., Pabst, M., Callewaert, N., and Weterings, K. (2011). Production of complex multiantennary N-glycans in Nicotiana benthamiana plants. *Plant Physiol.* 155, 1103–1112.

Schoberer, J., Liebminger, E., Vavra, U., Veit, C., Grünwald-Gruber, C., Altmann, F., Botchway, S. W., & Strasser, R. (2019). The Golgi Localization of GnTI Requires a Polar Amino Acid Residue within Its Transmembrane Domain. *Plant Physiology*, *180*(2), 859–873.

Vézina, L.-P., Faye, L., Lerouge, P., D'Aoust, M.-A., Marquet-Blouin, E., Burel, C., Lavoie, P.-O., Bardor, M., & Gomord, V. (2009). Transient co-expression for fast and high-yield production of antibodies with human-like N-glycans in plants. *Plant Biotechnology Journal*, *7*(5), 442–455.

2.4.4 Constructs for HuGEs with plant CTS replacements

The existing HuGE sequences previously obtained from the NCBI database were used as basis for CTS fusion design (Appendix 3). The region upstream of the HuGE catalytic subunit was replaced with the appropriate plant specific CTS region (Table 1). For the medial-Golgi HuGEs GNTIV and GNTV, MUR3 was used, and for the *trans*-Golgi targeting B4GALT1 and ST6GAL, FUT13 was used (Table 4). CTS-replacement constructs were synthesized by TWIST Biosciences (San Francisco, United States, https://www.twistbioscience.com/).

Table 2.5: Putative CTS regions (Schoberer et al., 2011) from medial- and *trans*-Golgi enzymes used as localisation strategies for novel HuGEs and fluorescent markers.

Golgi enzyme: localisation & source	Cytoplasmic-transmembrane-stem region (CTS) with transmembrane domain **underlined & bold**	Total length of CTS region (aa)
MNS1: Cis-Golgi, Arabidopsis thaliana, Golgi-α-mannosidase I	MARGSRSVGSSSSKWR**YCNPSYY LKRPKRLA**LLFIVFVCVSFVFWDR QT	49aa
MUR3: medial-Golgi, *Arabidopsis thaliana* β-1,2-galactosyltransferase	MFPRVSMRRRSAEVSPTEPMEKG NGKNQTNR**ICLLVALSLFFWALLL YFHFVVL**GTSNIDKQLQLQPSYA	70aa
FUT13: *trans*-Golgi, *Arabidopsis thaliana* α1,4-fucosyltransferase	MPMR**YLNAMAALLMMFFTLLILSF TGI**LEFPSASTSMEHSIDPEPKLSD STS	52aa
ST*: trans*-Golgi *Rattus norvegicus* α2,6-sialyltransferase.	MIHTNLKKK**FSLFILVFLLFAVICVW** KKGSDYEALTLQAKEFQMPKSQE KVA	52aa

2.5 Preparation of competent agrobacterium

2.5.1 Production of competent Agrobacteria

The *Agrobacterium tumefaciens* strain GV3101 was prepared as a large initial volume of 200 ml in LB (Lysogeny Broth media, with 25 µg mL^{-1} rifampicin) from a 28℃-overnight culture. The initial volume of agrobacterium was then pelleted by centrifugation at 3000 x g for 30 mins, at 4℃. The pellet was resuspended in cold 1M $CaCl_2$ and then spun again at 3500xg for 10 mins at 4℃. The pellet was re-suspended in 15 ml 1M $CaCl_2$ and the bacterial solution was split into 400 µl aliquots snap-frozen in liquid nitrogen and stored at -80℃.

2.5.2 Transformation of Agrobacteria

100 µl of competent Agrobacteria 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 15 ml Falcon tube with 1 ml LB and was then allowed to recover in a 28℃ incubator with shaking in 1 ml fresh LB for 2-3 hours. Agrobacteria were then spread onto an agar plate (25 µg mL^{-1} rifampicin, 50 µg mL^{-1} of gentamycin and either 50 µg 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 70% sterile glycerol.

2.5.3 Agrobacterium mediated transient protein expression in tobacco leaf epidermal cells

Transformed agrobacterium liquid cultures were pelleted by centrifugation at 2200g at room temperature for 5 mins. Infiltration buffer (5 mg 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 16hr light/dark cycle) with a blunt 1 ml 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

2.6 RT-PCR for screening stable plants

Seeds collected from regenerated plants (T0) were placed on sterile ½ MS plates (2.4 gL-1 Murashige and Skoog medium, 0.8% phytoagar in deionised water, pH 5.7), sealed with double-layered micropore tape for gas exchange, and placed in a plant growth chamber under 16hr light/dark cycle at 22°C. RNA was extracted from ~500mg of whole stable tobacco seedlings at 14 days old using the NEB Total RNA Miniprep Kit, following the manufacturer's instructions. DNA/RNA protection agent was added to the samples and the plant material was lysed by flash freezing in liquid nitrogen followed by mechanical disruption, after which RNA lysis buffer was added. The RNA was separated from most of the genomic DNA (gDNA) via a gDNA removal column and RNA was purified using 99% ethanol and an RNA purification column. Residual gDNA was removed by in column DNase I treatment. Following washes, the RNA was eluted in 50 µl nuclease-free water and the concentration was measured with a NanoDrop spectrophotometer. Additional gDNA removal was performed by in tube digestion with DNase I and samples were stored at -20 °C.

RT-PCR was performed with OneTaq® One-Step RT-PCR Kit (New England biolabs, E5315S). Primers were designed to target 300bp sequences specific to each HuGE

catalytic domain (table 5) to identify presence of insert in stably transformed seedlings. Bands were identified by gel electrophoresis (0.8% agarose gel) with 2 µl/100 ml SYBR Safe DNA Gel Stain (Invitrogen).

Table 2.6: Forward and reverse primers used for RT-PCR following extraction from stable lines expressing one or both of MUR3-GNT45-388/FUT13-B4ST6-389 vectors. All anneal at 60°C.

Primer name / HuGE for fragment	**Primer sequence**
MUR3-GNT4 fragment	
MURG4RTFWD	TACAGCCCTCATATGCGA
MURG4RTREV	TATGTAGCGTCTCAATCAAGTATGA
MUR3-GNT5 fragment	
MURG5RTFWD	TAGCTGAACCTGGC
MURG5RTREV	TAAACAGTCTTTGCACAGAGC
FUT13-B4GALT1 fragment	
FUTB4RTFWD	AAATTGAGTGATAGTACTAGTCTTCCA
FUTB4RTREV	ATCTCCCGCCTGATTAATAACG
FUT13-ST6GAL1 fragment	
FUTST6RTFWD	GTTTACCATAGCGATATCCCAAA
FUTST6RTREV	GAAGAACTTCTGATAGTAATAGCAAACG

2.7 Lectin binding assays

2.7.1 Preparation of plant extracts

Wild type and ΔXFT *Nicotiana benthamiana* plants, either untransformed or from stable lines, were infiltrated with constructs expressing proteins of interest and leaf tissue was collected 3 days after infiltration (~800mg). The tissue was flash frozen in liquid nitrogen in a mortar and pestle and immediately ground to a fine powder. The powder was homogenized with approximately 1.5 ml of non-denaturing lysis buffer (10 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.5mM EDTA ,0.5% (v/v) Nonidet P-40, 1 mM phenylmethylsulfonyl fluoride). Samples were incubated on ice for 30 minutes with pipetting every 10 and were centrifuged at 3000xg for 5 minutes. The supernatant could be used immediately for lectin binding assays, or was snap frozen and stored at -80°C.

2.7.2 Dot blot lectin binding assays

5 µl of each extract, either fresh or thawed on ice, were applied via 10 µl pipette to a sheet of 0.2 µm pore nitrocellulose membrane (Amersham Hybond™-N, GE healthcare) laid out in a grid pattern. These were allowed to air dry for five minutes, and 5 µl sample re-applied to subsequent dots to generate 5, 10, 15 and 20 µl dots for each extract across the membrane. The nitrocellulose sheet was then incubated in a 2% solution of bovine serum albumin (Sigma) and 0.05% Tween 20 (Sigma) in TBS pH 7, with constant agitation for 30min. The membrane was then rinsed once and washed twice in TBST with agitation for 5 minutes. It was then transferred to a solution of a given fluorescently labelled lectin at a concentration of 1 µg/ml in 0.05% Tween 20 in TBS, with constant agitation for 2.5 hours. Lectin used were one of either Fluorescein labelled *Ricinus communis* agglutin 1 (RCA1, RCA120, 1 mg/ml) from Vector labs, or *Sambucus Nigra* agglutin (SN, 1 mg/ml) lectin labelled with Cy3 (Bio world). The membrane was then washed in three five-minute changes of 0.05% Tween 20 in TBS. Following this the membrane was visualized according to appropriate collection parameters for the given fluorescent tag in a Syngene G:BOX system (Chemi XRQ).

2.7.3 SDS-PAGE lectin binding assays

Each sample was mixed in 1/5 volume to sample with 5x Laemmli sample loading buffer (50 mM Tris-HCl, pH 6.8, 100 mM DTT, 10% w/v SDS, 0.05% w/v bromophenol blue, 50% w/v glycerol) and boiled at 80°C for ten minutes, then left on ice. Leaf extract was separated by electrophoresis on a 12.5% SDS-PAGE (w/v) acrylamide gel (Biorad) under non-reducing conditions (Gallagher et al., 1998). For the initial run separated proteins were visualized by Coomassie blue staining (1 g/l Coomassie Brilliant Blue R-250, 50% v/v methanol, 10% v/v acetic acid) of the resulting gel. For subsequent runs separated proteins within gels were transferred onto a nitrocellulose membrane (Amersham Hybond™-N, GE Healthcare) for one hour at 60V (Biorad). The nitrocellulose sheet was then incubated in a 2% solution of bovine serum albumin (Sigma) and 0.05% Tween 20 (Sigma) in TBS, with constant agitation for 30min. The membrane was then rinsed once and washed twice in TBST with agitation for 5 minutes. It was then transferred to a solution of a given fluorescently labelled lectin at a concentration of 1 µg/ml in 0.05% Tween 20 in TBS, with constant agitation for 2.5 hours. Lectins used were one of either Fluorescein labelled *Ricinus communis* agglutin 1 (RCA1, RCA120, 1 mg/ml) from Vector labs, or *Sambucus Nigra* (SN, 1 mg/ml) lectin labelled with Cy3 (Bioworld). The membrane was then washed in three five-minute changes of 0.05% Tween 20 in TBS. Following this the membrane was visualized according to appropriate collection parameters for the given fluorescent tag in a Syngene G:BOX system (Chemi XRQ).

Chapter 3: Imaging the plant Golgi

3.1 The plant Golgi

The plant Golgi body is the major site of post-translational protein modification in plants, and most relevant for this project, N-linked glycosylation. This occurs in a stepwise fashion as products progress through a non-uniform distribution of glycosylation enzyme through the *cis,* medial and *trans*-Golgi (Figure 3.1), towards the trans-Golgi network (Schoberer et al., 2018). The basic structure of the individual Golgi body is a stack of disc-like cisternae characterized by ordering within the stack (Figure 3.1), with the *cis*-Golgi cisternae as the entry point for cargo exiting the ER, the medial-cisternae and the *trans*-cisternae. From the *trans*-cisternae the product is delivered to the Trans Golgi Network mediating intracellular delivery of ER product (Robinson et al 2020).

Fluorescent microscopy is a powerful tool for investigating cellular functions of protein, often by their specific association or distribution within intracellular structures. Co-localisation for proteins of interest can be performed by comparing the distribution of fluorescently labelled proteins with image analysis tools to determine their interaction (Dunn et al., 2011). To investigate sub-Golgi localisation, we can use known and well understood cisternae-specific markers for plant Golgi. These include the *cis*-Golgi marker MNS1, from *Arabidopsis thaliana* Golgi-α-mannosidase I (Liebminger et al., 2009; Schoberer et al., 2014) and the *trans*-Golgi marker ST, derived from a signal-anchor sequence at the N-terminal of *Rattus norvegicus* α-2,6-sialyl transferase (Runions et al., 2006; Saint-Jore et al., 2002)

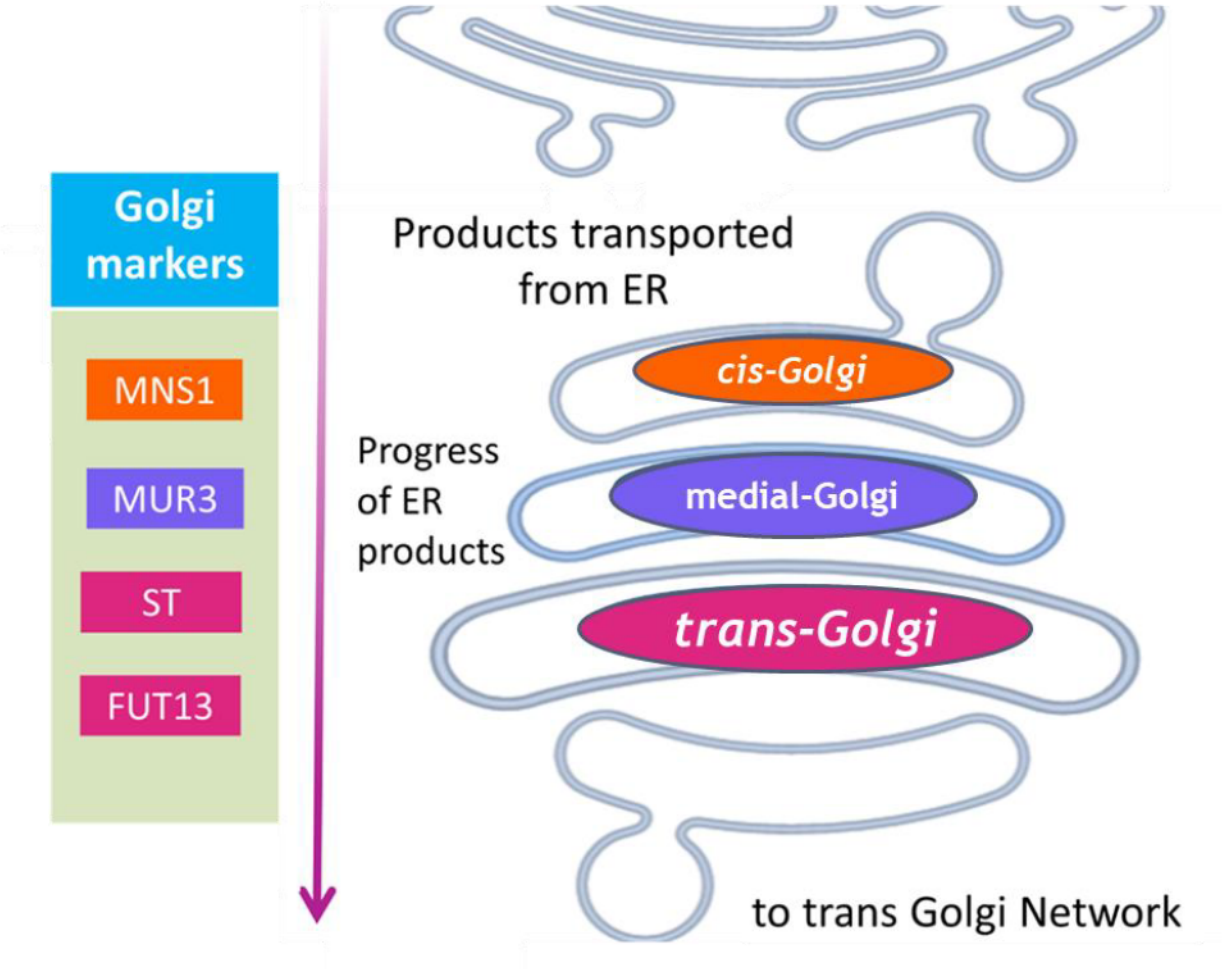


Figure 3.1: Schematic representation of sub-organellar organisation of the plant Golgi body cisternae, with relevant cisternae-specific markers colour-coded. Orange = *cis*-Golgi targeting (MNS1, from Golgi-α-mannosidase I), purple = medial-Golgi targeting (MUR3, from β-1,2-galactosyltransferase), magenta = *trans*-Golgi targeting (ST, from *Rattus norvegicus* α-2,6-sialyl transferase; FUT13, from α-1,4-fucosyltransferase).

In contrast to the singular, immobile perinuclear Golgi body typically found in mammalian cells, a plant cell can contain hundreds of dynamic (Figure 3.2, A) (Osterrieder et al., 2009), highly mobile Golgi bodies that rapidly move widely throughout the cell cytoplasm (Figure 3.2, B) transported along the ER (Robinson et al., 2020; Figure 3.2 C-D) by actin-myosin driven motility (Schoberer et al., 2011). As a result, imaging methods used were adapted to best capture Golgi in detail while accounting for rapid, dynamic movement. This consisted of recording time series of 50+ frames at high magnification, focused on an area of the cell in which Golgi were likely to pass through, to compensate for the trade-off in image quality usually arising in high-resolution acquisition of high-speed objects (Ito et al., 2022). Time series were recorded at the cell periphery, or in ER streams where Golgi would reliably pass in the correct view plane. From these time series individual frames were collected, taking advantage of the changing orientation of Golgi within the cell (Figure 3.2, B) to visualize Golgi body cisternae. This collection method best resembles a trail or game camera by employing a sit and wait approach to collecting high-magnification Golgi in areas of the cell known to have highly mobile Golgi. Imaging was developed with use of known cisternae-specific plant Golgi body markers MNS1 for the *cis*- Golgi and ST for the *trans*-Golgi. This was to visualize Golgi body movement and provide a separation between early and late Golgi cisternae to observe Golgi body mobility (Hawes et al., 2008).

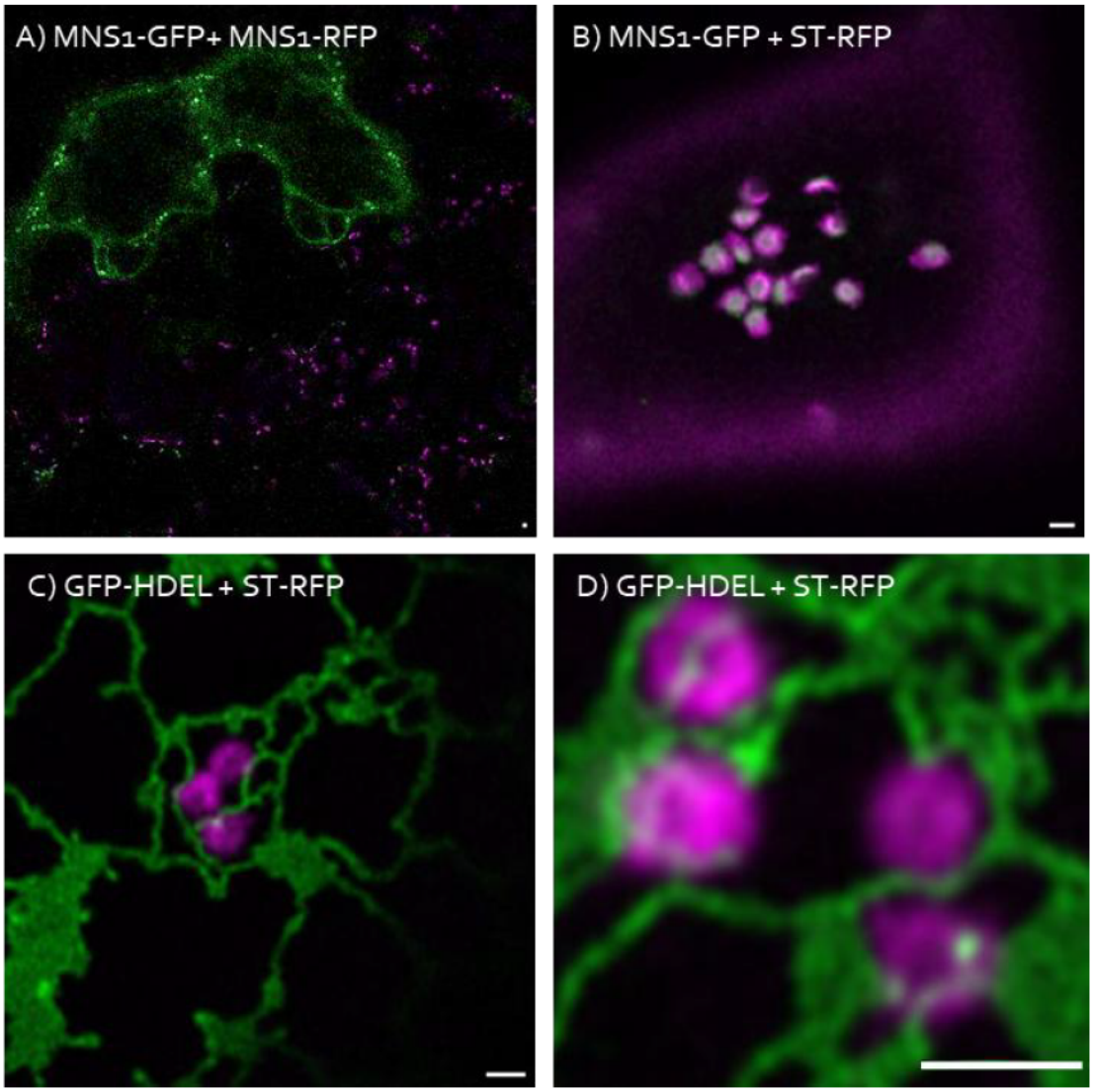


Figure 3.2: Representative confocal microscopy images for Golgi bodies labelled by standard Golgi body markers.

A) Low magnification image of double *cis*-Golgi marker MNS1-eGFP (green) + MNS1-mRFP (magenta) expression. B) High magnification image of *cis*-Golgi marker MNS1-eGFP (green) with *trans*-Golgi marker ST-mRFP (magenta). C) High magnification image of ER marker GFP-HDEL (green) and *trans*-Golgi marker ST-mRFP (magenta) showing positioning of Golgi bodies in the ER. D) Higher magnification image for the previous marker combination. Size bars = 1 µm.

Using this technique of high magnification and consistent zoom, time series were collected, from which individual frames could be identified and used to consistently observe the structure of the plant Golgi body. Most commonly Golgi within the plant cell can appear in a wide ring structure in close proximity to the ER (Figure 3.3, C-D). A typical view of Golgi within the cell (Figure 3.3, A) demonstrates the disc-like structure visible as both a flat and thin profile view depending on the rotation of the Golgi at the moment of image capture, with the small distance between the cisternae (>2 µm) visualized. Two-colour high resolution confocal microscopy using a *cis*-Golgi fluorescent marker protein (MNS1-eGFP, green) and a *trans*-Golgi fluorescent marker protein (ST-mRFP, magenta) was used to visualize the distinction or offset in localisation of fluorescence, and better illustrate Golgi body rotation. A high magnification of a Golgi in flat view (Figure 3.3, B) demonstrated marker localisation within separate cisternae of the stacked-disc structure of the Golgi body may overlap but illustrated the difficulty in quantifying localisation in this visual plane of the Golgi. This issue is lessened when plant Golgi rotate into a side-profile configuration (Figure 3.3, C and D) where physical overlap between cisternae is removed, and the stacked "pancakes" nature of the Golgi cisternae are visible (Ito et al., 2014). It is then simpler to differentiate cisternae-specific expression of two separately localised markers (Figure 3.3, C) from co-localisation of two markers within the same cisternae (Figure 3.3, D).

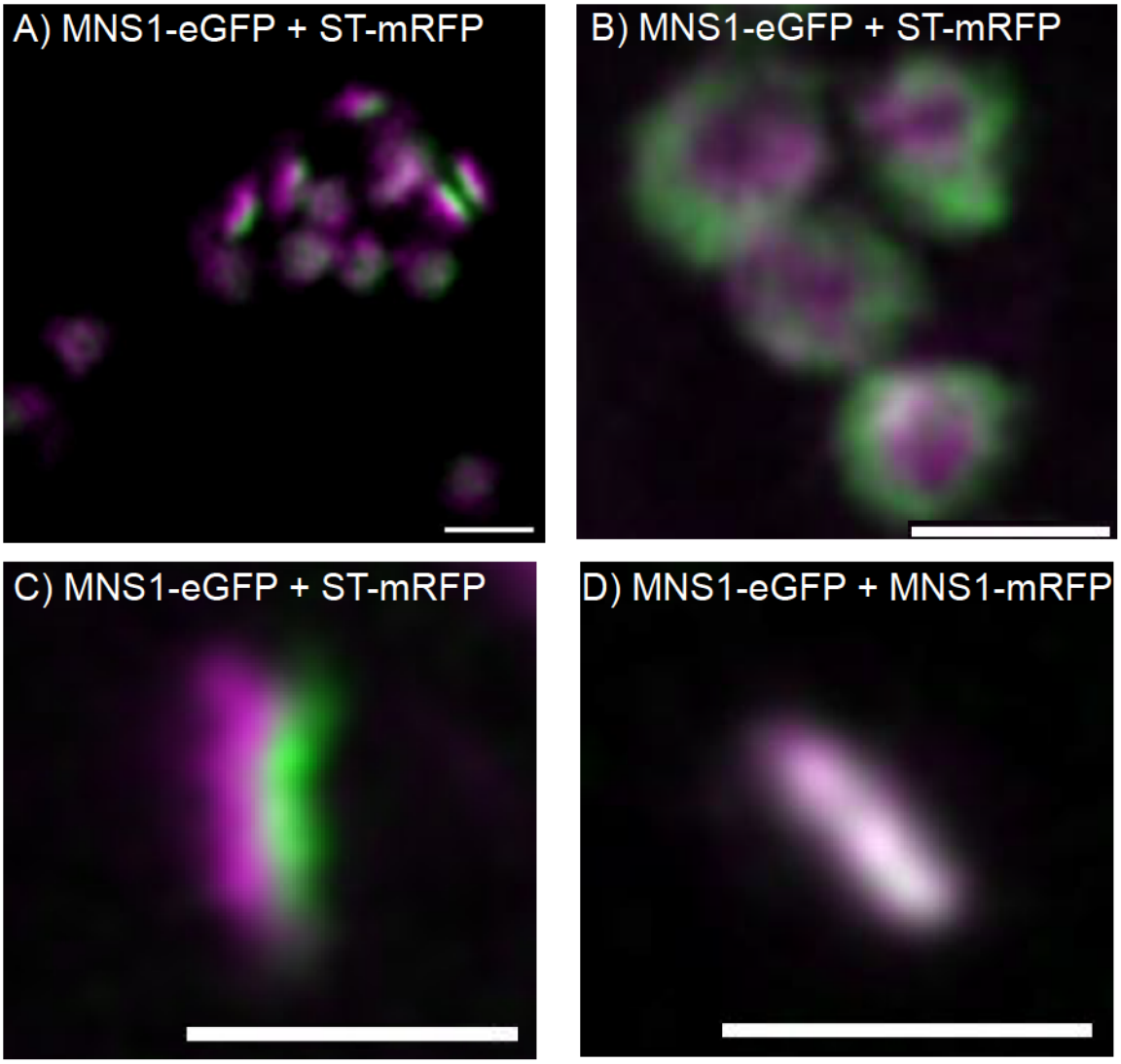


Figure 3.3 Representative confocal images showing Golgi bodies labelled with cisternae-specific markers.

A) Wide view of offset between a *cis*-Golgi body marker (MNS1-eGFP in green) and *trans*-Golgi marker (ST-mRFP in magenta; B) top-down view of MNS1-eGFP in green and ST-mRFP in magenta; C) Golgi body side profile suitable for analysis showing the MNS1-eGFP in green and ST-mRFP in magenta; D) side profile suitable for analysis of individual Golgi body with co-localisation of the *cis*-Golgi marker MNS1 in two separate colours (MNS1-eGFP in green and MNS1-mRFP in magenta). Size bars = 1 µm.

3.2 Aims:

1. To determine a suitable imaging methodology to visualize the plant Golgi for sub-organellar co-localisation analysis and to develop tools to assess fluorescent protein co-localisation between plant Golgi cisternae
2. To develop an accompanying data analysis pipeline to systematically quantify and elucidate sub-organellar localisation of proteins within the plant Golgi body

3.3 Results

3.3.1 Investigating Golgi co-localisation with JaCoPs

Progress on imaging of side-profile Golgi in observing cisternae localisation of fluorescent marker proteins (Figure 3.3, C and D) was promising for developing an analysis pipeline to quantify localisation of expression to specific Golgi body cisternae. It was necessary to determine the specific sub-Golgi localisation of HuGEs, as their correction localisation would directly impact the final glycan product and would need to be correct to achieve an effectively "humanised" glycosylation profile (Palacpac, et al 1999, Bakker et al 2001). A test dataset for MNS1-eGFP + ST-mRFP images of side-profile Golgi similar to those displayed in Figure 3.3 C-D were collected for analysis. The marker combination MNS1-eGFP + ST-mRFP were used as example of a maximum separation possible to observe within the Golgi, as in theory this reflects the possible separation distance between the *cis*- to *trans*-Golgi cisternae (Figure 3.1). A separate dataset was collected of side-profile Golgi expressing the marker combination MNS1-eGFP + MNS1-mRFP. This provided a value for the minimum separation of two markers within the same Golgi cisternae, representing maximum extent of co-localisation possible. This would be a control value for co-localisation. The two conditions serve as reference values for comparative analysis of HuGEs in sub-Golgi localisation. Any analysis methodology explored would need to reliably quantify and significantly differentiate between the two marker combinations.

In order to attempt co-localisation analysis of Golgi data, several approaches were trialled utilizing different image analysis software and methods. Initially the free image processing software Fiji was used (Schindelin et al., 2012), alongside an open-source co-localisation program accessed as a Fiji plugin called Just Another Co-localisation Plugin or JaCoP (Bolte et al., 2006). The JaCoP plug-in is described as a compilation of general co-localization indicators and more recently published methods. Depending on the ticked boxes, the plug-in will evaluate co-localization on two images according to the selected methods (Figure 3.4). According to accompanying literature for the utility it is described as suitable for co-localisation analysis in the Golgi with "The colocalization of two or more markers within these cellular structures (Golgi) may be defined as an overlap in the physical distribution of the molecular populations within a three-dimensional volume, where this may be complete or partial overlap." (Bolte 2006)

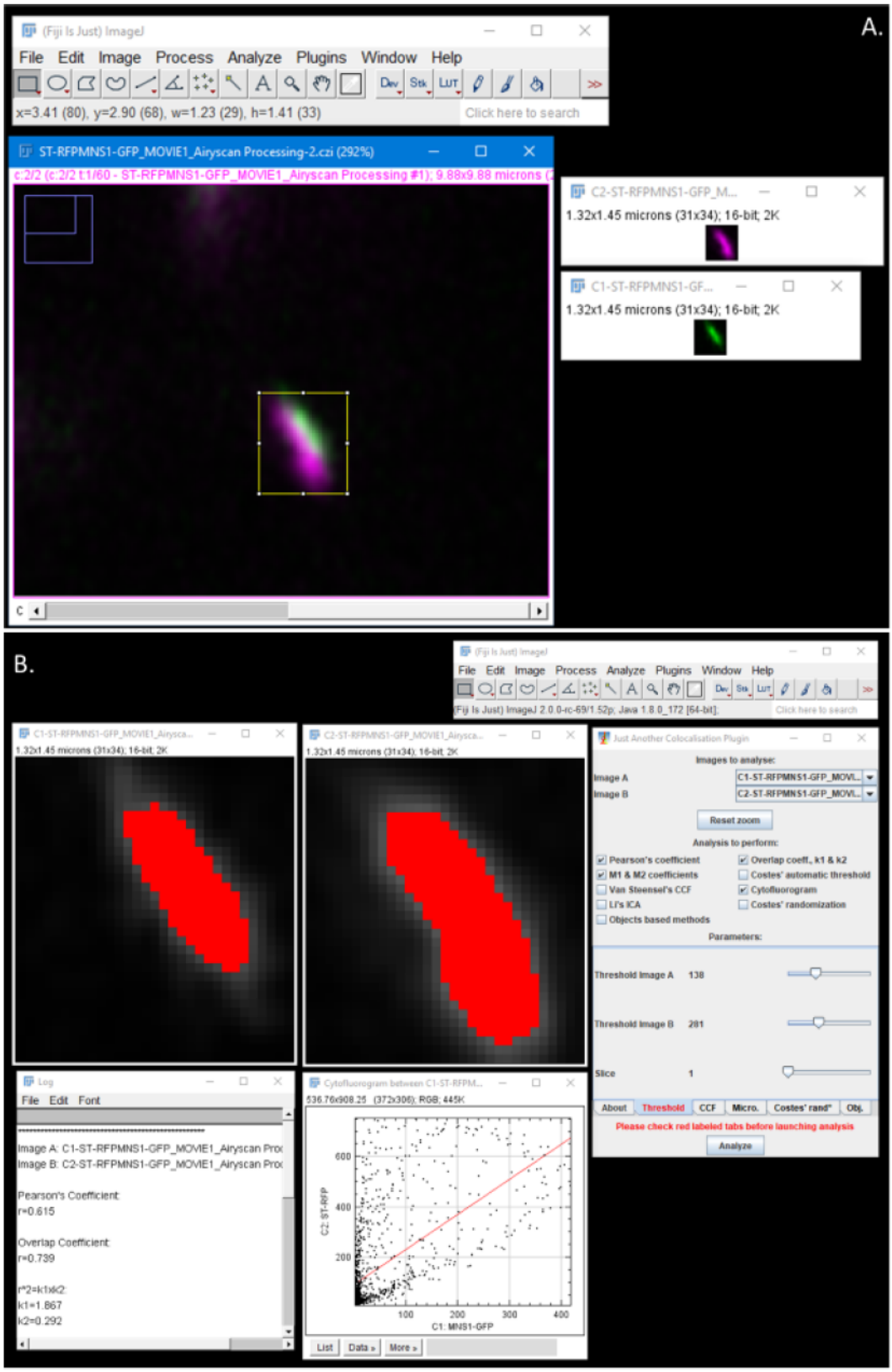


Figure 3.4: Workflow image process using the JaCoP plugin GUI in Fiji for processing a side-view image of a Golgi body. ST-mRFP (shown in magenta) and MNS1-eGFP (shown in green) marker image used as an example.

The program was used for both marker datasets described (MNS1-MNS1 or MNS1-ST) and required a rectangular selection region to be made (Figure 3.4, A). The colour channels are split to run analysis, as in theory the degree of overlap or separation between the two channels will determine the extent of co-localisation of the two signals. A variety of outputs can be generated, but deemed appropriate for this application included cytofluorograms, values for Pearson coefficient (McDonald et al., 2015), overlap coefficient and M1/M2 as a text log (Figure 3.4, B). It is possible to set a fluorescence intensity-based threshold for the analysis manually by use of sliders, or the program provides an option for Costes' automatic thresholding (Bolte 2006). The area of the given image is used for analysis by default, and so for this application the rectangular selection is cropped to an individual Golgi identified as side-profile. The process is repeated for each suitable Golgi identified.

The linear equation describing the relationship between the intensities in two images was calculated by linear regression. The slope of this linear approximation provides the rate of association of two fluorochromes. Pearson's coefficient provides an estimate of the goodness of this approximation. It does this by giving a correlation value from -1 to 1. Here 1 represents the strongest positive correlation (overlap), 0 indicates no correlation (no overlap) and -1 strongest inverse correlation. Here, R equals the Covariance divided by the standard deviation of X, multiplied by the standard deviation of Y (Bolte et al, 2006).

Results discriminated between two marker conditions by Pearson's coefficient (Figure 3.5). Co-localisation within the same cisternae (MNS1-MNS1, Figure 3.5) approached 1 with a result of 0.94, indicating a near complete overlap detected in this marker combination. For separation conditions between the *cis*- and *trans*-Golgi cisternae (MNS1-ST, Figure 3.5) the result was 0.49, indicating a weaker degree of correlation from non-overlapping marker conditions. The differing in Pearson's coefficient between the two combinations indicates a significant difference in localisation successfully indicated by the analysis used.

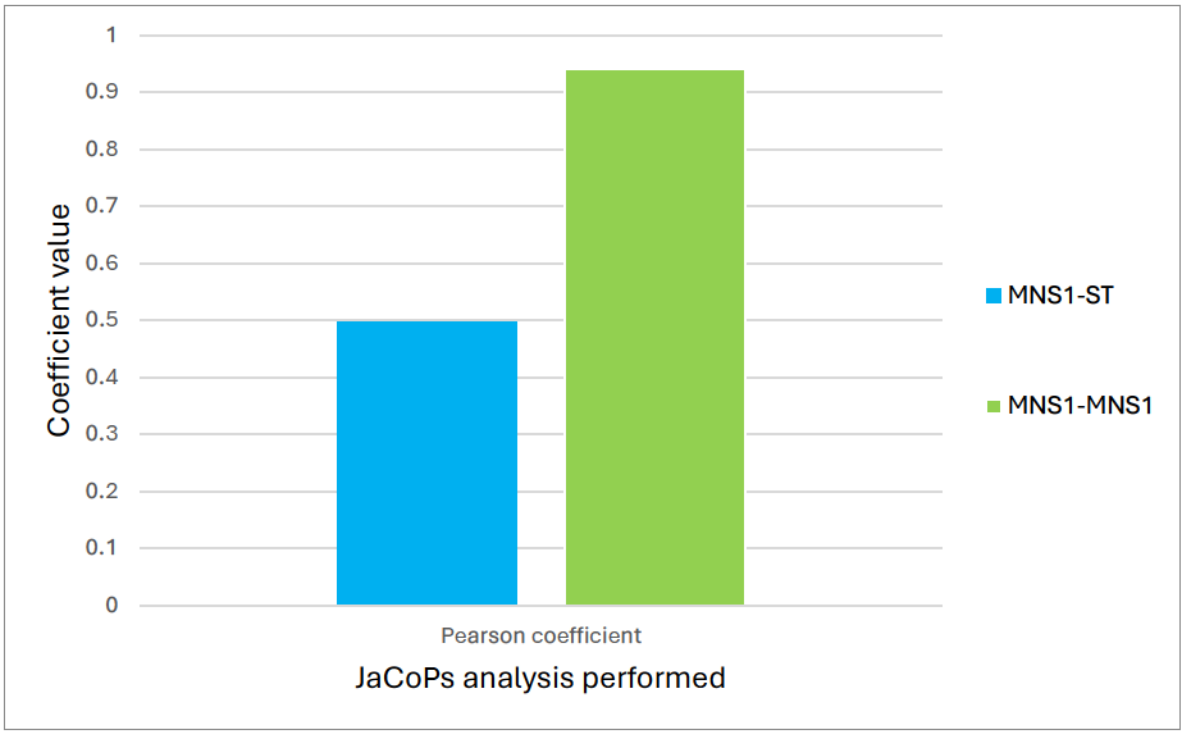


Figure 3.5: JaCoP co-localisation output of Pearson's coefficient for representative marker control with default threshold.

JaCoPs outputs for the test marker conditions for separation between *cis-trans*-Golgi cisternae (MNS1-ST, blue) and co-localisation within Golgi cisternae (MNS1-MNS1, green). Default thresholding applied automatically by JaCoPs. Dataset from Representative Golgi, results from N=1 Golgi biological repeats per condition and >3 technical repeats for each combination.

The overlap coefficient is calculated as the Pearson's coefficient with the mean intensity value of both channels being taken out of the expression. K1 and K2 are defined as two components of the overlap coefficient, the former being related to the first channel total intensity, the latter being related to the second channel total intensity (Manders et al. 1992). K1 and K2 allow identification of contribution of each channel to area colocalised, and have a high sensitivity to differences in signal intensities (Zinchuk et al., 2007). Values for K1 and K2 range above +1 in cases of high signal intensity in one channel. The overlap coefficient for co-localisation condition MNS1-MNS1 (Figure 3.6) was 0.967 (K1 = 0.878, K2 = 1.065), and for MNS1-ST was 0.683 (K1 = 0.507, K2 = 0.92). This is in line with the Pearson coefficient (Figure 3.5) and distinguishes between the two marker conditions, however the value for K1 and K2 in respective channels provided insight into the overlap proportions between channels. The Pearson's coefficient (Figure 3.6) for MNS1-MNS1 the K1 and K2 values are very similar, mostly overlapping due to same-cisternae localisation. In MNS1-ST results, these values are less proportionate, K2 having a much greater overlap than K1 indicating a lack of co-localised coverage between channels, in line with expectations from existing research for two separate cisternae-specific markers.

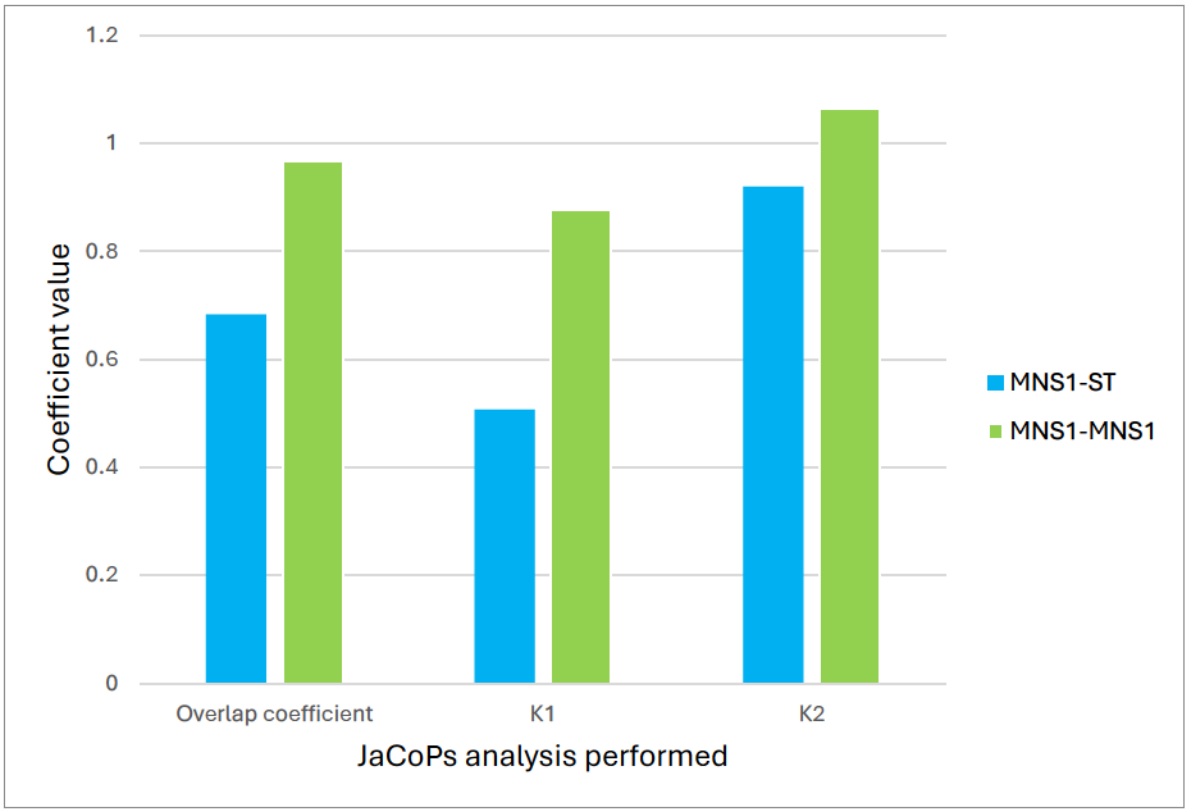


Figure 3.6: JaCoP co-localisation output of Overlap coefficient for representative marker control with default threshold.

JaCoP outputs for the test marker conditions for co-localisation within Golgi cisternae (MNS1-MNS1, green) and separation between *cis-trans*-Golgi cisternae (MNS1-ST, blue). Default thresholding applied automatically by JaCoP. Dataset from Representative Golgi, results from N=1 Golgi biological repeats per condition and >3 technical repeats for each combination.

JaCoPs provided options to threshold fluorescent intensity across an image manually by use of sliders, or automatically via Costes' automatic thresholding (Bolte et al, 2006) Costes' is described as: "limit values for each channel are initialised to the maximum intensity of each channel and progressively decremented. The Pearson's coefficient is concomitantly calculated for each increment. The final thresholds are then set to values which minimize the contribution of noise" (Costes' et al., 2004). An option to automatically threshold analysis was an attractive prospect to rapidly analysing a large dataset of Golgi in a reliable manner. The same dataset and analysis for the overlap coefficient (Manders et al., 1992) as performed in Figure 3.6 was repeated, this time using the option for Costes' automatic thresholding (Figure 3.7). The overlap coefficient for co-localisation condition MNS1-MNS1 was 0.983 (K1 = 0.864, K2 = 1.118), and for MNS1-ST was 0.918 (K1 = 0.44, K2 = 1.918). This showed a significant difference when compared to results obtained with default thresholding (Figure 3.5). Most critically there was a lack of significant difference obtained in the Overlap coefficient between the two marker conditions, both above 0.9 indicating complete overlap, which was not previously the case with MNS1-ST (Figure 3.5 and Figure 3.7). The use of Costes' thresholding made a significant difference to the values of K2 for MNS1-ST, with the second channel now considerably larger and likely causing overfitting of co-localisation due to the difference in fluorescent intensity interpreted, leading to the Overlap coefficient value output. Using these parameters, the analysis no longer delivers a significant difference between the two marker conditions. This may indicate that variation in intensity between the two channels could lead to inability to distinguish between marker conditions, which may cause a serious issue in considerations for experiment design.

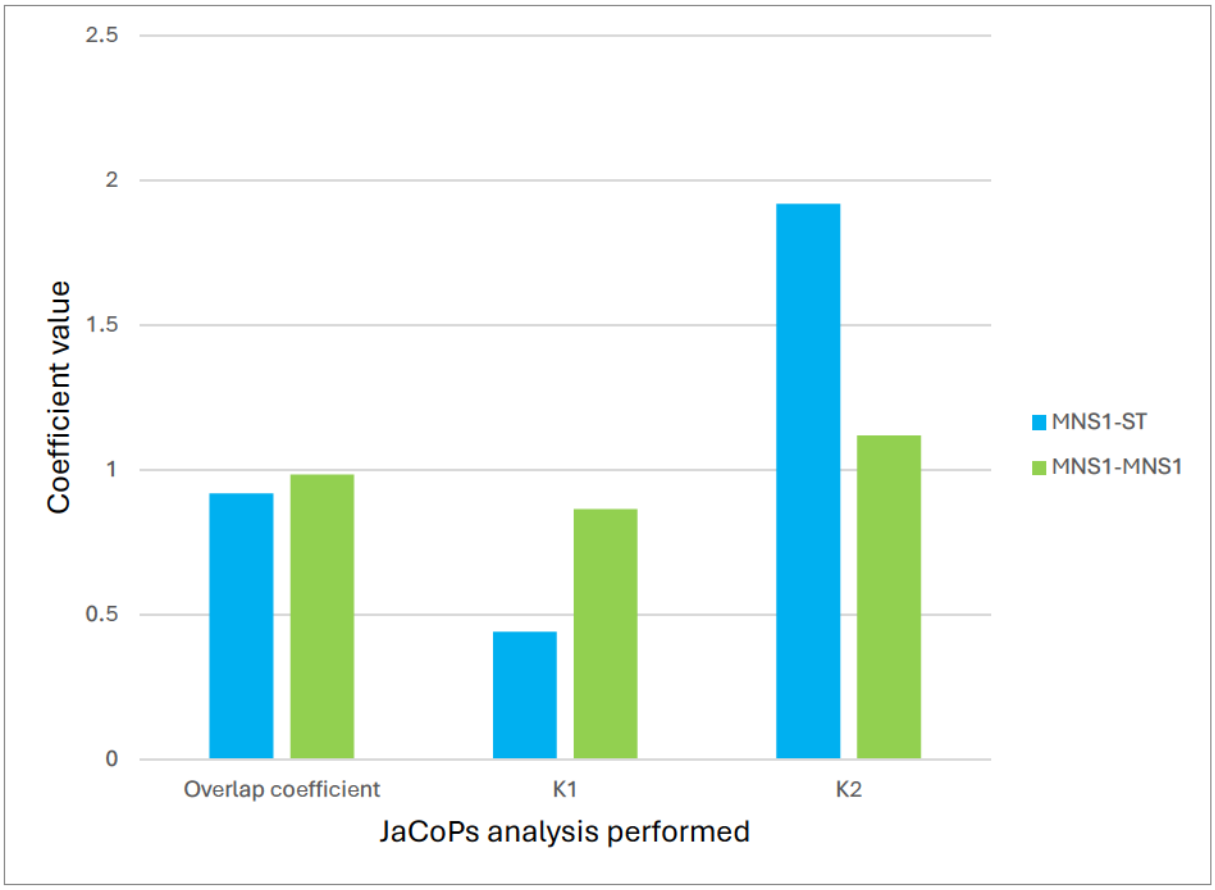


Figure 3.7: JaCoP co-localisation output of Overlap coefficient for representative marker control with Costes' automatic thresholding.

JaCoP outputs for the test marker conditions for co-localisation within Golgi cisternae (MNS1-MNS1, green) and separation between *cis-trans*-Golgi cisternae (MNS1-ST, blue). Costes' thresholding applied. Dataset from Representative Golgi, results from N=1 Golgi biological repeats per condition and >3 technical repeats for each combination.

The M1 and M2 Coefficient calcualted in JaCoPs (Bolte, 2006) will vary from 0 to 1, with 0 corresponding to non-overlapping images and 1 reflecting 100% co-localisation between both images (Manders et al., 1992). M1 is defined as the ratio of the "summed intensities of pixels from the “green” image for which the intensity in the “red” channel is above zero" to the "total intensity in the green channel" and M2 is defined conversely for “red”. This provides proportional overlap without accounting for intensity in either channel, however reliable thresholding is essential to ensure the validity of each analysis. For Costes’ thresholding on analysis of overlap conditions (Figure 3.8) (MNS1-MNS1) this was 0.932 for M1 and 0.852 for M2, the similarity of the two both approaching 1 indicating co-localisation between channels. Conversely, for offset conditions (MNS1-ST), M1 was 0.734 and M2 0.343, indicating a distinct difference in overlap of channels at a lower value than overlap conditions, with a sizeable difference between the M1 & M2 channels, indicating a greater and distinct difference between outputs for MNS1-MNS1 and MNS1-ST that discriminates between the two marker conditions.

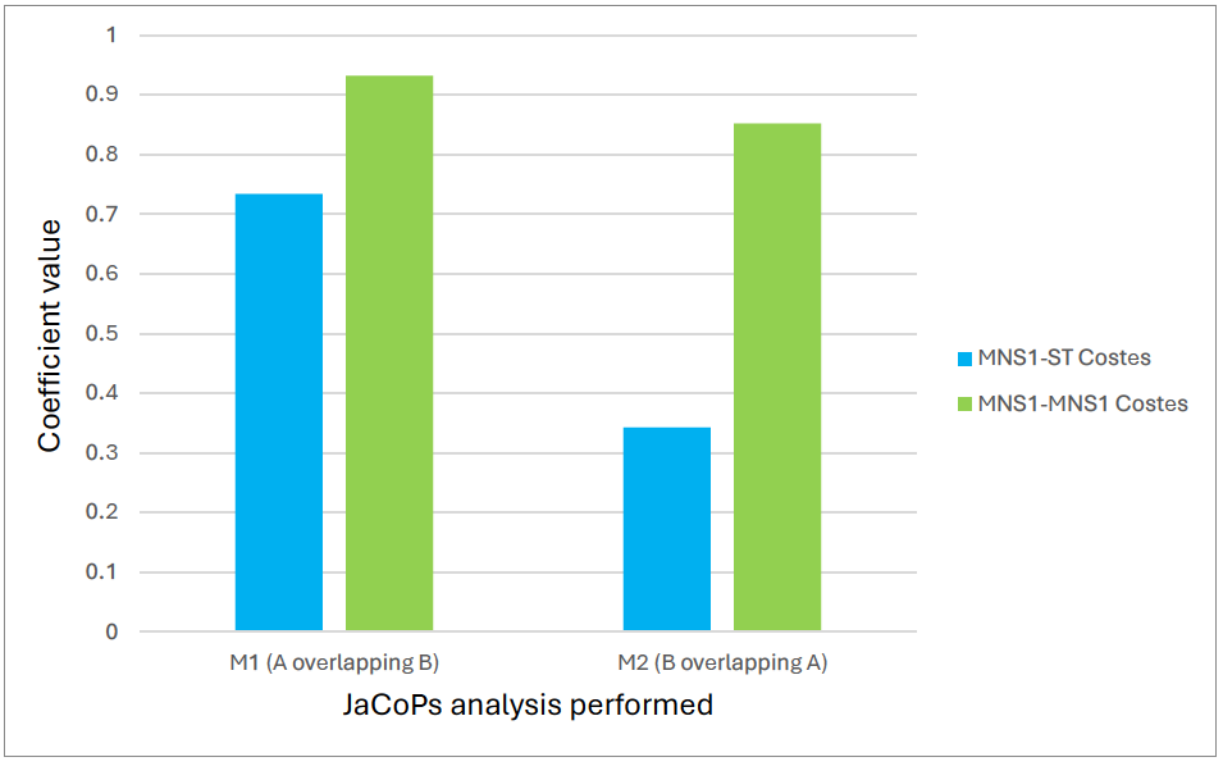


Figure 3.8: JaCoP co-localisation output of M1 & M2 coefficient for representative marker control with Costes' automatic thresholding.

JaCoP outputs for the test marker conditions for co-localisation within Golgi cisternae (MNS1-MNS1, green) and separation between *cis-trans*-Golgi cisternae (MNS1-ST, blue). Costes' thresholding applied. Dataset from Representative Golgi, results from N=1 Golgi biological repeats per condition and >3 technical repeats for each combination.

Overview of JaCoPs (Bolte, 2006) analysis methods evaluated (Figure 3.9). Pearson's coefficient alongside M1/M2 appear to be the most useful in distinguishing between the two marker conditions. Initial results were promising, issues were present in this method's ability to discriminate between marker combinations. Further limitations were present in standardizing the analysis to facilitate a systematic data analysis pipeline. In distinguishing the marker controls between two combinations with distinctly different localisation (complete co-localisation and complete separation) there was only a subtle difference in outputs obtained that had substantial variation between samples. As these marker combinations represented the most distinct differences in localisation, failure to clearly distinguish them was not promising for future analysis on more subtle differences in localisation. Existing research indicated a partial overlap was likely to occur incidentally when investigating fluorescent protein-glycosylation enzyme fusion localisation (Schoberer et al., 2011, Saint-Jore-Dupas et al., 2006) and so any methodology would need to be able to significantly discriminate between combinations. More problematic was reliably thresholding the analysis, with changes in the initial thresholding significantly affecting the output, and not easily standardized by JaCoP. The default threshold values assigned by the plugin varied significantly depending on fluorescent intensity of the sample, making it difficult to standardize systematically. Manual thresholding for each Golgi was an option but was time consuming and introduced greater human error and bias, with specific JaCoPs functions requiring suitable thresholding to function effectively. Applying Costes' automatic thresholding was one approach that was helpful in streamlining analysis, however results derived from this were not necessarily reliable between marker conditions. In addition, the time taken to analyse a single Golgi through this process would not have been feasible across a large dataset. This was a method requirement used in order to account for variation occurring from collected data of Golgi in motion, in addition to biological variation between samples.

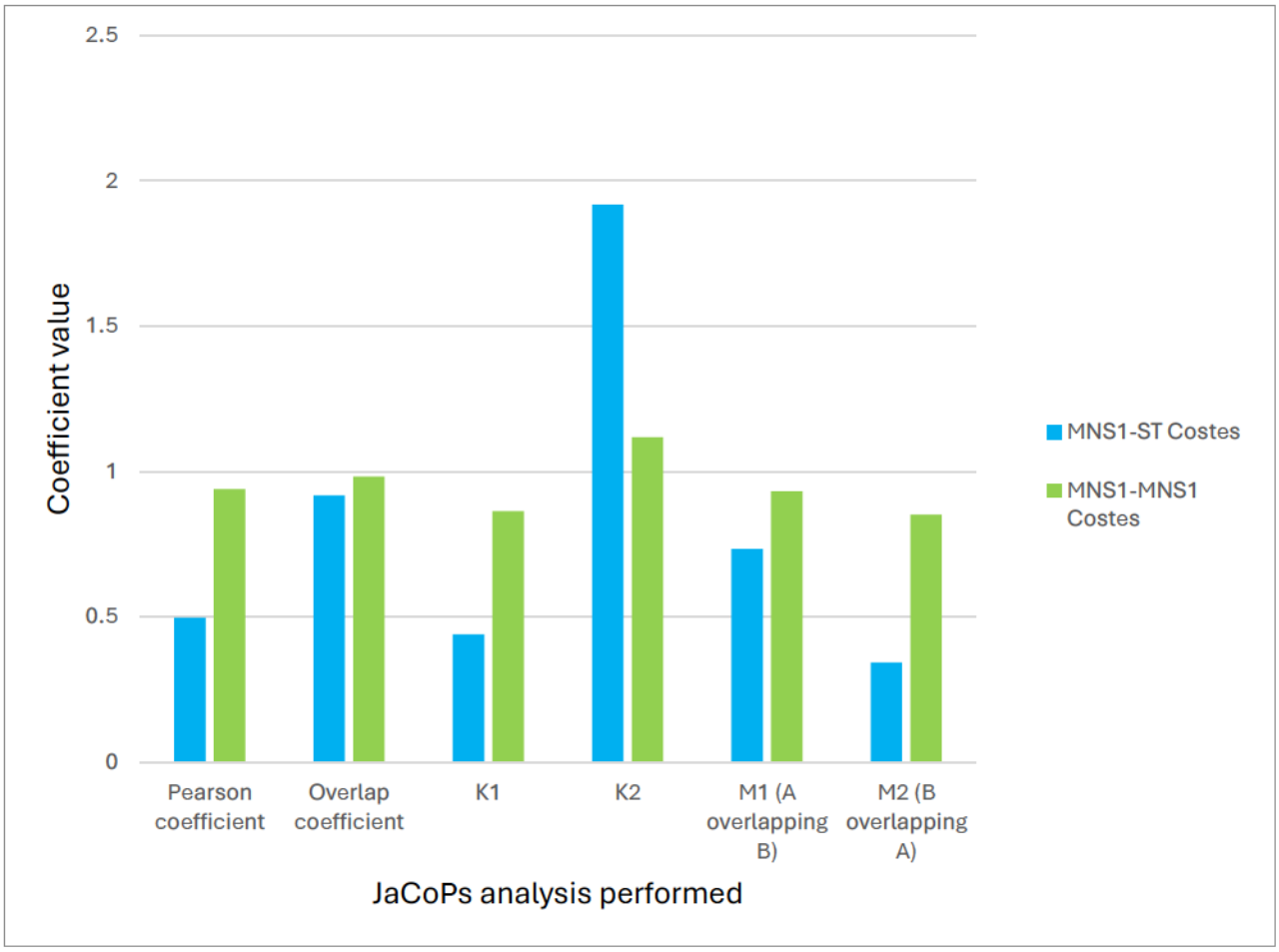


Figure 3.9: Summary JaCoP outputs for representative marker control with Costes' automatic thresholding.

Summary of possible JaCoP outputs for the test marker conditions for co-localisation (MNS1-MNS1, green) and separation between cisternae (MNS1-ST, blue). Costes' thresholding applied. Dataset from Representative Golgi, results from n=1 Golgi biological repeats per condition and >3 technical repeats for each combination.

3.3.2 Co-localisation by distance between peaks

In tandem with the work on attempting co-localisation by JaCoPs (Bolte, 2006), a different approach was tested based on measuring fluorescent intensity across line profiles and determining the distances between peak intensities of each channel. As side-profile Golgi enable view of cisternae localisation of fluorescent markers, it was theorized that it would be possible to determine degree of separation or co-localisation across the Golgi based on the distance between the peaks of fluorescent intensity between each channel. While there are no significant gaps identified between cisternae, this supports the description of localisation of endogenous glycosylation enzymes in literature as non-uniform, occurring in overlapping gradients with other enzymes (Schoberer et al., 2011). In this manner peak analysis would provide the localisation of greatest intensity of each marker or enzyme introduced, or at minimum a small comparative range of distribution. An advantage of this technique is that the fluorescent intensity of each channel does not significantly impact the output, and so thresholding the analysis or stoichiometry of each fluorescent construct's expression is not an issue in the same manner as it had been for JaCoPs.

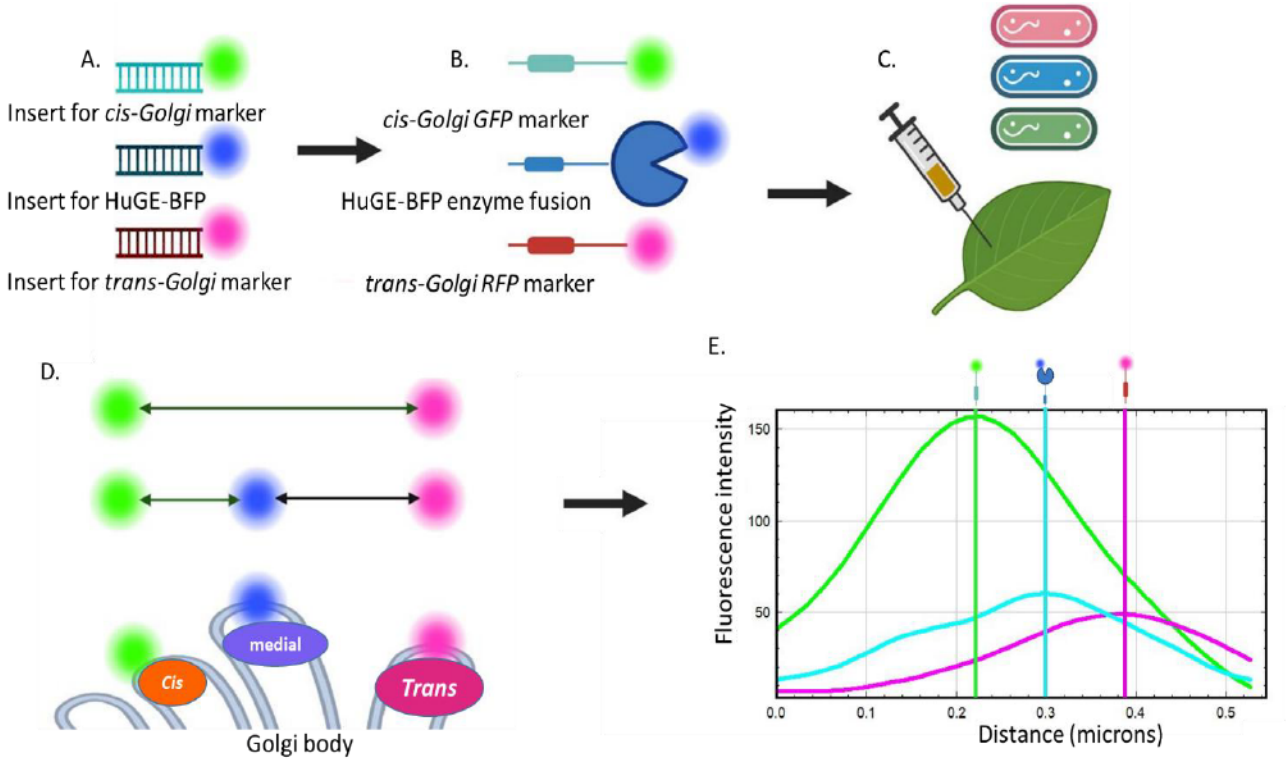


Figure 3.10: Conceptual overview of distance between peaks process for fluorescent protein sub-Golgi localisation analysis

A) DNA inserts for protein expression to generate B) fluorescently tagged Golgi markers as single or paired combinations (by self-cleaving peptides), fluorescently tagged HuGE constructs (blue). C) Agrobacterium-mediated expression of pairs / triplets of constructs in mature plant tissue. D) Cisternae-specific expression of markers (green and red in this instance) used to determine distance across Golgi body cisternae or compared with HuGE localisation. E) Example line profile data across Golgi body generated.

This approach was initially evaluated using the Fiji line profile utility (Schindelin et al., 2012) and conducted across representative Golgi bodies for the same marker conditions as previously described (Figures 3.5-3.9). These were MNS1-eGFP + ST-mRFP as example of a maximum separation between the *cis*- to *trans*-Golgi cisternae, MNS1-eGFP + MNS1-mRFP as minimum separation for localisation within the same cisternae. To perform with suitable side profile Golgi, a line was drawn perpendicular to the selected Golgi body (Figure 3.11) to generate a line intensity profile of the two channels.

This process easily distinguished the separation of the markers and gave a clear output based on physical distance across the Golgi of the two markers, with a single easily interpreted metric of distance with which to compare results between marker and HuGE combinations. This approach was selected for co-localisation analysis as it quickly gave interpretable data consistently across multiple Golgi, did not require any considerations for differences in threshold intensity. This could be rapidly and systematically conducted to generate averages from larger datasets of Golgi, providing reliable data for co-localisation analysis in this instance.

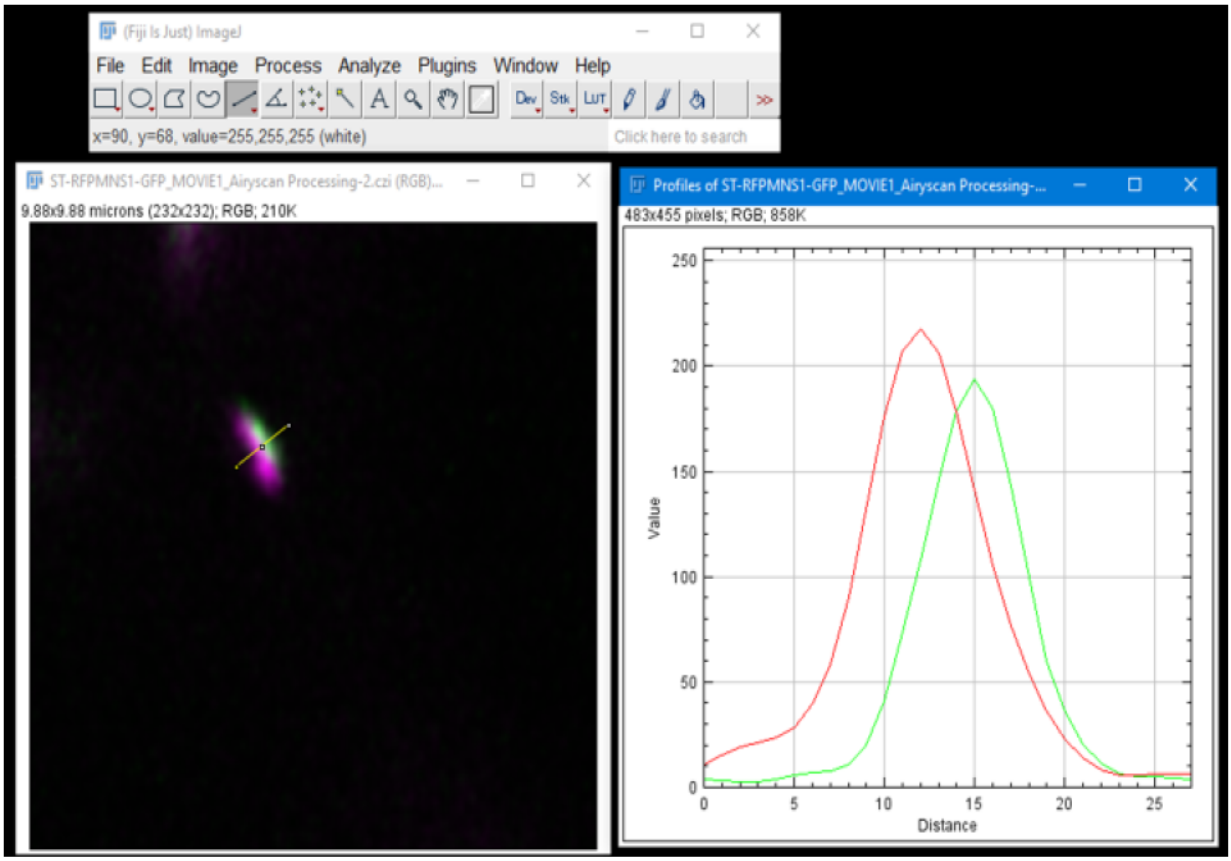


Figure 3.11: Workflow process image with GUI for use of line profile utility for Fiji. ST-mRFP (magenta/red) and MNS1-eGFP (green) marker image used as a representative example.

This process was optimised and standardized to make a data analysis pipeline for use in determining HuGE localisation with comparison to known plant Golgi body markers. ZEN blue software (ZEISS ZEN Microscopy Software RRID:SCR_013672; version 3.5.093.0000) was later used in the same manner to generate a line intensity profile (Figure 3.12) of channels as it provided better data export options.

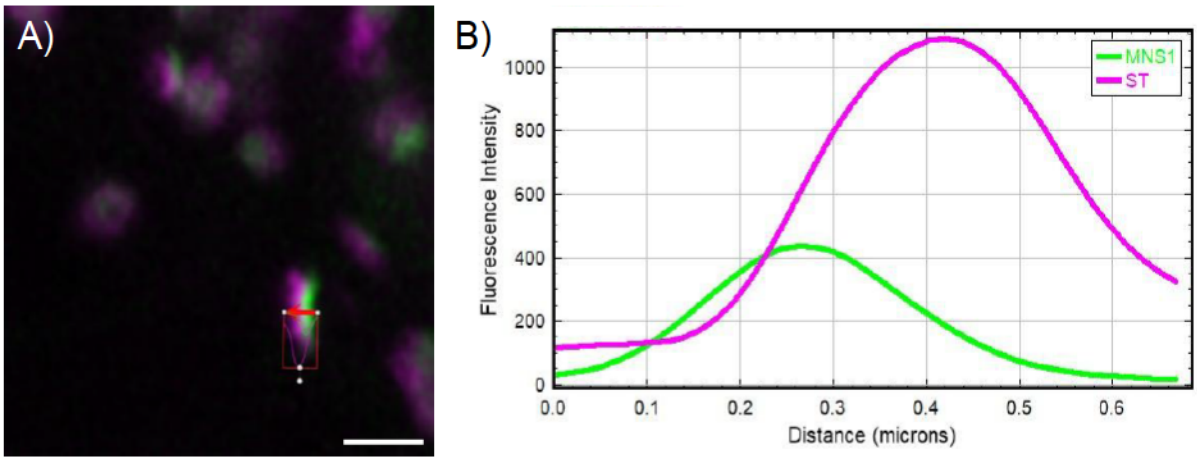
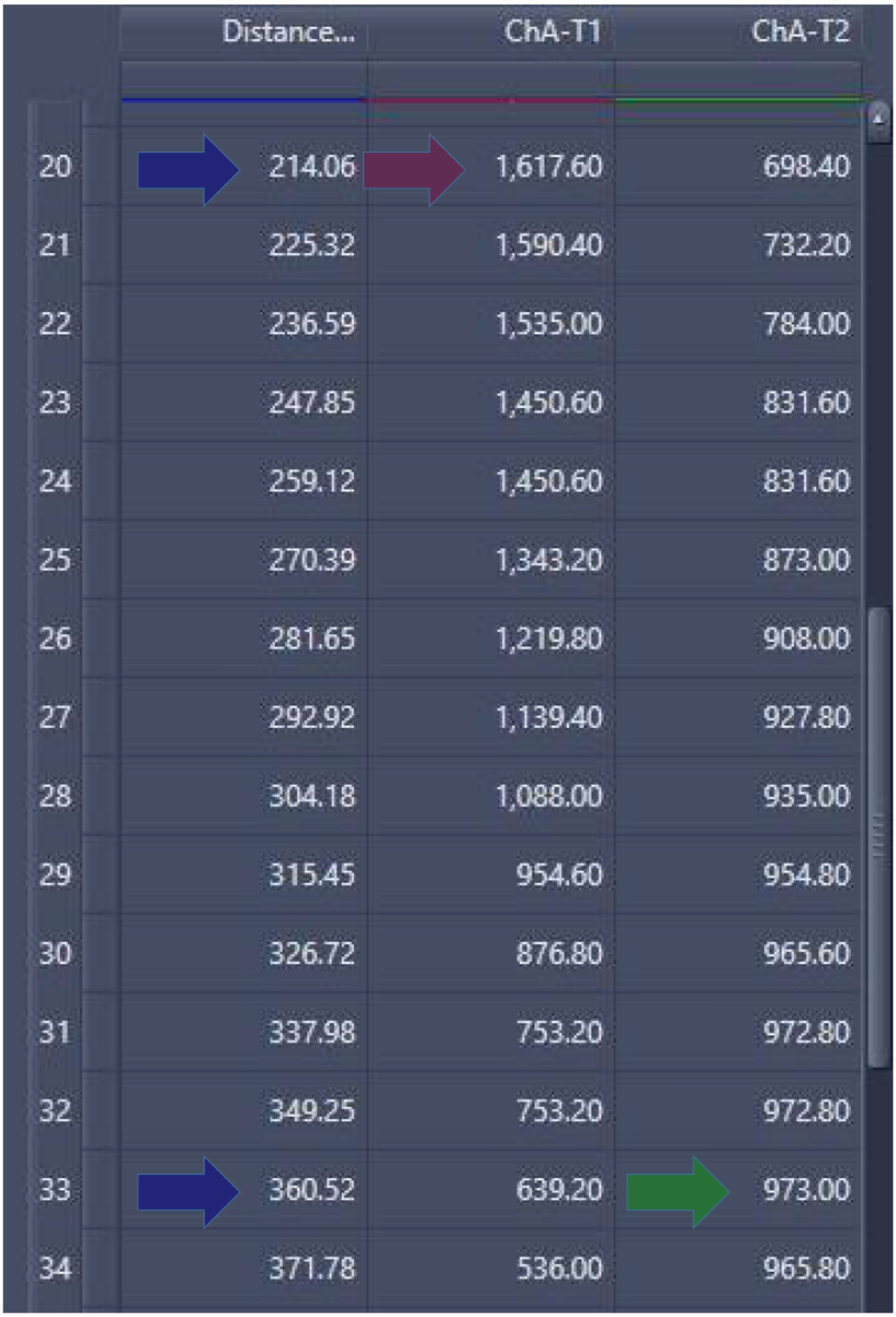


Figure 3.12: Example outputs of line profile analysis for MNS1-eGFP +ST-mRFP
A) Example confocal image of side profile Golgi with line profile drawn. B) Output of line intensity analysis from ZEN blue for line drawn, ST-mRFP (magenta) and MNS1-eGFP (green). Size bars = 1 µm.

The fluorescence intensity of each channel was collected across the distance of the line profile for each Golgi body analysed (Figure 3.13), with the data being used to determine the maximum peak intensity of each respective channel (Figure 3.13, red and green arrows). The distance value between the two peaks in nm (Figure 3.13, blue arrows) was recorded. The data for each Golgi analysed was exported as a .CSV file and accessed in Microsoft Excel to calculate averages and standard deviations for each combination. For test data this was performed manually. Each Excel produced for an individually assessed Golgi was opened, the distance at which each channel's peak occurs recorded (Figure 3.13, blue arrows). The larger value subtracted from the smaller to give the distance between the two peaks, and the process repeated for each Golgi analysed. This was then averaged in Excel to produce a mean and standard deviation for each combination. This process was labour and time intensive, and only conducted manually for the test dataset. When distance between peaks was decided as the most promising method of co-localisation, this process was automated by a Python script, the process of which is later described.



	Distance...	ChA-T1	ChA-T2
20	214.06	1,617.60	698.40
21	225.32	1,590.40	732.20
22	236.59	1,535.00	784.00
23	247.85	1,450.60	831.60
24	259.12	1,450.60	831.60
25	270.39	1,343.20	873.00
26	281.65	1,219.80	908.00
27	292.92	1,139.40	927.80
28	304.18	1,088.00	935.00
29	315.45	954.60	954.80
30	326.72	876.80	965.60
31	337.98	753.20	972.80
32	349.25	753.20	972.80
33	360.52	639.20	973.00
34	371.78	536.00	965.80

Figure 3.13: Example data output of line profiles in ZEN blue, arrows added for illustration.

Example line profile output shown. Maximum fluorescent intensity of channel 1 (red arrow) and maximum of channel 2 (green arrow) indicated, with the distance separating the two peaks along the line intensity profile indicated (blue arrows).

Using this approach, a test dataset was collected using the previous marker conditions (Figure 3.14).The dataset analysis was performed on side-profile Golgi expressing the marker combination MNS1-eGFP + MNS1-mRFP to provide a comparison for the minimum separation and maximum co-localisation possible between two markers. This provides a control value for co-localisation of expression within the same cisternae of the plant Golgi. Similarly, quantification of a dataset for MNS1-eGFP + ST-mRFP provided a control dataset for maximum separation and minimum co-localisation between most the distant Golgi cisternae (*cis*- to *trans*-Golgi). These two marker combinations served as controls for sub-Golgi co-localisation analysis, and any analysis methodology ultimately would need to reliably quantify and differentiate between them. The rationale for such being that HuGEs were unlikely to be so clearly defined in expression as the markers.

Peak distance analysis for the test dataset (Figure 3.14) was able to significantly distinguish between the two marker conditions. Average distance values of the localisation offset MNS1-eGFP+ST-mRFP were 121.3 ± 11.7 nm. For overlap conditions MNS1-eGFP + MNS1-mRFP the dataset average was 16.2 ± 10.5 nm. These were distinctly separated in terms of distance between peaks in a statistically significant manner. In comparison to analysis of the same marker conditions using JaCOPs (Figure 3.9) this methodology was able to generate a more precise distinction between conditions.

Results of the rest dataset analysis indicated peak distance in ZEN was the most appropriate methodology for analysing the sub organellar localisation of HuGEs with marker controls in the plant Golgi body. The methodology had set baseline values for maximum and minimum distances across the Golgi. Additionally, in this instance distance between peaks analysis was faster and easier to repeat than JaCoPs (Bolte et al., 2006) and more readily lent itself to use in larger datasets intended.

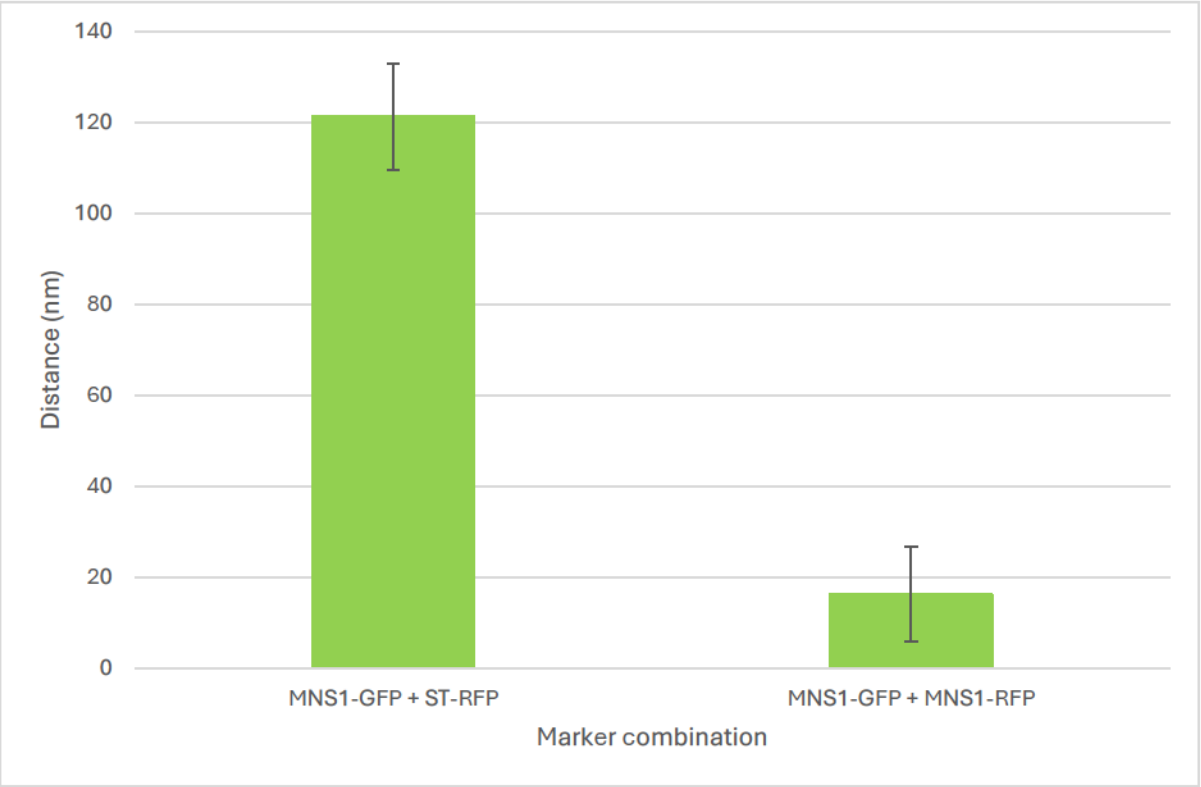


Figure 3.14: Distance between peaks analysis for fluorescent protein marker combinations MNS1-eGFP + MNS1-mRFP and MNS1-eGFP + ST-mRFP.

Peak distance graphs showing co-localisation analysis for overlap of localisation MNS1-eGFP + MNS1-mRFP and offset of localisation MNS1-eGFP + ST-mRFP, respectively. Markers MNS1-eGFP (*cis*-Golgi cisternae targeting) and ST-mRFP (*trans*-Golgi cisternae targeting). Results shown from n = 3 biological replicas and >28 technical repeats for each combination.

Modifications were made to optimize this process to the throughput necessary and to improve precision in the manner data was processed. To generate the analysis from the test dataset (Figure 3.14) the values for distance between peaks for each Golgi had to be individually calculated from ZEN output (Figure 3.15) in Excel. This process was time consuming and error prone process and was later automated by use of a Python script.

L	M	N	O	P
Channel 1	Channel 2	Gaussian I	Max Value	Distance
245.201	222.7861	22.41488	0	
233.696	212.5475	21.14854	17.57176	
254.2291	240.2529	13.97618	0	
217.8489	207.4693	10.37968	0	
169.1616	147.6828	21.47876	17.57176	
182.1005	166.3642	15.73625	17.7785	
233.7255	208.1534	25.57214	17.7785	
151.4259	121.3282	30.09769	52.93267	
196.5929	167.9806	28.61227	36.16588	
200.5003	187.8058	12.69444	17.57176	
	Mean	20.21108	17.73708	
	stdev	6.802474	16.74066	

Figure 3.15: Example screenshot of a data output subset from Python report for MNS1-eGFP + MNS1-mRFP.

Each row provides a single Golgi body line profile analysis from one combination showing maximum intensities and the distance between them, processed by use of a Python script (Figure 3.16). Table calculates mean and standard deviation (stdev) between two conditions, ZEN default processing (column O) or Gaussian fitted measurements (column N).

The script implemented (Figure 3.16) established a pipeline for data analysis, but had limitations based on output format of ZEN data (Figure 3.15). Default line profile data collected from ZEN formats distance measurements between the pixel grids, leading to a regular but imprecise scale for measuring distance that lacks small increments. E.g. close co-localisations have a distance between peaks of either 0nm or ~17nm with no values between (Figure 3.15, column O) that reduced the reliability of and averages and conclusions generated from this data. By adapting the Python script to fit data to a Gaussian distribution during processing (Figure 3.15, column N) this prevented the data binning occurring in default ZEN output data. This in turn significantly decreased artificial standard variation in the result. The modified Gaussian-fitted data and the unmodified max distance data output by ZEN was recorded side by side in the Python output report to compare for error and for transparent data handling (Figure 3.15).

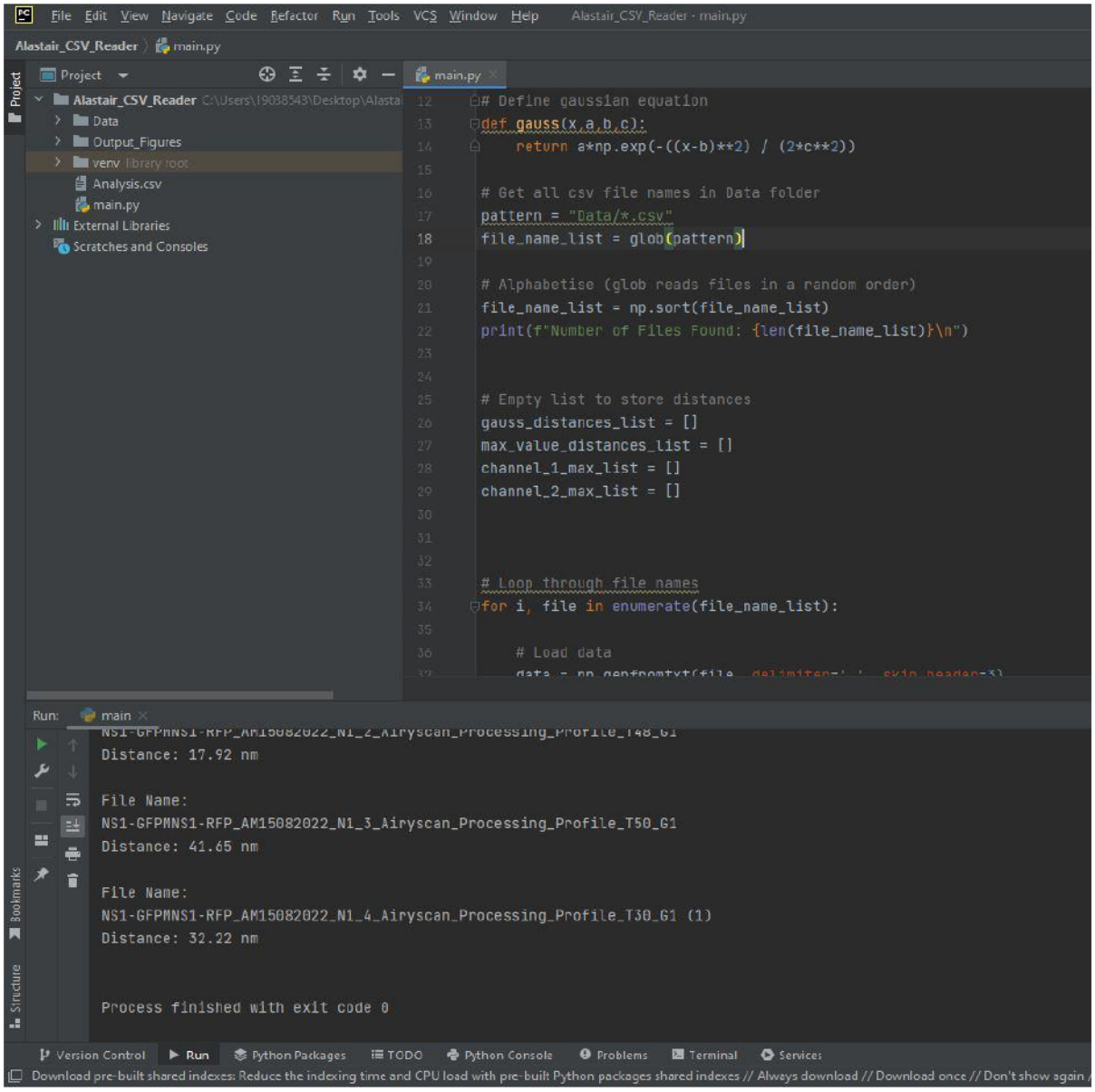


Figure 3.16: Example run of final Python script for peak distance calculation in PyCharm GUI window. Full code in Appendix 1.

Extracting peak distance values from ZEN output was automated through use of a Python script (Appendix 1, with code for Python version 3.9) to quickly iterate through the multiple .csv files generated by ZEN and obtain values for the distance between peaks. The original .csv files from the individual Golgi line profiles would be placed in a folder, accessed by the script sequentially, the maximum expression in each channel determined and values for distance between peaks calculated. These would then be written to a final .csv file as a report of the data processed, from which averages and standard deviations could be calculated.

Analysis results of the same Golgi dataset was expressed in both ZEN default output and Gaussian-fitted values (Figure 3.17). As the original test dataset for peak distance analysis (Figure 3.14) was individually calculated in Excel the same dataset of Golgi profiles was reused but processed by Python script (Figure 3.17). The mean values for co-localisation marker condition MNS1-eGFP + MNS1-mRFP were not significantly different between the two outputs (ZEN 16.2 ± 10.5 nm, Gaussian 20.4 ± 4.8 nm) however the standard deviation was significantly reduced in the Gaussian output, likely due to the imprecise nature of co-localisation data in the ZEN output. The values for both outputs for separation marker conditions MNS1-eGFP + ST-mRFP were also not significantly different (ZEN 121.3 ± 11.7 nm, 126.6 ± 12.3 nm) but did not have the difference in standard deviation as was observed in the MNS1-eGFP + MNS1-mRFP marker condition. The result obtained supported the use of a Python script to automate analysis of Golgi line profiles, as well as the rationale for using Gaussian-fitted values for reporting analysis.

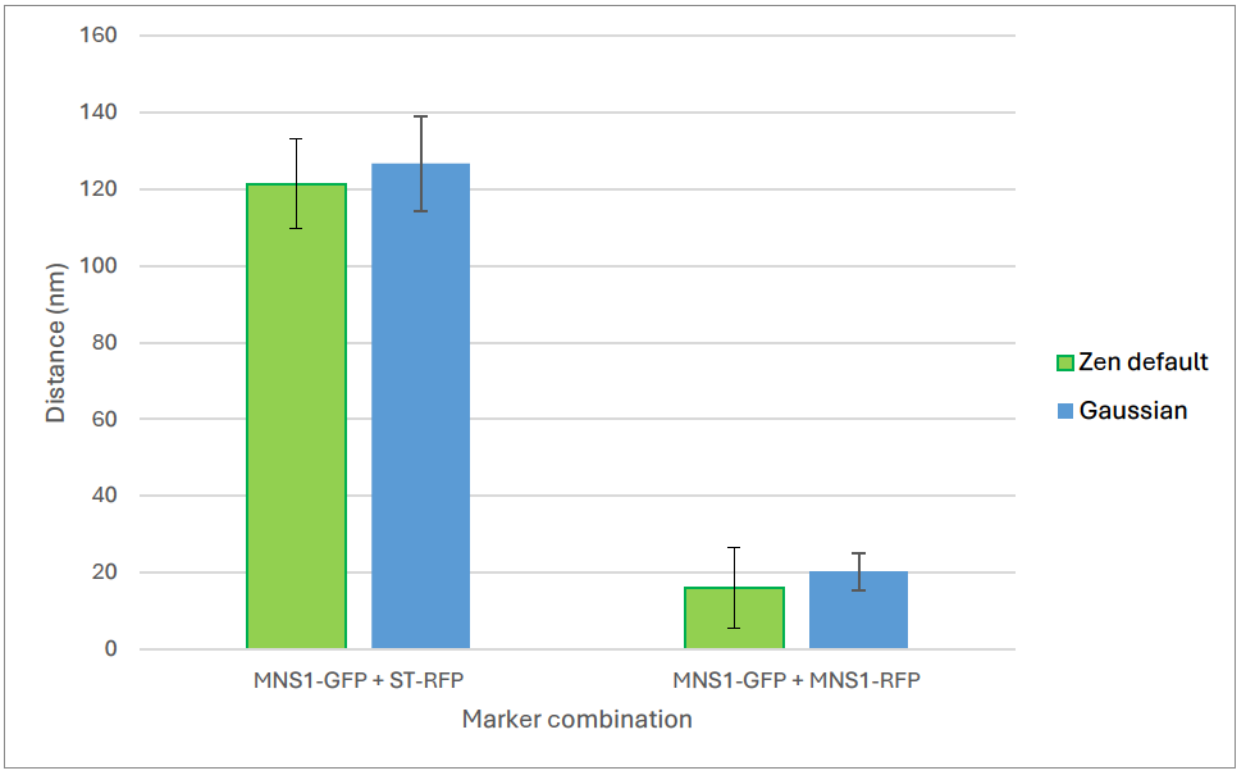


Figure 3.17: Gaussian distribution vs ZEN default in distance between peaks analysis for fluorescent marker control combinations MNS1-eGFP + MNS1-mRFP and MNS1-eGFP + ST-mRFP.
Peak distance graphs showing co-localisation analysis for overlap of localisation MNS1-eGFP + MNS1-mRFP and offset of localisation MNS1-eGFP + ST-mRFP, respectively. The same dataset shown with Gaussian fitting (blue) or with original ZEN data (green). Markers are MNS1-eGFP (*cis*-Golgi cisternae targeting) and ST-mRFP (*trans*-Golgi cisternae targeting). Results shown from n = 3 biological replicas and >28 technical repeats for each combination.

The project so far had established a method and analysis pipeline for analysing the sub-Golgi localisation of fluorescent protein constructs and present results (Figure 3.18), with the intention to use these tools to quantify HuGE localisation that was hopefully cisternae-specific to the medial-or *trans*-Golgi depending on the individual HuGE. The test dataset analysed had provide a maximum and minimum distance between peaks that reflected localisation between or within cisterna, which could act as a guideline for sub-organellar HuGE localisation across the Golgi body.

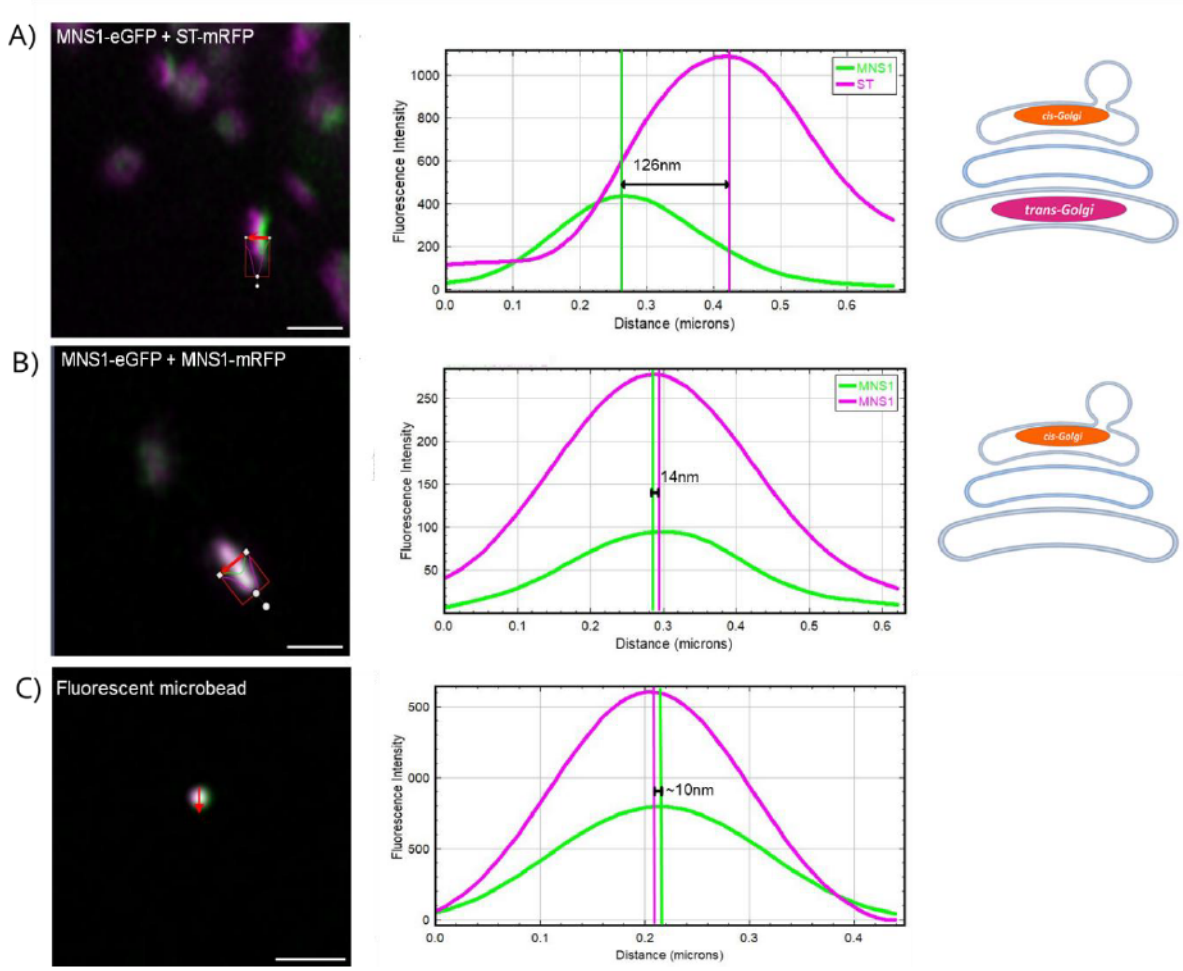


Figure 3.18: Presentation figure for peak distance analysis of test marker datasets.

Line profiles for fluorescent intensity analysis across a Golgi body, with known markers A) MNS1-eGFP+ST-mRFP localised to separate Golgi body cisternae to demonstrate the distance between peaks, or B) MNS1-eGFP+MNS1-mRFP double markers for the same Golgi body cisterna in two colours to demonstrate overlap of expression within one Golgi body cisterna. C) Fluorescent microbeads (ZEISS) used to baseline the configuration and provide a control for the minimum possible separation observable in one object between two channels. Example confocal images and graphs are representative of for the dataset. Size bars = 1 µm.

3.3.3 New marker analysis by distance between peaks

As a further proof of efficacy, a dataset was collected comparing existing Golgi markers to the localisation of two specific Golgi markers produced for this project: MUR3-FP for the Medial-Golgi and FUT13-FP for the *trans*-Golgi cisternae. Plant and human Golgi-resident N-glycosylation enzymes are all type II membrane proteins, each possessing at the N-terminal region a short cytoplasmic domain, a transmembrane domain, and a stem region (CTS region). The CTS region provides sub-Golgi localisation, offering an explanation to the non-uniform distribution of glycosylation enzymes in the Golgi (Uliana et al., 2006; Schoberer et al., 2011). Existing Golgi markers are based on the premise of using a CTS domain paired with a fluorescent protein to cause cisternae specific localisation of the fluorescent protein (Liebminger et al., 2009; Runions et al., 2006). New markers were designed with plant specific CTS domains MUR3, FUT13 (Figure 3.19) to provide cisternae specific localisation as a comparison to existing Golgi body markers (Schoberer et al., 2019) and provide a reliable medial-Golgi marker.

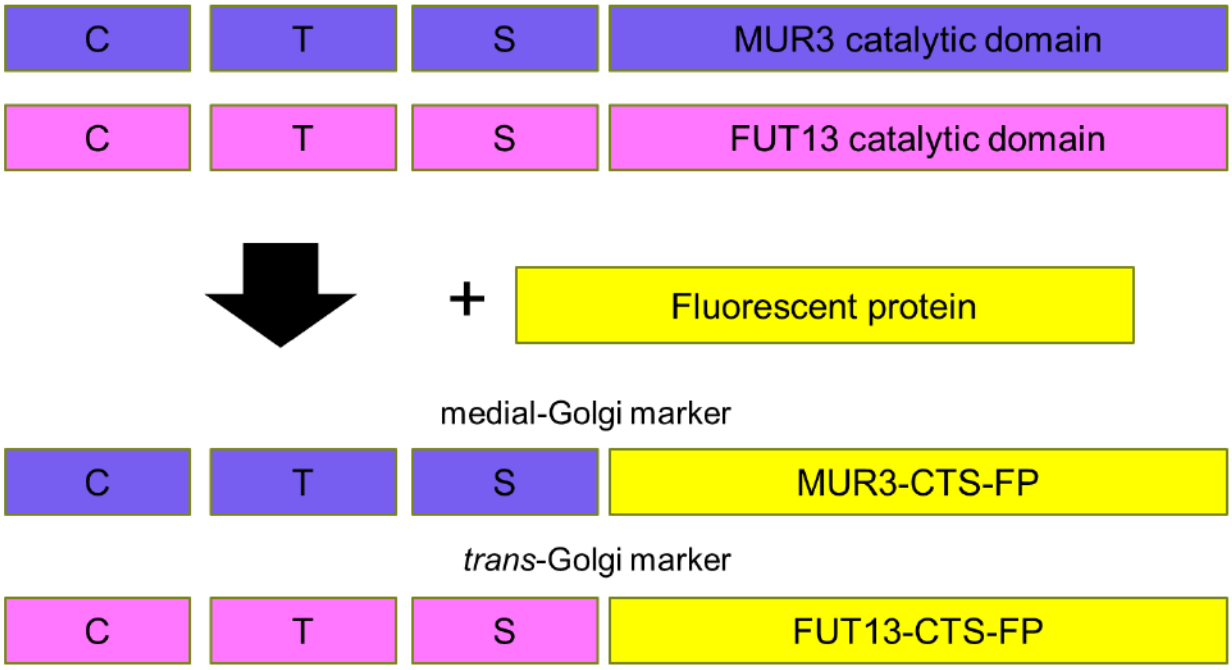


Figure 3.19: Schematic representation for the protein engineering of cisternae-specific markers from CTS fusions. Constructs outlining domain swaps between plant CTS domains and fluorescent proteins from endogenous plant glycosylation enzymes MUR3 and FUT13.

We analysed the CTS domains of MUR3 and FUT13 fused to fluorescent proteins (MUR3-CTS-eGFP and FUT13-CTS-eGFP; Figure 3.20). FUT13-CTS-eGFP co-localised with ST-eGFP (distance between maximum fluorescent intensities of 21.0 ± 14.6 nm) but not with MNS1-eGFP (distance between maximum fluorescent intensities of 160.4 ± 24.8 nm; statistically significant with $p < 0.001$). The distance between maximum fluorescent intensities between MUR3-CTS-eGFP and ST-eGFP was 80.9 ± 19.2 nm. The distance between maximum fluorescent intensities between MUR3-CTS-eGFP and MNS1-eGFP was 145.5 ± 34.9 nm (statistically significant with $p < 0.001$). This data is of interest, as MUR3-CTS is less closely aligned with ST than FUT13-CTS. This indicates a measurable separation between all three Golgi cisternae but also confirms ST is more specifically a *trans*-Golgi marker (Loos et al., 2014). Overall, the data showed that the generated markers appear effective in terms of localisation compared to existing Golgi markers, and that Golgi cisternae localisation is indeed dependent on the CTS domain. This validates the use for MUR3 and FUT13 for cisternae-specific targeting within the plant Golgi.

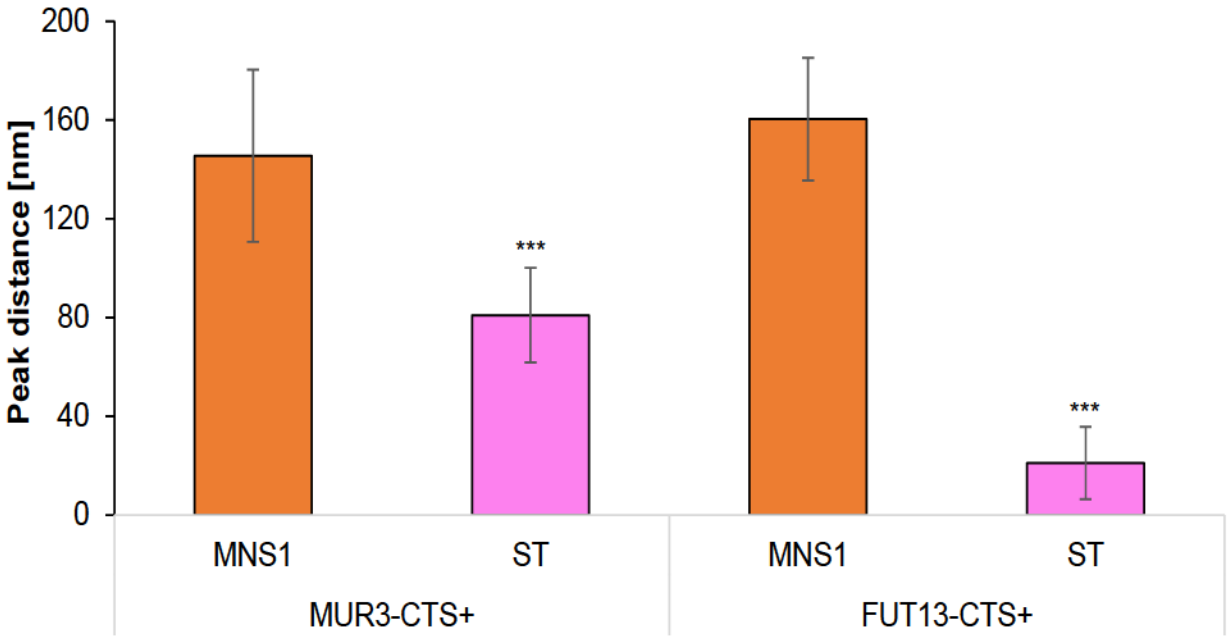


Figure 3.20: Peak distance analysis for fluorescent protein constructs with CTS domain only (MUR3-CTS and FUT13-CTS).

Peak distance analysis graphs showing co-localisation analysis for MUR3-CTS-mRFP and FUT13-CTS-mRFP, respectively, with the markers MNS1-eGFP (*cis*-Golgi cisternae, orange) and ST-eGFP (*trans*-Golgi cisternae, pink). Results shown from n = 3 biological replicas and >28 technical repeats for each combination. Significance was analysed by Kruskal-Wallis (***p< 0.001).

3.3.4 Designing self-cleaving peptides for dual marker constructs

Localisation analysis for this project would be based on the comparison of fluorescently tagged HuGEs with known, cisternae-specific markers for the plant Golgi body. To this end, the selection of fluorophores and their combinations was key to reliable and effective analysis. Confocal microscopy techniques using multiple fluorophores require consideration of the coverage, intensity and stoichiometry of fluorescent protein constructs observed. The method of co-localisation for fluorescent proteins investigated did not require fluorescent intensity between channels to be equivalent, only a suitably clear signal was required. Investigation of localisation was made between structures in close proximity (<2 μm) (Robinson et al., 2020) and so fluorophore combinations were selected for minimum overlap in emission wavelength to prevent signal crossover or bleed through between channels.

Sub-Golgi localisation of HuGEs can be assessed by direct comparison with one marker specific to Golgi cisterna (Figure 3.20). Localisation was initially investigated with two-colour confocal of individual tagged HuGE proteins and either a *cis*- or *trans*-Golgi marker (MNS1 or ST respectively) for co-localisation analysis. An incremental improvement to this approach would be comparison of HuGEs with multiple cisternae-specific markers simultaneously i.e. a *cis*- and *trans*-Golgi marker. This would produce 3-colour Golgi with both endmost cisternae labelled, facilitating comparative analysis of a HuGE with two reference cisternae labelled by Golgi markers i.e. distance from *cis*- and from *trans*-Golgi. However, co-transformation of multiple genes by stacking under multiple promoters can lead to silencing, imbalanced stoichiometry or poor expression of downstream cassettes (Spatola Rossi et al., 2023). Self-cleaving peptides were investigated as an alternative expression strategy to circumvent these issues (Ryan et al., 1994).

Use of self-cleaving peptides was investigated for simultaneous expression of multiple cisternae-specific Golgi marker proteins for distance between peaks analysis (Figure 3.21). Self -cleaving peptides are short amino acid sequences of viral origin that result in protein separation at the penultimate bond position in the sequence during translation of polyproteins (Szymczak & Vignali, 2005). This causes protein translation in one frame, but separation of polyprotein to produce multiple polypeptides in equal stoichiometry, with varying efficiencies for different 2A sequences employed in building larger multiple subunit expression cassettes (Wang et al., 2015; Zhang et al.2017). This enables separate expression of multiple genes under the same promoter off the same plasmid in one open reading frame, while maintaining expression levels.

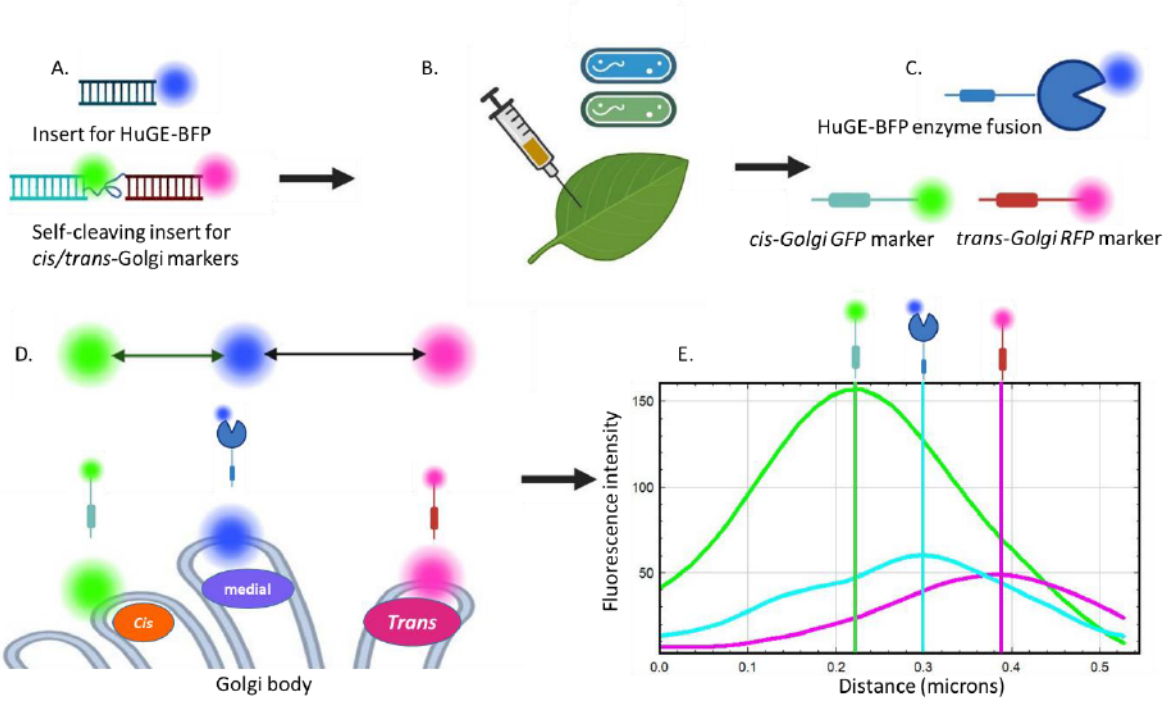


Figure 3.21: Overview of distance between peaks method for fluorescent protein sub-Golgi localisation using self-cleaving markers

A) Single human glycosylation enzyme (HuGE) protein and self-cleaving marker inserts for plant expression to generate B) Fluorescently tagged HuGE construct (blue) and fluorescently tagged individual Golgi markers (green and red) expressed in fixed stoichiometry by self-cleaving peptide. C) Paired agrobacterium mediated expression of constructs in mature plant tissue. D) *Cis*- or *trans*-Golgi lumenal expression of cisternae specfic markers (green and red) used to determine distance from HuGE localisation (blue). E) Line profile data across Golgi body generated and exported to use for analysis.

Self-cleaving peptide constructs were assembled (Figure 3.22) containing two cisternae-specific Golgi markers flanking a porcine teschovirus-1 viral 2A self-cleaving sequence (see Appendix 2 for sequences) as this had among the highest cleavage efficiencies (Wang et al., 2015). Successful expression of MNS1-ST (Figure 3.22; A) would provide *cis*- and *trans*-Golgi localised marker expression from a single insert, providing comparison between either end of the Golgi. Expression of MUR3-FUT13 (Figure 3.22; B) would provide medial- and *trans*-Golgi localised marker expression from a single insert. The MUR3-FUT13 insert was designed with an additional Intein self-excising domain, to remove additional C-terminal amino acids left on upstream protein post 2A peptide mediated cleavage (Zhang et al.2017).

All 2A peptides contain a C-terminal sequence and a similar mode of action. Ribosome skipping facilitates a break in translation around the ~17aa position, then continues translation of downstream sequence (Donnelly et al., 2001). Leaving ~17 amino acid residues added to C-terminus of protein upstream of the 2A subunit, and a single proline added to N-terminus of downstream protein (Liu et al., 2017). Protein expresses in 1:1 stoichiometry makes this approach useful in comparison of several marker proteins, as this provides equivalent expression levels while conserving targeting mechanisms (Samalova et al., 2006). Inteins are protein introns capable of autocatalytic splicing from host polypeptides to generate separate proteins (Lennon et al., 2017; Derbyshire et al., 1997). A functional synergy here is inclusion of an intein cassette upstream of a 2A sequence (Figure 3.22; B). This creates an expression cassette that cleaves into multiple protein by activity of the 2A domain, but the intein self-excises to remove the additional amino acids left behind by the P2A domain on the protein upstream of the 2A sequence.

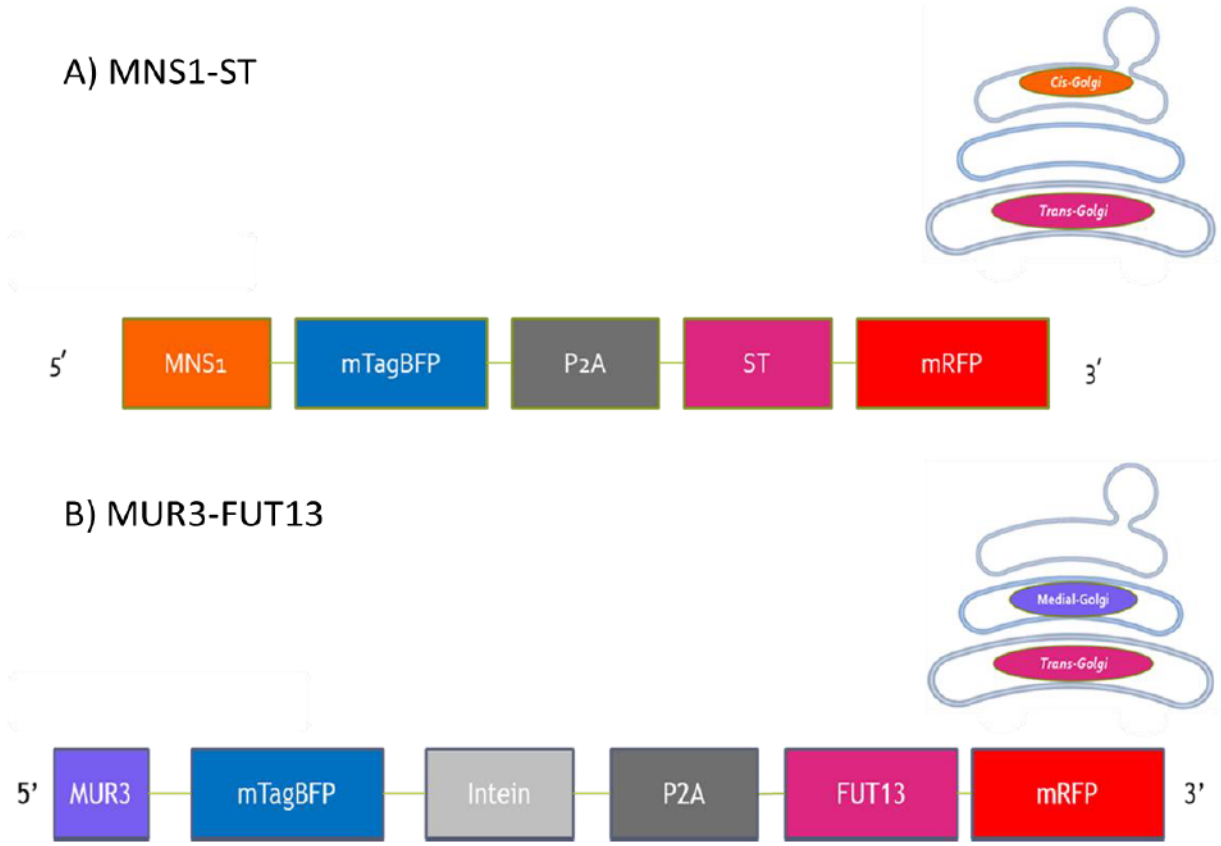


Figure 3.22: Example construction of self-cleaving inserts for dual marker constructs.
A) MNS1-ST construct design to express a MNS1-mTagBFP *cis*-Golgi marker in 1:1 stoichiometry with a ST-mRFP *trans*-Golgi marker via a linking P2A self-cleaving sequence. B) MUR3-FUT13 construct design to express a MUR3-mTagBFP medial-Golgi marker in 1:1 stoichiometry with a FUT13-mRFP *trans*-Golgi marker via a linking Intein self-excising sequence and P2A self-cleaving sequence. Inserts colour coded based on targeting towards *cis*- (Orange), medial- (Purple) and *trans*- (magenta) Golgi cisternae.

Single insert expression of the self-cleaving marker MNS1-ST (Figure 3.23) resulted in two-colour confocal displaying mRFP and mTagBFP that was Golgi-body associated. This was not directly co-localised, consistent with localisation for known markers MNS1 and ST. This suggests successful expression of the insert and successful cleavage into two distinct fluorescent protein markers with separate CTS-directed localisation.

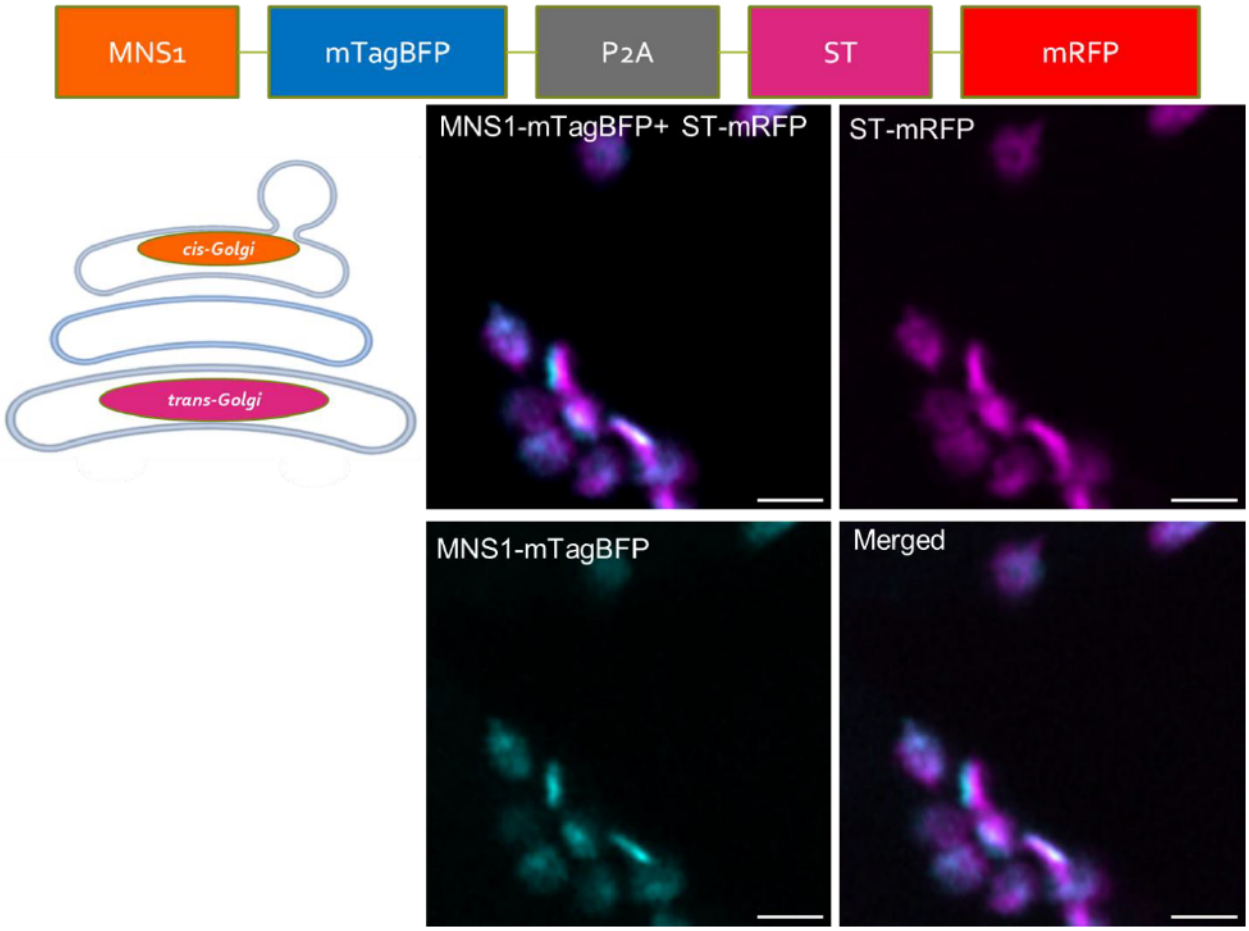


Figure 3.23: Insert design and confocal microscopy of *cis-/trans*-Golgi marker MNS1-ST expression. Single insert MNS1-ST construct expressing *cis*-Golgi marker MNS1-mTagBFP (blue) alongside *trans*-Golgi marker ST-mRFP via a linking P2A self-cleaving sequence. Representative images shown. Size bars = 1 µm.

Similarly, single insert expression of the self-cleaving marker MUR3-FUT13 (Figure 3.24) resulted in two-colour confocal displaying mRFP and mTagBFP that was Golgi-body associated, but not directly co-localised. This suggests successful expression of the insert and cleavage into two distinct fluorescent protein markers with separate CTS-directed localisation, and validates use of the MUR3 and FUT13 CTS domains in medial/*trans*-Golgi cisternae targeting markers.

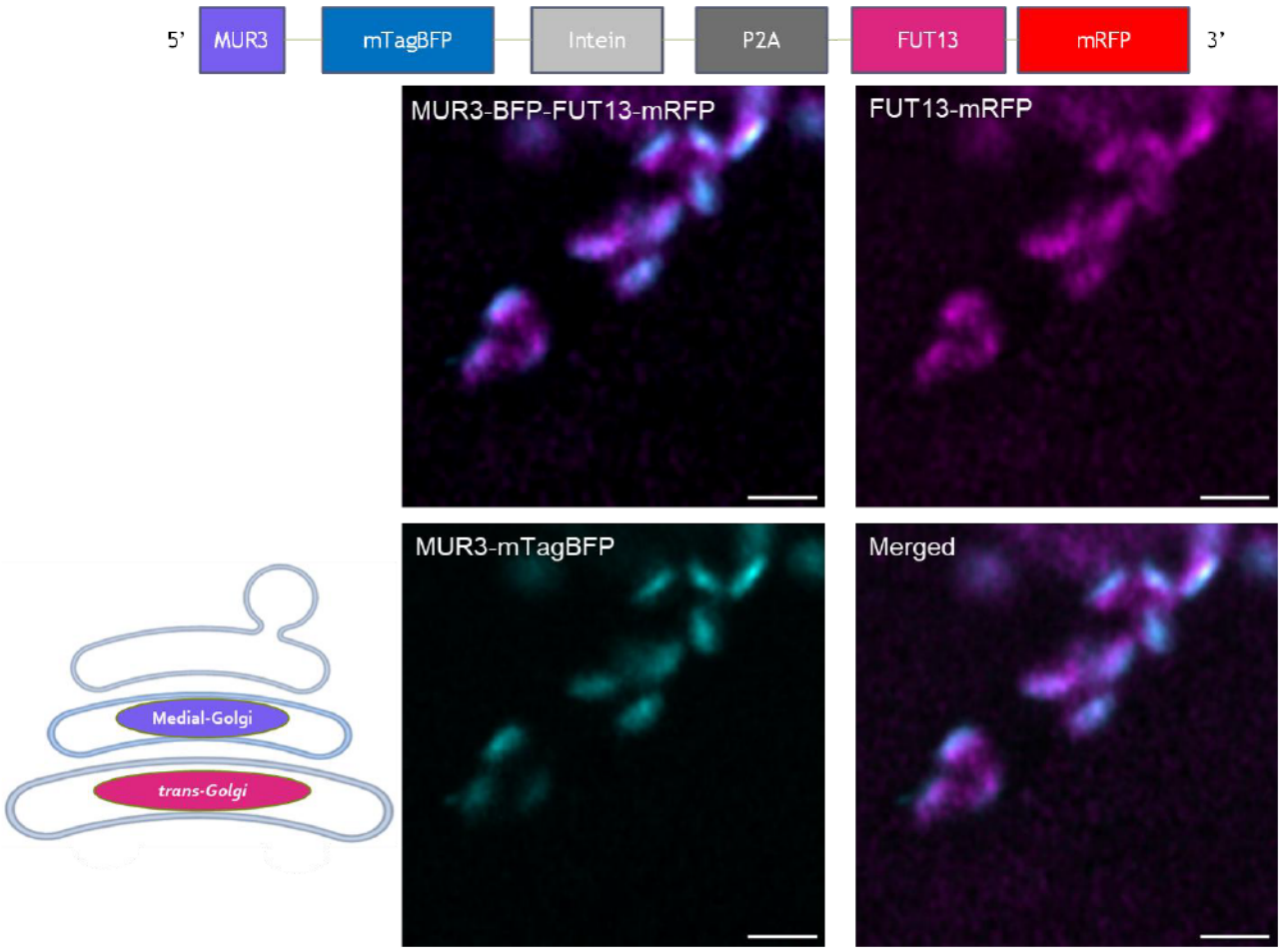


Figure 3.24: Insert design and Confocal microscopy of medial/*trans*-Golgi marker MUR3-FUT13 expression.

Single insert MUR3-FUT13 construct expressing medial-Golgi marker MNS1-mTagBFP (blue) alongside *trans*-Golgi marker FUT13-mRFP via a linking P2A self-cleaving sequence. Representative images are shown. Size bars = 1 µm.

Co-expression of two separate agrobactrium mediated plasmids for existing marker MNS1-eGFP alongside self-cleaving marker MNS1-ST (Figure 3.25) resulted in three-colour confocal from two inserts. This suggests successful expression of the insert and cleavage into two distinct fluorescent protein markers with partially overlapping CTS-directed localisation, and validates single insert stoichiometric expression of existing markers MNS1 and ST via self-cleaving peptide.

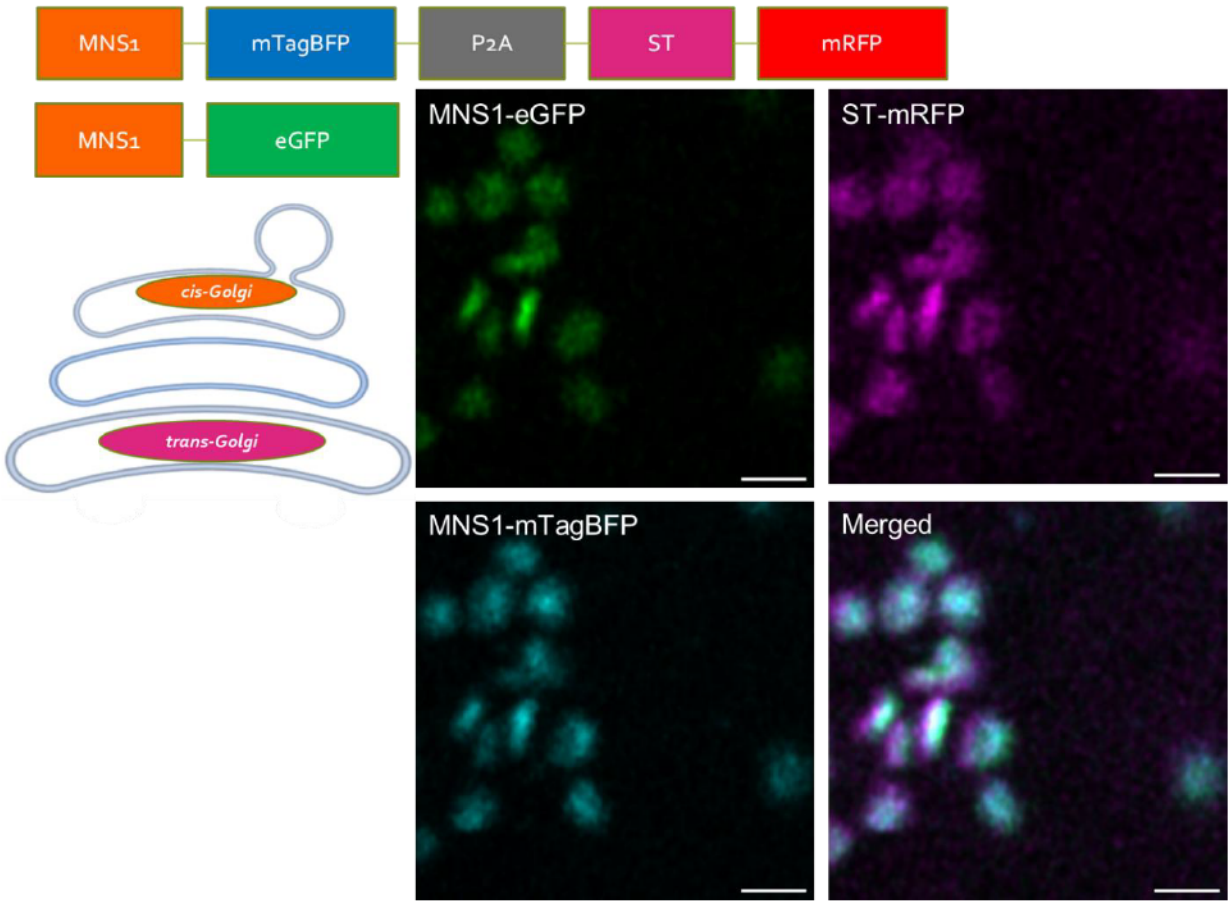


Figure 3.25: Insert design and confocal microscopy of *cis-/trans*-Golgi marker MNS1-ST in combination with *cis*-Golgi maker MNS1-eGFP expression.
Single insert MNS1-ST construct expressing *cis*-Golgi marker MNS1-mTagBFP (blue) alongside *trans*-Golgi marker ST-mRFP via a linking P2A self-cleaving sequence. Combined with separate MNS1-eGFP marker for three-colour Golgi body localisation with *cis*-Golgi co-localisation. Representative images are shown. Size bars = 1 µm.

Similarly, co-expression of two separate agrobactrium mediated plasmids for existing marker MNS1-eGFP alongside self-cleaving marker MUR3-FUT13 (Figure 3.26) resulted in three-colour confocal from two inserts. This suggests successful expression of the insert and cleavage into two distinct fluorescent protein markers with separate CTS-directed localisation, and validates use of the MUR3 and FUT13 CTS domains for use in medial/*trans*-Golgi cisternae targeting markers.

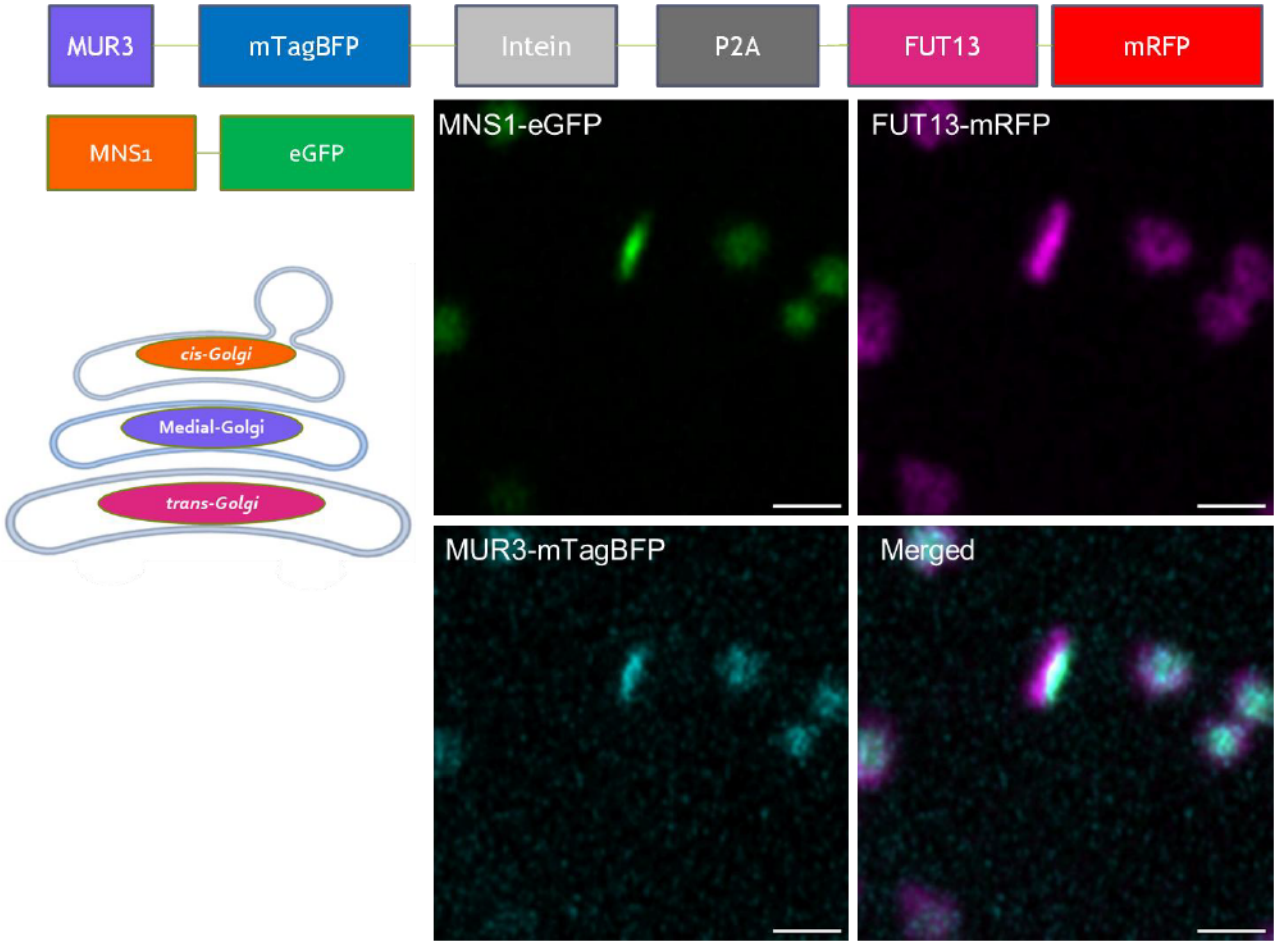


Figure 3.26: Insert design and confocal microscopy of medial-/*trans*-Golgi marker MUR3-FUT13 in combination with *cis*-Golgi maker MNS1-eGFP expression.
Single insert MUR3-FUT13 construct expressing medial-Golgi marker MUR3-mTagBFP (blue) alongside *trans*-Golgi marker FUT13-mRFP via a linking P2A self-cleaving sequence. Combined with separate MNS1-eGFP marker for three-colour Golgi body localisation with cisternae-specific localisation. Representative images shown. Size bars = 1 µm.

Development of self-cleaving constructs to express multiple cisternae-specific Golgi body markers from single inserts proved successful. A *cis*-/*trans*-Golgi MNS1-ST insert successfully produced two-colour confocal with separate expression of both markers from a single insertion. The same effect was shown with a medial-/trans-Golgi MUR3- FUT13 insert, developed utilizing novel plant CTS domains as fluorescent protein fusions for new markers. Both self-cleaving markers were successfully expressed alongside known cisternae-specific Golgi markers to produce 3-colour expression in the plant Golgi from two inserts, with individual cisternae-specific expression.

3.4 Discussion

3.4.1 Co-localisation in the plant Golgi

Current methodology for analysis of sub-organellar co-localisation takes advantage of a variety of techniques. For the specific application in this project a distance between the peaks methodology was considered most suitable. But similar research working within Golgi cisternae and the immediate vicinity have effectively analysed co-localisation using Pearson's correlation coefficient (Schoberer et al., 2019). In assessing similar sub-Golgi co-localisation Pearson's has been used to assess an Arabidopsis *cis*-Golgi protein (Schoberer et al., 2019) localised alongside the same *trans*-Golgi marker ST used in this research. Similar work made additional use of line profiles to visualise fluorescence overlap between channels in addition to analysis using Pearson's coefficient (Schoberer et al., 2014). Advanced super-resolution confocal live imaging microscopy (SCLIM) techniques can be combined with existing Pearson's-based co-localisation of fluorescent intensity. Research into Golgi resident enzyme 3D distribution utilized thin section Z-stacks (0.2-µm intervals) covering the full thickness of the Golgi and converted to 3D voxel-based fluorescence intensity maps. These were subjected to deconvolution and use of a point-spread function to elucidate and present 3D-localisation of Golgi resident glycosyltransferases with Pearson's based co-localisation analysis and reference enzyme marker constructs (Yagi et al., 2024).

The new markers made for medial-Golgi MUR3-CTS and *trans*-Golgi FUT13-CTS were promising in their localisation analysis (Figure 3.20). Showing a clear difference in localisation between the two when compared alongside the *trans*-Golgi marker ST, but failing to show a significant difference in comparison to MNS1. FUT13-CTS was only slightly further away from MNS1 than MUR3-CTS, and the difference in distance is not reflected in the results for ST. While there was more significant deviation in the values for markers with MNS1 (Figure 3.20) than those with ST, this still strongly overlapped and does not readily explain the lack of distinction. As type–II membrane proteins, Golgi resident enzymes are asymmetrically distributed, overlapping but non-uniform (Saint-Jore et al., 2006, Nilsson et al., 2009), t's possible the localisation of MNS1 is more broadly distributed within the Golgi as a cis/medial marker (Schoberer

et al., 2014) and less abundant that ST, leading to larger variability in where the maximum signal intensity occurs across the Golgi, and leading to this maximum fluorescence based method being less appropriate (Figure 3.21). MNS1 has significant roles intracellular trafficking and fundamental early pathway glycan modification and trimming (Liu et al., 2018) as the first commited step to Golgi N-glycan processing potentially relating to abundance and distribution. It may be possible that there is variability of MNS1 labelling as a *cis/*medial-Golgi marker that limits its effectiveness in this application, that is not present with the ST marker (Schoberer et al., 2011)

3.4.2 Applications of self-cleaving peptides

Co-expression under a single promoter by self-cleaving peptides, while achieving efficient expression, offers an incredible tool for a diverse range of research where coordinated co-expression of multiple factors is required (Liu et al., 2017). 2A peptides offer an unprecedented method for expression of bi-, tri- and quad-cistronic constructs, with optimized co-expression strategies based on modifying gene arrangements and improved Kozak regions (Wang et al., 2015). Early work into ratiometric fluorescence-imaging analysis by Samalova et al (2006) was fundamental in establishing the effective expression of multiple markers under single promoters via expression cassettes assembled with 2A self-cleaving peptides. Further, this work confirmed the successful localisation of multiple marker proteins with separate methods of directed localisation, each of which remained functional, and took advantage of the fixed stoichiometry 2A-mediated expression provided for ratiometric analysis (Samalova et al., 2006). This work provided research rationale for self-cleaving peptide constructs for the self-cleaving Golgi markers developed, suggesting CTS-directed markers such as MNS1 and ST being readily expressed with cisternae-specific localisation remaining functional.

Fixed stoichiometric expression of multiple transgenes is an attractive prospect for studying markers at equivalent intensities but lends itself to a range of applications. The ability to assemble multiple complex proteins at a fixed ratio (Zhang et al., 2017) makes this technology attractive for the assembly of difficult, multiple subunit protein

complexes or multiple enzyme pathways that benefit from regulated expression of separate protein components. Bioengineering of a particulate methane monooxygenase (pMMO) enzyme complex in rice as a potential solution to environmental methane emissions is one example that benefits from coordinated, stoichiometric expression of multiple subunits from self-cleaving peptides (Spatola Rossi et al., 2023). Creative use of the self-cleaving includes methods of regulating expression beyond multiple insert expression. Investigation of an adeno-associated virus mediated delivery for CRISPR-Cas9 gene editing has been modified to include a self-limiting feedback loop, utilizing self-cleaving peptides to improve the safety profile of in targeted gene editing therapeutics for human administration (Li et al., 2021).

In summary, this research undertaken was successful in investigating and developing an effective co-localisation methodology for sub-Golgi localisation, to be used in determining HuGE sub-Golgi localisation. This research established a data collection and analysis pipeline for super resolution confocal data of Golgi. With it sub-Golgi localisation of both well research and novel cisternae-specific markers were quantified for future comparison with HuGEs. Findings were consistent with existing literature on sub-Golgi localisation of markers used, supporting the use of this methodology (Boevink et al., 1998, Schoberer et al., 2011., Schoberer et al., 2018a; Schoberer et al., 2018b., Strasser et al., 2009). Additionally, an expression strategy to reduce insert number and provide fixed stoichiometric expression of Golgi markers was developed and tested through use of self-cleaving peptide constructs. This work represents the foundation of subsequent work into sub-Golgi expression and localisation of HuGEs.

Chapter 4: CTS design and HuGE localisation

Contents

4.1 Glycosylation enzyme distribution in plant Golgi stacks

N-linked glycosylation in plants occurs within the Golgi bodies, a significant site of protein, lipid and polysaccharide synthesis as well as glycomodification within the plant cell, and a focal point for production of recombinant proteins (Schoberer et al., 2018b; Gomord et al., 2020). In contrast to the immobile, perinuclear Golgi body found in mammalian cells, Golgi bodies in higher plants are discrete stacks that number several hundred per cell, and are extremely mobile throughout the cytoplasm (Robinson et al., 2020). The basic structure of the individual Golgi body (Figure 4.1) 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 mediating intracellular delivery of ER products (Robinson et al., 2020).

Plant and human Golgi-resident N-glycosylation enzymes are type II membrane proteins, possessing at the N-terminal region a short cytoplasmic domain, a transmembrane domain, and a stem region (collectively the CTS region). The CTS region is responsible both for orientation of the enzyme's catalytic domain in the Golgi lumen, and providing sub-Golgi localisation, an explanation to the non-uniform distribution of glycosylation enzymes in the Golgi (Schoberer et al., 2011). Existing markers (MNS1 for *cis*-Golgi cisternae (Liebminger et al., 2009; Schoberer et al., 2014) and ST for *trans*-Golgi cisternae (Runions et al., 2006) consist of fluorophore fusions to respective CTS domains, providing cisternae specific localisation as a comparison with the plant specific CTS domains MUR3, FUT13 (Schoberer et al., 2019) (Figure 4.1). The use of endogenous plant CTS domains can alter the localisation of human glycosylation enzymes (HuGEs), through fusion of a HuGE catalytic region with a plant CTS directed to the relevant sub-Golgi compartment (Strasser et al., 2014).

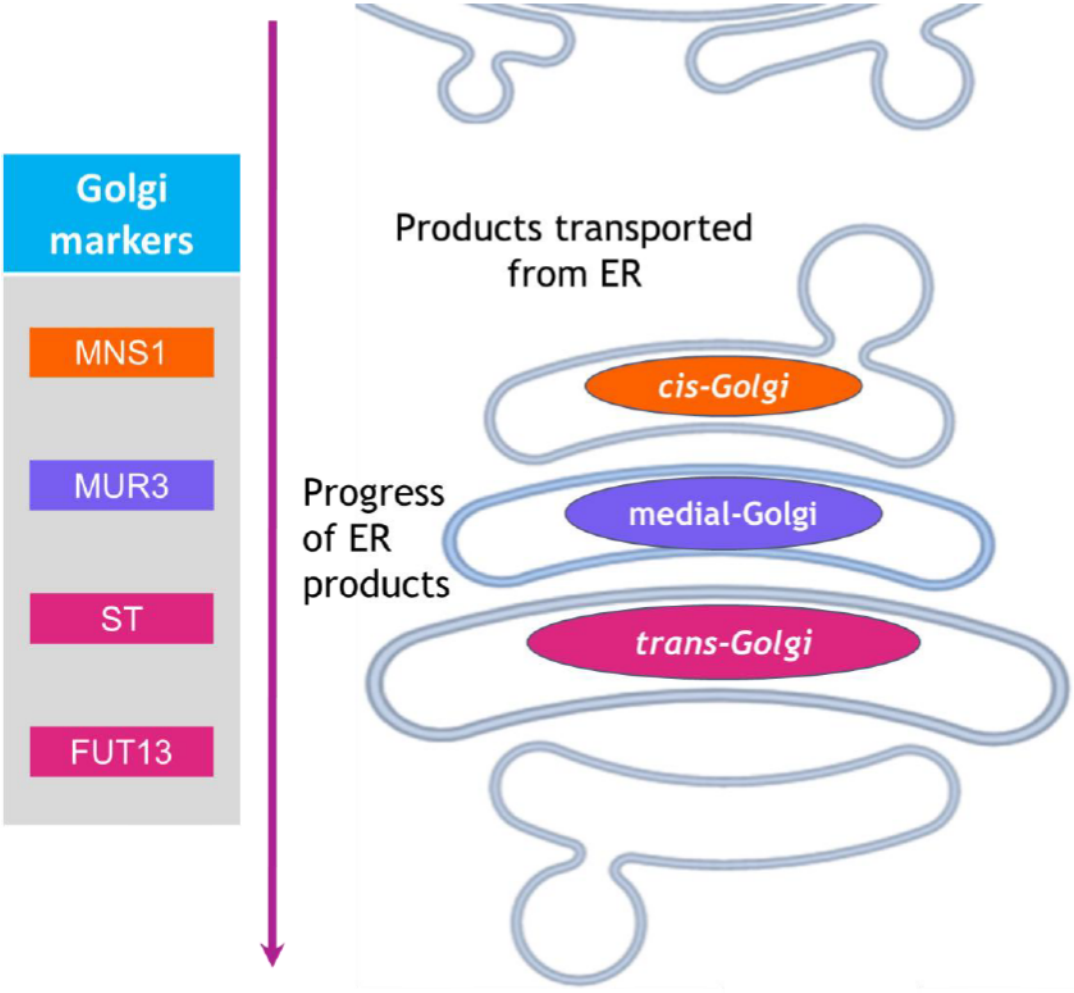


Figure 4.1: Schematic representation of sub-organellar organisation of the plant Golgi body cisternae, with relevant cisternae-specific markers colour-coded.
Orange = *cis*-Golgi targeting (MNS1, CTS of Golgi-α-mannosidase I), Purple = medial-Golgi targeting (MUR3, CTS of β-1,2-galactosyltransferase), Magenta = *trans*-Golgi targeting (ST, CTS of *Rattus norvegicus* α-2,6-sialyl transferase; FUT13, CTS of α-1,4-fucosyltransferase).

Here we describe the transient expression and localisation of human glycosylation enzymes (HuGEs) in tobacco plants as a model system, with localisation directed to medial- and *trans*-Golgi body cisternae, respectively. Sub-organellar localisation to Golgi body cisternae is analysed using high-resolution live-cell and dynamic confocal microscopy.

4.2 Aims:

1. Use the imaging approachs and data analysis developed (Chapter 3) to investigate and quantify the localisation of unmodified HuGEs in the plant Golgi
2. Develop and implement approaches for cisternae-specific localisation of HuGEs to optimise enzyme activity and resulting glycan profile

4.3 Results and Discussion

4.3.1 Investigating HuGE expression in the plant Golgi body

N-linked glycosylation requires stepwise addition of sugar moieties at each sequential Golgi body compartment. Therefore, individual glycosylation enzymes need to be located in specific Golgi cisternae to access substrate and achieve efficient glycosylation in order to produce a suitable yield of glycans (Castilho et al., 2011). In order to establish effective glycan processing, it was necessary to determine the sub-Golgi localisation of HuGEs, and where necessary to target the enzymes to the correct Golgi cisternae.

To investigate localisation of unmodified HuGEs (Appendix 4 for sequences), constructs were generated featuring the mammalian signal sequences, *in planta* constructs generated for HuGEs, fused to a C-terminal CLOVER (CLVR) fluorescent tag and expressed in *Nicotiana tabacum* (tobacco) leaf epidermal cells using Agrobacterium-mediated plant transformation HuGE localisation to Golgi cisternae was determined using statistical line profile analysis of fluorescent intensity for localisation analysis of closely associated proteins (McGinness et al., 2022).

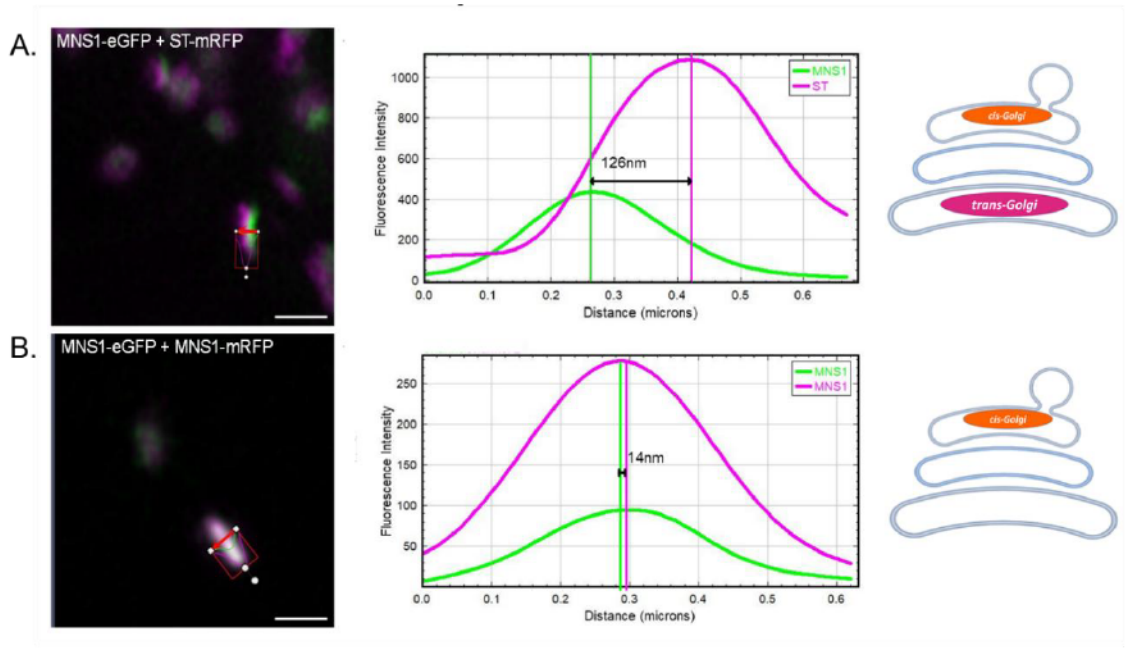


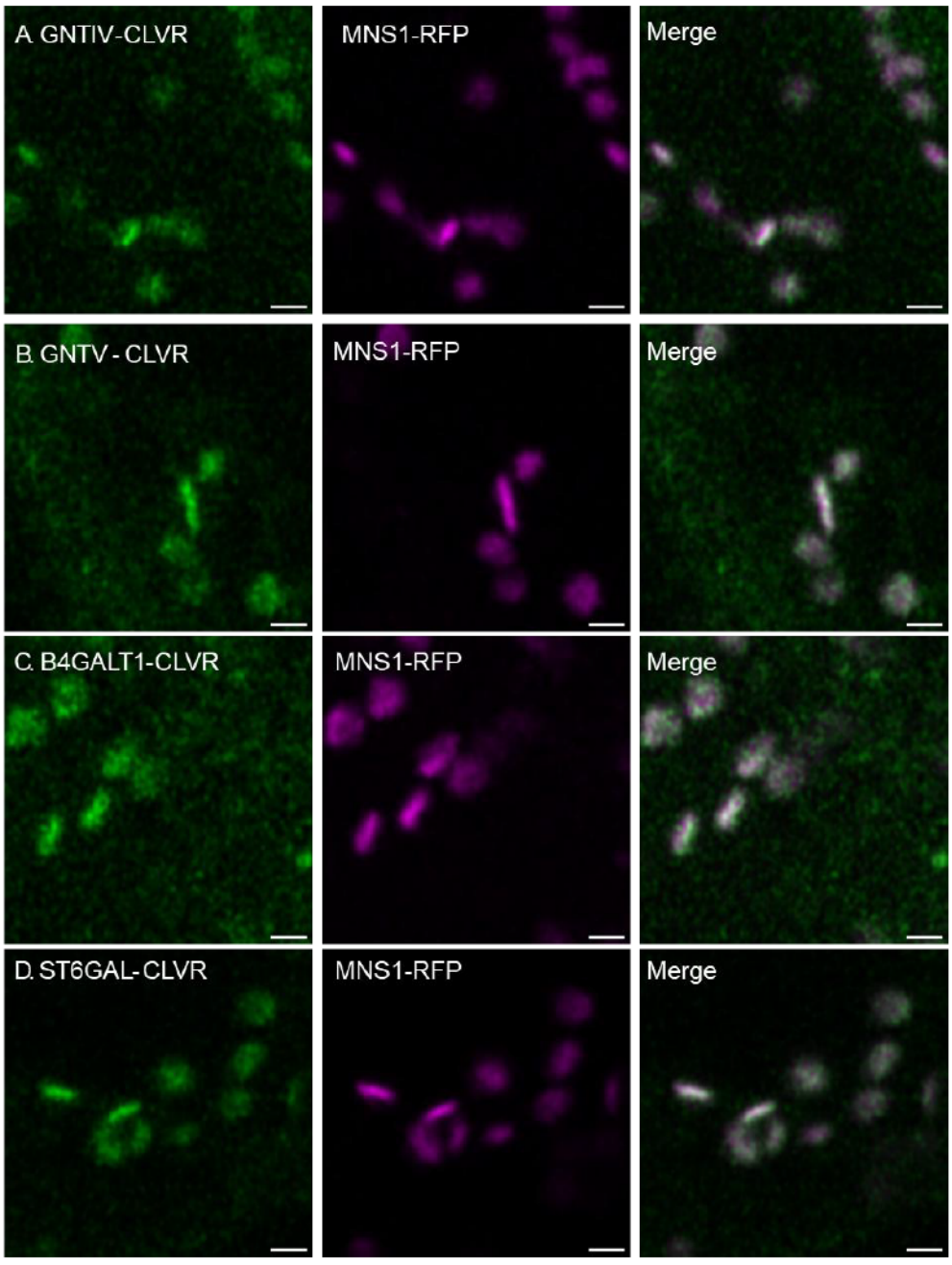
Figure 4.2: Co-localisation of known plant Golgi cisternae markers.

Line profiles for fluorescent intensity analysis across a Golgi body, with known markers A) MNS1-eGFP+ST-mRFP localised to separate Golgi body cisternae to demonstrate the distance between peaks, or B) MNS1-eGFP+MNS1-mRFP double markers for the same Golgi body cisterna in two colours to demonstrate overlap of expression within one Golgi body cisterna. Example confocal images and graphs are representative of for the dataset. Size bars = 1 µm.

For this approach, Golgi cisternae were imaged in side profile showing the distribution of HuGEs across the Golgi body cisternae (Figure 4.2). Analysis of *cis*- and *trans*-Golgi markers (MNS1-eGFP+ST-mRFP) provided a relative baseline for the maximum distance between Golgi cisternae (Figure 4.2A), when compared to markers which localise to the same Golgi body cisterna (MNS1-eGFP with MNS1-mRFP; Figure 4.2B). Fluorescent microbeads were used at the beginning of each collection session to error check and baseline the system, and serve to give minimum possible seperation between the two channels on one object (Figure 3.18)

4.3.2 Analysis of subcellular localisation of unmodified mammalian glycosylation enzymes in tobacco leaf cell

To evaluate initial expression of the unmodified enzymes, unmodified HuGEs were co-expressed with the *cis*-Golgi marker MNS1-mRFP (Schoberer et al., 2013) (Figure 4.3, A-D). Observed were low levels of expression and protein localisation did not appear specific to the plant Golgi bodies.



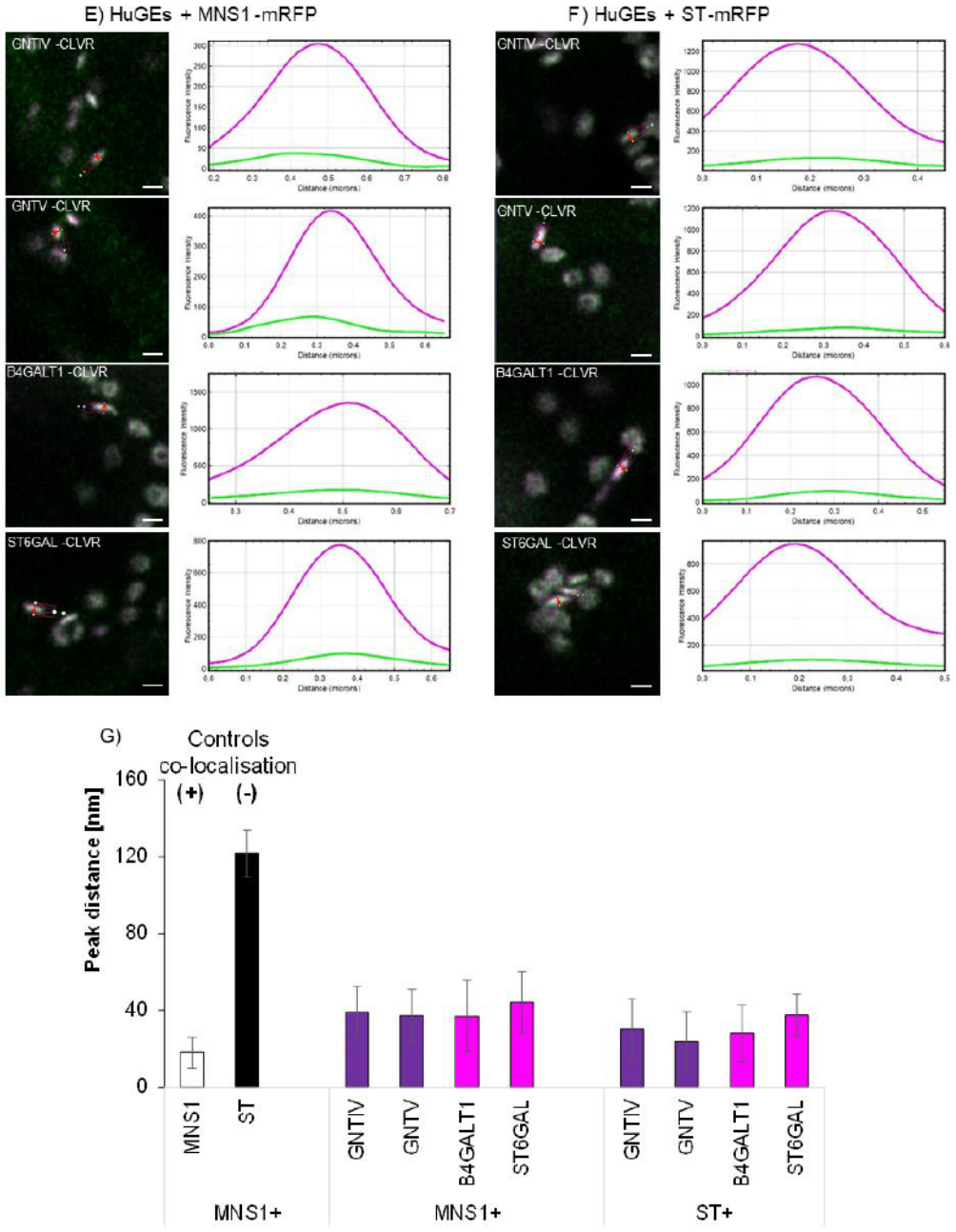


Figure 4.3: Example confocal images and co-localisation line profile analysis of unmodified HuGEs.

A) GNTIV-CLVR, B) B4GALT1-CLVR, C) GNTV-CLVR, D) ST6GAL-CLVR (green) at 3 days post infiltration alongside the *cis*-Golgi marker MNS1-mRFP (magenta). Representative images are shown. Size bars = 1 µm.

E) Example images and analysis for statistical line profile analysis of distances between peak fluorescent intensity for GNTIV-CLVR, GNTV-CLVR, B4GALT1-CLVR, ST6GAL-CLVR (green), with *cis*-Golgi marker MNS1 or

F) *trans*-Golgi body marker ST-mRFP, respectively (magenta).

G) Peak distance analysis graphs showing controls for co-localisation in the same Golgi cisternae (MNS1-eGFP+MNS1-mRFP, white bar) and for localisation in different Golgi cisternae (MNS1-eGFP+ST-mRFP, black bar) as well as comparison of the four HuGEs (medial HuGEs in purple and *trans*-Golgi HuGEs in magenta) with the respective markers MNS1-mRFP and ST-mRFP. Representative images are shown. Results shown from n = 3 biological replicas and >12 technical repeats for each combination. Size bars = 1 µm.

To further quantify localisation, HuGEs were co-expressed with both the *cis*-Golgi marker MNS1 and the *trans*-Golgi marker ST separately (Figure 4.3E) to determine their localisation within the Golgi stack. Markers for specific Golgi cisternae were used as control and a statistical analysis was carried out (Figure 4.3F). Co-expression of MNS1-eGFP with MNS1-mRFP represented approximate co-localisation in the same Golgi body cisternae and showed a distance between the maximum peak intensities of 18.1 ± 8.2 nm. Co-expression of the *cis*-Golgi marker MNS1-eGFP and the *trans*-Golgi marker ST-mRFP (Renna et al., 2005; Munro et al., 1995) showed a distance between maximum peak intensities of 121.6 ± 12.4 nm, which was considered as localisation to different Golgi body cisternae.

Line profiles generated across plant Golgi bodies for all native HuGEs (Figure 4.3) indicated a low level of fluorescence compared to the co-expressed marker control MNS1-mRFP. Medial-Golgi HuGEs GNTIV-CLVR and GNTV-CLVR (Figure 4.3) demonstrated similar distances between maximum peaks of intensity when expressed alongside MNS1-mRFP and ST-mRFP (ranging from 23 to 39 nm peak distance). This indicates localisation distributed across Golgi cisternae without the required specific localisation to either the medial or *trans*-Golgi. *Trans*-Golgi HuGEs also showed unspecific Golgi cisternae localisation; B4GALT1-CLVR and ST6GAL-CLVR (Figure 4.3) demonstrated similar distance between peaks with MNS1-mRFP and ST-mRFP (range 29-41 nm).

The lack of *trans*-Golgi specific localization for ST6GAL was especially surprising as the CTS regions of human ST6GAL and the marker (*Rattus norvegicus*) ST are highly conserved (Schoberer et al., 2011; Wee et al., 1998). Potential explanations could be differences due to the low expression levels, an impact of the ST6GAL catalytic domain on the localisation, or the presence of some specific amino acid changes in the CTS that could affect the localisation.

Without Golgi-specific localisation, HuGEs are unlikely to correctly glycosylate proteins. With all four HuGEs showing no specific localisation to medial or *trans*-Golgi cisternae as required for correct N-glycosylation to occur, endogenous plant signals to improve both expression levels and sub-organellar localisation were investigated.

4.3.3 Modifying HuGEs sub-Golgi localization using plant CTS domains

As type II membrane proteins, plant and human glycosylation enzymes both possess a single pass signal-anchor sequence provided by a short N-terminal CTS domain of about 40 to 70 amino acids, which provides specific sub-Golgi localisation and targets the enzyme's catalytic subunit to the Golgi lumen (Schoberer et al., 2013).

With the aim to achieve Golgi-cisternae specific targeting and to use plant-derived sequences, HuGEs were targeted *in planta* using two CTS domains from *Arabidopsis thaliana* resident glycosylation enzymes: a MUR3-CTS from β-1,2-galactosyltransferase as the medial-Golgi targeting domain and a FUT13-CTS from α-1,4-fucosyltransferase for *trans*-Golgi targeting (Figure 4.5A) (Schoberer et al., 2013). The mammalian CTS domains were located and replaced with the Arabidopsis CTS domains to create the following fusion proteins; MUR3-GNTIV, MUR3-GNTV, FUT13-B4GALT1 and FUT13-ST6GAL (Figure 4.5B).

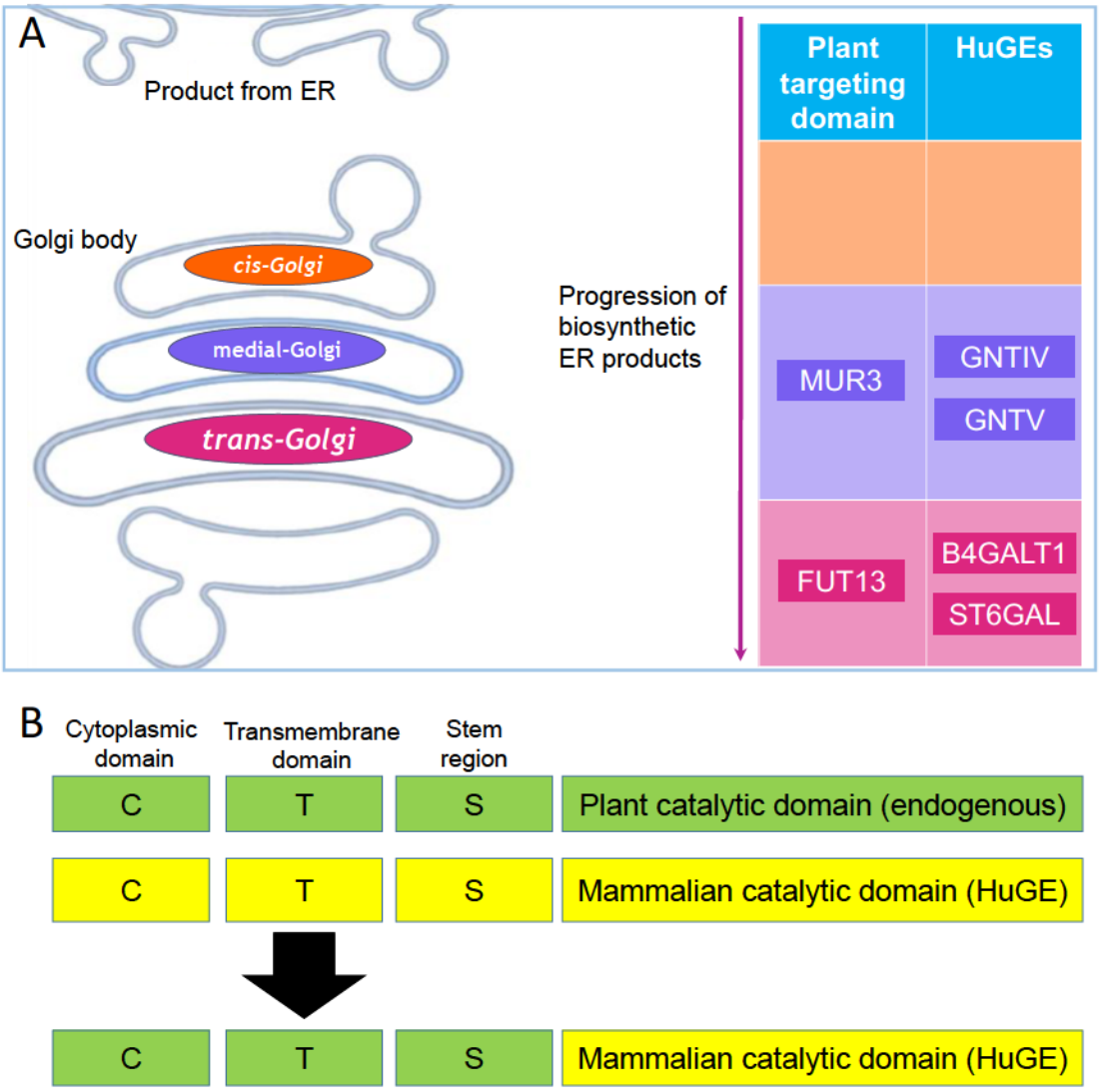


Figure 4.4: Schematic diagram of HuGE targeting approach.

A) HuGE targeting approach in relation to Golgi body structure. Cisternae localisation is colour-coded (*cis* = orange, medial = purple, *trans* = magenta). The required localisation of the HuGEs and the plant CTS used for targeting are indicated: MUR3-CTS to target GNTIV and GNTV to medial-Golgi cisternae; FUT13-CTS to target B4GALT1 and ST6GAL to *trans*-Golgi cisternae.

B) Protein engineering approach for modified CTS-HuGE constructs outlining domain swaps between mammalian and plant CTS domains.

The modified HuGE constructs were co-expressed with both the *cis*-Golgi marker MNS1 and the *trans*-Golgi marker ST in order to assess their sub-organellar localisation (Figure 4.5; Figure 4.7). Overall localisation was altered for HuGEs possessing plant CTS domains (Figure 4.5B). MUR3-GNTIV expression was further away from MNS1 (maximum peak intensity distance 93.0 ± 21.9 nm) but closer to *trans*-Golgi marker ST-mRFP (50.1 ± 14.4 nm). MUR3-GNTV was also further from MNS1 (95.3 ± 22.7 nm) but close proximity with the *trans*-Golgi cisternae (ST-mRFP; 55.9 ± 15.7 nm). This indicated that the MUR3-HuGEs were targeted to the medial Golgi cisternae.

Trans-Golgi HuGEs FUT13-B4GALT1 and FUT13-ST6GAL also showed targeting to the *trans*-Golgi cisternae. FUT13-B4GALT1 was a greater distance from MNS1 (93.2 ± 21.0 nm) but overlapped strongly with ST-mRFP (21.1 ± 15.8 nm). FUT13-ST6GAL was also further from MNS1 (100.7 ± 27.4 nm) but closer to ST-mRFP (24.1 ± 13.0 nm). Overall MUR3-HuGEs and FUT13-HuGEs demonstrated more distance from the *cis*-Golgi marker MNS1-mRFP (Figure 4.5B) than the native HuGEs (Figure 4.3). Interestingly, the FUT13-HuGEs demonstrated a shorter distance from ST-mRFP (Figure 4.5F) than the MUR3-HuGEs.

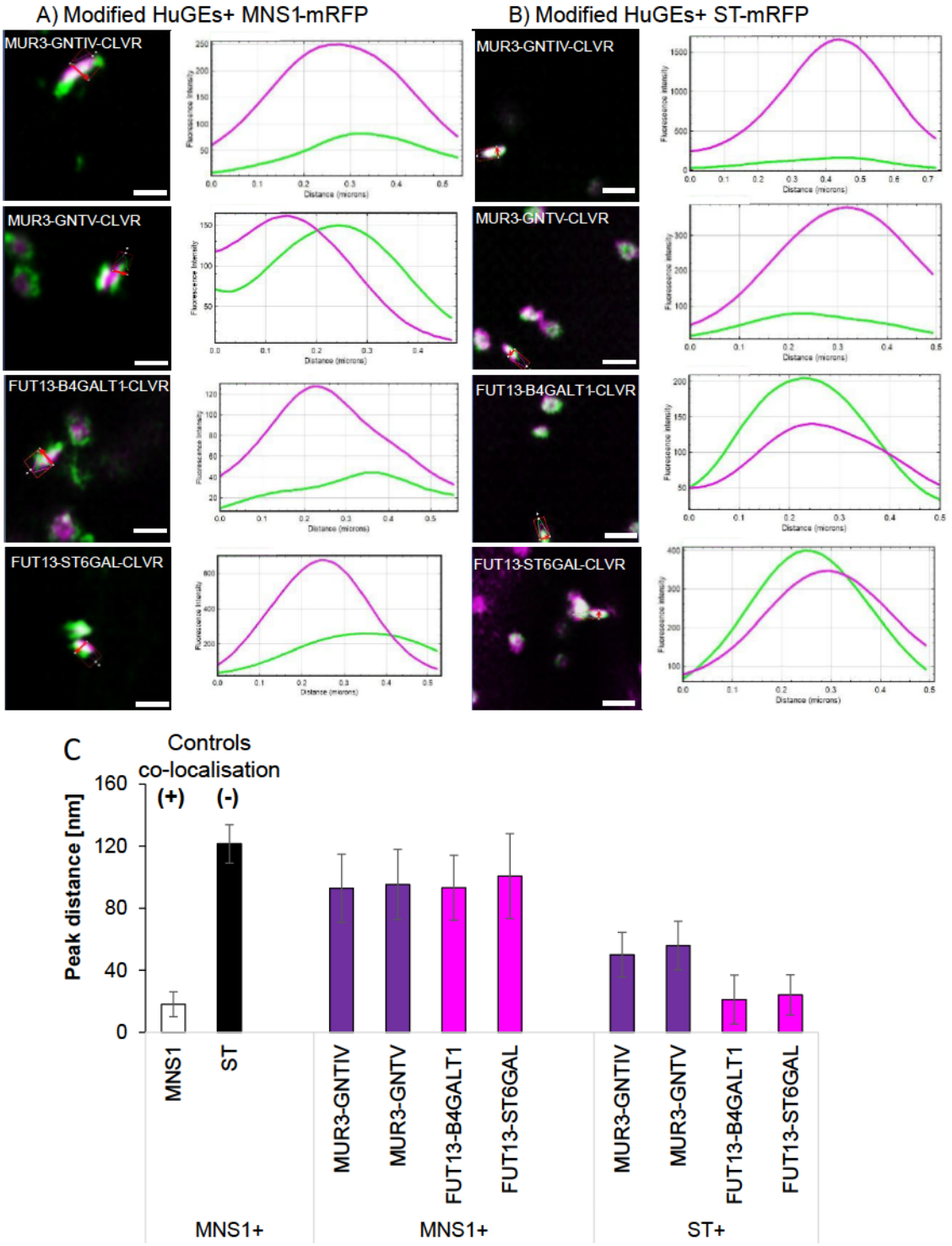


Figure 4.5: Co-localisation line profile analysis for modified HuGEs

A) Example images and analysis for distances between peak fluorescent intensities for MUR3-GNTIV-CLVR, MUR3-GNTV-CLVR, FUT13-B4GALT1-CLVR, FUT13-ST6GAL-CLVR (green), with *cis*-Golgi marker MNS1-mRFP or B) *trans*-Golgi body marker ST-mRFP, respectively (magenta).

C) Peak distance analysis graphs showing controls for co-localisation in the same Golgi cisternae (MNS1-eGFP+MNS1-mRFP, white bar) and for localisation in different Golgi cisternae (MNS1-

eGFP+ST-mRFP, black bar) as well as comparison of the four HuGEs (medial HuGEs in purple and *trans*-Golgi HuGEs in magenta) with the respective markers MNS1-mRFP and ST-mRFP. Representative images are shown. Results shown from n = 4 biological replicas and >40 technical repeats for each combination. Size bars = 1 µm.

Single channel and merged images from representative Golgi analysed in Figure 4.5 are shown below (Figure 4.6).

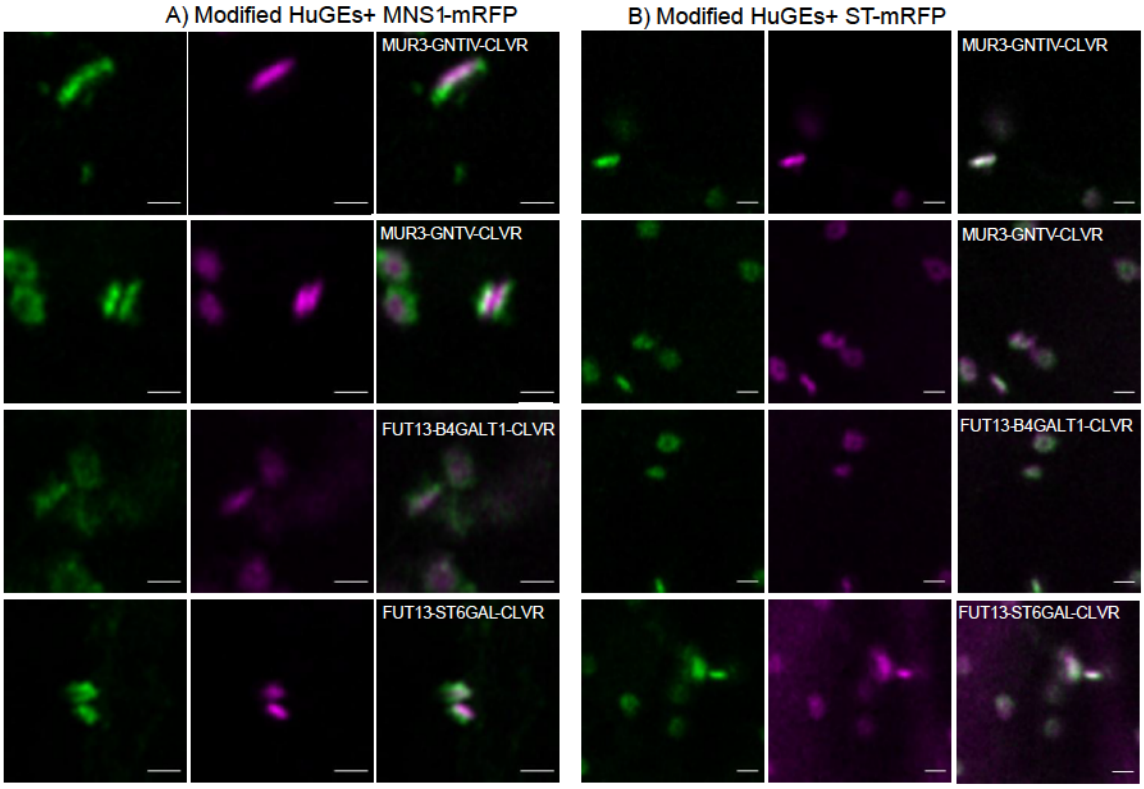


Figure 4.6: Representative single channel images for modified HuGEs from Figure 7.
A) Representative images (single channels green and red as well as merged channel) for MUR3-GNTIV-CLVR, MUR3-GNTV-CLVR, FUT13-B4GALT1-CLVR, FUT13-ST6GAL-CLVR (green), with *cis*-Golgi marker MNS1-mRFP or B) medial/*trans*-Golgi body marker ST-mRFP, respectively (magenta). Size bars = 1 µm.

This indicated that localisation to medial- and *trans*-Golgi cisternae can be distinguished based on the distance between peak intensities of fluorescent protein fusions, determined by statistical line profile analysis. To follow this up we created a double marker construct intended to label two distinct cisternae of the plant Golgi simultaneously on a single construct. This construct contained the CTS of the medial marker MUR3 fused to a blue fluorescent protein (mTagBFP) followed by an intein sequence and P2A self-cleaving peptide sequence (Spatola-Rossi et al., 2023; Spatola-Rossi et al., 2024) and the CTS of the *trans*-Golgi marker FUT13 fused to mRFP. The usage of two markers on the same plasmid but with MUR3-CTS-mTagBFP and FUT13-CTS-mRFP cleaved into separate proteins, enabled equal stoichiometry of marker expression in the same cell.

The modified HuGEs FUT13-B4GALT1-CLVR and FUT13-ST6GAL-CLVR, respectively, were co-expressed with the self-cleaving peptide construct to allow for simultaneous comparison with both a medial- and a *trans*-Golgi marker (Figure 4.7; Figure 4.8). FUT13-B4GALT1-CLVR and FUT13-ST6GAL-CLVR showed a distance between maximum fluorescent intensities with the medial marker MUR3 of 51.3 ± 14.4 nm and 66.4 ± 7.3 nm, respectively (Figure 4.7B, E). These values aligned closely with the difference between the CTS domain markers MUR3 and FUT13 (61.1 ± 18.7 nm and 69.4 ± 14.5 nm, respectively; Figure 4.7A, E). In contrast both FUT13-B4GALT1-CLVR and FUT13-ST6GAL-CLVR were a shorter distance from the *trans*-Golgi marker FUT13 (17.81 ± 10.2 nm and 26.7 ± 11.8 nm, respectively; Figure 4.7C, E). Overall, this analysis showed that FUT13-B4GALT1-CLVR and FUT13-ST6GAL-CLVR are located in *trans*-Golgi cisternae (Figure 4.7D, E).

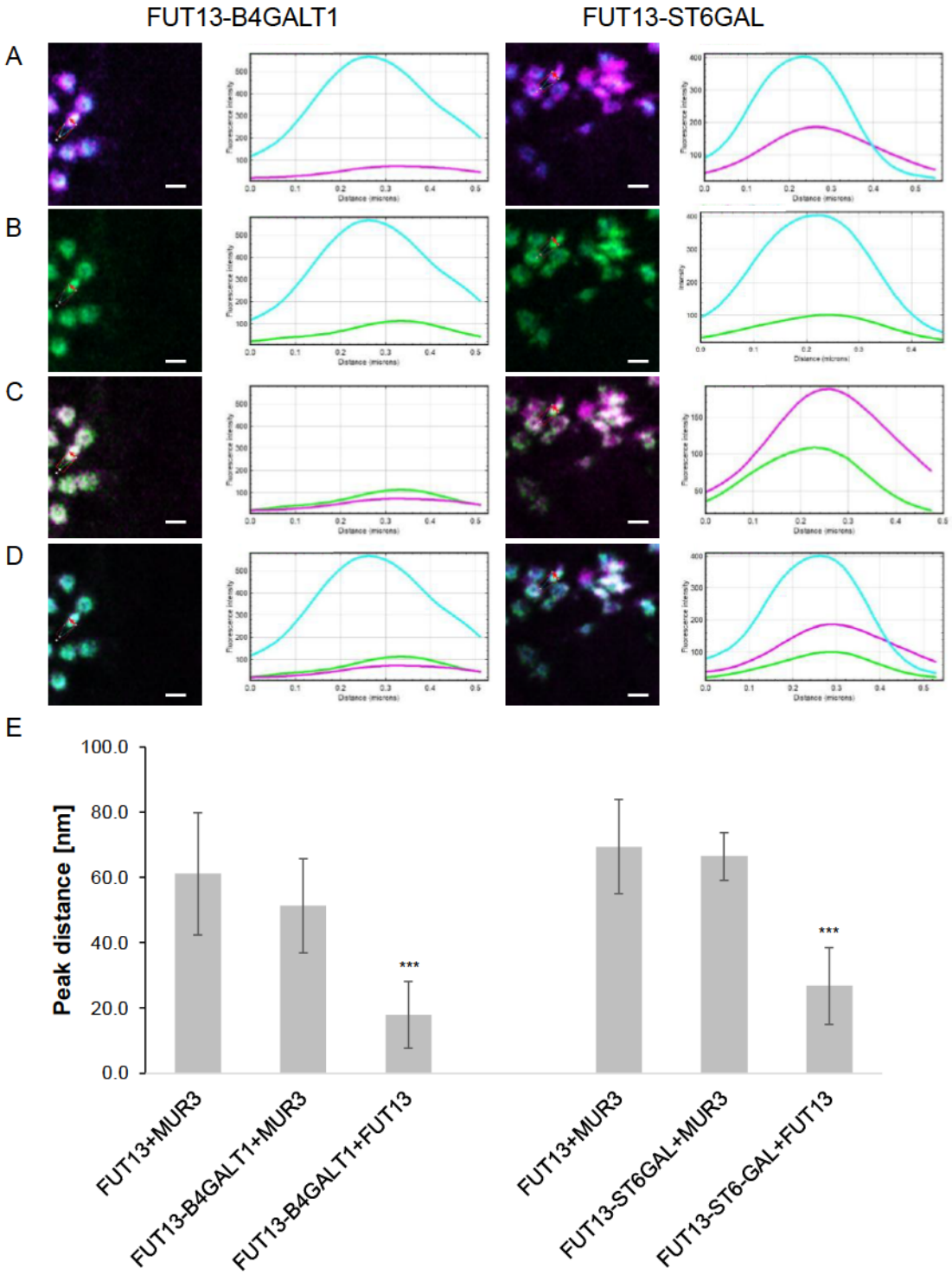


Figure 4.7: Co-localisation line profile analysis for FUT13-B4GALT1-CLVR, FUT13-ST6GAL-CLVR with medial/*trans*-Golgi marker construct.

A) Example images and line profile for the markers MUR3 (blue) and FUT13 (magenta), B) FUT13-B4GALT1-CLVR or FUT13-ST6GAL-CLVR (green) with the medial-Golgi marker MUR3 (blue), C) FUT13-B4GALT1-CLVR or FUT13-ST6GAL-CLVR (green) with the *trans*-Golgi marker FUT13 (magenta), D) shows the overlay of all three proteins and the respective line profiles. Size bars = 1 µm. Single channel images for Figure 4.7 with higher magnification are shown in Figure 4.8.

E) Peak distance analysis graphs showing the markers (MUR3+FUT13) and each marker with FUT13-B4GALT1-CLVR and FUT13-ST6GAL-CLVR, respectively. Significance was analysed by Kruskal-Wallis (***p< 0.001). Results shown from n = 3 biological replicas and 6 technical repeats.

Enlarged single channel and merged images of an individual representative Golgi used for analysis in Figure 4.7 are shown below (Figure 4.8) and demonstrate three-colour labelling of a Golgi body by the insertion of a self-cleaving marker and HuGE.

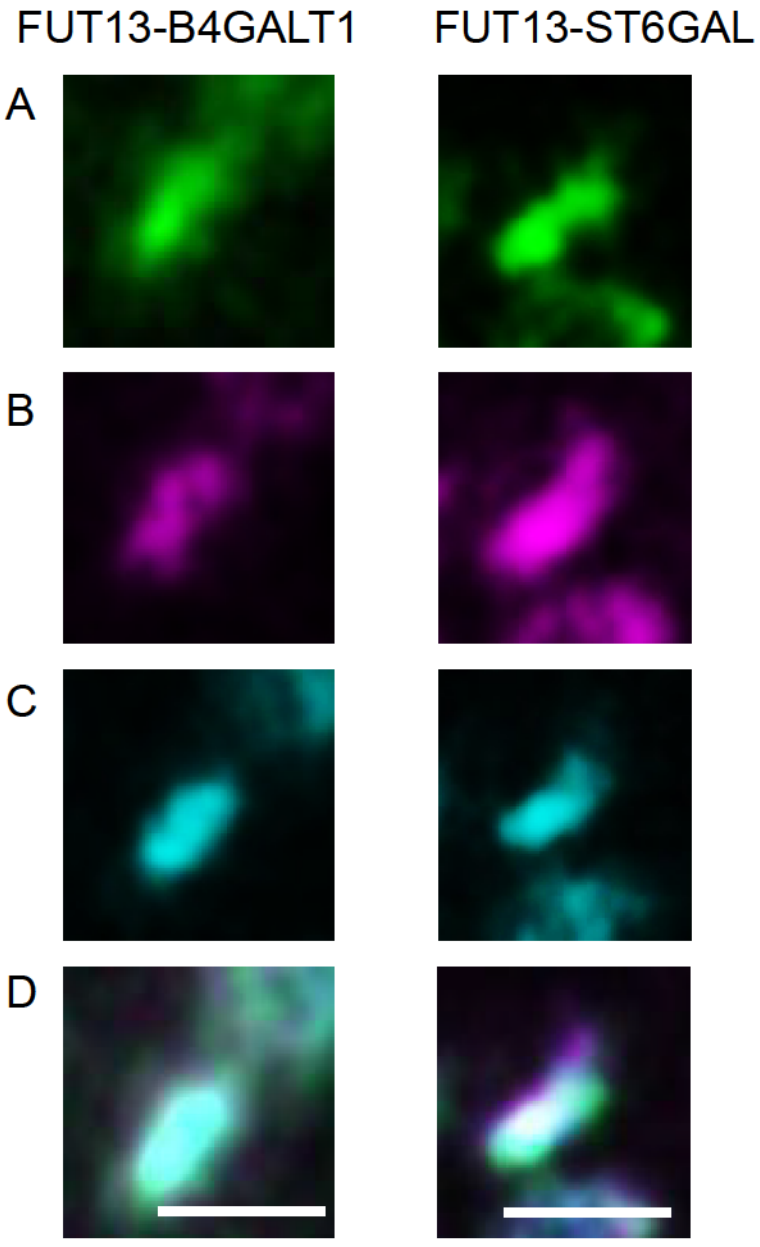


Figure 4.8: Representative and high-resolution images for FUT13-B4GALT1-CLVR, FUT13-ST6GAL-CLVR with medial/*trans*-Golgi marker construct from Figure 4.7.
A) Example images for FUT13-B4GALT1-CLVR or FUT13-ST6GAL-CLVR (green), B) the marker FUT13 (magenta), C) the medial-Golgi marker MUR3 (blue). D) Shows the overlay of all three proteins.

4.4 Discussion

Overall, the data described that HuGEs could be localised not just to the *trans*-Golgi, but also the medial-Golgi cisternae, and Golgi cisternae localisation is largely reliant on the structure of the CTS domain. This suggested that CTS play a significant role in directing localisation of the HuGEs, and show species specificity (Schoberer et al., 2013). Targeting of glycosylation enzymes to specific relevant sub-Golgi cisternae has been suggested to improve enzyme activity and to generate a greater yield of specific glycan structures (Strasser et al., 2009); for example, GNTIV and GNTV produce a more bi- and tri- antennary glycans when expressed in the medial-Golgi than when expressed in the *trans*-Golgi (Bosch et al., 2013; Castilho et al., 2011). A similar study expressed full length human β1,4-GalT in mature *N. benthamiana* plants (Bakker et al., 2001) indicated that resulting glycoforms were largely dictated by the sub-Golgi localisation of the insert, in which galactosylated glycans were scarce and GnGnXF was abundant. As such the concept that β1,4-GALT and ST should target different cisternae of the plant Golgi body in *N. benthamiana (Schoberer* et al., 2011; Strasser et al., 2009) was also factored in designing the localisation of HuGEs.

Replacing the mammalian CTS with plant-specific sequences has shown to be a valid approach for modifying targeting and enhancing expression levels of mammalian HuGEs in plant systems. These successful protein engineering approaches are the first steps towards a plant-based expression system for the production of therapeutic proteins of interest with modified N-linked glycosylation. In addition, high-resolution dynamic confocal microscopy and self-cleaving peptide markers have been adapted as a robust and effective methodology for sub organellar protein localisation studies in Golgi body cisternae. Hence, the results obtained supports existing research on fundamental conservation of targeting mechanisms, based on the properties of glycosylation enzymes as type II transmembrane proteins (Schoberer et al., 2010). Specificity of the localisation observed in our results also circumvents some risk that

random introduction of HuGEs could interfere with endogenous glycosylation and lead to generation of undesirable or incomplete N-glycan structures (Chen et al., 2016).

Precedent in the literature for insertion of HuGEs focuses primarily on the introduction of α-2,6-sialic acid into the glycan chain, this being a major motif of interest for the production of human therapeutics, and an early study observed activity of human α-2,6-sialyltransferase with *trans*-Golgi body localisation, demonstrating that glycosyltransferases can be subcompartmentalized to specific cisternae of the plant Golgi apparatus (Wee et al., 1998). Literature also provides precedent for effective *in planta* sialylation of a therapeutic protein achieved using a recombinant mAb as acceptor substrate observed after achieving proper subcellular localisation of an ST-GalT construct (Castilho et al., 2010) whereby all available acceptor substrates were sialylated, as would be the intended outcome of this system. However, with presence of some incompletely processed structures present attributed to overexpression of ST-GalT interfering with endogenous glycan processing.

Progress on optimising localisation of HuGEs to the late Golgi enabled the next steps for generating mammalian glycomotifs in plants as a viable production system. With cisternae-specific expression addressed, the next objectives were to express 4 HuGEs together, at sufficient expression levels, with as few separate vectors as possible.

Chapter 5: Protein production in stable plants

Contents

5.1 Plant systems for production of recombinant proteins

Plants offer the potential production of scalable, affordable, correctly folded protein in a host innately tolerated across human applications Plant systems lack serious contamination threats (Burnett et al, 2020), alongside typically reduced costs (Kizhner et al, 2015). Plant systems designed for protein production have been a fast-growing development and 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). Major benefits of plant-based systems for recombinant protein production are low-cost, sustainably scalable, offer reliable stable or transient expression, and flexibility of genetic modification (Chen et al., 2016; Nagels et al., 2011). Significant advances have been made in improving yields of soluble therapeutic proteins by diminishing proteolytic degradation in plant systems (Hehle et al., 2015) and manipulating ER expansion for complex protein assembly (Göritzer et al., 2025). The innate tolerance of mammals to plant proteins, through everyday contact and diet, makes plant-produced therapeutics highly desirable when compared to those produced in vectors, especially such as bacteria, because of their lack of a substantial immunogenic profile (Nagels et al., 2011). Similarly, yeast and bacterial systems can generate high risk immunogenic contaminants i.e. bacterial lipopolysaccharides and terminally mannosylated glycan structures; risk factors for therapeutic applications as they can trigger detrimental immune responses in patient (Burnett et al., 2020, Gomes et al., 2016).

In plant Golgi bodies, the glycosylation machinery differs from the mammalian equivalent and so is incapable of producing specific mammalian glycomotifs and a subsequent mammalian-type glycosylation profile (Fig. 1 (Strasser et al., 2014; Schoberer et al., 2018a; Schoberer et al., 2018b). Plant systems can perform the major post translational modification for eukaryotes, N-linked glycosylation, but do not generate mammal-specific motifs which can limit their effectiveness in some applications (Göritzer et al., 2021). The introduction or amplification of key mammalian glycans can significantly improve plasma half-life for therapeutic proteins which has been shown for example in a modified erythropoietin with enhanced end terminal

sialyation (Su et al., 2010). The applications of this in plant biopharming can be extended to reduce presence of problematic plant-specific glycan structures on synthesized therapeutic proteins (Kogelmann et al., 2023; Strasser et al., 2008).

Native plants enzymes can functionally N-glycosylate expressed heterologous human proteins, but do not possess specific medial- and trans-Golgi enzymes required for the generation of a multi-antennary N-glycan chain with end terminal sialic acid capping. Plant machinery will be adapted to produce human glycan patterns by the addition of mammalian enzymes GNTIV and GNTV (N-acetylglucosaminyltransferase IV and V), B4GALT1 (β-1,4-galactosyltransferase), and ST6GAL (α-2,6-sialyltransferase), and potential downregulation of the plant enzymes XylT (β-1,2-xylosyltransferase), FUT11 and FUT12 (α-1,3-fucosyltransferases) to avoid interference with plant glycosylation patterns. GNTIV, GNTV, ST6GAL and B4GALT1 have been targeted to medial or *trans*-Golgi cisternae (Figure 5.1) for effective enzymatic function on N-glycan intermediates (Schoberer et al., 2018b).

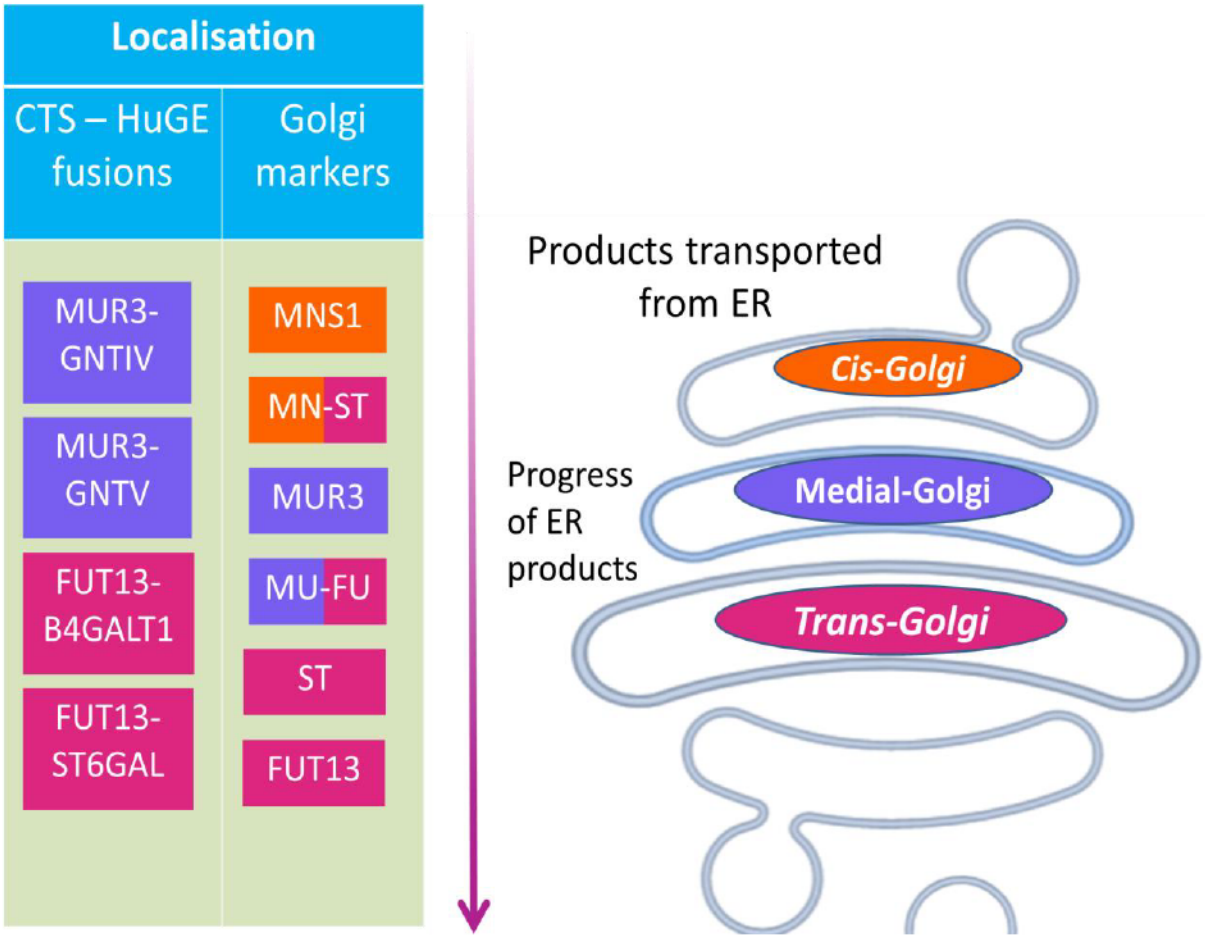


Figure 5.1: Schematic representation of sub-organellar organisation of the plant Golgi body cisternae, with relevant cisternae-specific markers colour-coded. Orange = *cis*-Golgi targeting (MNS1, from Golgi-α-mannosidase I), purple = medial-Golgi targeting (MUR3, from β-1,2-galactosyltransferase), magenta = *trans*-Golgi targeting (ST, from *Rattus norvegicus* α-2,6-sialyl transferase; FUT13, from α-1,4-fucosyltransferase).

Existing plant systems investigated for glycomodification in relation to the production of therapeutics have investigated a variety of growth forms and means of genetic modification. Stable (Bohlender et al., 2020) and transiently (Castilho et al., 2011) agrobacterium-mediated genetically modified mature plant systems have been successfully investigated in this manner. For the purposes of producing a generalizable production platform in *Nicotiana benthamiana* this project was directed to working towards a stably transformed line of glycomodified tobacco (Figure 5.2), within which a variety of proteins of interest could be transiently expressed.

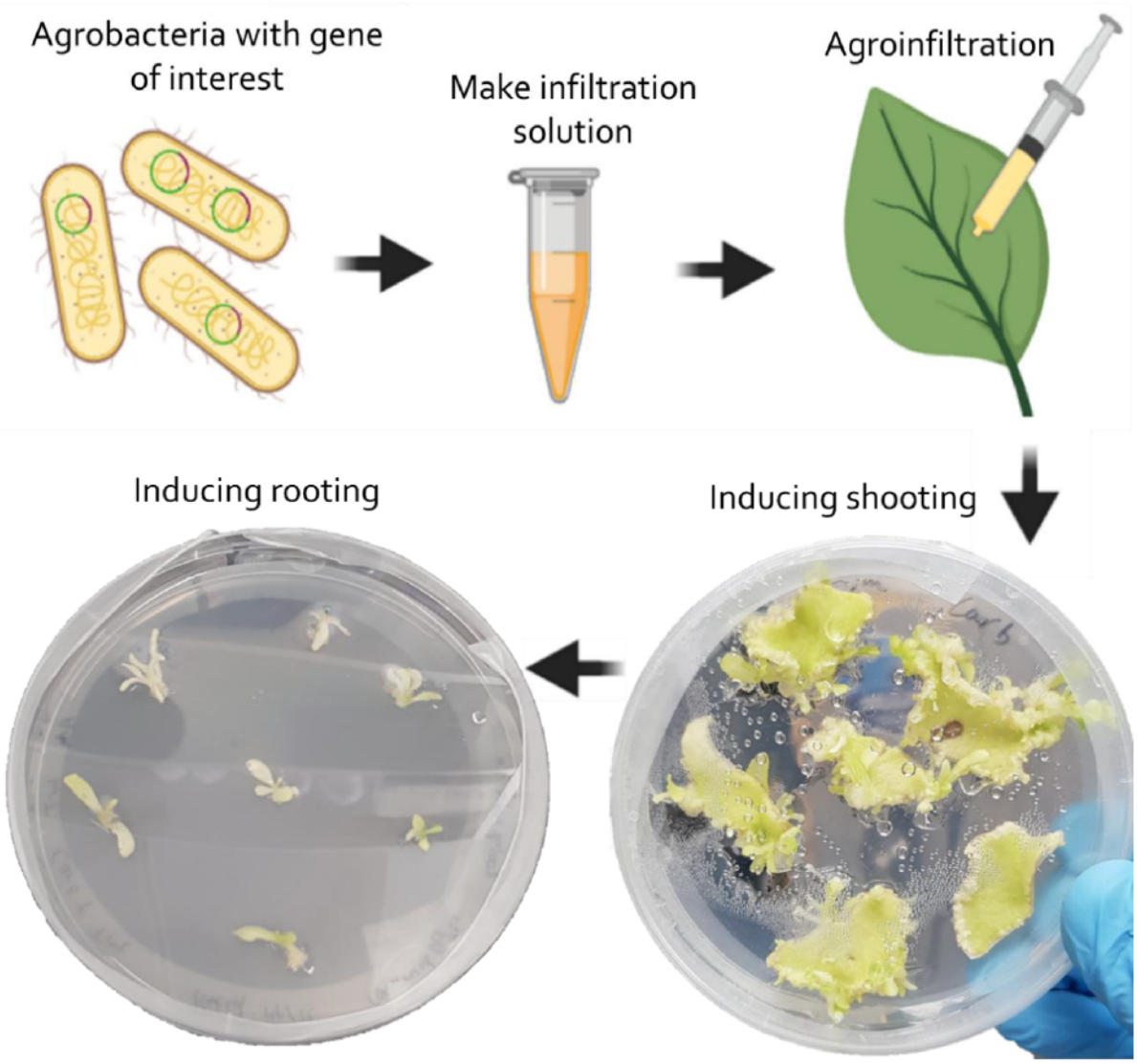


Figure 5.2: Basic workflow for agrobacterium-mediated expression for mature stable plant regeneration in tobacco. Shooting and rooting occurs under antibiotic selection in solid media. Images collected from successful regeneration of *Nicotiana benthamiana*.

Both plant and human glycosylation enzymes are type II membrane proteins possessing a single pass signal-anchor sequence provided by a short N-terminal CTS domain (~40-70aa), this provides specific sub-Golgi localisation and anchors the catalytic enzyme subunit into the Golgi lumen (Schoberer et al., 2011). The CTS region is responsible both for orientation of the enzyme's catalytic domain in the Golgi lumen, and providing sub-Golgi localisation, offering an explanation to the non-uniform distribution of glycosylation enzymes in the plant Golgi (Figure 5.1) (Schoberer et al., 2011).

To target HuGEs *in planta* two CTS domains were selected from *Arabidopsis thaliana* resident glycosylation enzymes, a MUR3 CTS from β-1,2-galactosyltransferase as the medial-Golgi targeting domain and a FUT13 CTS from α1,4-fucosyltransferase for *trans*-Golgi targeting. The mammalian CTS domains were replaced with the Arabidopsis CTS domains placed upstream of the mammalian catalytic domain redirecting HuGE localisation in a cisternae-specific manner (McGinness et al., 2024) and the following fusion proteins were created: MUR3-GNTIV, MUR3-GNTV, FUT13-B4GALT1, FUT13-ST6GAL.HuGE localisation to plant Golgi cisternae was assessed with fluorescent marker proteins and high-resolution confocal microscopy. Cisternae localisation of HuGEs can be both visualized and analysed using fluorescence intensity line profiles to determine the distance between individual fluorophore expression peaks of markers and HuGEs (McGinness et al, 2022). This research was prerequisite to the development of stably transformed HuGE expression tobacco lines and the production of human proteins of interest within. Stable lines were generated in wild type *Nicotiana benthamiana* (WT) and Richard Strasser's RNAi knock-down line (ΔXFT) of *Nicotiana benthamiana* with reduced activity of endogenous plant glycosylation enzymes β1,2-xylotransferase and α1,3-fucosyltransferase (Strasser et al., 2008).

5.2 aims:

1. Expression of HuGEs with directed localisation to the correct Golgi cisternae in tobacco plants in a stable manner, both in wild-type plants and tobacco lines with down regulated plant glycosylation (ΔXFT).
2. Analysis of glycosylation motifs of human proteins of interest expressed in these modified tobacco lines, including Lysosomal acid lipase (LAL) and immunoglobulins (IgG and IgE)

5.3 Results

5.3.1 Overlap extension PCRs for HuGE expression constructs

Existing sequences for CTS-modified HuGEs (Appendix 3) were used for an overlap-extension PCR protocol to generate constructs with both catalytic subunits of medial-Golgi targeting or *trans*-Golgi targeting constructs fused together. These use only one N-terminal CTS domain (MUR3/FUT13 respectively) and a linker domain between catalytic subunits. To begin, multiple fragments (Figure 5.3, PCRs 1-3) of the enzymes with linkers were generated via successive runs of overlap extension PCR to generate a first fragment of a CTS domain attached to the first enzymatic domain and a second fragment with a linker and the second enzyme.

The two resulting fragments (Figure 5.3, PCR 2 & 3) were combined to form a single complex consisting of a CTS domain, the first enzymatic domain, a linker and the second enzymatic domain (Figure 5.3, PCR4) to form the final insert. This resulted in medial-Golgi fusion MUR3-GNT45 (3535 bp) or *trans*-Golgi fusion FUT13-B4ST6 (1807 bp) (Appendix 5). Inserts were then cloned with Gateway overhangs for use in Gateway cloning. Vectors used (Karimi et al., 2002) for downstream stable plant expression were pH7WG2 (in the context of this research reffered to as 388) to generate a MUR3-GNT45-388 expression vector encoding a Hygromycin resistance for plant stable transformation. Similarly the vector pK7WG2 (in the context of this research reffered to as 389) was used to generate a FUT13-B4ST6-389 expression vector encoding a Kanamycin resistance.

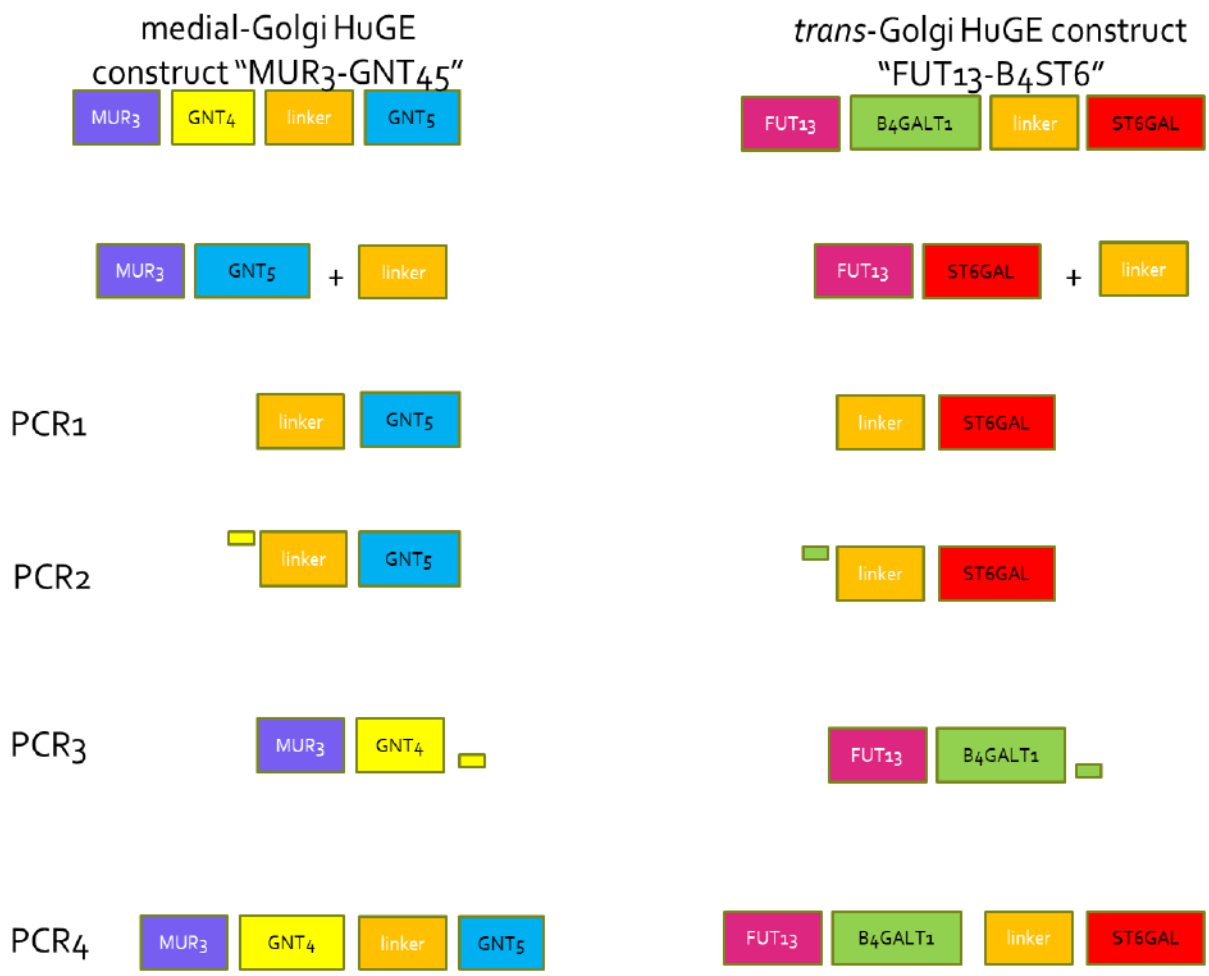


Figure 5.3 Schematic representation of workflow for overlap-extension PCR to generate HuGE fusion proteins. PCR number (left) indicates stage at which intermediate fragments generated. The medial-Golgi fusion MUR3-GNT45 consists of an N-terminal MUR3-CTS domain targeting to the medial-Golgi (purple), attached to the catalytic domain from GNTIV (yellow), which is subsequently attached to the catalytic domain of GNTV (blue) with a linker, generating a medial-targeting (purple) two-enzyme construct. The same rationale is applied for *trans*-Golgi fusion FUT13-B4ST6, an N-terminal FUT13-CTS domain (magenta) targeting the *trans*-Golgi (magenta) fused to the catalytic domain for both B4GATL1 (B4, green) and ST6GAL (ST6, red).

To study the subcellular and sub organellar localisation HuGEs in chapter 2, the design included fluorophore fusions for visualization. All mammalian glycosylation enzymes were preferentially given C-terminal fluorophore tags to assess localisation. As type II membrane proteins they natively possess a single pass 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 limit interference with localisation that an N-terminal fluorescent tag could cause. HuGEs were initially expressed as fluorescent protein fusions for the purpose of assessing localisation. Quantification of CTS-HuGEs indicated that human enzymes can be modified with plant CTS domains for redirected localisation within the plant Golgi stack. Following this success the dual enzyme fusion constructs MUR3-GNT4/5 and FUT13-B4/ST6 (Figure 5.4) used for N-glycosylation modification were expressed without fluorophores in transient and stable plants to prevent interference with protein or glycan production as an industry readiness consideration. In an established expression system for protein production, absence of fluorophores is preferred to prevent any interactions with tagging.

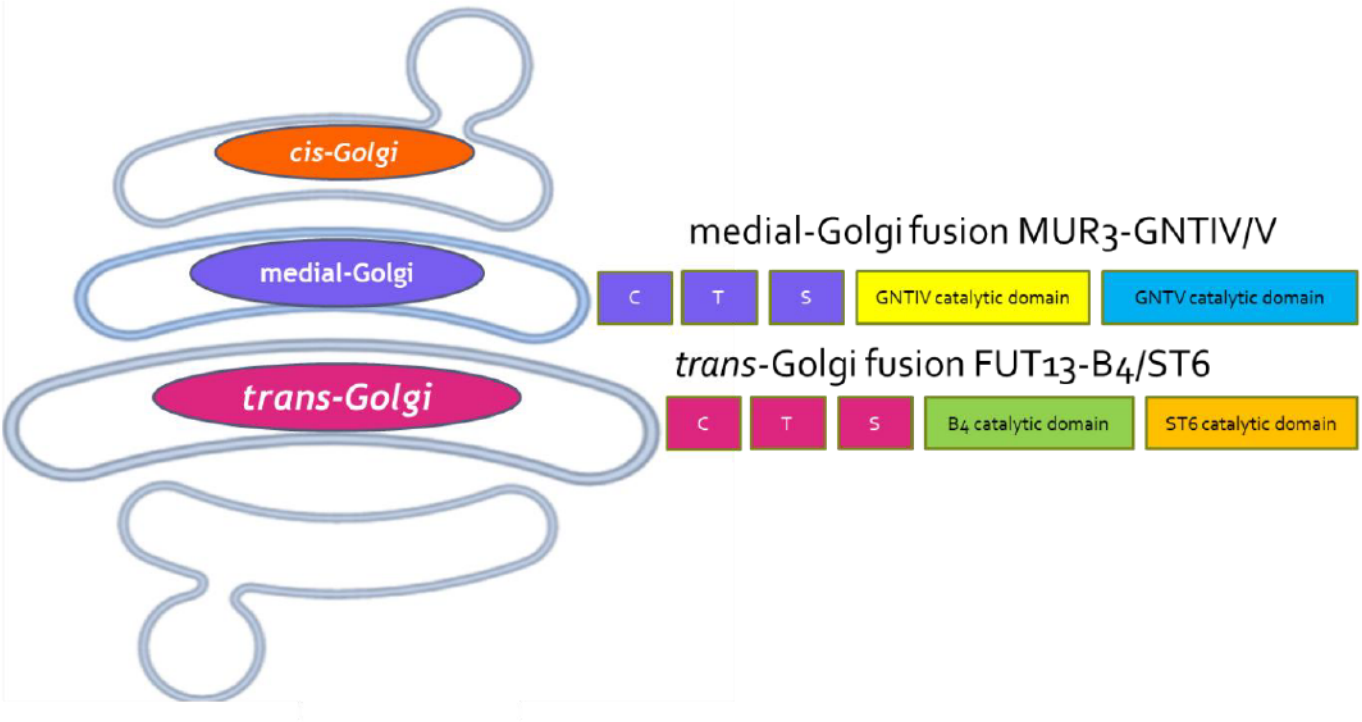


Figure 5.4: Schematic representation of HuGE fusion proteins. The medial-Golgi fusion MUR3-GNT45 consists of an N-terminal MUR3-CTS domain targeting to the medial-Golgi (purple), attached to the catalytic domain from GNTIV (yellow), which is subsequently attached to the catalytic domain of GNTV (blue), generating a medial-targeting (purple) two-enzyme construct. The same rationale is applied for *trans*-Golgi fusion FUT13-B4ST6, an N-terminal FUT13-CTS domain (magenta) targeting the *trans*-Golgi (magenta) fused to the catalytic domain for both B4GATL1 (B4, green) and ST6GAL (ST6, orange).

5.4 Plant regeneration

Successful development and testing of the MUR3-GNT45-388 and FUT13-B4ST6GAL-389 vectors were the prerequisite steps necessary for the development of stably transformed tobacco lines expressing HuGEs for N-linked glycomodification. These lineages provided a supply of plant tissue to assess for HuGE activity by screening for effectiveness of modification made in terms of the mammalian specific N-Glycan structures produced. Plant regeneration methods utilised ensure selection pressure on all transformed plant cells (Figure 5.5) from the beginning as mature tissue into shoot development and root formation, ensuring plants maintain transgenes. Presence of inserts were initially screened by reverse transcription PCR (RT-PCR) of 300bp sequences specific to each HuGE.

Figure 5.5: Example shoot growth for mature stable plant transformation. Left: Successful growth and shoot development of transformed tissue in solid shooting media, white arrows indicate likely shoots for transfer. Right: Successfully developed shoots transferred onto rooting media to stimulate root development, will be grown into stably transformed mature plants.

Generating stable lines in this manner takes several months, so proper controls were essential. This assisted effective methodological practice to reduce the number of unsuccessful attempts, which were limited by project time constraints. Positive controls with wild type tissue and no selection agent (Figure 5.6, A) were used to generate non-specific growth, and so indicate effectiveness of housekeeping antibiotics (that did not select against plant tissue) at preventing agrobacterium overgrowth or other contamination by microbes that would hinder successful regeneration, and appropriate bleaching that does not cause unnecessary tissue death. These controlled for effectiveness of growth conditions, in particular the plant hormone effectiveness in each media batch and regeneration run for the successful production of shoots or roots. This helped to troubleshoot any general method issues not associated with the test lines specifically, including the effectiveness of the bleaching protocol, general sterile environment and timing of plant regeneration in absence of selection.

Negative control plates (Figure 5.6, B) of wild type tissue in presence of selection were useful specifically to indicate effectiveness of the selection agent at preventing non-specific growth. Absence of healthy leaf tissue on these plates indicates successful elimination of non-specific growth, indicating only transformed tissue encoding a resistance to selection could grow. Eliminating non-specific growth in this manner prevents time wastage in screening plant tissue that may not have been successfully transformed initially and prevents transformed tissue competing for growth with wild type tissue.

Successful test plates (Figure 5.6, C) typically show large zones of tissue death with small areas of healthy growth (Figure 5.6, C, white arrows), indicating presence of selection and transformed tissue with resistance to it. Absence of contamination on the plate indicates success of initial handling, bleaching and housekeeping antibiotics. Taken together, this method was successfully used to generate *Nicotiana benthamiana* lines that would likely to express one or both vectors designed to express HuGEs in plants but required further screening to ensure successful transformation. A total of 85 plants were generated across all lines (Table 5.1).

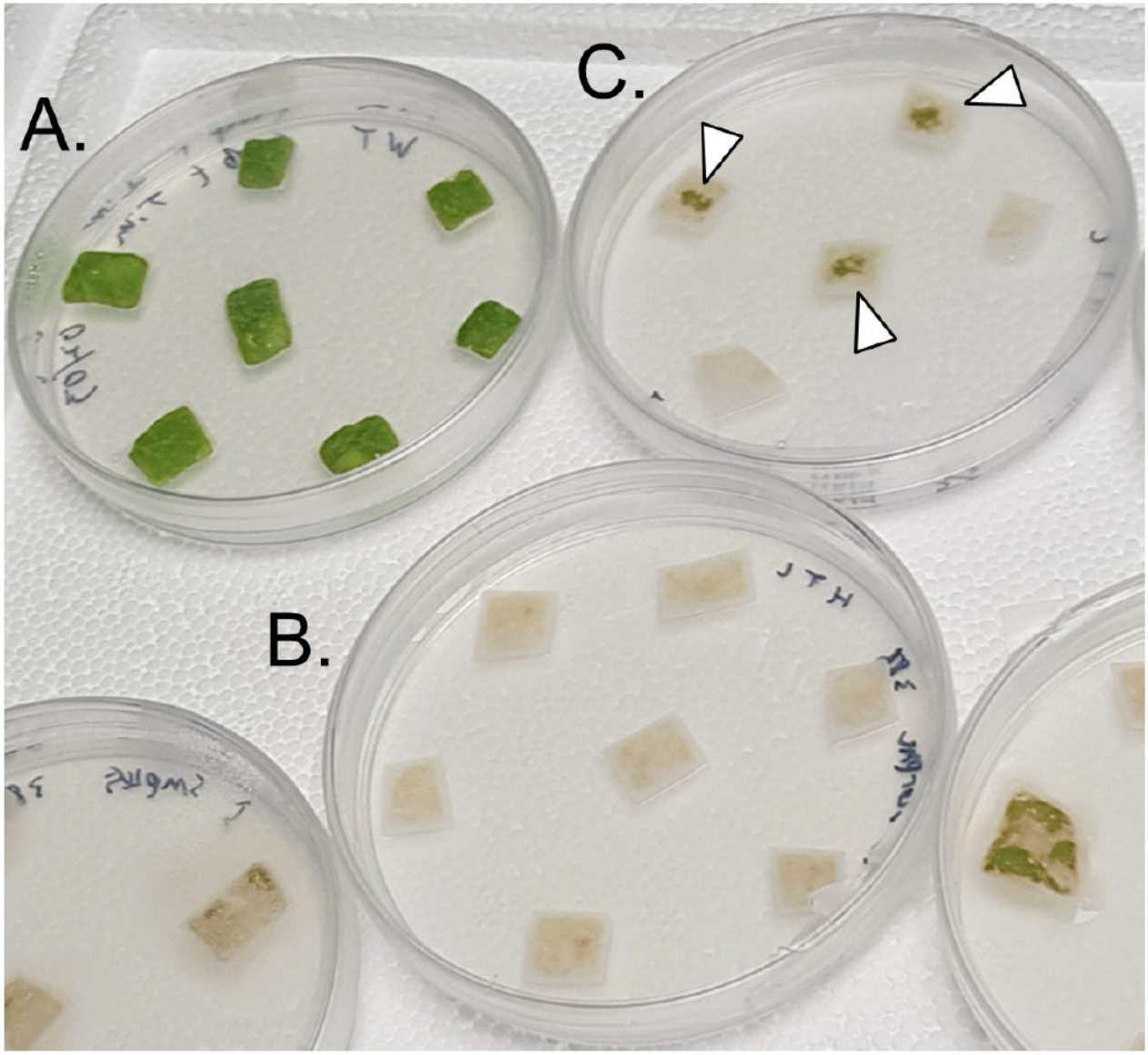


Figure 5.6: Mature stable plant transformation plates at 14 days post regeneration. A: WT tissue positive control without selection agent, living tissue with non-specific growth in absence of selection. B: Negative control with selection agent, selection against tissue without insert containing resistance gene leads to loss of colour and cell death. C: Test plate with transformed tissue and selection agent, white zones of selection occur alongside green areas of specific growth with selection resistance due to presence of transformation insert (white arrows).

Stable lines (Figure 5.7) were initially generated expressing only the MUR3-GNT45-388 in WT *Nicotiana benthamiana* plants as a test case. Later lines expressed both MUR3-GNT45-388 and FUT13-B4ST6GAL-389 vectors in ΔXFT N.B lines.

	A	B	C	D	E	F	G
1	Construct	Vector + plant selection	Number	Background	Date potted	Date regenerated	screened?
2	MUR3-GNT45 T0	388-hygro	#01	N.B WT	17/10/2023	20/06/2023	T1 seedlings confirmed by RT pcr for GNT4 and 5 on gel 03/05/2024
3	MUR3-GNT45 T0	388-hygro	#02	N.B WT	24/10/2023	20/06/2023	T1 seedlings confirmed by RT pcr for GNT4 and 5 on gel 03/05/2024
4	MUR3-GNT45 T0	388-hygro	#03	N.B WT	17/10/2023	23/06/2023	T1 seedlings confirmed by RT pcr for GNT4 and 5 on gel 03/05/2024
5	MUR3-GNT45 T0	388-hygro	#04	N.B WT	24/10/2023	20/06/2023	T1 seedlings confirmed by RT pcr for GNT4 and 5 on gel 03/05/2024
6	MUR3-GNT45 T0	388-hygro	#05	N.B WT	24/10/2023	20/06/2023	
7	MUR3-GNT45 T0	388-hygro	#06	N.B WT	06/10/2023	06/07/2023	
8	MUR3-GNT45 T0	388-hygro	#07	N.B WT	3/11/2023	20/08/2023	
9	MUR3-GNT45 T0	388-hygro	#08	N.B WT	7/11/2023	06/07/2023	
10	MUR3-GNT45 T0	388-hygro	#09	N.B WT	3/11/2023	20/08/2023	
11	MUR3-GNT45 T0	388-hygro	#10	N.B WT	17/11/2023	27/07/2023	
12	MUR3-GNT45 T0	388-hygro	#11	N.B WT	7/11/2023	06/07/2023	Store without screening, surplus
13	MUR3-GNT45 T0	388-hygro	#12	N.B WT	7/11/2023	23/06/2023	Store without screening, surplus
14	MUR3-GNT45 T0	388-hygro	#13	N.B WT	7/11/2023	06/07/2023	Store without screening, surplus
15	MUR3-GNT45 T0	388-hygro	#14	N.B WT	7/11/2023	20/08/2023	Store without screening, surplus
16	MUR3-GNT45 T0	388-hygro	#15	N.B WT	17/11/2023	07/07/2023	Store without screening, surplus
17	MUR3-GNT45 T0	388-hygro	#16	N.B WT	03/11/2023	03/08/2023	Store without screening, surplus
18	MUR3-GNT45 T0	388-hygro	#17	N.B WT	03/11	03/08	Store without screening, surplus
19	MUR3-GNT45 T0	388-hygro	#18	N.B WT	06/07	01/12	Store without screening, surplus
20	MUR3-GNT45 T0	388-hygro	#19	N.B WT	01/12	10/08	Store without screening, surplus
21	MUR3-GNT45 T0	388-hygro	#20	N.B WT	20/6/2023	1/6	Store without screening, surplus
22	MUR3-GNT45 T0	388-hygro	#21	N.B WT	08/12	20/8/2023	Store without screening, surplus
23	MUR3-GNT45 T0	388-hygro	#22	N.B WT	3/11	6//7	Store without screening, surplus
24	MUR3-GNT45 T0	388-hygro	#23	N.B WT	8/11	23/6	Store without screening, surplus
25	MUR3-GNT45 T0	388-hygro	#24	N.B WT	17/11	10/6	Store without screening, surplus
26	MUR3-GNT45 T0	388-hygro	#25	N.B WT	7/11	20/6	Store without screening, surplus
27	MUR3-GNT45 T0	388-hygro	#26	N.B WT	7/11	20/6	Store without screening, surplus
28	MUR3-GNT45 T0	388-hygro	#27	N.B WT	8/12	23/8	Store without screening, surplus
29	MUR3-GNT45 T0	388-hygro	#28	N.B WT	17/11	6/8	Store without screening, surplus
30	MUR3-GNT45 T0	388-hygro	#29	N.B WT	3/1	3/8	Store without screening, surplus
31	MUR3-GNT45 T0	388-hygro	#30	N.B WT	14/12	22/7	Store without screening, surplus
32	MUR3-GNT45 T0	388-hygro	#31	N.B WT	8/12	6/7	Store without screening, surplus
33	MUR3-GNT45 T0	388-hygro	#32	N.B WT	3/1	22/7	Store without screening, surplus
34	MUR3-GNT45 T0	388-hygro	#33	N.B WT	20/12	6/7	Store without screening, surplus

WT-388 | dXF-388:2+389-6A | dXF-388:2+389-6AC | Kaede-RFP

Figure 5.7: Layout example of recorded stable plant lineages generated. Individual plant growth and collections were carefully recorded and grouped by vector used. Dates regenerated and potted were recorded both to track lineages and provide a timeline. 35 plants were successfully grown transformed with the first vector MUR3-GNT45-388, 31 expressing both MUR3-GNT45-388 and FUT13-B4ST6-389. T1 seedlings (seeds from T0 plants) of T0 generated plants (regenerated plants) were later screened by RT-PCR to confirm presence of transcript.

Total number of stably transformed plants that reached maturity and produced seeds for screening were logged (Figure 5.7) with sufficient numbers produced to account for potential variation in expression intensity when screened by RT-PCR, and to account for low survival rate of plants throughout regeneration protocol (Table 5.1).

Table: 5.1: Mature regenerated plants produced. Original plant lineage, insert for transformation and total number of stable plants generated that reached maturity and produced seeds suitable for HuGE enzyme expression screening.

Plant lineage	Insert & Selection antibiotic	Number of mature plants generated
WT *N. Benthamiana*	388- Hygromycin	35
ΔXFT *N. Benthamiana*	388-Hygromycin & 389-Kanamycin	50

5.5 Confirming stables: RT-PCR

Removal of fluorophores from the HuGE constructs necessitated a different method of identifying successful insert expression. Generating stable plants with the given methodology was not guaranteed to succeed, and there were likely variation of transcription between individual plants. Therefore, stable plants were screened for construct expression by RT-PCR to confirm presence of HuGE-specific transcripts (Figure 5.8). 14-day seedlings of the T1 generation of each line were selected as appropriate to screen. Primers were designed against 300 bp fragments for each of the HuGE inserts within the two expression cassettes (MUR3-GNT4, MUR3-GNT5, FUT13-B4GALT1, FUT13-ST6GAL) to identify transcripts *in planta*. As these sequences do not occur natively in plants, this would effectively indicate presence of the transcript for each HuGE insert, and results were compared against wild-type tissue as a control (Figure 5.8).

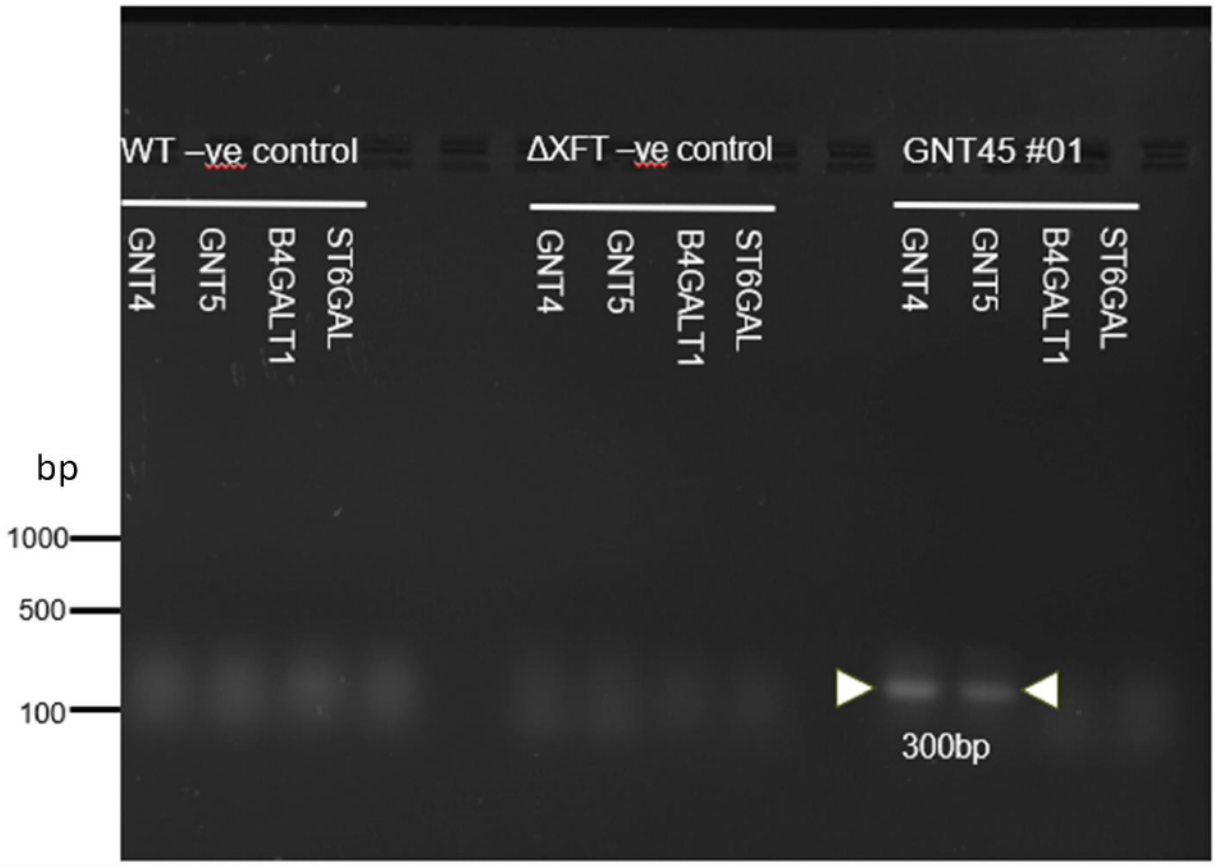


Figure 5.8: Example RT-PCR gel of three T1 seedlings assessed for expression of all four HuGE inserts. Arrow indicates 300BP banding. WT= wild type plant without transformation screened as negative control. ΔXFT= Xylose/Fucose deficient RNAi tobacco line without transformation screened as negative control. GNT45 #01= stable plant line generated with the MUR3-GNT45-388 construct, where N is lineage number. No bands are visible for WT or ΔXFT control, nor on PCR control. 300BP bands present for both GNT4 and GNT5 for MUR3-GNT45#01 seedlings. Black lines indicate sizes in bp from ladder.

The first stable lineages were generated expressing the MUR3-GNT45-388 vector for MUR3-GNT4 and MUR-GNT5 using hygromycin selection. RT-PCR of 14-day T1 seedlings for the first four stable plant lineages generated (GNT45#01-04) identified positive bands for both constructs (MUR3-GNT4 + MUR3-GNT5) across all plants screened (Figure 5.9). No bands were seen in the wild-type negative control or the PCR control. This indicates HuGE transcript presence for MUR3-GNT45-388 that were specific to the modified plants in these four lines.

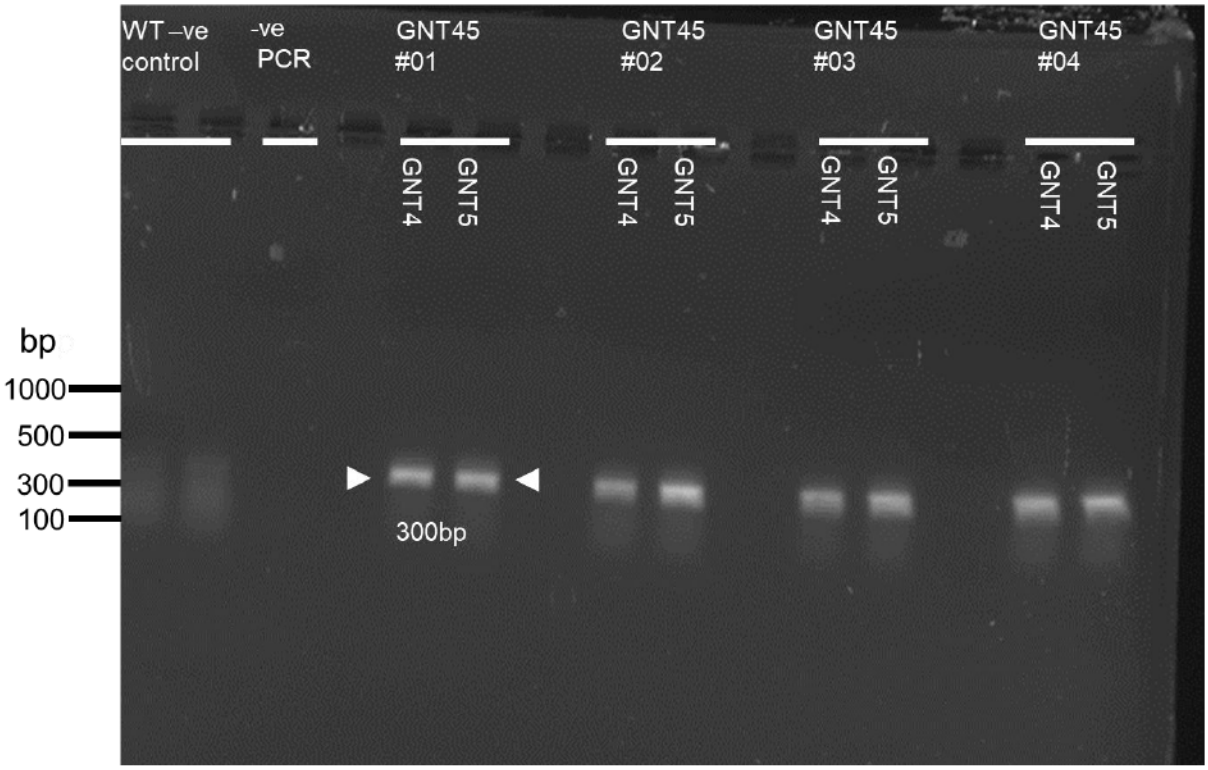


Figure 5.9: RT-PCR gel of four T1 seedlings assessed for expression of a GNT4 & GNT5 insert. Arrow indicates 300BP banding. WT= wild type plant without transformation screened. GNT45 #NN= stable plants generated with the MUR3-GNT45-388 construct, where N is lineage number. No bands were visible for WT control or on PCR control. 300BP bands were present for both GNT4 and GNT5 across four stable plants. Black lines indicate sizes in bp from ladder.

Following generation and positive screening of the GNT45#NN lines, GNT45#01 was used as a successful stable lineage for the expression of MUR3-GNT4 & MUR3-GNT5. This established a successful transformation methodology. To achieve glycosylation intended still required further modification by the stable co-expression of the 389 vector to introduce the inserts FUT13-B4GALT1 and FUT13-ST6GAL into the plant *trans*-Golgi (Figure 5.4). Two strategies were tested for expression of all four HuGEs; transient expression of the FUT13-B4ST6-389 vector into generated stable GNT45#NN lines, or co-expression of the MUR3-GNT45-388 and FUT13-B4ST6-389 cassettes into ΔXFT plants for stable lines.

Transient expression of the FUT13-B4ST6-389 cassette into generated stable GNT45#NN lines were first tested (Figure 5.10) as the fastest approach for achieving expression of four HuGEs, with the intention to provide rapid results and test expression before committing to the lengthy plant regeneration protocol with both MUR3-GNT45-388 and FUT13-B4ST6-389 cassettes. RT-PCR was performed against 300BP sequences individual to each of the 4 HuGEs for wild type tissue, mature GNT45#01 tissue, and two repeats of GNT45#01 mature plants that had undergone transient agrobacterium mediated expression of the FUT13-B4ST6-389 cassette. No bands were seen on either the WT or the PCR negative control. Bands for MUR3-GNT4 and MUR3-GNT5 were visible on each of the GNT45#01 plants tested, however no bands were present for FUT13-B4GALT1 or FUT13-ST6GAL in any condition.

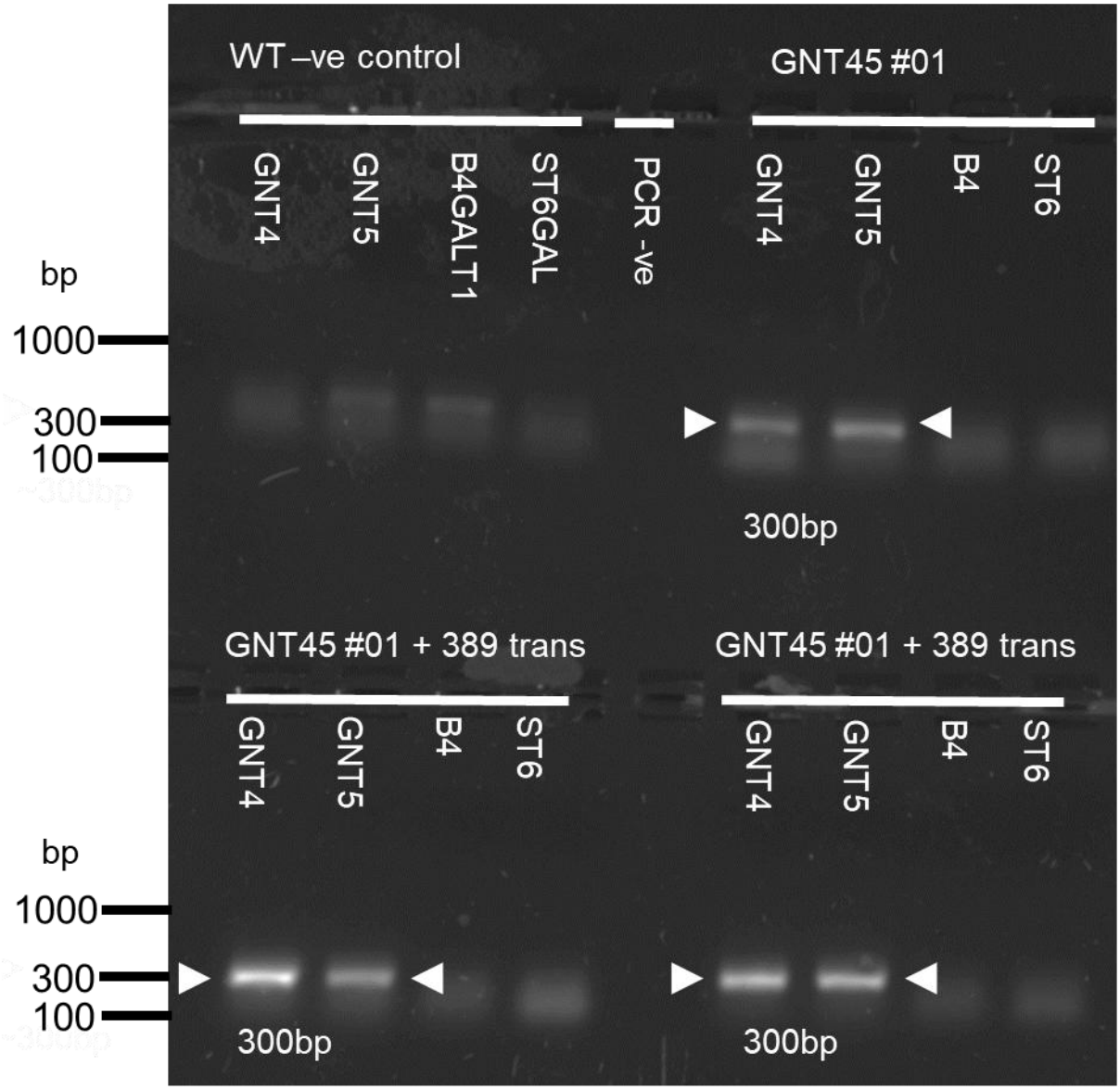


Figure 5.10: RT-PCR gel of four T1 seedlings assessed for mixed stable and transient expression of all 4 HuGE inserts. Arrow indicates 300BP banding on ladder. WT= wild type plant screened. GNT45 #NN= stable plants generated with the MUR3-GNT45-388 construct, where N is lineage number. Samples were labelled + 389 trans indicates mature plant transient agrobacterium mediated expression of the FUT13-B4ST6-389 into stable GNT45# line. No bands were visible for WT control or on PCR control. 300BP bands were present for both GNT4 and GNT5 across the three stable plant lines. No bands were present in any condition for B4GALT1 or ST6GAL1. Black lines indicate sizes in bp from ladder.

Following unsuccessful RT-PCR for attempts to transiently express the FUT13-B4ST6-389 vector into stable GNT45 lines (Figure 5.11), an alternative option was

explored. MUR3-GNT45-388 and FUT13-B4ST6-389 vectors were co-expressed under dual kanamycin and hygromycin selection into ΔXFT plants to generate stable lines. RT-PCR of 14-day T1 identified positive bands for all four HuGE transcripts (MUR3-GNT4, MUR3-GNT5, FUT13-B4GALT1, FUT13-ST6GAL) in FUT13-B4ST6-389 seedlings (Figure 5.11). This indicates HuGE transcript presence for MUR3-GNT45-388 and FUT13-B4ST6-389, and so in theory successful generation of stable lines expressing both 388 and 389 vectors.

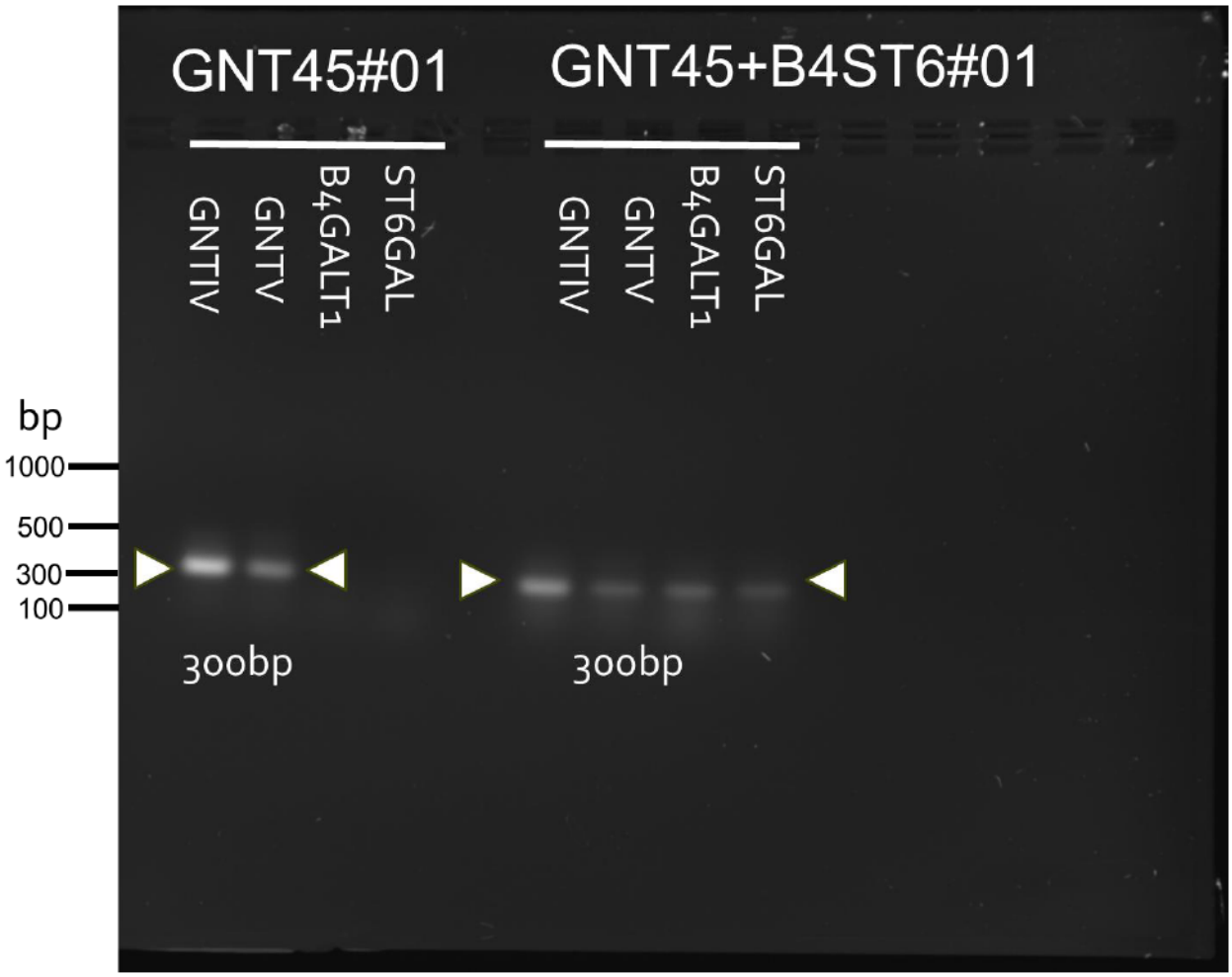


Figure 5.11: RT-PCR gel for T1 seedlings assessed for stable expression of HuGE inserts. Arrow indicates 300BP banding on ladder. GNT45 #01= stable plant generated with the MUR3-GNT45-388 construct. GNT45+B4ST6#01= stable plant generated with co-expression ofthe MUR3-GNT45-388 and FUT13-B4ST6-389 constructs. 300BP bands are present for both GNT4 and GNT5 in the GNT45 #01 line, with bands present for all 4 HuGE inserts in GNT45+B4ST6#01 seedlings. Black lines indicate sizes in bp from ladder.

5.4 Expressing human proteins of interest

Initial infiltrations of human LAL as a fluorescent protein fusion had shown successful expression and localisation to Golgi bodies, demonstrating apparent co-localisation with the *trans*-Golgi marker ST-eGFP (Figure 5.12). However, low transformation efficiency of the fluorophore fusion was an issue, with expression found in an extremely low proportion of tobacco cells tested, with a minimal number of cells strongly expressing the construct.

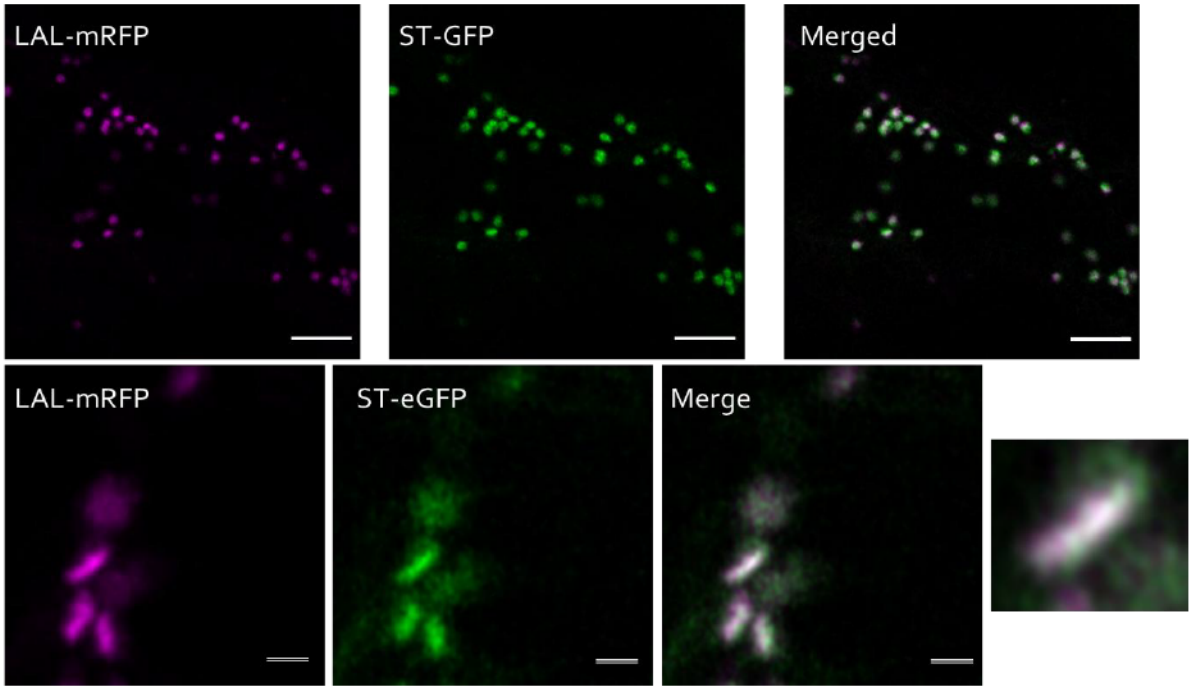


Figure 5.12: Co-localisation of LAL-mRFP with the Golgi marker ST-eGFP. LAL-mRFP (magenta) is co-expressed with the *trans*-Golgi marker ST-eGFP (shown in green). Displays a subcellular view of LAL-mRFP with ST-eGFP to indicate distinct localisation to the Golgi body. Inset shows enlarged structures indicating co-localisation. Size bars = 1 µm. Representative images are shown; n=3 with 5 technical replicas each.

To identify issues 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 5.13). 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.

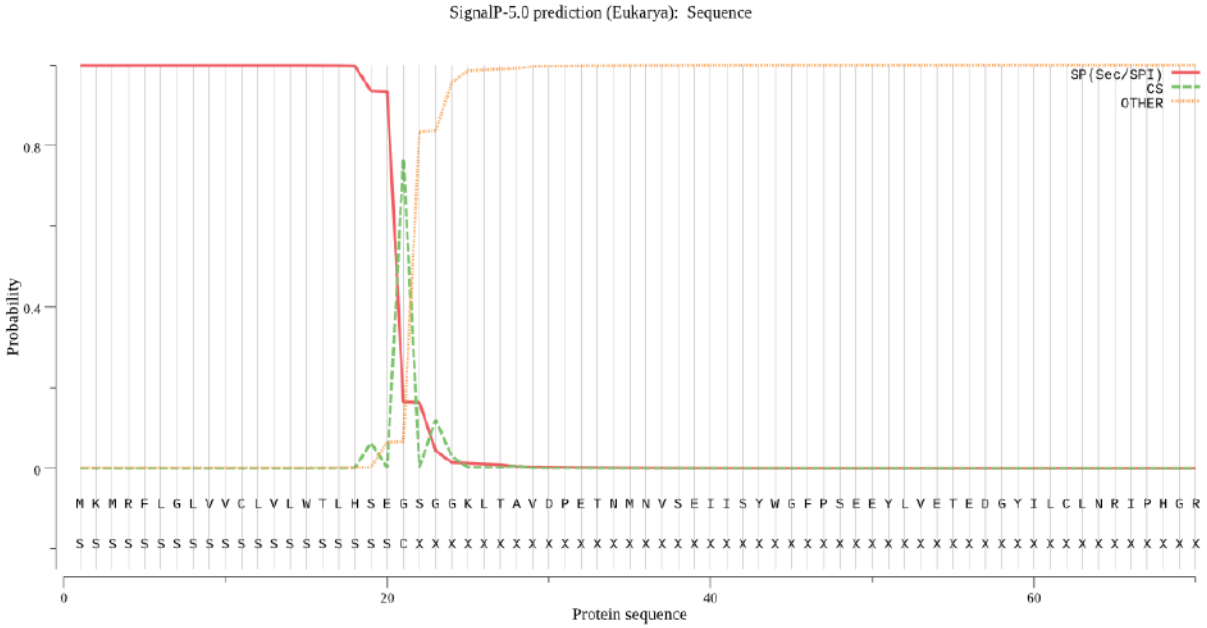


Figure 5.13: Sequence analysis of LAL using SignalP. SignalP output of LAL sequence indicates a predicted lipoprotein signal peptide (Sec/SPII) site position in the sequence between 1-22aa with a probability of 0.7682.

In response, a modified LAL construct lacking the endogenous signal peptide was created and termed LALΔ1-22, based on predicted signal peptide position (Figure 5.13). LALΔ1-22 was fused to a C-terminal mRFP and co-expressed with the *trans*-Golgi marker ST-GFP (Figure 5.14). LALΔ1-22-mRFP showed the same general localisation as LAL-mRFP and co-localised with the trans-Golgi marker ST-eGFP. Concluding that the removal of the signal peptide did not seem to drastically impact on localisation.

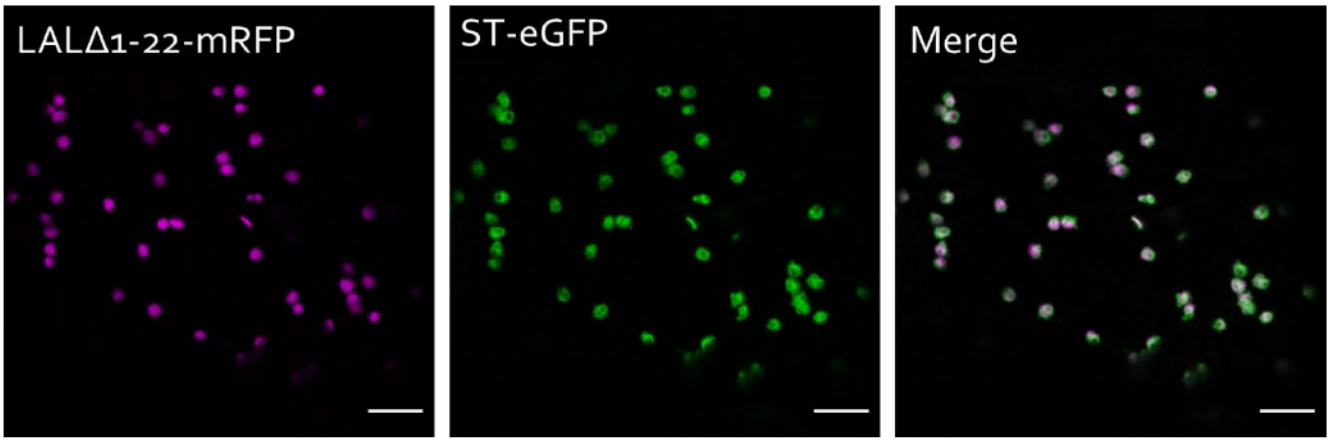


Figure 5.14: Co-localisation of LALΔ1-22-mRFP with the Golgi marker ST-eGFP. LAL-mRFP (magenta) is co-expressed with the *trans*-Golgi marker ST-eGFP (shown in green). Displays a subcellular view of LALΔ1-22-mRFP with ST-eGFP to indicate distinct localisation to the Golgi body. Size bars = 5 µm. Representative images are shown; n=3 with 5 technical replicas each.

To quantitatively assess improvements in expression, cell counts were performed using a 10x objective lens to assess the proportion of cells expressing LALΔ1-22 (Figure 5.15). Visual cell counts were undertaken of cells clearly expressing the LALΔ1-22 construct (Figure 5.15, magenta bordered) and of cells presenting the marker protein ST-GFP (Figure 5.15, 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 (below 1%), surface cell counts could not be accurately conducted to enable comparison. Of 568 cells expressing infiltrated constructs, 447 expressed only ST-GFP and 121 expressed also LALΔ1-22, an average of approximately 27% of cells actively expressing LALΔ1-22. The removal of the signal peptide therefore results in increased transformation efficiency and so protein expression levels.

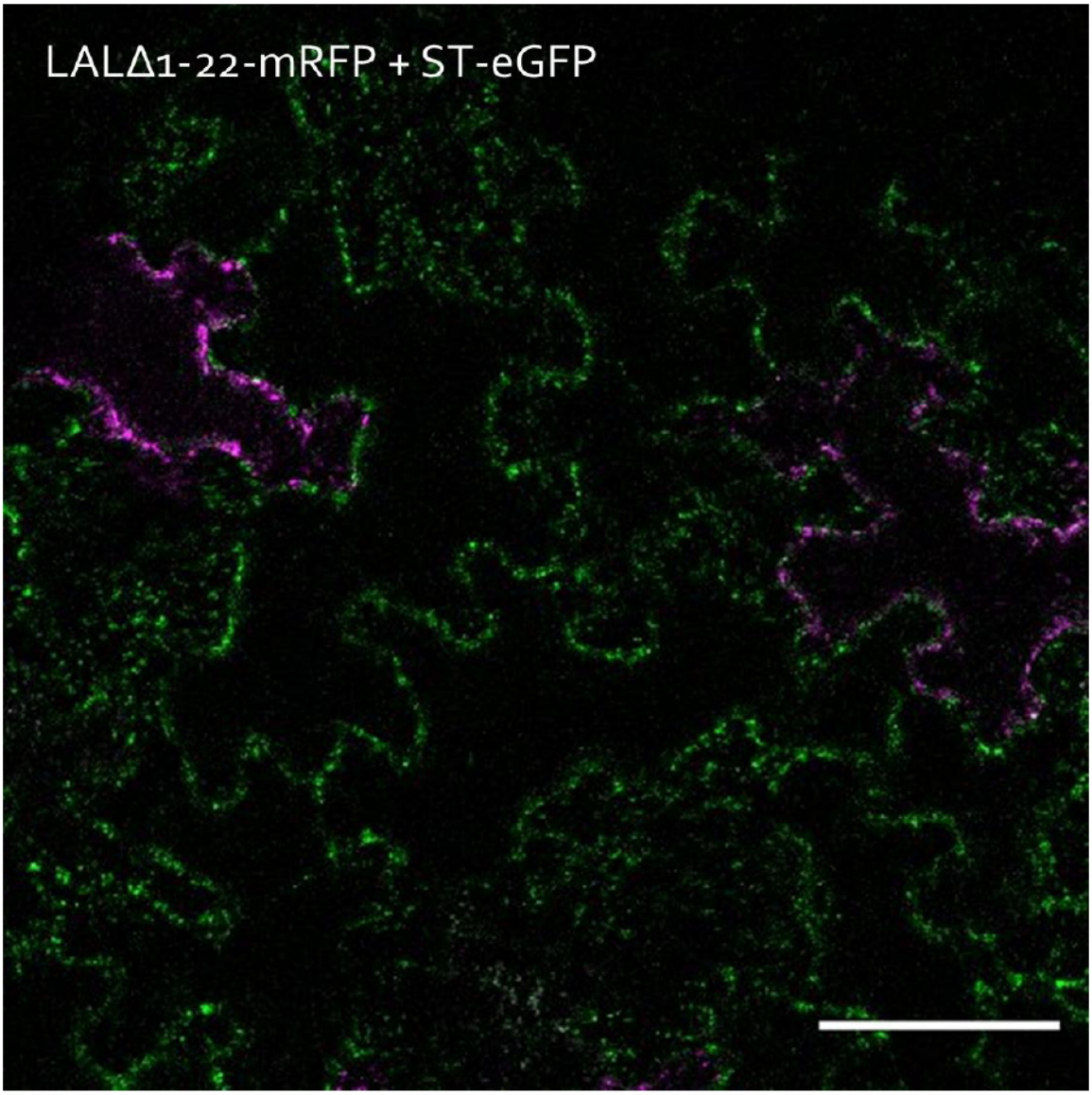


Figure 5.15: Representative image for leaf surface under 10x objective zoom. Merged image of cells expressing LALΔ1-22-mRFP (magenta) construct alongside ST-eGFP (green) trans-Golgi marker. Scale bar = 100 µm. n=3 with 10 technical replicas each. 568 cells expressing infiltrated constructs, 447 expressing marker only and 121 expressing LALΔ1-22, approximately 27% of cells expressing construct of interest.

The use of immunoglobulins E and G as test proteins relates to their diverse applications as a key class of recombinant biopharmaceuticals within research and medicine, but additionally their effectiveness as a reporter. *In planta* expression of both c-terminally fluorescently tagged immunoglobulins G and E (Figure 5.16, A and B respectively) revealed a similar localisation of expression, with Ig's localisation appearing to be associated with the plant endoplasmic reticulum. Literature on the glycosylation and production of Igs in tobacco indicates that Igs pass through the Golgi body and further to the secretory pathway (Göritzer et al., 2025), but progress slowly through the ER (Frigerio et al., 2000). As time of imaging is 3 days post infiltration for the Golgi markers, they may either still be progressing through the ER or like some immunoglobins heavy chains expressed in plants, remain in the ER (Irons et al., 2008).

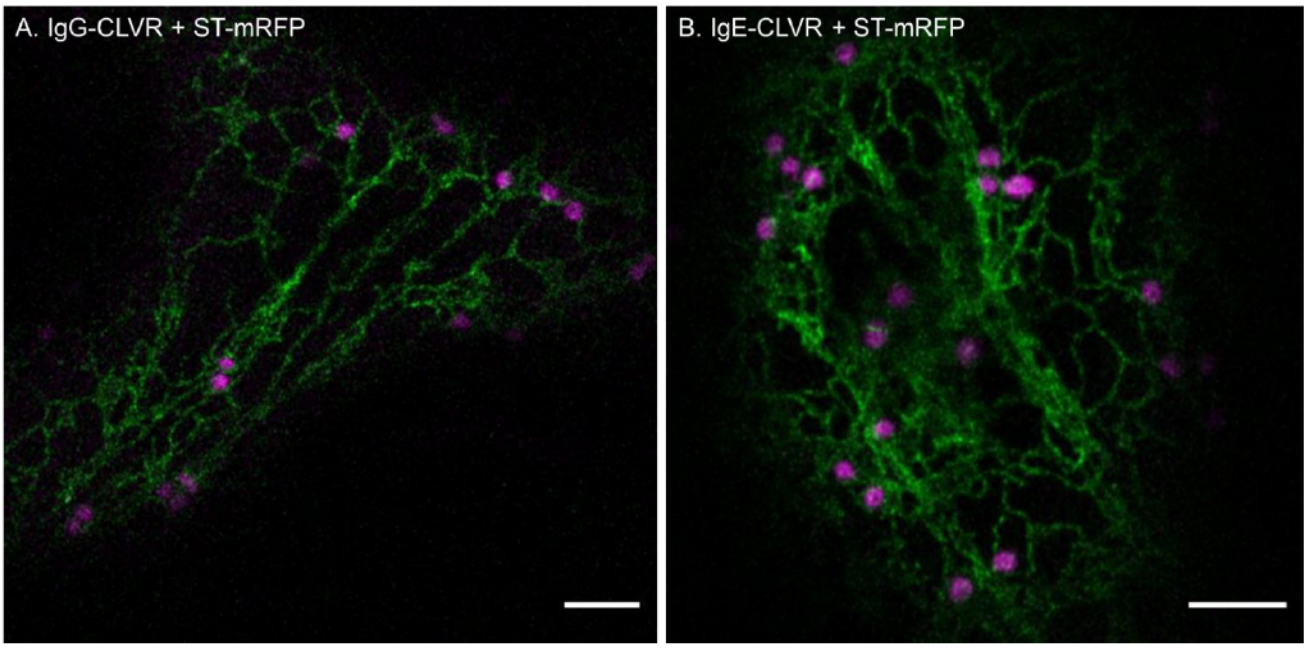


Figure 5.16: Co-expression of immuoglobins with the Golgi marker ST-eGFP. CLVR was added to the HC of the antibody. A. IgG-CLVR (green) is co-expressed with the *trans*-Golgi marker ST-mRFP (red). B. IgE-CLVR (green) is co-expressed with the *trans*-Golgi marker ST-mRFP (red). Displays a subcellular view of immuoglobins with ST-mRFP to indicate localisation distinct from the Golgi body, mostly ER associated. Size bars = 5 µm. Representative images are shown; n=3 with 5 technical replicas each.

5.5 Confirming glycosylation: lectin binding assays

RT-PCR performed indicated that the generation of stable lines had been successful, so far as presence of a transcript for each of the HuGEs was identified. This, however, does not confirm presence of a functional HuGE with N-glycosylation activity or successful production of specific glycomotifs. As such the final confirmation of successful implementation would be the generation of HuGE-specific glycan structures and their subsequent detection and quantification. These are defined as an increase in multi-antennary galatosylated N-glycans through the activity of MUR3-GNT4 and MUR3-GNT5 HuGEs, and the formation of α-2,6-Sialic acid by FUT13-B4GALT1 and FUT13-ST6GAL (Schoberer et al., 2018(b), a mammalian specific glycan structure not found in plants.

Identifying and quantifying glycans usually requires significant time and infrastructure, (Chia et al., 2023) both of which were constrained during this project. Lectin binding assays were selected as a suitable reporter for rapid detection of HuGE specific N-glycosylation within our facilities. Lectins binding assays were suggested to account for the variety of structured needed to identify, speed and simplicity of performing the assay, and their affinity for detection of highly specific glycan structures (Brooks et al., 2024). Lectins are naturally occurring proteins with the capability to bind specific sugars with high specificity, which make them an effective reporter of glycosylation when fused with a reporter such as a fluorescent probe (Mishra et al., 2019). There is long standing precedent with their use in detecting glycosylation changes in immunoglobins (Sumar et al., 1990). Two commercially available fluorescently labelled lectins were ultimately selected to perform analysis. Firstly, a *Ricinus Communis* agglutin 1 (RCA I) (Wu et al., 2006) to bind branched galactosylated glycans produced by the medial-Golgi MUR3-GNT45-388 insert. A similar dot blot protocol has previously been successfully applied to detect β1,4-galactosylation in glycomodified recombinant proteins produced in tobacco using a similar RCA lectin (Kogelmann et al., 2025). Secondly, a *Sambucus Nigra* (SN) lectin that specifically binds α-2,6-sialic acid, this would detect a specific glycan structure produced by the FUT13-B4ST6-389 insert. Assays were performed on mature leaf tissue of respective lines expressing each of the human proteins of interest LAL (LALΔ1-22-mRPF), IgG

(IgG-CLVR) and IgE (IgE-CLVR) via agroinfiltration. Tissue underwent a lysis protocol to release protein into solution and so generate extracts for testing.

The first protocol was a dot-blot grid on a nitrocellulose membrane for each extract in ascending volume, to present a range of signal intensities (Figure 5.17). The process was duplicated to generate two identical membranes, then processed separately for each respective lectin. The purposed of this test was to determine if any difference could be observed between the stable plant lines and controls, in theory glycomodified HuGE lines would non-specifically glycosylate plant proteins. Unfortunately, the results of this assay lacked distinction. No clear differences were seen in signal intensity between test and control values, and it was not possible even after several attempts to meaningfully interpret results due to a high background signal, likely from other cellular components of the plant tissue. As such an alternative methodology was explored.

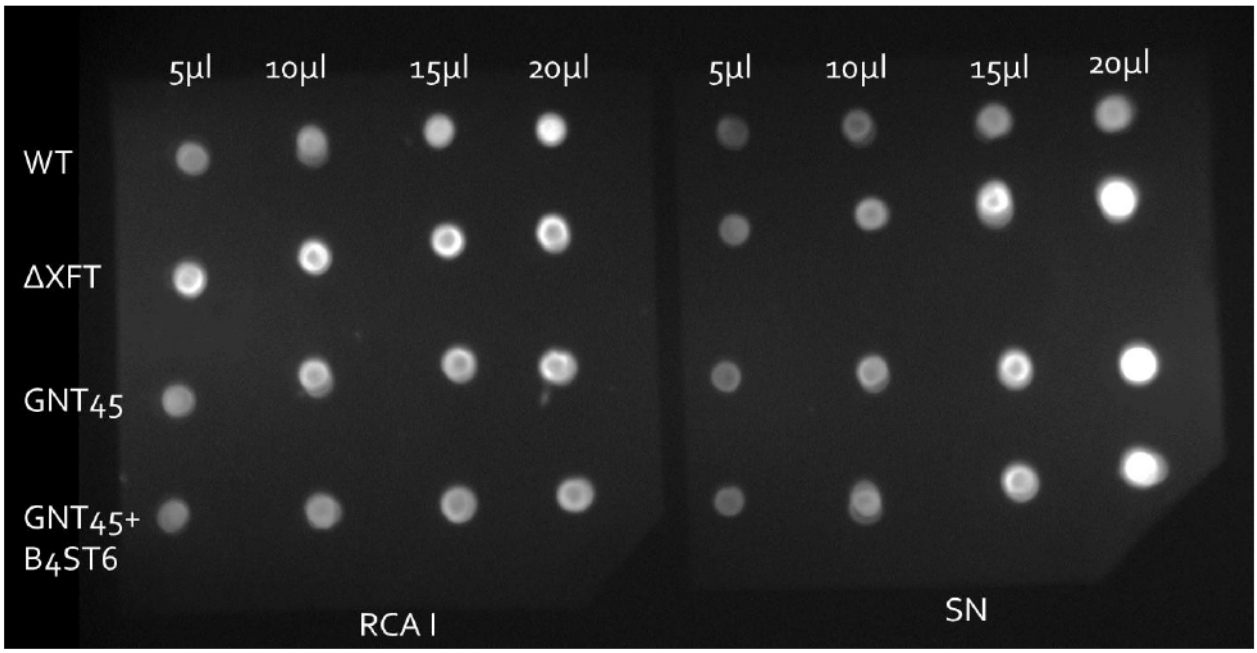


Figure 5.17: Example lectin binding assay conducted in duplicate for RCA I lectin (left) and SN lectin (Right). Columns indicate 5, 10, 15 and 20 µl of each plant tissue extract used for assay. Rows denote the source of each extract, WT and ΔXFT in theory as control for glycomodification. GNT45 represents stable plant line generated with MUR3-GNT45-388 vector. GNT45+B4ST6 represents stable plant line generated with both MUR3-GNT45-388 and FUT13-B4ST6-389 vectors.

After several attempts failed to produce a clean signal, a separate protocol involving sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) for extracts was pursued. This expressed human proteins of interest (LAL, IgE and IgG) in the modified HuGE and control lines, then performing extraction followed by SDS-PAGE to separate protein bands (Figure 5.18). The gel was transferred to a nitrocellulose membrane, underwent the same lectin binding protocol as previously. The purpose of this test was to separate proteins with glycan acceptor sites onto which modified glycans were likely to attach, in theory providing a clearer signal for lectin assays. As a first attempt to confirm equal loading, the SDS-PAGE was performed across all lines and proteins of interest, the resulting gel undergoing Coomassie blue staining (Figure 5.18).

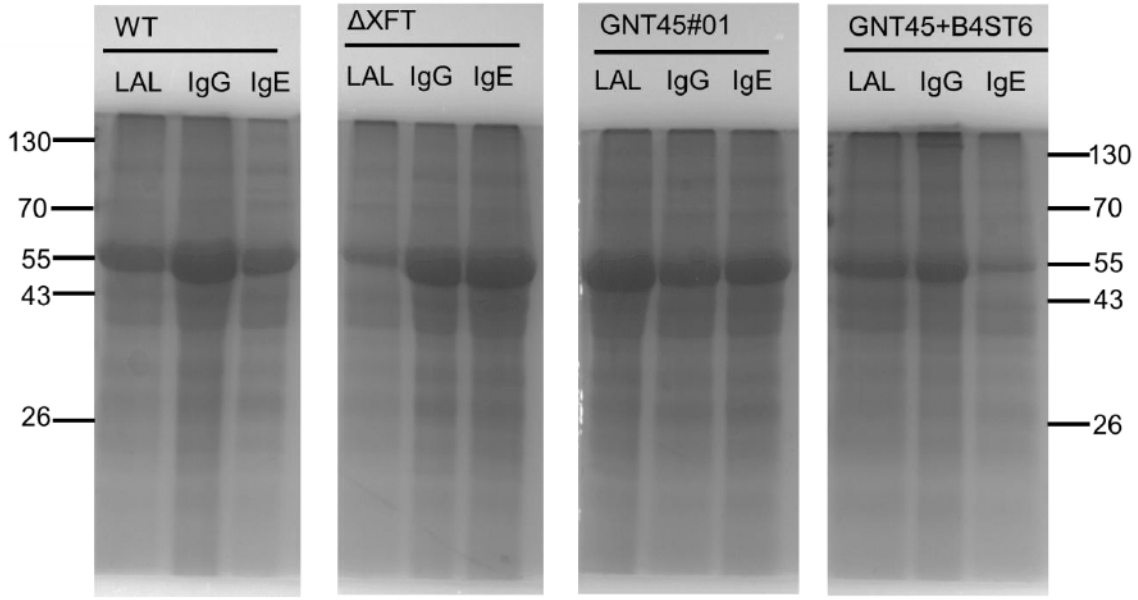


Figure 5.18: SDS-PAGE protein gel after Coomassie blue staining for mature Plant leaf extracts. Lines used for protein expression are wild type N.B (WT), ΔXFT N.B (ΔXFT), GNT45#01 N.B lines (GNT45) and GNT45+B4ST6#01 N.B lines (GNT45+B4ST6). Human proteins of interest expressed are LAL (~50 KDA), IgG (~55KDA) and IgE (~73 KDA). Molecular mass calculated based on uniport entries for Ig heavy chains, and LAL predicted based on protein sequence used. Demonstrates roughly equivalent loading for most samples. White arrows indicate ~70 KDA for approximate size of human POIs. Black lines indicate protein sizes in kDa from ladder. Samples were reduced prior to loading.

This protocol was attempted as previous (Figure 5.18) but with gel transferred onto a nitrocellulose membrane, and lectin binding protocol performed. The results failed to provide clear banding for individual POI but showed bright fluorescent streaking for specific samples (Figure 5.19). This may still indicate a positive result, as within these lines activity of the HuGEs would modify the N-glycosylation of all proteins, affecting endogenous plant proteins in addition to expressed POI, as such we would expect the entire sample to display a signal rather than individual bands therein. A signal indicating significant presence of galactosylation was observed in several lines. Less intense examples were observed in IgE expressed in both WT and ΔXFT lines, with IgG expressing in ΔXFT also present. A significantly more intense signal was detected from IgG expressed in GNT45+B4ST6. However, the lack of clarity of signal (streaking rather than proper banding) from each of these discounts their use in forming conclusions but suggests promise for this protocol's application in future.

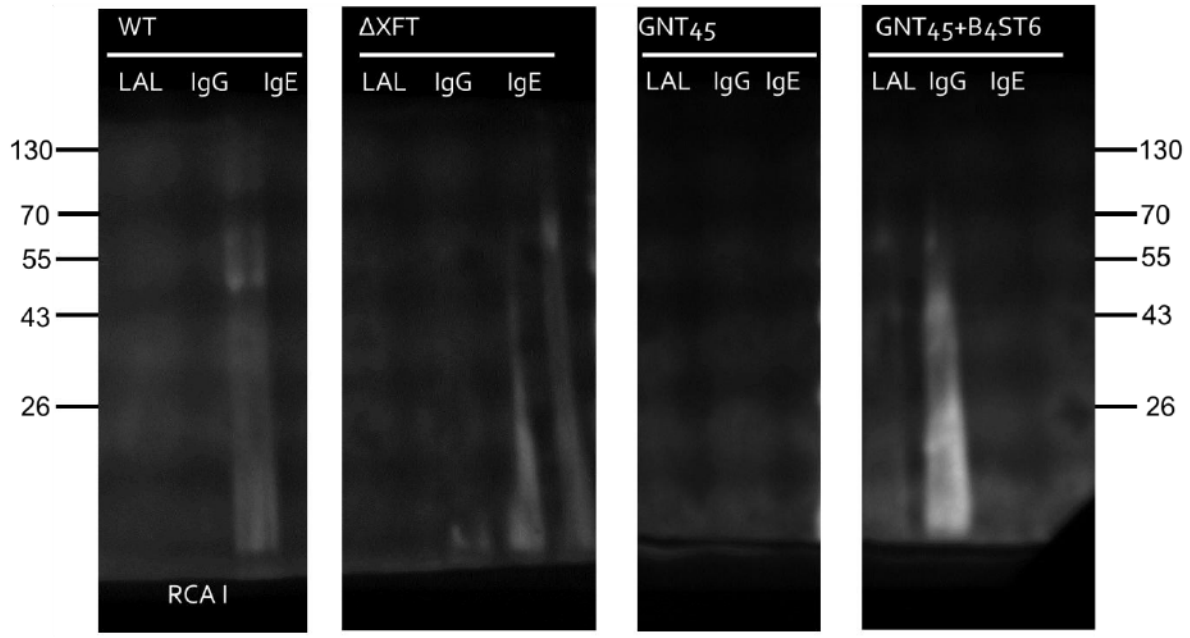


Figure 5.19: SDS-PAGE for RCA I lectin binding assay. Imaged post gel transfer to nitrocellulose membrane and lectin binding assay using RCA I lectin for mature plant leaf extracts. Lines used for protein expression are wild type N.B (WT), ΔXFT N.B (ΔXFT), GNT45#01 N.B lines (GNT45) and GNT45+B4ST6#01 N.B lines (GNT45+B4ST6). Human proteins of interest expressed are LAL (~50 KDA), IgG (~55KDA) and IgE (~73 KDA). Molecular mass calculated based on uniport entries for Ig heavy chains, and LAL predicted based on protein sequence used. Black lines indicate protein sizes in kDa from ladder. Lectin used is *Ricinus communis* agglutin 1 (RCA1, RCA120) Fluorescein 1 mg/ml. Samples were reduced prior to loading.

This protocol was performed in duplicate as previously, but in this instance with the *Sambucus nigra* lectin and visualized for Cy3 (Figure 5.20). Unfortunately, again there was no clear banding on the visualized membrane, however two streaks of significant intensity occurred for both LAL and IgG expressed in GNT45+B4ST6. This may indicate presence of α-2,6-Sialic acid and so a successful outcome for this project, but results lack the clarity from which to draw reliable conclusions. This approach holds promise for future work but was not successfully repeated due to time constraints under the approaching project conclusion.

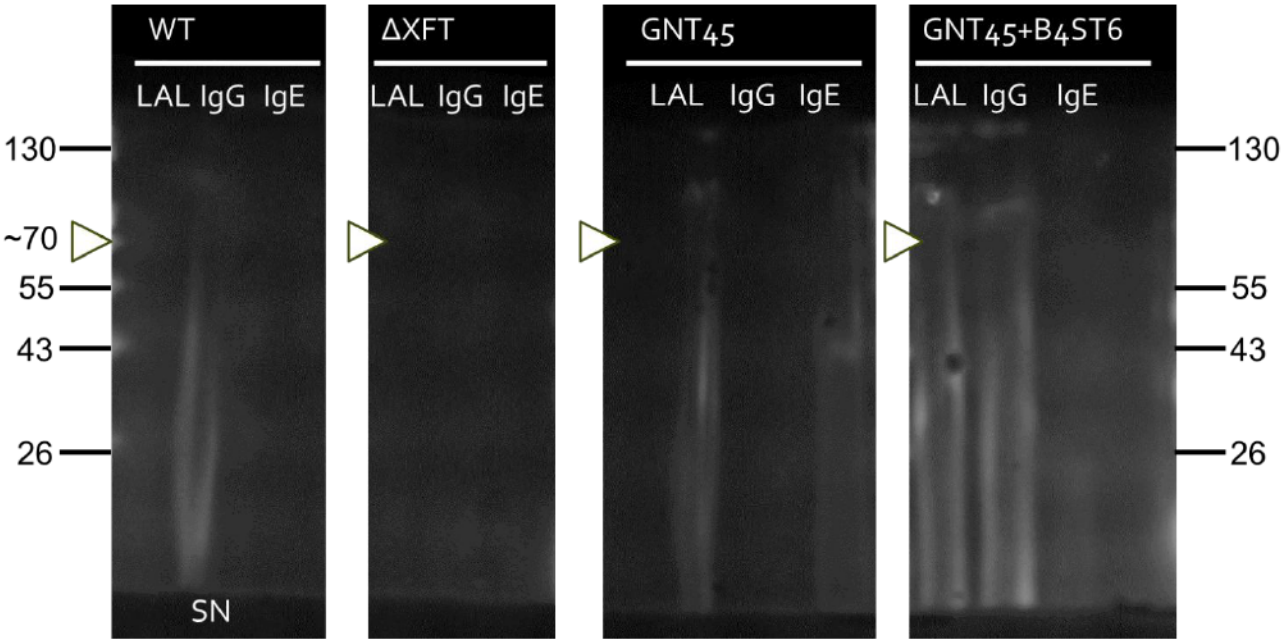


Figure 5.20: SDS-PAGE for SN lectin binding assay. Imaged post gel transfer to nitrocellulose membrane and lectin binding assay using SN lectin for mature plant leaf extracts. Lines used for protein expression are wild type N.B (WT), ΔXFT N.B (ΔXFT), GNT45#02 N.B lines (GNT45) and GNT45+B4ST6#01 N.B lines (GNT45+B4ST6). Human proteins of interest expressed are LAL, IgG and IgE. White arrows indicate ~70 KDA for approximate size of human POIs. Black lines indicate protein sizes in kDa from ladder. Lectin used is *Sambucus nigra* (SN) lectin Cy3 1 mg/ml

In summary, expression cassettes were designed and constructed to co-express HuGE pairs with the vectors MUR3-GNT45-388 and FUT13-B4ST6-389, with the intention to use these generate stably transformed tobacco lines. Two lines were successfully generated under selection, GNT45#01 expressing only MUR3-GNT45-388, and GNT45+B4ST6#01 expressing both 388 and 389 vectors. Presence of

transcript in these lines was confirmed by RT-PCR. The human proteins of interest IgG, IgE and LALΔ1-22 were then produced in these lines as well as in wild-type and ΔXFT tobacco to enable comparison of N-linked glycosylation structures generated. Lectin binding assays were then conducted using extract from transformed plants by both dot-blot and SDS-PAGE with lectin binding assays. These lectins were specific against glycan structures that would be generated by the activity of the 388 and 389 vectors. While the lectin binding assays conducted by SDS-PAGE against the GNT45#01 and GNT45+B4ST6#01 lines began to show promise, they did not produce conclusive data during the time limitations of this project.

5.6 Discussion

5.6.1 Engineering glycosylation in plants

Core fucose, xylose and Lewis-a bodies are all features of plant glycosylation and are conserved in most of the higher plants 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). 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. A drug production system with a lack of host organism endogenous contaminants, alongside improved general production parameters offered by plant systems, would be scientifically and commercially significant.

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, used in long term risk reduction in several applications. Use of the *Nicotiana benthamiana* ΔXFT line for the generation of stable plant lineages was greatly appreciated and of significant downstream interest for structural analysis of glycans produced. Acceptance of plant production systems has in the past been hampered by concerns on immunogenic hypersensitivity reactions to these glycan structures (Altmann et al.,

2007), in the context of therapeutic recombinant proteins (Kizhner et al., 2015). Approaches for the removal or diminished production of such glycan structures remains a priority in research concerning therapeutic protein production in plants (Kogelmann et al., 2023). Some 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).

Plasma half-life of human therapeutic proteins is essential in many drugs for optimal therapeutic potency, lack thereof would significantly diminish effectiveness of treatment (Bosch et al., 2013). Absence of complex mammalian type N-glycans in an anti-rabies monoclonal antibody produced in plant systems has reduced efficacy by diminished plasma half-life (Ko et al., 2003). Furthermore, the introduction or amplification of key mammalian glycans can significantly improve plasma half-life for therapeutic proteins (Chia et al., 2023) which has been shown in a modified erythropoietin with enhanced end terminal sialyation (Su et al., 2010). Existing therapeutics have shown enhanced potency in animal trials from the addition of mammalian N-linked glycoforms obtained through plant glycoengineering (Chen et al., 2016) as have cutting edge therapeutics used against SARs-CoV-2 (Castilho et al., 2021).

5.6.2 Investigating glycosylation and future work

Overall, glycoengineering in plants has been shown to be not only possible, but of increasing interest and application. Production of mammalian glycan motifs in plants can be achieved by the targeted heterologous expression of mammalian glycosylation enzymes (Strasser et al., 2014a; Strasser et al., 2014b). Earlier studies demonstrating that a mammalian glycosyltransferase can convert N-glycans in plants in a similar way as in human cells (Palpac et al., 1999; Bakker et al., 2001) paved the way for reconstructing mammalian glycosylation pathways in plant systems. The simultaneous overexpression of six mammalian glycosylation proteins in *N. benthamiana* resulted in the synthesis of the activated sialic acid and in planta protein sialylation (Castilho et

al., 2010). This project has been an interesting exploration into the glycoengineering of tobacco to produce human therapeutics, but not an exhaustive study, and limited by the given time frame. Unfortunately, this study ends before the detection and quantification of glycans produced in the stable lines engineered can be accomplished.

As many plant proteins are N-glycosylated, glycomodification can be assessed on plant proteins without tandem expressing of a human protein of interest as a prerequisite. Glycan structures produced by the MUR3-GNT45-388 insert (bi- and tri-antennary glycans) are mammalian specific and so otherwise would not be present on plant N-glycosylation sites, which would indicate effective expression and glycomodification of the insert. Similarly, galactose and terminal sialic acid produced by the FUT13-B4ST6-389 insert would otherwise not be found in plants and would indicate effective expression and activity. As some mammalian motifs require mammalian specific N-glycan sites (Kallolimath et al., 2016), use of human proteins of interest expressed in modified stable lines has provided a platform for validation of these.

Proteins of interest selected would serve as useful reporters of both production efficiency and glycomodification capacity. Presence and quantification assays for proteins of interest, such as Western blots or quantitative ELISA for protein abundance, could be used to ensure a detectable level of synthesis with the downstream goal of producing sufficient protein for activity assays. IgE and IgGs are useful as ELISA assays where binding affinity will be a metric of correct protein structure production, and the heavily glycosylated nature of immunoglobulins also serves as a site at which N-Glycan modification can be tested (Ree et al., 2000; Shade et al., 2020). More detailed glycosylation analysis would ideally have been conducted by mass spectrometry with supervision at BOKU, to offer a greater depth as to glycan composition on individual proteins rather than simple quantification, and further confirm correct activity of introduced glycans. Production and isolation of proteins of interest in these plant systems would ideally be isolated and undergo mass spectrometry previously performed for therapeutic proteins produced in plant systems (Göritzer et al 2025). This could determine the occupancy of N-glycosylation sites on the protein, the individual glycan structures in these and their proportion of total glycans present, as modification will likely produce hybrids of mammalian and plant

glycan structures, and will give a clearer impression of the underlying enzyme activity (Castilho et al., 2021)

For LAL, a combination of embedded enzyme assays to measure objective catalytic activity on cholesterol esters, and structural analysis via mass spectrometry (Vézina et al., 2009) could be conducted. However, animal studies would likely be needed to certify effectiveness in vivo. 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.

Chapter 6: Discussion

Contents

6.1 Imaging the plant Golgi body

The first phase of this project was centered on methodology to systematically visualize and quantify localisation of fluorescent protein fusions in the plant Golgi, by super resolution airyscan light microscopy. This was to assess selective localisation of HuGEs in a cisternae-specific manner, ultimately considered successful. Existing research with diverse analysis methodology was fundamental in establishing the methods and their success used in this thesis. Early work on localisation of resident glycosylation enzymes and their distribution by immune-electron microscopy in HeLa cells (Rabouille et al., 1995) fundamentally established sub-Golgi localisation of these enzymes, including ST to the *trans*-Golgi. When these findings were translated into plant studies, they established use of critical markers (ST) and understanding utilized in this work (Boevink et al., 1998). More recent research on the specific ultrastructure of the plant Golgi body (Robinson et al., 2020) and the formation of the Golgi and cisternae (Ito et al., 2014) was indispensable for fundamental understanding of the maturation, structure and transport of the plant Golgi body, and its many roles in protein and lipid modification, intercellular trafficking and key polysaccharide synthesis.

Existing research into subcellular localisation by Bolte et al. (2006) informed the decisions and techniques used in this project, especially parameters used when collecting Golgi data. Sub-organellar localisation has significant downstream consequences on resulting glycans, and imaging a small highly dynamic anisotropic structure such as the plant Golgi makes co-localisation challenging. The 3D orientation and physical distribution of the plant Golgi is at the working limits of the optical microscopy resolution we possess, meaning any reliable analysis would have to operate consistently within tightly constrained definitions of co-localisation/separation while also being rapid enough to enable sufficient repeats, in order to establish a consistent mean among variation between images. Analysis relied on the quality and consistency of the markers used to identify Golgi cisternae alongside the defined dimensions of image-acquisition within the constraints of the optical system (Bolte et al., 2006). The time taken to establish a reliable data analysis pipeline and marker

integrity, through use of known markers ST and MNS1, provided an initial background to establishing a baseline. But further the MNS1-ST combination was continually used in individual sessions alongside HUGE marker combinations to establish consistency in the use of these markers in terms of localisation, and their consistent separation determined by the analysis pipeline used to check the system's ability to perform colocalization measurements. This used alongside fluorescent microbeads controls used at the onset of sessions (Figure 3.18) ensured consistency of imaging parameters and identified errors, especially in deconvolution presenting the greatest issue in super resolution work (Wallace & Swedlow, 2001). The use of self-cleaving peptide marker constructs for MNS1-ST and MUR3-FUT13 HuGE marker combinations was in part inspired by studies in ratiometric imaging of plant membranes conducted by Samalova et al (2006), and expertise within our group on the use of 2A self-cleaving peptides in tandem with of self-excising intein domains for use of multiple subunit self-cleaving constructs for assembly of complex polyproteins (Rossi et al., 2023) in addition to their diverse uses including separately localising fluorescent markers (Rossi et al., 2024).

While the fixed stoichiometry of self-cleaving peptides is helpful in ratiometric imaging for studying proportional relationship between proteins as fluorescent protein fusion, it was not as directly helpful for this application due to the less-critical need for equivalent fluorescence intensity. Instead, this provided a proof of concept for using self-cleaving markers for localisation studies with different sub-organellar targets from a single insert and proved efficacy of this approach. This was useful in establishing guaranteed expression of two marker constructs within the same cell, which was observed as otherwise infrequent, and facilitated three-colour imaging of modified HuGEs with two marker constructs from minimal insert numbers.

6.2 Effect of sub-Golgi localisation on glycan production

The second phase of this project was implementation of methodology to determine the sub-Golgi localisation of existing HuGES, and that can be used to direct localisation to quantify and deliver cisternae-specific localisation of modified HuGES. Investigation of CTS-directed localisation and their potential for use as both marker and enzyme localisation strategy began with excellent work directed by Schoberer et al. (2011, 2013) identifying plant resident glycosylation enzyme CTS domains and establishing the use of an α-mannosidase I (MNS1) as basis for FP fusion to use as a *cis*-Golgi body marker (Schoberer et al., 2013). Alongside the same rationale for ST's use as a *trans*-Golgi marker (Boevink et al., 1998) this set precedent for development of the MUR3/FUT13 markers and ultimately their use as a localisation signal for larger constructs, being the modified HuGEs and 388/389 stable plant transformation cassettes. The CTS of ST from *Rattus norvegicus* has previously been utilized for the selective localisation of β1,4-galactosyltransferase (Strasser et al., 2009) to the plant (*N. benthamiana*) *trans*-Golgi and despite mammalian origin, indicated successful localisation through the production of di-galactosylated N-glycan structures. This hints at some fundamental conservation of underlying structural mechanism (Schoberer et al., 2010) that was not maintained in this research with localisation of full length human ST6GAL. However, while expression of targeted glycosylation enzymes generally does not have associated growth defects, overexpression of an ST-GalT trans-Golgi targeted construct showed stunted growth and developmental retardation (Schneider et al., 2015).

Precedent in the literature for insertion of HuGEs focuses primarily on the introduction of α-2,6-sialic acid into the glycan chain, this being a major motif of interest for the production of human therapeutics, and an early study observed activity of human α-2,6-sialyltransferase with trans-Golgi body localisation, demonstrating that glycosyltransferases can be subcompartmentalized to specific cisternae of the plant

Golgi apparatus (Wee et al., 1998). Literature also provides precedent for effective *in planta* sialylation of a therapeutic protein achieved using a recombinant mAb as acceptor substrate observed after achieving proper subcellular localisation of an ST-GalT construct (Castilho et al., 2010) whereby all available acceptor substrates were sialylated, as would be the intended outcome of this system. Some incompletely processed structures remained, but these were attributed to the overexpression of ST-GalT interfering with endogenous glycan processing (Castilho et al., 2010).

Even with modern predictive tools, existing research suggests experimental testing is necessary to confirm CTS-directed localisation (Loos et al., 2014). While not fully understood, the basis of the CTS mechanism seems to relate to the underlying nature as resident glycosylation enzymes as type II membrane proteins (Schoberer et al., 2010), but with potentially separate functional roles for the individual components of the CTS. While the structure operates as a single pass signal-anchor sequence that provides lumenal orientation of the enzyme (Schoberer et al., 2013) the cytoplasmic domain is proposed in existing research to provide modified ER export and orientation of the enzyme catalytic subunit (Schoberer et al., 2009). The length, and with that the size of the transmembrane is thought to relate to varying membrane thickness of Golgi body cisternae (Pagny et al., 2003). With these factors in mind, manipulation of the CTS can be an effective tool in directing sub-Golgi localisation, but only if properly understood (Loos et al., 2014) and offer explanations to the distribution of the enzymatic environment in the plant Golgi body.

This project was concerned with achieving the cisternae-specific localisation of exogenous HuGEs into the late Golgi for the introduction of mammalian specific glycans (Schoberer et al, 2018b). The specific sub-Golgi localisation of enzymes impacts on the final glycan structures produced. Targeting of glycosylation enzymes to specific relevant sub-Golgi cisternae was indicated in literature to improve enzyme activity and to generate a greater yield of specific glycan structures; for example, GNTIV and GNTV produce a greater yield of bi- and tri- antennary glycans when expressed in the medial-Golgi than when expressed in the *trans*-Golgi (Bosch et al., 2013; Castilho et al., 2011). These are essential as a prerequisite to produce complex human-type glycans, and a variety of localisation strategies have previously been

successfully applied (Nagels et al., 2011), indicating that this was a reasonable methodology for optimising glycan yield. However, while the previously mentioned use of an ST-β1,4-galactosyltransferase (Strasser et al., 2009) construct successfully produced of di-galactosylated N-glycan structures due to *trans*-Golgi directed localisation, the same approach using a fusion of XylT CTS domain alongside β1,4-galactosyltransferase failed to produce any di-galactosylated structures. It produced only mono-galactosylated or hybrid, indicating that the earlier medial-Golgi localisation prevented or limited formation of di-galactosylated N-glycan structures, and that the degree to which localisation affected yield included absence of intended structures (Strasser et al., 2009). *Cis*-Golgi targeting of a GalT fusion similarly failed to produce significant galactosylation (Vézina et al., 2009). Earlier research in both tobacco BY2 cells and mature plants (Palacpac et al., 1999; Bakker et al., 2001) expressing β1,4-galactosyltransferase indicated that sub-Golgi localisation of the insert similarly dictated resulting glycan structures, but the degree of galactosylation differed significantly between growth forms (Loos et al., 2014) again indicating the complexity and need for experimental testing for modifying glycosylation. While the choice to use plant derived CTS domains MUR3 (medial-Golgi) and FUT13 (*trans*-Golgi) (Schoberer et al., 2011) in this project may be less tested in modifying HuGE localisation, it was to understand a further level of optimisation by using native plant localisation signals. Specificity of the localisation observed in our results also circumvents possibility that random introduction of HuGEs could interfere with endogenous glycosylation and lead to generation of undesirable or incomplete N-glycan structures (Chen et al., 2016).

Conversely, specific localisation of exogenous mammalian glycosylation enzymes can reduce activity of endogenous plant glycosylation (Frey et al., 2008). In this instance rat GNTIII either as the native sequence or affixed with an *Arabidopsis thaliana* mannosidase II CTS domain increased the fraction of *N*-glycans devoid of β-1,2-xylose and α-1,3-fucose (Frey et al., 2008), which could be advantageous in the production of therapeutics. Similarly, literature indicates that human β1,4-galactosyltransferase 1 shows targeting in the plant Golgi that interferes with biantennary complex N-glycan formation, leading to increased N-glycan heterogeneity (Bakker et al., 2001; Bakker et al. 2006; Strasser et al. 2009; Strasser et al., 2016). As this project similarly directs localisation of exogenous HuGEs into the medial- and *trans*-Golgi, the sites for β-1,2-

xylose and α-1,3-fucose attachment, it is possible these could similarly lead to reduction in formation of these glycans.

6.3 Immunogenicity of plant produced glycans

Attachment of β-1,2-xylose and core α-1,3 fucose onto complex N-glycans in addition to formation of Lewis-a bodies are all common features of endogenous plant glycosylation in species currently used to produce therapeutics (Gomord et al., 2004; Strasser et al., 2014). These motifs have raised concerns for their potential immunogenic profile when applied to use for human therapeutics (Bardor et al., 2003; Koppelman et al., 2000; Ree et al., 2000) being the structural basis of cross-reactive carbohydrate determinants (Altmann et al., 2007). High risk applications such as enzyme replacement therapy, could significantly limit the long-term efficacy of treatment with increased risk of hypersensitivity reactions. β-1,2-xylose and core α-1,3 fucose plant-glycans can induce rapid clearance from circulation and in addition to allergic reactions because of the presence of IgE antibodies directed against these epitopes (Gomord et al., 2010) through a poorly understood mechanism (Stelter et al., 2019) further limiting the scope of any therapeutic proteins produced. Lewis-a epitopes are biantennary complex-type N-glycans that are nearly ubiquitous in plants, and N-glycans containing them are particularly abundant in tobacco plants, and are known to be highly immunogenic in mammals (Fitchette et al., 1999).In addition to direct risk associated with Lewis-a glycan motif**s**, a debated source of immunoreactivity(Ree et al., 2000), these potentially interfere with β1,4-galactosylation (Strasser et al., 2008; Göritzer et al., 2021) diminishing both potential safety and intended glycomodification of expressed proteins.

Immunogenic N-glycans are not specific to plant production systems, however. Yeasts add highly immunogenic terminal N-glycan mannosylation to recombinant proteins, and mammalian expression systems including Chinese hamster ovary cell lines are known produce several highly immunogenic glycan structures, including N-glycosylneuraminic acid sialic acid isoform and terminal α-1,3-galactose that may elicit immune responses (Gomord et al., 2004). While this means plant systems are not in a unique position, these glycans are still a smaller barrier to a low immunogenic profile

production system than other popular host organisms. This is the significant reason the generation of the Strasser et al (2008) ΔXFT RNAi *Nicotiana benthamiana* line and its use in this project, in part to work towards creation of a such a system. Similar approaches have been developed in α1,3-fucosyltransferase (Fuc-T) and β1,2-xylosyltransferase (Xyl-T) in *Physcomitrella patens* for α1,3-fucosyltransferase) and β1,2-xylosyltransferase single or double knockout mutant lines (Koprivova et al., 2004). In the context of this project it would be interesting to see if HuGE insertions in the WT N.B lines are in any way comparable to the ΔXFT lines generated due to the downregulation of these problematic glycoepitopes, resulting from interference as discussed previously (Frey et al., 2008), and from attachment of Lewis-a bodies chiefly occurring in the *trans*-Golgi (Fitchette et al., 1999; Strasser et al., 2007). Therefore, a drug production system with a lack of endogenous contaminants from the host organism, alongside improved general production parameters offered by plant systems, would be scientifically and commercially significant.

6.4 Producing human POIs in plants

Recent history has seen plant-based platforms to produce 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 (Rup et al., 2017; Shaaltiel et al., 2007; Montero-Morales et al., 2018). The diversity of applications and administrations can make use of otherwise less-desirable properties of plants, such as utilizing endogenous plant glycans to stimulate greater immune response for immunisation delivered by virus-like particles (Ward et al., 2014).

Use of Agrobacterium-mediated transformation of *Nicotiana benthamiana* 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 like mammalian cells being an essential requirement for the purposes of this project. Complete mammalian glycan profiled IgG can be in part produced in glycoengineered *N. benthamiana* expressing human B4GALT1 in the late Golgi (Göritzer et al., 2021), as these are diverse in application and IgEs are the most heavily glycosylated antibodies, these were of great interest as both a reporter of glycan synthesis and receptor binding affinity in this project. Additionally, plant-produced mAbs also have a degree of glycan homogeneity that cannot be produced by mammalian cells (Chen et al, 2016).

Use of well-documented crop species as a vector is useful to produce 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 benthamiana* in various forms allows the possibility to investigate HuGE localisation and production of recombinant proteins from transiently transformed plants and also in stably transformed plants, including the possibility of manipulating this system in future to pursue better industrially scalable production methods (Xu et al., 2012; Kogelmann et al., 2025; Shaaltiel et al., 2007). Producing human therapeutics requires a final product that is capable of meeting necessary demands: sufficient quantity, with correct structure, activity, and pure as possible. 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 and contaminants are concerned. Many current enzyme replacement therapies utilized have recorded cases of allergic

reactions to contaminants or immunogenic structures which prevent patients from receiving consistent effective treatment options (Zimran et al., 2011).

Plasma half-life of human therapeutic proteins is essential in many drugs for optimal therapeutic potency, lack thereof would significantly diminish effectiveness of treatment (Bosch et al., 2013). The sialic acids on N-glycans prevent rapid clearance of therapeutics from the bloodstream by the Gal-specific hepatic asialoglycoprotein receptor (Stockert, 1995). Absence of complex mammalian type N-glycans in an anti-rabies monoclonal antibody produced in plant systems has reduced efficacy by diminished plasma half-life (Ko et al., 2003). Furthermore, the introduction or amplification of key mammalian glycans can significantly improve plasma half-life for therapeutic proteins (Chia et al., 2023) which has been shown in a modified erythropoietin with enhanced end terminal sialyation (Castilho et al., 2013; Su et al., 2010). Existing therapeutics have shown enhanced potency in animal trials from the addition of mammalian N-linked glycoforms obtained through plant glycoengineering (Chen et al., 2016) as have cutting edge therapeutics used against SARs-CoV-2 (Castilho et al., 2021).

6.5 LAL as a plant-produced ERT

LAL is a protein of interest both for the application and as a reporter of the system's efficacy in protein production and glycosylation. LAL is a lysosomal glycoprotein that specifically requires presence of N-Glycans including Mannose-6-phosphate and N-acetylglucosamine to allow targeted delivery of LAL to the lysosomal compartment, and therefore requires proper glycosylation to be functionally active (Su et al., 2016). This is an essential feature of the current LAL ERT Sebelipase Alfa© for LAL-D (Balwani et al., 2014), and any recombinant LAL would need to replicate this effect 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).

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. 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 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 intravenous ERT. 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, and highly desired further N-linked glycans bearing terminal a2,6-sialic acid. Activity of the characteristic mannose-targeting mechanism in a plant produced LAL has been successfully demonstrated in in tobacco (Du et al., 2008), however this contained plant-specific xylose and fucose residues, and lacked mammalian end-terminal sialic acid. Mature human LAL has six putative N-glycosylation sites, Asn-15, Asn-51, Asn-80, Asn-140, Asn-252, and Asn-300 (Zschenker et al., 2005, Rajamohan et al., 2020). These readily accept glycans and their structure will largely depend on the enzymatic environment present, e.g. all sites

can be saturated with high-mannose oligosaccharides when produced in an insect host (Rajamohan et al., 2020), but in mammalian cells will be glycosylated with sialic acid structures (Du et al., 2005). The intent of the production system explored in this project would be to generate a plant produced LAL with all six glycosylation sites completely occupied by end-terminal a2,6-sialic acid structures, a functioning M6P for directed localisation and without potentially problematic plant-specific xylose and fucose residues.

6.6 Future work

Future work for this project would be continued production of human proteins of interest in glycomodified lines, with development of a successful lectin binding methodology for detection of mammalian glycomodification. Ideally mass spectrometry could be used after this point to assess presence and abundance of glycan structures on candidate proteins, with further assays to assess the binding affinities of lgs and enzyme activity of LAL (Ohira et al., 2022). Effective production of terminal sialylated proteins in this system may 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, possibly through use of further 2A self-cleaving peptide approaches.

Current research into plant production systems have taken the basic concepts of this project to a new level. Industry ready lines *Nicotiana benthamiana* capable of mammalian type glycoprotein production with negligible immunogenic profile through endogenous glycan depletion already exist (Kogelmann et al., 2025). Engineering for production of therapeutic proteins in plants now looks beyond the plant Golgi into modifying the dimensions of the ER to produce larger and more complicated proteins or enhance quality and yield of well know therapeutic proteins (Göritzer et al., 2025). Significant advances have been made in improving yield and protein quality by preventing proteolytic degradation of recombinant proteins (Hehle et al., 2015). Plants

are now used to produce cutting edge therapeutics that include mammalian N-glycosylated SARs-CoV-2 drugs (Castilho et al., 2021) rather than bulk simple proteins, and use of CRISPR/Cas9 technology has yielded plant produced therapeutics on par with those produced in mammalian CHO systems (Jansing et al., 2019). Downstream extraction of recombinant protein remains a major challenge, with many strategies difficult to transfer from lab scale to commercial setting (Buyel et al., 2015). In short, production of therapeutic proteins in plant systems is a popular domain of rapid enhancement that increasingly promises to outpace traditional approaches to generating recombinant proteins with mammalian post-translational modifications.

7. References

Aguisanda, F., Thorne, N., & Zheng, W. (2017). Targeting Wolman Disease and Cholesteryl Ester Storage Disease: Disease Pathogenesis and Development. *Current Chemical Genomics and Translational Medicine*, *11*(1), Article 1.

Altmann, F. (2007). The role of protein glycosylation in allergy. International Archives of Allergy and Immunology, 142(2), Article 2.

Apweiler, R., Hermjakob, H., & Sharon, N. (1999). On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochimica et biophysica acta*, *1473*(1), 4–8.

Bakker, H., Bardor, M., Molthoff, J. W., Gomord, V., Elbers, I., Stevens, L. H., Jordi, W., Lommen, A., Faye, L., Lerouge, P., & Bosch, D. (2001) Galactose-extended glycans of antibodies produced by transgenic plants. Proceedings of the National Academy of Sciences of the United States of America, 98(5), 2899–2904.

Bakker, H., Rouwendal, G. J. A., Karnoup, A. S., Florack, D. E. A., Stoopen, G. M., Helsper, J. P. F. G., van Ree, R., van Die, I., & Bosch, D. (2006). An antibody produced in tobacco expressing a hybrid beta-1,4-galactosyltransferase is essentially devoid of plant carbohydrate epitopes. Proceedings of the National Academy of Sciences of the United States of America, 103(20), 7577–7582.

Balwani, M., Breen, C., Enns, G. M., Deegan, P. B., Honzík, T., Jones, S., Kane, J. P., Malinova, V., Sharma, R., Stock, E. O., Valayannopoulos, V., Wraith, J. E., Burg, J., Eckert, S., Schneider, E., & Quinn, A. G. (2013). Clinical effect and safety profile of recombinant human lysosomal acid lipase in patients with cholesteryl ester storage disease. Hepatology (Baltimore, Md.), 58(3), 950–957.

Bardor, M., Faveeuw, C., Fitchette, A. C., Gilbert, D., Galas, L., Trottein, F., Faye, L., & Lerouge, P. (2003). Immunoreactivity in mammals of two typical plant glyco-epitopes, core α(1,3)-fucose and core xylose. Glycobiology, 13(6), Article 6.

Beihammer, G., König-Beihammer, J., Kogelmann, B., Ruocco, V., Grünwald-Gruber, C., D'Aoust, M.-A., Lavoie, P.-O., Saxena, P., Gach, J. S., Steinkellner, H., & Strasser, R. (2023). An oligosaccharyltransferase from Leishmania donovani

increases the N-glycan occupancy on plant-produced IgG1. Frontiers in Plant Science, 14, 1233666.

Besler, K. J., Blanchard, V., & Francis, G. A. (2022). Lysosomal acid lipase deficiency: A rare inherited dyslipidemia but potential ubiquitous factor in the development of atherosclerosis and fatty liver disease. Frontiers in genetics, 13, 1013266.

Boevink P., Oparka K., Cruz SS., Martin B., Betteridge A., Hawes C., (1998). Stacks on tracks: the plant Golgi apparatus traffics on an actin/ER network. The Plant Journal.;15(3):441-447.

Bohlender, L. L., Parsons, J., Hoernstein, S. N. W., Rempfer, C., Ruiz-Molina, N., Lorenz, T., Rodríguez Jahnke, F., Figl, R., Fode, B., Altmann, F., Reski, R., & Decker, E. L. (2020). Stable Protein Sialylation in Physcomitrella. Frontiers in Plant Science,

Bolte S., Cordelieres FP. (2006). A guided tour into subcellular colocalization analysis in light microscopy. Journal of Microscopy.;224(3):213-232.

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

Brooks S. (2024) Lectins as versatile tools to explore cellular glycosylation. Eur J Histochem. Jan 29;68(1):3959.

Burnett, M. J. B., & Burnett, A. C. (2020). Therapeutic recombinant protein production in plants: Challenges and opportunities. PLANTS, PEOPLE, PLANET, 2(2), 121–132.

Burnett, M. J. B., & Burnett, A. C. (2020). Therapeutic recombinant protein production in plants: Challenges and opportunities. PLANTS, PEOPLE, PLANET, 2(2), 121–132.

Buyel, J. F., Stöger, E., & Bortesi, L. (2021). Targeted genome editing of plants and plant cells for biomanufacturing. Transgenic Research, 30(4), 401–426.

Buyel, J. F., Twyman, R. M., & Fischer, R. (2015). Extraction and downstream processing of plant-derived recombinant proteins. Biotechnology Advances, 33(6), Article 6.

Castilho, A., Beihammer, G., Pfeiffer, C., Göritzer, K., Montero-Morales, L., Vavra, U., Maresch, D., Grünwald-Gruber, C., Altmann, F., Steinkellner, H., & Strasser, R. (2018). An oligosaccharyltransferase from Leishmania major increases the N-glycan occupancy on recombinant glycoproteins produced in Nicotiana benthamiana. Plant biotechnology journal, 16(10), 1700–1709.

Castilho, A., Bohorova, N., Grass, J., Bohorov, O., Zeitlin, L., Whaley, K., Altmann, F., & Steinkellner, H. (2011). Rapid High Yield Production of Different Glycoforms of Ebola Virus Monoclonal Antibody. PLOS ONE, 6(10), Article 10.

Castilho, A., Gattinger, P., Grass, J., Jez, J., Pabst, M., Altmann, F., Gorfer, M., Strasser, R., & Steinkellner, H. (2011). N-Glycosylation engineering of plants for the biosynthesis of glycoproteins with bisected and branched complex N-glycans. *Glycobiology*, *21*(6), Article 6.

Castilho, A., Neumann, L., Gattinger, P., Strasser, R., Vorauer-Uhl, K., Sterovsky, T., Altmann, F., & Steinkellner, H. (2013). Generation of biologically active multi-sialylated recombinant human EPOFc in plants. PloS one, 8(1), e54836.

Castilho, A., Schwestka, J., Kienzl, N. F., Vavra, U., Grünwald-Gruber, C., Izadi, S., Hiremath, C., Niederhöfer, J., Laurent, E., Monteil, V., Mirazimi, A., Wirnsberger, G., Stadlmann, J., Stöger, E., Mach, L., & Strasser, R. (2021). Generation of enzymatically competent SARS-CoV-2 decoy receptor ACE2-Fc in glycoengineered Nicotiana benthamiana. *Biotechnology journal*, *16*(6), e2000566.

Castilho, A., Strasser, R., Stadlmann, J., Grass, J., Jez, J., Gattinger, P., Kunert, R., Quendler, H., Pabst, M., Leonard, R., Altmann, F., & Steinkellner, H. (2010). In Planta protein sialylation through overexpression of the respective mammalian pathway. *Journal of Biological Chemistry*, *285*(21), Article 21.

Chen Q., Davis KR., The potential of plants as a system for the development and production of human biologics. F1000, 2016 19;5: F1000 Faculty Rev-912.

Chia, S., Tay, S. J., Song, Z., Yang, Y., Walsh, I., & Pang, K. T. (2023). Enhancing pharmacokinetic and pharmacodynamic properties of recombinant therapeutic proteins by manipulation of sialic acid content. Biomedicine & Pharmacotherapy, 163, 114757.

Costes SV., Daelemans D., Cho EH., Dobbin Z., Pavlakis G., Lockett S., Automatic and quantitative measurement of protein-protein colocalization in live cells. Biophysics Journal. 2004;86(6):3993-4003.

Denecke, J., Aniento, F., Frigerio, L., Hawes, C., Hwang, I., Mathur, J., Neuhaus, J.-M., & Robinson, D. G., Secretory pathway research: The more experimental systems the better. The Plant Cell, 2012, 24(4), 1316–1326.

Derbyshire, V., Wood, D. W., Wu, W., Dansereau, J. T., Dalgaard, J. Z., & Belfort, M., Genetic definition of a protein-splicing domain: Functional mini-inteins support structure predictions and a model for intein evolution. Proceedings of the National Academy of Sciences, 1997, 94(21), 11466–11471.

Dicker, M., & Strasser, R. (2015). Using glyco-engineering to produce therapeutic proteins. *Expert opinion on biological therapy*, *15*(10), 1501–1516.

Donnelly, MLL., Luke, G., Mehrotra, A., Li, X., Hughes, LE., Gani, D., & Ryan, M. D. Analysis of the aphthovirus 2A/2B polyprotein “cleavage” mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal “skip”. J Gen Virol. 2001;82(Pt 5):1013-1025.

Du, H., Levine, M., Ganesa, C., Witte, D. P., Cole, E. S., & Grabowski, G. A. (2005). The role of mannosylated enzyme and the mannose receptor in enzyme replacement therapy. American journal of human genetics, 77(6), 1061–1074.

Du, H., Cameron, T. L., Garger, S. J., Pogue, G. P., Hamm, L. A., White, E., Hanley, K. M., & Grabowski, G. A. (2008). Wolman disease/cholesteryl ester storage disease: efficacy of plant-produced human lysosomal acid lipase in mice. Journal of lipid research, 49(8), 1646–1657.

Dünser, K., & Schoberer, J. (2025). Trafficking and localization of Golgi-resident N-glycan processing enzymes in plants. Frontiers in plant science, 16, 1624949

Dunn, K. W., Kamocka, M. M., & McDonald, J. H., A practical guide to evaluating colocalization in biological microscopy. American Journal of Physiology. Cell Physiology, 2011, 300(4), C723-742.

Fitchette, A.-C., Cabanes-Macheteau, M., Marvin, L., Martin, B., Satiat-Jeunemaitre, B., Gomord, V., Crooks, K., Lerouge, P., & Hawes, C. (1999). Biosynthesis and Immunolocalization of Lewis a-Containing N-Glycans in the Plant Cell 1.

Frey, A.D., Karg, S.R. and Kallio, P.T. (2009), Expression of rat β(1,4)-*N*-acetylglucosaminyltransferase III in *Nicotiana tabacum* remodels the plant-specific *N*-glycosylation. Plant Biotechnology Journal, 7: 33-48.

Frigerio, L., Vine, N. D., Pedrazzini, E., Hein, M. B., Wang, F., Ma, J. K., & Vitale, A. (2000). Assembly, secretion, and vacuolar delivery of a hybrid immunoglobulin in plants. Plant physiology, 123(4), 1483–1494.

Gallagher, S. and Sasse, J. (1998), Protein Analysis by SDS-PAGE and Detection by Coomassie Blue or Silver Staining. Current Protocols in Pharmacology, 2: A.3B.1-A.3B.10.

Gomaraschi, M., Bonacina, F., & Norata, G. D. (2019). Lysosomal Acid Lipase: From Cellular Lipid Handler to Immunometabolic Target. *Trends in pharmacological sciences*, *40*(2), 104–115.

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., & Faye, L. (2004). Posttranslational modification of therapeutic proteins in plants. Current Opinion in Plant Biology, 7(2), Article 2.

Gomord, V., Fitchette, A. C., Menu-Bouaouiche, L., Saint-Jore-Dupas, C., Plasson, C., Michaud, D., & Faye, L. (2010). Plant-specific glycosylation patterns in the context of therapeutic protein production. *Plant Biotechnology Journal*, *8*(5), Article 5.

Göritzer, K., & Strasser, R. (2021). Glycosylation of Plant-Produced Immunoglobulins. In M. Pezer (Ed.), Antibody Glycosylation (pp. 519–543). Springer International Publishing.

Göritzer, K., Melnik, S., Schwestka, J., Arcalis, E., Drapal, M., Fraser, P., Ma, J.K.-C. and Stoger, E. (2025), Enhancing quality and yield of recombinant secretory IgA antibodies in Nicotiana benthamiana by endoplasmic reticulum engineering. Plant Biotechnol. J.

Hawes, C., Osterrieder, A., Hummel, E. and Sparkes, I., The Plant ER–Golgi Interface. Traffic, 2008, 9: 1571-1580.

Hawes, C., P. Wang, and V. Kriechbaumer, Make It Shine: Labelling the ER for Light and Fluorescence Microscopy. Methods Mol Biol, 2024. 2772: p. 1-14.

Hehle, V.K., Lombardi, R., van Dolleweerd, C.J., Paul, M.J., Di Micco, P., Morea, V., Benvenuto, E., Donini, M. and Ma, J.K.-C. (2015), Site-specific proteolytic degradation of IgG monoclonal antibodies expressed in tobacco plants. Plant Biotechnol J, 13: 235-245.

Irons, S.L., Nuttall, J., Floß, D.M., Frigerio, L., Kotzer, A.M. and Hawes, C. (2008), Fluorescent protein fusions to a human immunodeficiency virus monoclonal antibody reveal its intracellular transport through the plant endomembrane system. Plant Biotechnology Journal, 6: 649-662.

Ito, Y., & Uemura, T., Super resolution live imaging: The key for unveiling the true dynamics of membrane traffic around the Golgi apparatus in plant cells. Frontiers in Plant Science, 2022, 13.

Ito, Y., Uemura, T., & Nakano, A. Chapter Six—Formation and Maintenance of the Golgi Apparatus in Plant Cells. In K. W. Jeon (Ed.), International Review of Cell and Molecular Biology, 2014, 310, 221–287

Jansing, J., Sack, M., Augustine, S. M., Fischer, R., & Bortesi, L. (2019). 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 journal*, *17*(2), 350–361.

Kallolimath, S., Castilho, A., Strasser, R., Grünwald-Gruber, C., Altmann, F., Strubl, S., Galuska, C. E., Zlatina, K., Galuska, S. P., Werner, S., Thiesler, H., Werneburg, S., Hildebrandt, H., Gerardy-Schahn, R., & Steinkellner, H. (2016). Engineering of complex protein sialylation in plants. Proceedings of the National Academy of Sciences of the United States of America, 113(34), Article 34.

Kallolimath, S., Castilho, A., Strasser, R., Grünwald-Gruber, C., Altmann, F., Strubl, S., Galuska, C. E., Zlatina, K., Galuska, S. P., Werner, S., Thiesler, H., Werneburg, S., Hildebrandt, H., Gerardy-Schahn, R., & Steinkellner, H. (2016). Engineering of

complex protein sialylation in plants. Proceedings of the National Academy of Sciences of the United States of America, 113(34), Article 34.

Karimi, M., B. De Meyer, and P. Hilson, Modular cloning in plant cells. Trends Plant Sci, 2005. 10(3): p. 103-5.

Karimi, M., Inzé, D., & Depicker, A. (2002). GATEWAY™ vectors for Agrobacterium-mediated plant transformation. Trends in Plant Science, 7(5), 193–195.

Khosla A., Rodriguez-Furlan C., Kapoor S., Van Norman JM., Nelson DC., A series of dual-reporter vectors for ratiometric analysis of protein abundance in plants. Plant Direct. 2020;4(6):e00231.

Kim, H. S., Jeon, J. H., Lee, K. J., & Ko, K. (2014). N-glycosylation modification of plant-derived virus-like particles: An application in vaccines. BioMed Research International.

Kizhner, T., Azulay, Y., Hainrichson, M., Tekoah, Y., Arvatz, G., Shulman, A., Ruderfer, I., Aviezer, D., & Shaaltiel, Y. (2015). Characterization of a chemically modified plant cell culture expressed human α-Galactosidase-A enzyme for treatment of Fabry disease. Molecular Genetics and Metabolism, 114(2), Article 2.

Ko, K., Tekoah, Y., Rudd, P. M., Harvey, D. J., Dwek, R. A., Spitsin, S., Hanlon, C. A., Rupprecht, C., Dietzschold, B., Golovkin, M., & Koprowski, H. (2003). Function and glycosylation of plant-derived antiviral monoclonal antibody. Proceedings of the National Academy of Sciences of the United States of America, 100(13), 8013–8018.

Ko, S. Y., Pegu, A., Rudicell, R. S., Yang, Z. Y., Joyce, M. G., Chen, X., Wang, K., Bao, S., Kraemer, T. D., Rath, T., Zeng, M., Schmidt, S. D., Todd, J. P., Penzak, S. R., Saunders, K. O., Nason, M. C., Haase, A. T., Rao, S. S., Blumberg, R. S., Mascola, J. R., … Nabel, G. J. (2014). Enhanced neonatal Fc receptor function improves protection against primate SHIV infection. Nature, 514(7524), 642–645.

Kogelmann, B., Melnik, S., Bogner, M., Kallolimath, S., Stöger, E., Sun, L., Strasser, R., D'Aoust, M.-A., Lavoie, P.-O., Saxena, P., Gach, J. S., & Steinkellner, H. (2024). A genome-edited N. benthamiana line for industrial-scale production of recombinant glycoproteins with targeted N-glycosylation. Biotechnology Journal, 19, e2300323.

Kogelmann, B., Melnik, S., Keshvari, T., Bogner, M., Lavoie, P.-O., D`Aoust, M.-A., Hermle, A., Lux, A., Strasser, R., Stöger, E., & Steinkellner, H. (2025). An industrial-grade Nicotiana benthamiana line for the production of glycoproteins carrying fucose-free galactosylated N-glycans. New Biotechnology, 85, 23–30.

Koprivova, A., Stemmer, C., Altmann, F., Hoffmann, A., Kopriva, S., Gorr, G., Reski, R., & Decker, E. L. (2004). Targeted knockouts of Physcomitrella lacking plant-specific immunogenic N-glycans. Plant Biotechnology Journal, 2(6), Article 6.

Lennon, C. W., & Belfort, M., Inteins. Current Biology, 2017, 27(6), R204–R206.

Li, F., & Zhang, H. (2019). Lysosomal Acid Lipase in Lipid Metabolism and Beyond. Arteriosclerosis, Thrombosis, and Vascular Biology, 39(5), Article 5.

Li, Q., Su, J., Liu, Y., Jin, X., Zhong, X., Mo, L., Wang, Q., Deng, H., & Yang, Y., In vivo PCSK9 gene editing using an all-in-one self-cleavage AAV-CRISPR system. Molecular Therapy - Methods & Clinical Development, 2021, 20, 652–659.

Liebminger E., Hüttner S., Vavra U., Fischl R., Schoberer J., Grass J., Blaukopf C., Seifert GJ., Altmann F., Mach L., Strasser R., Class I α-Mannosidases Are Required for N-Glycan Processing and Root Development in Arabidopsis thaliana. Plant Cell. 2009;21(12):3850.

Liu Z., L,Olivia C., Wall JB., Michael Z.,Yang Z., Li W., Vaseghi, HR., Qian, Li Q., Jiandong L., Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. Scientific Reports. 2017;7(1):2193.

Liu, C., Niu, G., Zhang, H., Sun, Y., Sun, S., Yu, F., Lu, S., Yang, Y., Li, J., & Hong, Z. (2018). Trimming of N-Glycans by the Golgi-Localized α-1,2-Mannosidases, MNS1 and MNS2, Is Crucial for Maintaining RSW2 Protein Abundance during Salt Stress in Arabidopsis. *Molecular plant*, *11*(5), 678–690.

Loos A., Steinkellner H., Plant glyco-biotechnology on the way to synthetic biology. Front Plant Sci. 2014;5:523.

MALIGA, P. (2005). Molecular farming: Plant-made pharmaceuticals and technical proteins. Fischer, R. and Schillberg, S. (eds) 2004. Weinheim: Wiley-VCH Verlag GmbH & Co. KgaA. €145 (hardback). 375 pp. Annals of Botany, 96(2), 343–343.

Manders EMM., Verbeek FJ., Aten JA., Measurement of co-localization of objects in dual-colour confocal images. Journal of Microscopy. 1993;169(3):375-382.

Margolin, E., Crispin, M., Meyers, A., Chapman, R., & Rybicki, E. P. (2020). A Roadmap for the Molecular Farming of Viral Glycoprotein Vaccines: Engineering Glycosylation and Glycosylation-Directed Folding. Frontiers in plant science, 11, 609207.

McDonald JH., Dunn K W., Statistical tests for measures of colocalization in biological microscopy. Journal of Microscopy, 2013, 252(3), 295–302.

McDonald, A. G., Hayes, J. M., Bezak, T., Głuchowska, S. A., Cosgrave, E. F. J., Struwe, W. B., Stroop, C. J. M., Kok, H., van de Laar, T., Rudd, P. M., Tipton, K. F., & Davey, G. P. (2014). Galactosyltransferase 4 is a major control point for glycan branching in N-linked glycosylation. *Journal of Cell Science*, *127*(23), Article 23.

McGinness AJ, Schoberer J, Pain C, Brandizzi F, Kriechbaumer V. On the nature of the plant ER exit sites. Front Plant Sci. 2022, 13:1010569

McGinness, A. J., Brooks, S. A., Strasser, R., Schoberer, J., & Kriechbaumer, V. (2024). Suborganellar resolution imaging for the localisation of human glycosylation enzymes in tobacco Golgi bodies. Journal of Microscopy, 1–16.

Mishra, A., Behura, A., Mawatwal, S., Kumar, A., Naik, L., Mohanty, S. S., Manna, D., Dokania, P., Mishra, A., Patra, S. K., & Dhiman, R. (2019). Structure-function and application of plant lectins in disease biology and immunity. Food and chemical toxicology 134, 110827.

Montero-Morales, L., & Steinkellner, H. (2018). Advanced Plant-Based Glycan Engineering. Frontiers in Bioengineering and Biotechnology, 6.

Montero-Morales, L., Maresch, D., Crescioli, S., Castilho, A., Ilieva, K. M., Mele, S., Karagiannis, S. N., Altmann, F., & Steinkellner, H. (2019). In Planta Glycan Engineering and Functional Activities of IgE Antibodies. Frontiers in Bioengineering and Biotechnology, 7.

Moon, K. B., Park, J. S., Park, Y. I., Song, I. J., Lee, H. J., Cho, H. S., Jeon, J. H., & Kim, H. S. (2019). Development of Systems for the Production of Plant-Derived Biopharmaceuticals. *Plants* , *9*(1), 30.

Munro, S., An investigation of the role of transmembrane domains in Golgi protein retention. EMBO J, 1995. 14(19): p. 4695-704.

Nagels, B., van Damme, E. J. M., Pabst, M., Callewaert, N., & Weterings, K. (2011). Production of complex multiantennary N-glycans in Nicotiana benthamiana plants. Plant Physiology, 155(3), Article 3.

Neelamegham, S., Aoki-Kinoshita, K., Bolton, E., Frank, M., Lisacek, F., Lütteke, T., O'Boyle, N., Packer, N. H., Stanley, P., Toukach, P., Varki, A., & Woods, R. J. (2019). Updates to the Symbol Nomenclature for Glycans guidelines. *Glycobiology*, *29*(9), Article 9.

Nilsson T., Au CE., Bergeron J., Sorting out glycosylation enzymes in the Golgi apparatus. FEBS Letters, 2009, 583

Ocampo, C.G., Vignolles, F., Pombo, M.A., Colombo, M.L., Rosli, H.G. and Petruccelli, S. (2024), AtLEC2-Mediated Enhancement of Endoplasmic Reticulum-Targeted Foreign Protein Synthesis in Nicotiana benthamiana Leaves: Insights From Transcriptomic Analysis. Biotechnology and Bioengineering.

Ohira, M., Barr, M., Okuyama, T., & Mashima, R. (2022). LC-MS/MS-based enzyme assay for lysosomal acid lipase using dried blood spots. *Molecular genetics and metabolism reports*, *33*, 100913.

Osterrieder, A., Carvalho, C.M., Latijnhouwers, M., Johansen, J.N., Stubbs, C., Botchway, S. and Hawes, C., Fluorescence Lifetime Imaging of Interactions between Golgi Tethering Factors and Small GTPases in Plants. Traffic, 2009, 10: 1034-1046.

Pagny, S., Bouissonnie, F., Sarkar, M., Follet-Gueye, M. L., Driouich, A., Schachter, H., Faye, L., & Gomord, V. (2003). Structural requirements for Arabidopsisβ1,2-xylosyltransferase activity and targeting to the Golgi. The Plant Journal, 33(1), Article 1.

Palacpac N,Yoshida S., Sakai H., Kimura Y., Fujiyama K., Yoshida T., Seki T., Stable Expression of Human 1,4-Galactosyltransferase in Plant Cells Modifies N-Linked Glycosylation Patterns. Vol 96.; 1999:4692-4697.

Parsons H., Stevens T., McFarlane H., Vidal-Melgosa S., Griss J., Lawrence N., Butler R., Sousa M., Salemi M., Willats W., Petzold C., Heazlewood J., & Lilley K.,

Separating Golgi Proteins from Cis to Trans Reveals Underlying Properties of Cisternal Localization. The Plant Cell. 2019;31(9):2010-2034.

Rabouille C., Hui N., Hunte F., Kieckbusch R., Berger E. G., Warren G., & Nilsson T.,Mapping the distribution of Golgi enzymes involved in the construction of complex oligosaccharides. Journal of Cell Science, 1995, 108(4), 1617–1627.

Rajamohan, F., Reyes, A. R., Tu, M., Nedoma, N. L., Hoth, L. R., Schwaid, A. G., Kurumbail, R. G., Ward, J., & Han, S. (2020). Crystal structure of human lysosomal acid lipase and its implications in cholesteryl ester storage disease. *Journal of lipid research*, *61*(8), 1192–1202.

Ree, R., Cabanes-Macheteau, M., Akkerdaas, J., Milazzo, J.-P., Loutelier-Bourhis, C., Rayon, C., Villalba, M., Koppelman, S., Aalberse, R., Rodriguez, R., & Lerouge, P. (2000). Fucose Residues Have a Strong Contribution in IgE Binding to Plant Glycoallergens* (Vol. 3, Issue 1). http://www.jbc.org/

Renna, L., Hanton, S. L., Stefano, G., Bortolotti, L., Misra, V., & Brandizzi, F. (2005). Identification and characterization of AtCASP, a plant transmembrane Golgi matrix protein. *Plant molecular biology*, *58*(1), 109–122.

Richard Strasser, Plant protein glycosylation, Glycobiology, Volume 26, Issue 9, September 2016, Pages 926–939,

Robinson DG. Plant Golgi ultrastructure. Journal of Microscopy. 2020;280(2):111-121.

Runions J., Brach T., Kühner S., Hawes C., Photoactivation of GFP reveals protein dynamics within the endoplasmic reticulum membrane. J Exp Bot. 2006;57(1):43-50.

Rup, B., Alon, S., Amit-Cohen, B. C., Almon, E. B., Chertkoff, R., Tekoah, Y., & Rudd, P. M. (2017). Immunogenicity of glycans on biotherapeutic drugs produced in plant expression systems—The taliglucerase alfa story. *PLoS ONE*, *12*(10), Article 10.

Ryan MD., Drew J., Foot-and-mouth disease virus 2A oligopeptide mediated cleavage of an artificial polyprotein. EMBO J. 1994;13(4):928-933.

Sack, M., Hofbauer, A., Fischer, R., & Stoger, E. (2015). *The increasing value of plant-made proteins*. Elsevier Ltd.

Saint-Jore CM., Evins J., Batoko H., Brandizzi,F., Moore I. and Hawes C., Redistribution of membrane proteins between the Golgi apparatus and endoplasmic reticulum in plants is reversible and not dependent on cytoskeletal networks. The Plant Journal, 2002, 29: 661-678.

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

Samalova M., Fricker M., Moore I., Ratiometric Fluorescence-Imaging Assays of Plant Membrane Traffic Using Polyproteins. Traffic. 2006;7(12):1701-1723.

Scalettar, B. A., Swedlow, J. R., Sedat, J. W., & Agard, D. A. (1996). Dispersion, aberration and deconvolution in multi-wavelength fluorescence images. *Journal of microscopy*, *182*(Pt 1), 50–60.

Schindelin J., Arganda-Carreras I., Frise E., Kaynig V., Longair M., Pietzsch T., Preibisch S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez J.-Y., White D. J., Hartenstein V., Eliceiri K., Tomancak P., Cardon, A., Fiji: An open-source platform for biological-image analysis. Nature Methods, 2012, 9(7), 676–682.

Schneider, J., Castilho, A., Pabst, M., Altmann, F., Gruber, C., Strasser, R., Gattinger, P., Seifert, G. J., & Steinkellner, H. (2015). Characterization of plants expressing the human β1,4-galactosyltrasferase gene. Plant physiology and biochemistry : PPB, 92, 39–47.

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(4):1737-1754.

Schoberer J., Liebminger E., Vavra U., Veit C., Castilho A., Dicker M., Maresch D., Altmann F., Hawes C., Botchway S.W. and Strasser R., The transmembrane domain of N –acetylglucosaminyltransferase I is the key determinant for its Golgi subcompartmentation. The Plant Journal. 2014;80(5):809-822.

Schoberer J., Liebminger E., Vavra U., Veit C., Grünwald-Gruber C., Altmann F., Botchway S. W., Strasser R., The Golgi Localization of GnTI Requires a Polar Amino Acid Residue within Its Transmembrane Domain. Plant Physiology, 2019, 180(2), 859–873.

Schoberer J., Strasser R., Sub-Compartmental Organization of Golgi-Resident N-Glycan Processing Enzymes in Plants. Molecular Plant. 2011;4(2):220-228.

Schoberer, J. and R. Strasser, Plant glyco-biotechnology. Semin Cell Dev Biol, 2018(b). 80: p. 133-141.

Schoberer, J., König, J., Veit, C., Vavra, U., Liebminger, E., Botchway, S. W., Altmann, F., Kriechbaumer, V., Hawes, C., & Strasser, R. (2019). A signal motif retains Arabidopsis ER-α-mannosidase I in the cis-Golgi and prevents enhanced glycoprotein ERAD. *Nature Communications*, *10*(1), Article 1.

Schoberer, J., Liebminger, E., Vavra, U., Veit, C., Castilho, A., Dicker, M., Maresch, D., Altmann, F., Hawes, C., Botchway, S. W., & Strasser, R. (2014). The transmembrane domain of N –acetylglucosaminyltransferase I is the key determinant for its Golgi subcompartmentation. *The Plant Journal*, *80*(5), Article 5

Schoberer, J., Runions, J., Steinkellner, H., Strasser, R., Hawes, C. and Osterrieder, A., Sequential Depletion and Acquisition of Proteins during Golgi Stack Disassembly and Reformation. Traffic, 2010. 11: 1429-1444.

Schoberer, J., Shin, Y. J., Vavra, U., Veit, C., & Strasser, R. (2024). Analysis of Protein Glycosylation in the ER. *Methods in molecular biology (Clifton, N.J.)*, *2772*, 221–238.

Shaaltiel, Y., Bartfeld, D., Hashmueli, S., Baum, G., Brill-Almon, E., Galili, G., Dym, O., Boldin-Adamsky, S. A., Silman, I., Sussman, J. L., Futerman, A. H., & Aviezer, D. (2007). Production of glucocerebrosidase with terminal mannose glycans for enzyme replacement therapy of Gaucher's disease using a plant cell system. Plant Biotechnology Journal, 5(5), Article 5.

Shade, K. C., Conroy, M. E., Washburn, N., Kitaoka, M., Huynh, D. J., Laprise, E., Patil, S. U., Shreffler, W. G., & Anthony, R. M. (2020). Sialylation of immunoglobulin E is a determinant of allergic pathogenicity. *Nature*, *582*(7811), 265–270.

Sparkes, I. A., Runions, J., Kearns, A., & Hawes, C. (2006). Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nature protocols*, *1*(4), 2019–2025

Spatola Rossi, T., M. Fricker, and V. Kriechbaumer, Gene Stacking and Stoichiometric Expression of ER-Targeted Constructs Using "2A" Self-Cleaving Peptides. Methods Mol Biol, 2024. 2772: p. 337-351.

Spatola-Rossi T., Tolmie AF., Nichol T., Pain C., Harrison P., Smith TJ., Fricker M., Kriechbaumer V., Recombinant expression and subcellular targeting of the particulate methane monooxygenase (pMMO) protein components in plants. Scientific Reports, 2023, 13(1), 15337.

Stockert R. J. (1995). The asialoglycoprotein receptor: relationships between structure, function, and expression. *Physiological reviews*, *75*(3), 591–609.

Strasser R., Castilho A., Stadlmann J., Kunert R., Quendler H., Gattinger P., Jez J., Rademacher T., Altmann F., Mach L., Steinkellner H., Improved virus neutralization by plant-produced anti-HIV antibodies with a homogeneous beta1,4-galactosylated N-glycan profile. J Biol Chem, 2009. 31;284(31):20479-85.

Strasser, R., Biological significance of complex N-glycans in plants and their impact on plant physiology. Front Plant Sci, 2014. 5: p. 363.

Strasser, R., Bondili, J. S., Schoberer, J., Svoboda, B., Liebminger, E., Glössl, J., Altmann, F., Steinkellner, H., & Mach, L. (2007). Enzymatic properties and subcellular localization of Arabidopsis beta-N-acetylhexosaminidases. Plant physiology, 145(1), 5–16.

Strasser, R., Castilho, A., Stadlmann, J., Kunert, R., Quendler, H., Gattinger, P., Jez, J., Rademacher, T., Altmann, F., Mach, L., & Steinkellner, H. (2009). Improved virus neutralization by plant-produced anti-HIV antibodies with a homogeneous β1,4-galactosylated N-glycan profile. Journal of Biological Chemistry, 284(31), Article 31.

Strasser, R., F. Altmann, and H. Steinkellner, Controlled glycosylation of plant-produced recombinant proteins. Curr Opin Biotechnol, 2014. 30: p. 95-100.

Strasser, R., Stadlmann, J., Schähs, M., Stiegler, G., Quendler, H., Mach, L., Glössl, J., Weterings, K., Pabst, M., & Steinkellner, H. (2008). Generation of glyco-

engineered Nicotiana benthamiana for the production of monoclonal antibodies with a homogeneous human-like N-glycan structure. Plant Biotechnology Journal, 6(4), 392–402.

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

Stelter, S., Paul, M.J., Teh, A.Y.-.-H., Grandits, M., Altmann, F., Vanier, J., Bardor, M., Castilho, A., Allen, R.L. and Ma, J.K.-C. (2020), Engineering the interactions between a plant-produced HIV antibody and human Fc receptors. Plant Biotechnol J, 18: 402-414

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

Su, D., Zhao, H., & Xia, H. (2010). Glycosylation-modified erythropoietin with improved half-life and biological activity. International journal of hematology, 91(2), 238–244.

Sumar, N., Bodman, K. B., Rademacher, T. W., Dwek, R. A., Williams, P., Parekh, R. B., Edge, J., Rook, G. A. W., Isenberg, D. A., Hay, F. C., & Roitt, I. M. (1990). Analysis of glycosylation changes in IgG using lectins. Journal of Immunological Methods, 131(1), 127–136.

Szymczak A., Vignali DA., Development of 2A peptide-based strategies in the design of multicistronic vectors. Expert Opin Biol Ther. 2005;5(5):627-638.

Tekoah, Y., Shulman, A., Kizhner, T., Ruderfer, I., Fux, L., Nataf, Y., Bartfeld, D., Ariel, T., Gingis-Velitski, S., Hanania, U., & Shaaltiel, Y. (2015). Large-scale production of pharmaceutical proteins in plant cell culture-the protalix experience. Plant Biotechnology Journal, 13(8), Article 8.

Tekoah, Y., Tzaban, S., Kizhner, T., Hainrichson, M., Gantman, A., Golembo, M., Aviezer, D., & Shaaltiel, Y. (2013). Glycosylation and functionality of recombinant β-glucocerebrosidase from various production systems. Bioscience Reports, 33(5), Article 5.

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

Uliana A., Crespo P., Martina J., Daniotti J., Maccioni H., Modulation of GalT1 and SialT1 Sub-Golgi Localization by SialT2 Expression Reveals an Organellar Level of Glycolipid Synthesis Control. Journal of Biological Chemistry, 2006, 281(43), 32852–32860.

Varki, A., Cummings, R. D., Aebi, M., Packer, N. H., Seeberger, P. H., Esko, J. D., Stanley, P., Hart, G., Darvill, A., Kinoshita, T., Prestegard, J. J., Schnaar, R. L., Freeze, H. H., Marth, J. D., Bertozzi, C. R., Etzler, M. E., Frank, M., Vliegenthart, J. F., Lütteke, T., Perez, S., Kornfeld, S. (2015). Symbol Nomenclature for Graphical Representations of Glycans. *Glycobiology*, *25*(12), 1323–1324.

Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Darvill, A. G., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L., & Seeberger, P. H. (Eds.). (2015). *Essentials of Glycobiology*. (3rd ed.). Cold Spring Harbor Laboratory Press, ch56

Vézina, L.-P., Faye, L., Lerouge, P., D'Aoust, M.-A., Marquet-Blouin, E., Burel, C., Lavoie, P.-O., Bardor, M., & Gomord, V. (2009). Transient co-expression for fast and high-yield production of antibodies with human-like N-glycans in plants. Plant Biotechnology Journal, 7(5), 442–455.

Wallace, W., Schaefer, L. H., & Swedlow, J. R. (2001). A Workingperson's Guide to Deconvolution in Light Microscopy. BioTechniques, 31(5), 1076–1097.

Walsh, G., & Jefferis, R. (2006). Post-translational modifications in the context of therapeutic proteins. *Nature biotechnology*, *24*(10), 1241–1252.

Wang Y., Wang F., Wang R., Zhao P., Xia Q., 2A self-cleaving peptide-based multi-gene expression system in the silkworm Bombyx mori. Scientific Reports. 2015;5(1):16273.

Ward, B. J., Landry, N., Trépanier, S., Mercier, G., Dargis, M., Couture, M., D'Aoust, M. A., & Vézina, L. P. (2014). Human antibody response to N-glycans present on plant-made influenza virus-like particle (VLP) vaccines. Vaccine, 32(46), Article 46.

Wee E., Sherrier D., Prime T., and Dupree P., Targeting of active sialyltransferase to the plant Golgi apparatus. Plant Cell, 1998, 10, 1759–1768

Wu, A. M., Wu, J. H., Singh, T., Lai, L.-J., Yang, Z., & Herp, A. (2006). Recognition factors of Ricinus communis agglutinin 1 (RCA1). Molecular Immunology, 43(10), 1700–1715.

Yagi H., Tateo S., Saito T., Ohta Y., Nishi E., Obitsu S., Suzuki T., Seetaha S., Hellec C., Nakano A., Tojima T., Kato K., Deciphering the sub-Golgi localization of glycosyltransferases via 3D super-resolution imaging. Cell Structure and Function, 2024, 49(2), 47–55.

Zhang B., Rapolu M., Kumar S., Gupta M., Liang Z., Han Z., Williams P., Su W. W., Coordinated protein co-expression in plants by harnessing the synergy between an intein and a viral 2A peptide. Plant Biotechnology Journal. 2017;15(6):718-728.

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

Zschenker, O., Bähr, C., Hess, U. F., & Ameis, D. (2005). Systematic mutagenesis of potential glycosylation sites of lysosomal acid lipase. Journal of biochemistry, 137(3), 387–394.

8. Appendices

Contents

Appendix 8.1: Python script for peak intensity distance

Code for python version 3.9

```
import numpy as np
from glob import glob
import pandas as pd
from scipy.optimize import curve_fit
import matplotlib.pyplot as plt

# How to use:
# Put all csv files in Data Folder
# Run main.py
# Output Figures will appear in Output Figures folder

# Define gaussian equation
def gauss(x,a,b,c):
    return a*np.exp(-((x-b)**2) / (2*c**2))

# Get all csv file names in Data folder
pattern = "Data/*.csv"
file_name_list = glob(pattern)

# Alphabetise (glob reads files in a random order)
file_name_list = np.sort(file_name_list)
print(f"Number of Files Found: {len(file_name_list)}\n")



# Empty list to store distances
gauss_distances_list = []
```

```
max_value_distances_list = []
channel_1_max_list = []
channel_2_max_list = []



# Loop through file names
for i, file in enumerate(file_name_list):

    # Load data
    data = np.genfromtxt(file, delimiter=',', skip_header=3)

    # Get channel data (unnecessary, but i'm making an effort to change variables into something less abstract)
    channel_1 = data[:,1]
    channel_2 = data[:,2]
    distances = data[:,0]

    # The fit doesn't have to have the same number of datapoints, it looks smoother with 1000 points
    x_values = np.linspace(0, np.max(distances), 1000)



    # Fit Channel 1 Data to Gaussian

    # Guesses 1
    height = np.max(channel_1)
    position = distances[int(np.argmax(channel_1))]
    stdev = 100
    guesses = [height, position, stdev]
```

```
    # Get residuals and create fit function
    fit_values_1, _ = curve_fit(gauss, xdata=distances, ydata=channel_1,
p0=guesses)
    y_fit = gauss(x_values, fit_values_1[0], fit_values_1[1], fit_values_1[2])

    # Plot data as scatter and fit as line
    plt.figure(i, figsize=(7,5))
    plt.scatter(distances, channel_1, color="#ca0cf5", label="Channel 1 Data")
    plt.plot(x_values, y_fit, color="#ca0cf5", label="Channel 1 Fit")



    # Fit Channel 2 Data to Gaussian

    # Guesses 2
    height = np.max(channel_2)
    position = distances[int(np.argmax(channel_2))]
    stdev = 100
    guesses = [height, position, stdev]

    # Get residuals and create fit function
    fit_values_2, _ = curve_fit(gauss, xdata=distances, ydata=channel_2,
p0=guesses)
    y_fit = gauss(x_values, fit_values_2[0], fit_values_2[1], fit_values_2[2])

    # Plot data as scatter and fit as line
    plt.scatter(distances, channel_2, color="#2bff00", label="Channel 2 Data")
    plt.plot(x_values, y_fit, color="#2bff00", label="Channel 2 Fit")

    # Graph formatting
    plt.legend()
    plt.xlabel("Distance (nm)")
```

```
    plt.ylabel("Intensity")
    plt.title(file[5:-4])
    plt.tight_layout()
    fig_title = f"./Output_Figures/{file[5:-4]}.png"
    plt.savefig(fig_title)



    # Calculate distance between peak locations using gaussian distribution
    gaus_separation = abs(fit_values_1[1] - fit_values_2[1])
    gauss_distances_list.append(gaus_separation)

    # Calculate distance between peak locations using maximum value
    max_value_sep = abs(distances[np.argmax(channel_1)] -
distances[np.argmax(channel_2)])
    max_value_distances_list.append(max_value_sep)



    channel_1_max_list.append(fit_values_1[1])
    channel_2_max_list.append(fit_values_2[1])

    print("File Name:")
    print(file[5:-4])
    print(f"Distance: {np.round(gaus_separation, 2)} nm\n")



# Store it in a CSV that excel can open
dictionary = {"File Name": file_name_list,
              "Channel 1 Peak Location": channel_1_max_list,
              "Channel 2 Peak Location": channel_2_max_list,
               "Gaussian Distance (nm)": gauss_distances_list,
              "Max Value Distance (nm)": max_value_distances_list}
```

```
df = pd.DataFrame(dictionary)
df.to_csv("Analysis.csv")
```

Appendix 8.2: Sequences for self-cleaving peptide markers

8.2.1 Synthesised and codon-optimised medial/trans Golgi marker with self-cleaving peptide

MUR3-linker-mTagBFP-Intein-linker-P2A-FUT13/CTS-marker

ATGTTCCCCAGAGTCAGTATGCGAAGGAGAAGTGCAGAGGTCAGCCCCACTGA
ACCCATGGAAAAGGGCAACGGTAAGAACCAAACAAACCGCATTTGCCTCTTGG
TCGCGCTGAGCTTGTTCTTTTGGGCGCTTCTATTATATTTTCACTTCGTAGTACT
TGGTACCAGCAATATAGACAAGCAATTACAGCTTCAGCCATCCTACGCATGCAG
TTGCGGGTCCGGAAGTAGGATGGTAAGTAAAGGAGAGGAATTAATCAAAGAAA
ATATGCATATGAAGCTTTATATGGAAGGAACAGTTGATAATCACCATTTTAAATG
TACTTCTGAAGGGGAAGGGAAACCTTATGAGGGTACTCAAACAATGCGCATAAA
AGTTGTTGAAGGTGGTCCATTACCATTTGCGTTTGATATTCTCGCCACAAGTTTT
CTATATGGGTCTAAAACATTTATTAATCATACACAGGGAATCCCAGATTTCTTTA
AACAGTCATTTCCAGAAGGTTTTACTTGGGAACGGGTTACAACTTATGAGGATG
GTGGTGTTTTAACAGCCACACAAGATACTAGTTTGCAAGACGGGTGTTTAATTTA
TAACGTTAAAATTCGGGGAGTAAATTTTACCTCAAATGGACCAGTAATGCAGAA
GAAGACTTTGGGGTGGGAAGCTTTTACTGAAACATTGTATCCGGCAGATGGTG
GGCTCGAGGGACGTAATGATATGGCATTGAAATTGGTCGGTGGTAGTCACCTT
ATTGCCAATGCTAAAACAACGTACAGGAGCAAGAAGCCAGCTAAGAATTTGAAA
ATGCCCGGTGTTTATTACGTGGATTATAGGCTTGAGAGGATTAAAGAAGCAAAT
AATGAAACATATGTTGAACAACATGAAGTCGCTGTCGCTAGGTATTGTGATCTTC
CATCCAAGTTGGGTCATAAATTAAACTGCCTGTCTTTTGGTACAGAAATCCTCAC
TGTGGAGTACGGCCCTTTGCCAATAGGGAAAATCGTTTCTGAGGAAATTAATTG
TTCCGTGTACTCTGTAGACCCCGAAGGCAGGGTCTATACGCAGGCTATAGCAC
AATGGCACGACCGAGGGGAGCAAGAAGTCCTGGAGTACGAACTCGAGGACGG
CTCAGTTATCCGTGCAACGTCTGACCATAGATTCTTGACAACTGACTATCAGTT
GTTAGCTATAGAGGAAATATTTGCACGCCAACTTGACCTGCTGACCCTCGAGAA
CATAAAACAGACTGAGGAAGCCCTCGACAATCACAGACTGCCTTTCCCACTACT
CGACGCGGGGACAATTAAAATGGTCAAGGTTATTGGTCGGCGTTCTCTCGGAG
TACAGAGGATTTTCGACATTGGTTTACCCCAAGACCATAACTTCTTGTTGGCGA
ATGGCGCAATAGCAGCGGCCTGCAGTTGCGGCTCAGGCAGTCGTGGCTCTGG

CGCAACTAACTTTTCCCTTCTCAAGCAAGCAGGTGATGTAGAAGAAAACCCAGG
CCCGATGCCGATGCGATATTTGAATGCCATGGCTGCCCTGCTTATGATGTTCTT
CACGCTTTTAATTTTAAGTTTTACTGGAATACTTGAGTTTCCTTCCGCAAGCACC
TCTATGGAACACAGTATTGATCCAGAGCCCAAGTTAAGCGACAGCACTAGTC

8.2.2 Synthesised and codon-optimised *cis*/*trans*-Golgi marker with self-cleaving peptide

MNS1-mTagBFP-P2A-ST-marker

MNS1/CTS-linker-mTagBFP-P2A-ST/CTS-marker

ATGGCGAGAAGTAGATCGATTAGTGGTTATGGGATATGGAAATATTTGAATCCT
GCGTATTATCTTAGAAGACCGAGACGTTTGGCTTTGCTTTTCATTGTCTTCGTCT
CTGTTTCTATGCTTGTCTGGGATCGTATTAATCTTGCCTGTTCTTGTGGTTCTGG
TTCTAGAATGGTGTCTAAGGGCGAAGAGCTGATTAAGGAGAACATGCACATGAA
GCTGTACATGGAGGGCACCGTGGACAACCATCACTTCAAGTGCACATCCGAGG
GCGAAGGCAAGCCCTACGAGGGCACCCAGACCATGAGAATCAAGGTGGTCGA
GGGCGGCCCTCTCCCCTTCGCCTTCGACATCCTGGCTACTAGCTTCCTCTACG
GCAGCAAGACCTTCATCAACCACACCCAGGGCATCCCCGACTTCTTCAAGCAG
TCCTTCCCTGAGGGCTTCACATGGGAGAGAGTCACCACATACGAAGACGGGGG
CGTGCTGACCGCTACCCAGGACACCAGCCTCCAGGACGGCTGCCTCATCTACA
ACGTCAAGATCAGAGGGGTGAACTTCACATCCAACGGCCCTGTGATGCAGAAG
AAAACACTCGGCTGGGAGGCCTTCACCGAGACGCTGTACCCCGCTGACGGCG
GCCTGGAAGGCAGAAACGACATGGCCCTGAAGCTCGTGGGCGGGAGCCATCT
GATCGCAAACGCCAAGACCACATATAGATCCAAGAAACCCGCTAAGAACCTCAA
GATGCCTGGCGTCTACTATGTGGACTACAGACTGGAAAGAATCAAGGAGGCCA
ACAACGAGACCTACGTCGAGCAGCACGAGGTGGCAGTGGCCAGATACTGCGA
CCTCCCTAGCAAACTGGGGCACAAGCTTAATGGTTCGGGGGCTACGAACTTCA
GTCTGCTGAAACAGGCGGGAGACGTTGAGGAGAATCCTGGTatgattcataccaactt
gaagaaaaagttcagcctcttcatcctggtctttctcctgttcgcagtcatctgtgtttggaagaaagggagcgactatga
ggcccttacactgcaagccaaggaattccagatgcccaagagccaggagaaagtggccc

Appendices 8.3. Sequences for constructs with HUGE CTS replacements

8.3.1 MUR3-GNTIV

MFPRVSMRRRSAEVSPTEPMEKGNGKNQTNRICLLVALSLFFWALLLYFHFVVLGT
SNIDKQLQLQPSYATIVQQFKRVGAETNGSKDALNKFSDNTLKLLKELTSKKSLQVP
SIYYHLPHLLKNEGSLQPAVQIGNGRTGVSIVMGIPTVKREVKSYLIETLHSLIDNLYP
EEKLDCVIVVFIGETDIDYVHGVVANLEKEFSKEISSGLVEVISPPESYYPDLTNLKET
FGDSKERVRWRTKQNLDYCFLMMYAQEKGIYYIQLEDDIIVKQNYFNTIKNFALQLS
SEEWMILEFSQLGFIGKMFQAPDLTLIVEFIFMFYKEKPIDWLLDHILWVKVCNPEKD
AKHCDRQKANLRIRFRPSLFQHVGLHSSLSGKIQKLTDKDYMKPLLLKIHVNPPAEV
STSLKVYQGHTLEKTYMGEDFFWAITPIAGDYILFKFDKPVNVESYLFHSGNQEHPG
DILLNTTVEVLPFKSEGLEISKETKDKRLEDGYFRIGKFENGVAEGMVDPSLNPISAF
RLSVIQNSAVWAILNEIHIKKATN

8.3.2 MUR3-GNTV

MFPRVSMRRRSAEVSPTEPMEKGNGKNQTNRICLLVALSLFFWALLLYFHFVVLGT
SNIDKQLQLQPSYATNSTNSTTAVPSLVALEKINVADIINGAQEKCVLPPMDGYPHC
EGKIKWMKDMWRSDPCYADYGVDGSTCSFFIYLSEVENWCPHLPWRAKNPYEEA
DHNSLAEIRTDFNILYSMMKKHEEFRWMRLRIRRMADAWIQAIKSLAEKQNLEKRK
RKKVLVHLGLLTKESGFKIAETAFSGGPLGELVQWSDLITSLYLLGHDIRISASLAELK
EIMKKVVGNRSGCPTVGDRIVELIYIDIVGLAQFKKTLGPSWVHYQCMLRVLDSFGT
EPEFNHANYAQSKGHKTPWGKWNLNPQQFYTMFPHTPDNSFLGFVVEQHLNSSD
IHHINEIKRQNQSLVYGKVDSFWKNKKIYLDIIHTYMEVHATVYGSSTKNIPSYVKNH
GILSGRDLQFLLRETKLFVGLGFPYEGPAPLEAIANGCAFLNPKFNPPKSSKNTDFFI
GKPTLRELTSQHPYAEVFIGRPHVWTVDLNNQEEVEDAVKAILNQKIEPYMPYEFT
CEGMLQRINAFIEKQDFCHGQVMWPPLSALQVKLAEPGQSCKQVCQESQLICEPS
FFQHLNKDKDMLKYKVTCQSSELAKDILVPSFDPKNKHCVFQGDLLLFSCAGAHPR
HQRVCPCRDFIKGQVALCKDCL

8.3.3 FUT13-B4GALT1

MPMRYLNAMAALLMMFFTLLILSFTGILEFPSASTSMEHSIDPEPKLSDSTSLPACPE
ESPLLVGPMLIEFNMPVDLELVAKQNPNVKMGGRYAPRDCVSPHKVAIIIPFRNRQE
HLKYWLYYLHPVLQRQQLDYGIYVINQAGDTIFNRAKLLNVGFQEALKDYDYTCFVF
SDVDLIPMNDHNAYRCFSQPRHISVAMDKFGFSLPYVQYFGGVSALSKQQFLTING
FPNNYWGWGGEDDDIFNRLVFRGMSISRPNAVVGRCRMIRHSRDKKNEPNPQRF
DRIAHTKETMLSDGLNSLTYQVLDVQRYPLYTQITVDIGTPS

8.3.4 FUT13-ST6GAL

MPMRYLNAMAALLMMFFTLLILSFTGILEFPSASTSMEHSIDPEPKLSDSTSASFQV
WNKDSSSKNLIPRLQKIWKNYLSMNKYKVSYKGPGPGIKFSAEALRCHLRDHVNVS
MVEVTDFPFNTSEWEGYLPKESIRTKAGPWGRCAVVSSAGSLKSSQLGREIDDHD
AVLRFNGAPTANFQQDVGTKTTIRLMNSQLVTTEKRFLKDSLYNEGILIVWDPSVYH
SDIPKWYQNPDYNFFNNYKTYRKLHPNQPFYILKPQMPWELWDILQEISPEEIQPNP
PSSGMLGIIIMMTLCDQVDIYEFLPSKRKTDVCYYYQKFF

Appendices 8.4. Sequences for Native HuGE constructs

8.4.1 GNTIVa

>sp|Q9UM21|MGT4A_HUMAN Alpha-1,3-mannosyl-glycoprotein 4-beta-N-acetylglucosaminyltransferase A OS=Homo sapiens OX=9606 GN=MGAT4A PE=1 SV=1

MRLRNGTVATALAFITSFLTLSWYTTWQNGKEKLIAYQREFLALKERLRIAEHRISQR
SSELNTIVQQFKRVGAETNGSKDALNKFSDNTLKLLKELTSKKSLQVPSIYYHLPHLL
KNEGSLQPAVQIGNGRTGVSIVMGIPTVKREVKSYLIETLHSLIDNLYPEEKLDCVIV
VFIGETDIDYVHGVVANLEKEFSKEISSGLVEVISPPESYYPDLTNLKETFGDSKERV
RWRTKQNLDYCFLMMYAQEKGIYYIQLEDDIIVKQNYFNTIKNFALQLSSEEWMILE
FSQLGFIGKMFQAPDLTLIVEFIFMFYKEKPIDWLLDHILWVKVCNPEKDAKHCDRQ
KANLRIRFRPSLFQHVGLHSSLSGKIQKLTDKDYMKPLLLKIHVNPPAEVSTSLKVYQ
GHTLEKTYMGEDFFWAITPIAGDYILFKFDKPVNVESYLFHSGNQEHPGDILLNTTV
EVLPFKSEGLEISKETKDKRLEDGYFRIGKFENGVAEGMVDPSLNPISAFRLSVIQN
SAVWAILNEIHIKKATN

8.4.2 GNTV

>sp|Q09328|MGT5A_HUMAN Alpha-1,6-mannosylglycoprotein 6-beta-N-acetylglucosaminyltransferase A OS=Homo sapiens OX=9606 GN=MGAT5 PE=1 SV=1

MALFTPWKLSSQKLGFFLVTFGFIWGMMLLHFTIQQRTQPESSSMLREQILDLSKR
YIKALAEENRNVVDGPYAGVMTAYDLKKTLAVLLDNILQRIGKLESKVDNLVVNGTG
TNSTNSTTAVPSLVALEKINVADIINGAQEKCVLPPMDGYPHCEGKIKWMKDMWRS
DPCYADYGVDGSTCSFFIYLSEVENWCPHLPWRAKNPYEEADHNSLAEIRTDFNIL
YSMMKKHEEFRWMRLRIRRMADAWIQAIKSLAEKQNLEKRKRKKVLVHLGLLTKES
GFKIAETAFSGGPLGELVQWSDLITSLYLLGHDIRISASLAELKEIMKKVVGNRSGCP
TVGDRIVELIYIDIVGLAQFKKTLGPSWVHYQCMLRVLDSFGTEPEFNHANYAQSKG
HKTPWGKWNLNPQQFYTMFPHTPDNSFLGFVVEQHLNSSDIHHINEIKRQNQSLV
YGKDSFWKNKKIYLDIIHTYMEVHATVYGSSTKNIPSYVKNHGILSGRDLQFLLRETK
LFVGLGFPYEGPAPLEAIANGCAFLNPKFNPPKSSKNTDFFIGKPTLRELTSQHPYA
EVFIGRPHVWTVDLNNQEEVEDAVKAILNQKIEPYMPYEFTCEGMLQRINAFIEKQD
FCHGQVMWPPLSALQVKLAEPGQSCKQVCQESQLICEPSFFQHLNKDKDMLKYKV

TCQSSELAKDILVPSFDPKNKHCVFQGDLLLFSCAGAHPRHQRVCPCRDFIKGQVA
LCKDCL

8.4.3 B4GALT1

>sp|P15291|B4GT1_HUMAN Beta-1,4-galactosyltransferase 1 OS=Homo sapiens OX=9606 GN=B4GALT1 PE=1 SV=5
MRLREPLLSGSAAMPGASLQRACRLLVAVCALHLGVTLVYYLAGRDLSRLPQLVGV
STPLQGGSNSAAAIGQSSGELRTGGARPPPPLGASSQPRPGGDSSPVVDSGPGP
ASNLTSVPVPHTTALSLPACPEESPLLVGPMLIEFNMPVDLELVAKQNPNVKMGGR
YAPRDCVSPHKVAIIIPFRNRQEHLKYWLYYLHPVLQRQQLDYGIYVINQAGDTIFNR
AKLLNVGFQEALKDYDYTCFVFSDVDLIPMNDHNAYRCFSQPRHISVAMDKFGFSL
PYVQYFGGVSALSKQQFLTINGFPNNYWGWGGEDDDIFNRLVFRGMSISRPNAVV
GRCRMIRHSRDKKNEPNPQRFDRIAHTKETMLSDGLNSLTYQVLDVQRYPLYTQIT
VDIGTPS

8.4.4 ST6GAL1

>sp|P15907|SIAT1_HUMAN Beta-galactoside alpha-2,6-sialyltransferase 1 OS=Homo sapiens OX=9606 GN=ST6GAL1 PE=1 SV=1
MIHTNLKKKFSCCVLVFLLFAVICVWKEKKKGSYYDSFKLQTKEFQVLKSLGKLAMG
SDSQSVSSSSTQDPHRGRQTLGSLRGLAKAKPEASFQVWNKDSSSKNLIPRLQKI
WKNYLSMNKYKVSYKGPGPGIKFSAEALRCHLRDHVNVSMVEVTDFPFNTSEWE
GYLPKESIRTKAGPWGRCAVVSSAGSLKSSQLGREIDDHDAVLRFNGAPTANFQQ
DVGTKTTIRLMNSQLVTTEKRFLKDSLYNEGILIVWDPSVYHSDIPKWYQNPDYNFF
NNYKTYRKLHPNQPFYILKPQMPWELWDILQEISPEEIQPNPPSSGMLGIIIMMTLCD
QVDIYEFLPSKRKTDVCYYYQKFFDSACTMGAYHPLLYEKNLVKHLNQGTDEDIYLL
GKATLPGFRTIHC

Appendices 8.5. Sequences for stable plant multiple HuGE subunit fusion constructs

8.5.1 MUR3-GNT4-linker-GNT5 3535 bp

ATGTTTCCAAGAGTTAGTATGAGACGTCGTAGCGCTGAAGTGAGTCCTA
CCGAGCCTATGGAGAAAGGTAATGGTAAAAATCAAACTAACCGTATTTGCCTCC
TGGTTGCACTCTCACTCTTCTTTTGGGCTCTTCTTCTTTACTTTCACTTCGTTGTG
TTGGGTACTTCCAATATTGATAAACAACTACAATTACAGCCCTCATATGCGACTA
TTGTCCAACAGTTCAAGCGAGTAGGTGCTGAGACTAATGGGTCAAAAGATGCTC
TTAACAAATTCTCAGACAACACTTTGAAACTATTGAAGGAACTTACGAGCAAGAA
GTCTTTGCAAGTTCCGTCCATTTACTACCACTTGCCGCATCTACTTAAGAATGAA
GGCAGTCTTCAGCCAGCAGTACAGATTGGAAACGGGCGGACTGGTGTCTCAAT
TGTGATGGGTATACCAACTGTGAAGCGCGAAGTTAAGTCATACTTGATTGAGAC
GCTACATAGTTTGATAGATAATTTGTATCCAGAGGAGAAGCTTGACTGCGTTATT
GTTGTTTTTATCGGTGAAACTGACATTGATTACGTTCATGGAGTCGTTGCTAATC
TCGAAAAGGAATTTTCCAAAGAGATATCTAGTGGTCTTGTTGAAGTAATCTCACC
TCCAGAATCATATTATCCAGACTTAACTAATTTAAAGGAGACATTTGGGGATTCA
AAAGAAAGAGTTCGATGGAGGACCAAGCAGAATCTTGACTACTGCTTTTTGATG
ATGTATGCGCAAGAAAAAGGAATATACTACATACAACTTGAAGACGATATTATTG
TAAAACAAAATTACTTTAACACAATTAAGAACTTCGCATTACAACTTTCTTCTGAG
GAGTGGATGATATTAGAGTTTTCTCAACTTGGCTTTATTGGAAAGATGTTTCAAG
CGCCTGACCTTACTCTTATTGTTGAGTTCATTTTCATGTTTTATAAAGAAAAACCA
ATAGATTGGCTTCTTGATCACATTTTGTGGGTCAAGGTATGTAATCCTGAGAAAG
ATGCAAAGCATTGTGATCGCCAAAAGGCTAATCTCAGGATTAGGTTTAGGCCTT
CACTTTTTCAGCATGTGGGACTGCATTCTTCTCTGTCTGGGAAGATACAAAAGC
TTACAGATAAAGACTACATGAAGCCGCTCCTGCTAAAAATCCATGTTAATCCACC
TGCTGAGGTCTCTACGTCCTTGAAGGTTTATCAAGGACACACACTCGAGAAGAC
TTACATGGGCGAAGATTTTTTTTGGGCCATAACACCTATAGCTGGGGATTATATA
CTATTCAAGTTCGACAAGCCAGTCAACGTCGAATCATACTTATTTCATTCAGGGA
ATCAAGAACACCCAGGGGATATACTGCTGAATACTACAGTTGAGGTTCTGCCAT
TTAAGTCTGAGGGTCTTGAAATAAGTAAGGAAACTAAGGATAAAAGATTGGAAG
ATGGTTATTTTCGCATTGGAAAGTTTGAGAATGGCGTTGCCGAAGGAATGGTTG
ATCCTTCTCTGAACCCTATTTCCGCCTTCAGGCTCAGCGTCATACAAAATTCTGC

TGTGTGGGCAATTCTTAATGAGATTCATATTAAGAAGGCCACAAACTGTTCTTGT
GGTTCTGGTTCTAGAACGAATAGTACAAATTCTACTACAGCAGTTCCCTCCCTG
GTTGCATTGGAGAAAATTAACGTAGCAGACATTATTAATGGCGCACAAGAAAAG
TGTGTTCTTCCTCCAATGGATGGCTATCCACACTGCGAAGGTAAAATAAAATGG
ATGAAGGATATGTGGCGGTCCGACCCATGTTATGCTGACTATGGTGTTGATGG
GAGTACATGTAGCTTTTTTATCTACTTGAGCGAGGTCGAGAATTGGTGCCCTCA
TTTGCCCTGGCGTGCCAAAAATCCTTATGAAGAAGCGGATCATAACTCATTAGC
TGAGATCCGTACTGATTTCAATATTCTGTACTCTATGATGAAAAAACACGAAGAG
TTTAGATGGATGAGGCTTAGAATTAGACGGATGGCCGATGCCTGGATCCAGGC
AATCAAAAGCCTTGCTGAGAAGCAGAATTTAGAGAAGCGTAAGCGTAAAAAGGT
GCTCGTTCATTTGGGACTGCTAACTAAAGAAAGCGGTTTCAAGATTGCAGAAAC
TGCCTTCTCTGGAGGTCCTTTGGGTGAATTGGTTCAGTGGTCTGACCTTATTAC
CTCATTATACCTGTTAGGACACGACATTCGCATCTCAGCCTCTTTGGCAGAGTT
GAAGGAAATTATGAAAAAGGTTGTGGGAAATCGAAGCGGATGCCCTACCGTAG
GTGATAGAATAGTTGAGCTTATATATATCGATATAGTTGGCCTTGCCCAGTTTAA
GAAAACACTCGGTCCTAGCTGGGTACACTATCAATGTATGTTGAGGGTTCTTGA
TAGTTTTGGGACTGAACCAGAGTTCAATCATGCAAACTATGCTCAATCTAAAGG
CCATAAAACTCCATGGGGTAAATGGAATCTGAATCCTCAACAGTTTTATACTATG
TTTCCTCACACTCCCGATAATTCATTTCTTGGATTCGTGGTAGAACAGCATCTCA
ATTCCTCAGACATTCATCACATAAATGAAATTAAGAGACAAAACCAAAGTTTGGT
GTACGGGAAAGTTGATTCTTTTTGGAAAAACAAGAAGATTTACCTTGACATTATT
CACACTTATATGGAAGTTCACGCAACTGTTTACGGCTCTTCAACTAAAAACATCC
CAAGTTACGTGAAGAATCACGGGATTTTGAGCGGACGAGACCTCCAGTTTTTGT
TACGCGAAACTAAGCTCTTTGTAGGTCTCGGTTTCCCCTACGAGGGCCCCGCT
CCATTGGAAGCTATTGCTAACGGATGTGCCTTTCTTAACCCAAAGTTCAACCCT
CCAAAAAGTAGTAAAAACACAGATTTTTTTATAGGGAAGCCAACTCTGAGGGAG
TTGACATCCCAACATCCGTACGCGGAAGTCTTCATTGGACGACCACATGTGTGG
ACTGTCGATCTTAACAATCAAGAAGAAGTCGAAGATGCTGTCAAGGCCATATTA
AATCAGAAAATAGAACCATATATGCCTTACGAATTTACATGTGAGGGTATGCTTC
AACGAATAAACGCATTTATAGAGAAGCAGGATTTCTGCCACGGACAGGTAATGT
GGCCGCCACTAAGCGCATTACAGGTTAAGCTAGCTGAACCTGGCCAATCTTGT
AAGCAGGTGTGTCAAGAATCTCAATTAATATGTGAGCCAAGTTTTTTTCAGCATT
TGAACAAAGATAAAGACATGCTAAAATACAAAGTTACATGCCAAAGTTCTGAATT
AGCTAAGGATATTCTAGTGCCATCATTTGATCCTAAGAATAAGCACTGCGTCTTC

CAAGGTGACCTATTGTTATTCAGCTGTGCCGGGGCACACCCTAGACACCAACG
AGTATGCCCGTGCCGTGATTTCATTAAAGGCCAGGTAGCTCTGTGCAAAGACTG
TTTA

8.5.2 FUT13-B4GALT1-linker-ST6GAL 1807 bp

ATGCCAATGAGATATTTAAATGCTATGGCAGCACTGTTAATGATGTTTTTTACTC
TTCTTATTCTATCATTTACAGGAATTCTAGAGTTCCCCTCTGCCTCTACATCAAT
GGAGCACTCTATAGATCCTGAGCCTAAATTGAGTGATAGTACTAGTCTTCCAGC
TTGTCCAGAAGAATCTCCTCTCCTTGTAGGACCTATGCTGATTGAATTTAACATG
CCTGTAGACCTTGAACTTGTTGCAAAGCAAAATCCAAATGTGAAGATGGGTGGA
AGGTACGCTCCAAGAGATTGCGTCTCCCCTCATAAAGTGGCAATAATAATTCCG
TTTAGAAACCGGCAAGAACATCTAAAGTACTGGCTATATTATCTTCACCCGGTTT
TACAGAGACAACAATTAGATTACGGAATTTACGTTATTAATCAGGCGGGAGATA
CTATATTTAACAGAGCCAAGTTACTGAATGTGGGGTTTCAGGAGGCATTGAAGG
ATTACGACTACACTTGTTTTGTCTTTTCAGATGTTGACCTTATACCAATGAATGAT
CATAATGCTTATCGATGTTTCTCTCAACCAAGACATATCTCTGTGGCCATGGATA
AGTTCGGGTTTTCCTTACCGTACGTGCAGTATTTTGGTGGCGTATCAGCACTGA
GCAAACAACAATTCCTCACTATCAACGGTTTTCCCAACAATTATTGGGGATGGG
GAGGAGAGGATGATGACATATTCAATAGATTGGTGTTTCGTGGTATGTCCATTT
CCAGACCAAATGCAGTGGTAGGAAGATGCCGAATGATCAGGCACTCTAGAGAT
AAAAAAAATGAACCAAATCCTCAAAGATTTGACAGGATCGCTCATACCAAAGAAA
CTATGCTAAGTGATGGTTTAAACTCTCTGACGTACCAAGTTCTGGATGTCCAGA
GGTACCCATTGTATACACAAATAACAGTCGATATTGGAACACCTTCATGTTCTTG
TGGTTCTGGTTCTAGAGCATCTTTCCAGGTTTGGAATAAAGACTCTTCTTCTAAG
AATTTGATCCCTAGATTACAGAAAATTTGGAAAAATTACCTTAGCATGAATAAGT
ACAAAGTGTCATATAAAGGACCAGGTCCTGGAATTAAGTTCTCAGCCGAAGCTC
TTAGATGCCACTTGCGTGATCATGTAAATGTGAGTATGGTGGAGGTTACCGATT
TTCCTTTCAATACTAGTGAGTGGGAAGGCTACCTTCCAAAAGAATCTATAAGGA
CTAAAGCAGGACCATGGGGAAGGTGTGCCGTGGTGAGTTCTGCTGGAAGTTTG
AAGTCATCCCAGCTGGGTAGAGAAATTGATGATCATGACGCAGTCCTTAGATTT
AATGGTGCGCCAACAGCAAACTTTCAACAGGATGTAGGTACAAAAACAACAATT
AGACTTATGAACTCACAACTGGTTACTACCGAGAAGAGATTTTTGAAGGACAGC
TTGTACAATGAAGGAATATTAATAGTATGGGATCCCTCAGTTTACCATAGCGATA
TCCCAAAGTGGTATCAAAACCCGGATTACAATTTTTTCAACAATTATAAAACTTAT
AGAAAGCTCCACCCCAATCAACCATTTTATATTCTTAAACCTCAAATGCCATGGG
AGCTTTGGGATATTCTACAAGAGATTTCCCCTGAAGAAATCCAACCTAATCCAC
CCAGCTCCGGAATGCTAGGTATCATTATTATGATGACTCTGTGTGATCAAGTTG
ATATCTATGAATTCCTTCCATCAAAACGTAAGACAGACGTTTGCTATTACTATCA
GAAGTTCTTC

Appendices 8.6. Sequences for proteins of interest

8.6.1 LAL

ATGAAAATGCGGTTTTTGGGTCTTGTTGTGTGCTTAGTACTCTGGACTCTGCAC
AGTGAGGGCTCAGGTGGAAAGTTGACAGCAGTGGACCCTGAAACTAATATGAA
TGTTTCTGAAATTATAAGTTACTGGGGGTTCCCATCAGAAGAATATCTTGTTGAG
ACTGAGGATGGATACATTCTTTGTTTGAATCGTATTCCACACGGGCGTAAGAAT
CACTCCGATAAGGGACCAAAACCTGTGGTGTTTCTCCAACATGGGTTGCTCGCT
GATTCATCTAATTGGGTTACCAACCTCGCGAACAGTTCCCTTGGTTTCATCTTAG
CAGATGCAGGTTTCGACGTCTGGATGGGCAATAGTCGGGGAAACACCTGGTCT
AGAAAGCACAAGACACTTTCCGTAAGCCAGGATGAGTTTTGGGCTTTTTCATAC
GATGAGATGGCTAAGTATGACTTGCCGGCTTCAATAAATTTTATCTTAAATAAGA
CTGGTCAAGAACAGGTTTACTATGTTGGGCACTCCCAAGGAACTACTATTGGAT
TTATTGCTTTCTCTCAGATTCCAGAGCTTGCAAAACGCATCAAAATGTTCTTCGC
ACTGGGACCTGTAGCTTCAGTTGCTTTCTGTACCAGCCCGATGGCAAAATTGGG
ACGTTTGCCTGACCACTTGATCAAAGATCTTTTTGGGGACAAGGAGTTTCTACC
TCAATCCGCTTTTCTTAAGTGGCTTGGTACTCATGTGTGCACTCATGTTATTTTG
AAAGAACTCTGCGGGAATCTTTGTTTCCTCTTGTGCGGATTTAATGAAAGGAATT
TAAATATGTCAAGAGTTGACGTTTATACAACACATTCACCAGCCGGAACCAGTG
TTCAAAATATGCTTCACTGGAGCCAAGCCGTAAAATTCCAGAAGTTCCAAGCAT
TCGATTGGGGGTCAAGTGCAAAGAATTATTTTCATTACAATCAGTCTTATCCACC
AACGTATAACGTTAAGGACATGTTGGTACCTACCGCTGTTTGGAGTGGAGGTCA
TGACTGGCTGGCCGACGTTTATGATGTTAACATCTTGCTTACACAAATTACTAAT
CTAGTTTTCCATGAATCAATTCCGGAGTGGGAACACCTTGATTTCATCTGGGGC
CTTGATGCTCCATGGCGCCTGTACAATAAAATTATTAACCTTATGCGCAAGTATC
AG

8.6.2 LALΔ1-22

GGTGGAAAGTTGACAGCAGTGGACCCTGAAACTAATATGAATGTTTCTGAAATT
ATAAGTTACTGGGGGTTCCCATCAGAAGAATATCTTGTTGAGACTGAGGATGGA
TACATTCTTTGTTTGAATCGTATTCCACACGGGCGTAAGAATCACTCCGATAAG
GGACCAAAACCTGTGGTGTTTCTCCAACATGGGTTGCTCGCTGATTCATCTAAT
TGGGTTACCAACCTCGCGAACAGTTCCCTTGGTTTCATCTTAGCAGATGCAGGT
TTCGACGTCTGGATGGGCAATAGTCGGGGAAACACCTGGTCTAGAAAGCACAA
GACACTTTCCGTAAGCCAGGATGAGTTTTGGGCTTTTTCATACGATGAGATGGC
TAAGTATGACTTGCCGGCTTCAATAAATTTTATCTTAAATAAGACTGGTCAAGAA
CAGGTTTACTATGTTGGGCACTCCCAAGGAACTACTATTGGATTTATTGCTTTCT
CTCAGATTCCAGAGCTTGCAAAACGCATCAAAATGTTCTTCGCACTGGGACCTG
TAGCTTCAGTTGCTTTCTGTACCAGCCCGATGGCAAAATTGGGACGTTTGCCTG
ACCACTTGATCAAAGATCTTTTTGGGGACAAGGAGTTTCTACCTCAATCCGCTTT
TCTTAAGTGGCTTGGTACTCATGTGTGCACTCATGTTATTTTGAAAGAACTCTGC
GGGAATCTTTGTTTCCTCTTGTGCGGATTTAATGAAAGGAATTTAAATATGTCAA
GAGTTGACGTTTATACAACACATTCACCAGCCGGAACCAGTGTTCAAAATATGC

TTCACTGGAGCCAAGCCGTAAAATTCCAGAAGTTCCAAGCATTCGATTGGGGGT
CAAGTGCAAAGAATTATTTTCATTACAATCAGTCTTATCCACCAACGTATAACGT
TAAGGACATGTTGGTACCTACCGCTGTTTGGAGTGGAGGTCATGACTGGCTGG
CCGACGTTTATGATGTTAACATCTTGCTTACACAAATTACTAATCTAGTTTTCCAT
GAATCAATTCCGGAGTGGGAACACCTTGATTTCATCTGGGGCCTTGATGCTCCA
TGGCGCCTGTACAATAAAATTATTAACCTTATGCGCAAGTATCAG

8.6.3 IgG heavy chain

ATGGATTGGACCTGGAGGTTTCTATTTGTAGTGGCCGCTGCTACTGGAGTTCAG
TCTCAAATGCAAGTGGTTCAATCTGGCGCTGAGGTAAAGAAACCGGGAAGTAG
TGTCACTGTAAGCTGCAAGGCTAGTGGGGGAACATTCTCAAACTATGCTATAAG
CTGGGTCCGCCAGGCACCTGGCCAAGGCCTTGAATGGATGGGAGGTATAATCC
CGCTTTTTGGAACCCCAACCTATAGCCAAAATTTCCAAGGCCGAGTCACAATTA
CTGCAGATAAGTCTACTAGCACTGCTCATATGGAATTAATTTCATTGCGTTCTGA
GGATACCGCAGTATACTACTGCGCTACGGATAGGTATCGCCAGGCAAACTTCG
ACAGAGCAAGGGTTGGTTGGTTTGATCCATGGGGCCAGGGTACGTTAGTGACA
GTTTCATCTGCTTCAACTAAAGGACCCTCTGTTTTTCCATTAGCGCCATCTTCTA
AGTCTACATCTGGAGGTACAGCCGCTCTTGGATGTTTGGTAAAGGATTACTTTC
CAGAACCTGTGACTGTATCTTGGAACTCTGGAGCACTTACTAGCGGTGTCCACA
CTTTTCCTGCAGTACTCCAAAGTTCCGGGTTATATTCACTCTCAAGTGTAGTGAC
CGTGCCTTCTTCTTCATTAGGAACGCAGACATATATCTGTAATGTAAACCACAAA
CCAAGTAACACAAAGGTGGATAAGAAGGTTGAACCAAAGTCTTGCGACAAAACT
CATACTTGTCCTCCATGTCCAGCGCCGGAACTCCTGGGAGGCCCATCTGTTTTT
TTGTTCCCTCCAAAGCCTAAAGATACCTTAATGATATCAAGAACGCCAGAGGTG
ACTTGTGTAGTTGTTGATGTTAGTCACGAGGATCCTGAAGTTAAGTTTAATTGGT
ACGTCGATGGTGTAGAAGTGCATAATGCAAAGACAAAGCCTAGGGAAGAACAA
TACAATAGCACATATAGGGTGGTCTCTGTCCTAACTGTTCTTCACCAAGATTGG
CTTAATGGAAAGGAATATAAATGTAAAGTTTCAAATAAAGCTCTACCTGCTCCTA
TTGAGAAAACTATTTCTAAGGCTAAGGGTCAACCTAGGGAGCCACAAGTCTATA
CCCTACCTCCAAGCAGGGATGAGTTGACTAAAAATCAAGTTTCTCTCACCTGTC
TGGTGAAGGGATTTTATCCGTCTGATATCGCTGTAGAGTGGGAGTCTAACGGC
CAACCTGAAAATAATTATAAGACAACACCACCCGTTTTGGATAGCGATGGTTCTT
TTTTTTTGTACAGTAAGTTGACAGTCGACAAGTCTCGGTGGCAACAGGGTAACG
TTTTCTCCTGCTCAGTAATGCACGAAGCTTTACATAATCATTATACACAAAAGTC
TCTTTCACTTTCCCCTGGAAAGC

8.6.4 IgE heavy chain

GCCTCAACTCAGAGTCCCTCAGTTTTTCCGCTTACTAGATGCTGTAAGAATATTC
CATCTAATGCAACTTCTGTCACACTAGGATGTTTAGCAACAGGATACTTTCCCGA
ACCTGTAATGGTTACTTGGGACACGGGTTCTTTAAATGGGACTACAATGACCCT
ACCTGCTACAACATTGACACTTTCTGGCCATTACGCAACTATCTCATTACTCACA
GTGTCCGGAGCTTGGGCAAAACAAATGTTTACTTGTCGTGTTGCACACACTCCT
TCATCAACTGACTGGGTGGACAATAAAACTTTTTCTGTGTGTTCAAGAGATTTTA
CTCCACCAACTGTGAAGATCCTGCAATCCAGTTGTGATGGTGGGGGACATTTCC
CACCTACGATACAATTGTTATGTCTTGTATCAGGATACACACCCGGAACAATTAA
TATTACTTGGCTCGAAGACGGTCAAGTAATGGATGTCGATCTTTCCACCGCATC
AACCACTCAAGAAGGAGAGTTAGCTTCCACACAAAGTGAGCTTACTCTTTCTCA
AAAACATTGGCTATCCGACCGAACATATACATGTCAAGTTACATATCAAGGTCAC
ACTTTTGAAGATAGTACTAAAAAGTGCGCTGATTCTAATCCTCGGGGGGTTAGT
GCGTATCTTTCACGTCCAAGTCCTTTCGATCTCTTCATTAGGAAGTCCCCTACCA
TAACATGTCTTGTGGTTGATCTTGCCCCTTCCAAGGGAACTGTGAATCTCACGT
GGTCACGAGCTTCAGGGAAGCCAGTGAACCATTCTACTAGGAAGGAAGAAAAG
CAGCGCAATGGTACTCTTACAGTTACATCAACCCTTCCCGTAGGGACAAGAGAT
TGGATCGAAGGCGAAACATACCAATGTCGAGTGACTCACCCACATTTACCAAGA
GCCTTGATGCGCAGTACAACAAAAACATCAGGGCCTCGAGCCGCACCAGAGGT
TTATGCCTTCGCAACACCAGAGTGGCCTGGTAGTAGGGATAAACGCACTCTTG
CGTGCTTGATTCAAAACTTTATGCCTGAGGACATATCAGTGCAATGGCTTCATAA
CGAAGTGCAGTTGCCTGATGCTAGACACTCCACAACCCAACCAAGAAAGACTAA
AGGATCTGGATTTTTCGTTTTCAGTAGACTCGAAGTAACACGGGCCGAATGGGA
GCAGAAGGATGAGTTTATATGTAGAGCCGTACATGAGGCTGCCAGTCCAAGTC
AAACTGTTCAAAGAGCTGTCTCTGTAAACCCAGGGAAAC

9. Publications

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 07 October 2022
DOI 10.3389/fpls.2022.1010569

OPEN ACCESS

EDITED BY
Allison Van De Meene,
The University of Melbourne, Australia

REVIEWED BY
Miharu Maeda,
Akita University, Japan
Ulla Neumann,
Max Planck Institute for Plant Breeding Research, Germany

*CORRESPONDENCE
Verena Kriechbaumer
vkriechbaumer@brookes.ac.uk

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to Technical Advances in Plant Science, a section of the journal Frontiers in Plant Science

RECEIVED 03 August 2022
ACCEPTED 14 September 2022
PUBLISHED 07 October 2022

CITATION
McGinness AJ, Schoberer J, Pain C, Brandizzi F and Kriechbaumer V (2022) On the nature of the plant ER exit sites.
Front. Plant Sci. 13:1010569.
doi: 10.3389/fpls.2022.1010569

On the nature of the plant ER exit sites

Alastair J. McGinness[1†], Jennifer Schoberer[2†], Charlotte Pain[1], Federica Brandizzi[3,4,5] and Verena Kriechbaumer[1*]

[1]Endomembrane Structure and Function Research Group, Department of Biological and Medical Sciences, Oxford Brookes University, Oxford, United Kingdom, [2]Department of Applied Genetics and Cell Biology, Institute of Plant Biotechnology and Cell Biology, University of Natural Resources and Life Sciences, Vienna, Austria, [3]MSU-DOE Plant Research Lab, Michigan State University, East Lansing, MI, United States, [4]Great Lakes Bioenergy Research Center, Michigan State University, East Lansing, MI, United States, [5]Department of Plant Biology, Michigan State University, East Lansing, MI, United States

In plants, the endoplasmic reticulum (ER) and Golgi bodies are not only in close proximity, but are also physically linked. This unique organization raises questions about the nature of the transport vectors carrying cargo between the two organelles. Same as in metazoan and yeast cells, it was suggested that cargo is transported from the ER to Golgi cisternae *via* COPII-coated vesicles produced at ribosome-free ER exit sites (ERES). Recent developments in mammalian cell research suggest, though, that COPII helps to select secretory cargo, but does not coat the carriers leaving the ER. Furthermore, it was shown that mammalian ERES expand into a tubular network containing secretory cargo, but no COPII components. Because of the close association of the ER and Golgi bodies in plant cells, it was previously proposed that ERES and the Golgi comprise a secretory unit that travels over or with a motile ER membrane. In this study, we aimed to explore the nature of ERES in plant cells and took advantage of high-resolution confocal microscopy and imaged ERES labelled with canonical markers (Sar1a, Sec16, Sec24). We found that ERES are dynamically connected to Golgi bodies and most likely represent pre-*cis*-Golgi cisternae. Furthermore, we showed fine tubular connections from the ER to Golgi compartments (ERGo tubules) as well as fine protrusions from ERES/Golgi cisternae connecting with the ER. We suggest that these tubules observed between the ER and Golgi as well as between the ER and ERES are involved in stabilizing the physical connection between ER and ERES/Golgi cisternae, but may also be involved in cargo transport from the ER to Golgi bodies.

KEYWORDS

endoplasmic reticulum, Golgi body, exit site, ERES, Sec16, Sar1a, Sec24, ER exit sites

Introduction

In plants, the Golgi apparatus (GA) comprises many distinct stacks of membrane-bounded cisternae, approximately 1 micron in diameter that tend to be distributed around the cytoplasm. In many cell types, especially highly vacuolated cells, Golgi stacks are motile (Boevink et al., 1998) and closely associated with the endoplasmic reticulum (ER). This close association has been demonstrated by the application of optical laser tweezers, which can trap individual Golgi bodies and pull strands of attached ER behind them (Sparkes et al., 2009). Attachment to the ER however may not occur *via* membrane connections, such as hemifusion, as it has been shown that AtCASP, a member of the golgin family of proteins, may act as a proteinaceous tether between the ER and Golgi (Osterrieder et al., 2017). Due to this very close association between the ER and Golgi bodies, it was proposed that the cargo exit sites on the ER surface (ER exit sites, ERES) and the Golgi comprise a secretory unit that travels over or with a motile ER membrane (daSilva et al., 2004). The close association between the ER and Golgi bodies raises the question of the nature of the transport vectors carrying cargo between the two organelles.

It is generally accepted that most cargo from the ER reaches the Golgi *via* mechanisms dependent on the COPII machinery (for recent reviews see (Brandizzi, 2018; Kurokawa and Nakano, 2019; Robinson, 2020; Kriechbaumer and Brandizzi, 2020). COPII proteins are a set of functionally conserved proteins such as the cytosolic small GTPase Sar1, the ER membrane protein guanine nucleotide exchange factor (GEF) Sec12 and components of the COPII coat Sec23/24 and Sec13/31. Retrograde transport back from the Golgi bodies to the ER well as retrograde transport within the Golgi apparatus is facilitated by the COPI machinery (Oprins et al., 1993; Arakel and Schwappach, 2018; Robinson, 2020). COPI vesicles were identified by immunogold labelling of cryosections using antibody-labelling of COPI components such as the GTPase ARF1 and different coatomer subunits (Orci et al., 2000; Pimpl et al., 2000; Donohoe et al., 2007; Langhans et al., 2012) but COPII vesicles have been rarely observed (Langhans et al., 2012), raising the question on the identity of the carriers operating in ER-to-Golgi transport.

Indeed, it appears that several types of carriers (tubules, saccules, tunnels, and coated vesicles) may coexist and operate along the ER-to-Golgi route allowing for a diverse, flexible and robust trafficking system. For instance in mammalian cells, using 3D electron-tomography (Zeuschner et al., 2006) it was shown that ERES are continuous with the ER and capable of adapting in size depending on the secretory load (Farhan et al., 2008; Boncompain et al., 2012; Weigel et al., 2021; Shomron et al., 2021). It was previously suggested by the metazoan and yeast research communities that the anterograde vectors for general cargo are COPII-coated vesicles produced at ribosome-free ERES and COPII coated tubules could carry larger proteins (Gorur et al., 2017). COPII vesicles and tubules would then deliver cargo directly to the *cis*-Golgi in yeast (Matsuura-Tokita et al., 2006) or to an intermediate vesiculo-tubular compartment that is transported over varying distances to the *cis*-Golgi in animal cells (Tang et al., 2005). However, various reports suggested that ER-to-Golgi transport requires both vesicular and tubular compartments to sort the ERES cargo to the Golgi apparatus (Saraste and Svensson, 1991; Bannykh et al., 1996; Xu and Hay, 2004; Watson and Stephens, 2005). Assays with temperature-sensitive fluorescently tagged cargo also showed that cargo colocalizes with COPII components at the ERES but does not exit the ERES with the COPII coat component Sec24, suggesting that at least parts of Golgi trafficking might be COPII independent (Presley et al., 1997; Stephens et al., 2000). Indeed, it was reported that COPII components remain on the ERES and that cargo is transported in Rab1-dependent carriers that are not labelled by COPII proteins (Westrate et al., 2020). These data highlight that vesicle carriers are not the only means for anterograde transport and highlight the possibility that COPII is rather involved in secretory cargo selection than its actual transport.

In support of this, recent data in mammalian cells showed that ERES expand into a tubular network containing secretory cargo but no COPII components. COPII components were found to decorate the neck of these tubules, implicating that the departing transport carrier is not coated (Weigel et al., 2021; Shomron et al., 2021). Based on the findings that COPII does not coat the carriers leaving the ER, it was suggested that COPII helps solely to select secretory cargo.

The plant trafficking community has generally embraced the concept of COPII vesicle vectors, due to the presence of all the structural proteins required for COPII coat construction (Kang and Staehelin, 2008). However, there is little microscopy evidence for the existence of such vesicles (Langhans et al., 2012), and due to the limited space between the ER and *cis*-Golgi an alternative hypothesis based on direct tubular connections between the two organelles has been proposed and debated over the years (Hawes et al., 2008; Hawes, 2012; Robinson et al., 2015; Kriechbaumer and Brandizzi, 2020).

Here we show that by utilizing high-resolution microscopy in live plant cells and by considering geometry of the donor and acceptor membranes, ERES are associated with the *cis*-face of Golgi cisternae rather than associated with the ER surface. Moreover, multiple small ER-derived tubules are associated with most Golgi cisternae. Thus, in plants, the nature of the ER-Golgi interface and even the nature of ERES in their currently accepted form should be re-evaluated.

Materials and methods

Plant growth and Agrobacterium-mediated transient protein expression

Nicotiana tabacum (SR1 cv Petit Havana) plants were grown for transient Agrobacterium-mediated expression, as previously reported (Sparkes et al., 2006). In brief, transformed agrobacteria were pelleted by centrifugation at 1800 g at room temperature for 5 min. Infiltration buffer (5 mg ml^{-1} glucose, 50 mM MES, 2 mM $Na_3PO_4{\cdot}12H_2O$ and 0.1 mM acetosyringone) was used to wash the pellet once and then to resuspend the agrobacteria in 1 ml. The bacterial suspension was diluted in the infiltration buffer to an OD_{600} of 0.1. The final dilution of the infiltration medium was injected through the stomata on the underside of the tobacco leaf using a 1 ml syringe. Infiltrated plants were kept at 22°C for 72 hours prior to imaging.

Arabidopsis thaliana (Col-0) plants expressing GFP-HDEL and ST-mRFP in a stable manner were briefly sterilized with 70% ethanol and placed on plates with ½ strength Murashige and Skoog medium. The plates were stratified at 4°C in the dark for 3 days before being transferred to a plant incubator at 22°C with 16 hours light and 8 hours dark.

Confocal microscopy

Transformed leaf epidermal samples were imaged using a Zeiss PlanApo ×100/1.46 NA oil immersion objective on a Zeiss LSM880 confocal equipped with Airyscan detector. Typically 512 × 512 images and time-lapse sequences were collected in 8-bit with 2-line averaging at an (x,y) pixel spacing of 20–80 nm with excitation at 488 nm (GFP) and 561 nm (RFP), and emission at 495–550 nm and 570–615 nm, respectively.

Distances between markers were analyzed using Zen line profiles, measuring the distance between the maximum of the corresponding fluorescence peaks. Significance was calculated by Kruskal-Wallis (*p< 0.05; **p< 0.01; ***p< 0.001).

Drug treatments

Segments (roughly 5 mm^2) of transformed leaves were used for drug treatment, confocal imaging, and analysis. For inhibition of actin polymerization, leaf tissue was submerged in 25 µM latrunculin B (stock solution: 10 mM in DMSO) for 30 mins. BFA (stock solution: 10 mg ml^{-1} in DMSO) was used at a concentration of 100 µg ml^{-1} (Brandizzi et al., 2002). BFA washes were performed as described earlier (Brandizzi et al., 2002). All stock solutions were kept at -20°C, working solutions were prepared fresh just before use.

Results

High-resolution imaging of the ER-Golgi interface in live plant cells

To obtain high-resolution imaging, we implemented a Zeiss LSM880 confocal microscope equipped with Airyscan detector, which enables up to 120 nm lateral resolution. We imaged cells coexpressing markers for the ER lumen (GFP-HDEL; (Brandizzi et al., 2003)) and the medial/*trans* cisternae of Golgi bodies (sialyltransferase (ST)-mRFP; (Runions et al., 2006)). Proteins were first expressed transiently in tobacco leaf epidermal cells (Sparkes et al., 2006).

We found that Golgi bodies are motile and associated with ER tubules (Figure 1A; Supplementary movie S1) or curved

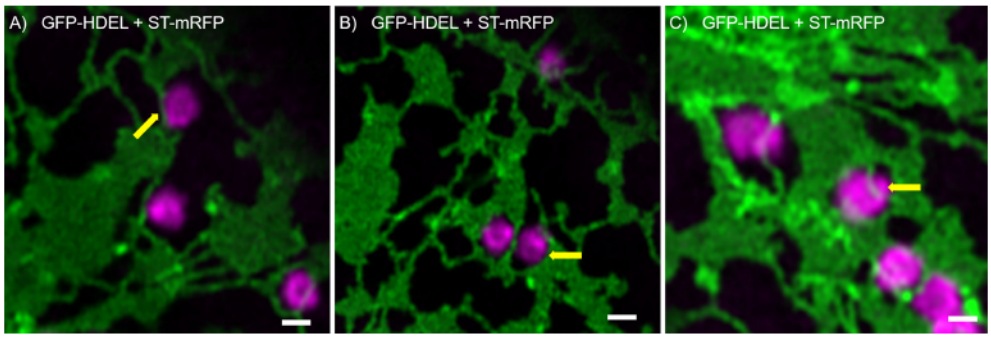


FIGURE 1
ER and Golgi are in close association as demonstrated by high-resolution microscopy. The ER marker GFP-HDEL (in green) and the Golgi body marker ST-mRFP (in magenta) are transiently expressed in tobacco leaf epidermal cells. Golgi bodies are associated with ER tubules (**A**), yellow arrow) and cisternal edges (**B**, yellow arrow) and can sit in small lacunae within ER cisternae (**C**, white arrow). Size bars = 1 µm.

edges of cisternae (Figure 1B) or may be trapped within small lacunae in ER cisternae (Figure 1C). Although connections between ER and Golgi bodies have been described before (daSilva et al., 2004; Sparkes et al., 2009; Osterrieder et al., 2017; Vieira et al., 2020), our results indicate that we can visualize the spatial distribution of the ER and Golgi in live cells at a level of resolution that has not been reported before.

Fine dynamic tubules connect ER and Golgi bodies

Looking at the dynamics and the shaping of Golgi bodies with high-resolution live-cell microscopy, we found that all Golgi bodies visualized produce small membrane protrusions from their surface that are very dynamic but always make contact with or are wrapped around ER tubules (Figure 2), allowing for a constant physical but highly dynamic connection between the two organelles.

Furthermore, fine ER tubules were seen to thread through Golgi bodies (Figure 3A time series, Figure 3B) and wrap themselves around Golgi bodies (Figures 3A, C). These thin tubules average about 20 nm diameter, considerably smaller compared to 40 nm diameter for the average ER tubule (Pain et al., 2019). Almost every Golgi body imaged appears to be attached to the ER *via* one or two of these ER-derived tubules. These tubules can also connect up to Golgi body protrusions (Figure 3D), but these are rare event to image and were only observed in approximately 1:1000 Golgi bodies. The morphology and location of these fine ER-derived and HDEL-labelled tubules distinguishes them from the tubules of the bulk ER network. To clearly refer to them, we termed them ERGo (ER-Golgi) tubules in the further text. A schematic representation of ERGo tubules and Golgi body protrusions is depicted in Figure 4A.

To confirm that these tubules were not limited to transient expression in tobacco leaf cells, we also imaged the ER and Golgi body markers HDEL and ST in stable expression in *Arabidopsis thaliana* plants. Also in this system, the dynamic morphology with ERGo tubules (Figure 5A) and Golgi protrusions (Figure 5B) was observed (Figure 5; Supplementary movie S2).

To investigate if ER dynamics impacts on ER-Golgi tubules, the drug latrunculin B (LatB) was applied (Figure 6). LatB inhibits actin polymerization and interrupts ER dynamics, resulting in immobile Golgi bodies. LatB treatment also resulted in the enhanced formation of ER cisternae (Pain et al., 2019). After LatB treatment, ERGo tubules were clearly visible but we could not determine if they remained through LatB treatment or were formed *de novo*. *De novo* formation of ERGo tubules would indicate that they are not only a physical connection required when ER and Golgi bodies are rapidly moving in the cell.

Distribution of ER exit site markers and Golgi glycosylation enzymes

To further investigate the physical link between ER and Golgi bodies, we wanted to look at a structure located between the ER and Golgi, the ERES. Interestingly, protrusions similar to those previously shown with Golgi markers (Figures 2, 5) were also observed when ERES markers/COPII components were expressed together with the ER marker RFP-HDEL (Wang et al., 2014) (Figures 4B, 7). We used functional fluorescent fusion markers of the COPII coat proteins GFP-Sec24 (Faso et al., 2009), GFP-Sec16 (Takagi et al., 2013) and Sar1a-GFP (Hanton et al., 2008). We found that the structures labelled with these markers showed dynamic protrusions in most cases wrapped around ER tubules or connecting with ER tubules (Figure 7) comparable to the protrusions labelled with the Golgi marker (Figures 1–5).

It is also of great interest that ERES markers/COPII components did not label the ER as one would expect for ER

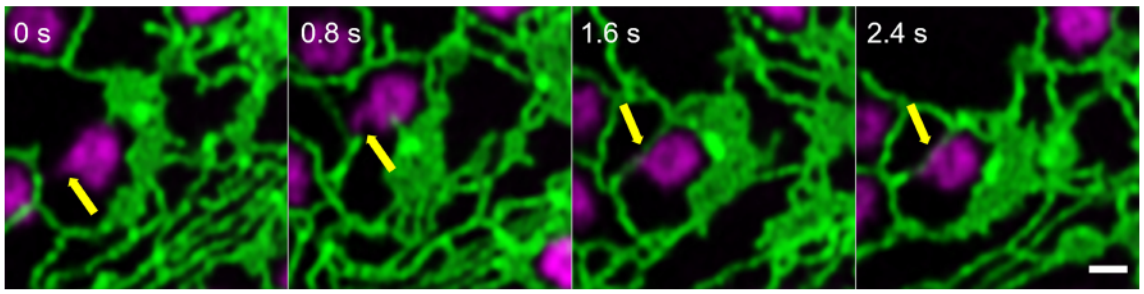


FIGURE 2
Identification of dynamic Golgi-derived tubules that connect with the ER. Golgi bodies labelled with the medial/*trans*-Golgi marker ST-mRFP (in magenta) form protrusions (yellow arrow) that can also be wrapped around ER tubules (labelled with GFP-HDEL in green). Markers are expressed transiently in tobacco leaf epidermal cells. Images are shown in time intervals of 0.8 s between the frames and time points are indicated. Size bar = 1 µm.

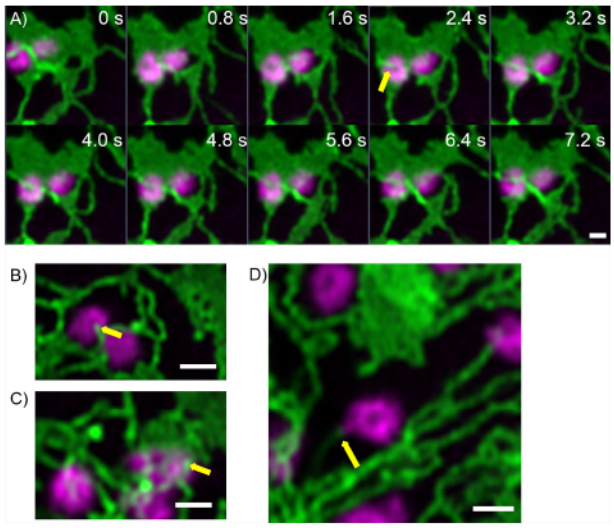


FIGURE 3
Uniquely shaped ER tubules connect with Golgi bodies. Fine ER tubules thread through (**A**, **B** yellow arrows) and around the Golgi bodies (**C** yellow arrow) and can connect with Golgi protrusions (**D**, yellow arrow). Golgi bodies are labelled with the medial/*trans*-Golgi marker ST-mRFP (in magenta) and the ER is labelled with GFP-HDEL in green. Markers are expressed transiently in tobacco leaf epidermal cells. Images for **A**) are shown in time intervals of 0.8 s between the frames and time points are indicated. Size bars = 1 μm.

vesicle budding sites but are closely and dynamically aligned with Golgi cisternae. We next aimed to gain deeper insights on the spatial relationship of the *cis*-Golgi and the COPII markers using high-resolution microscopy. When coexpressed with the *cis*-Golgi body marker MNS1 (Liebminger et al., 2009; Schoberer et al., 2014), the ERES markers GFP-Sec24 (Figure 8A), GFP-Sec16 (Figure 8B) and Sar1a-GFP (Figure 8C) labelled a compartment close to MNS1-mRFP, however their localization did not fully overlap (Figure 8). As a control, MNS1-mRFP was coexpressed with MNS1-eGFP (Figure 8D). The distance between the three tested ERES markers and the *cis*-Golgi marker MNS1 was statistically significantly larger than the distance between MNS1-mRFP and MNS1-eGFP (Figure 8E). This indicates that in plant cells the labelled ERES structures may represent a pre-*cis*-Golgi compartment with ERGo tubules and protrusions allowing for physical connections to the ER (Figure 4B).

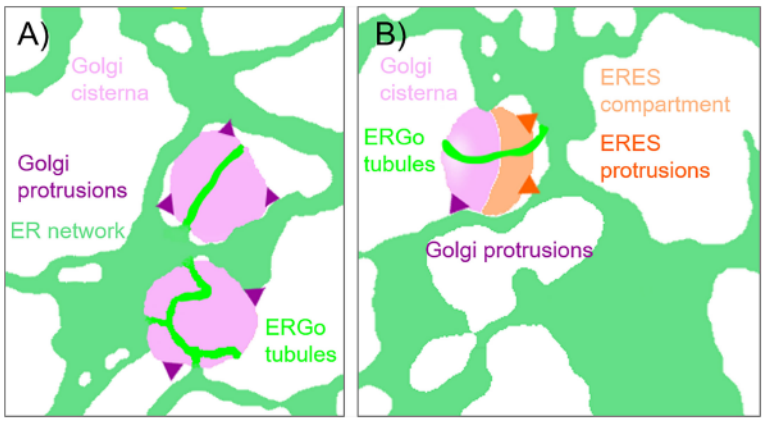


FIGURE 4
Schematic representation of ER-Golgi connections. The diagram indicates the various connections observed between ER and Golgi bodies: **(A)** ERGo tubules (dark green) labeled by the ER lumenal marker GFP-HDEL are connecting ER (light green) and Golgi cisternae (light purple). Dynamic protrusions from Golgi bodies (purple) are labelled with the medial/*trans*-Golgi body marker ST-RFP. **(B)** Protrusions from ERES to ER (dark orange) can also be observed with the ERES/COPII components Sec24, Sec16 and Sar1a labelling a pre-*cis*-Golgi compartment (light orange).

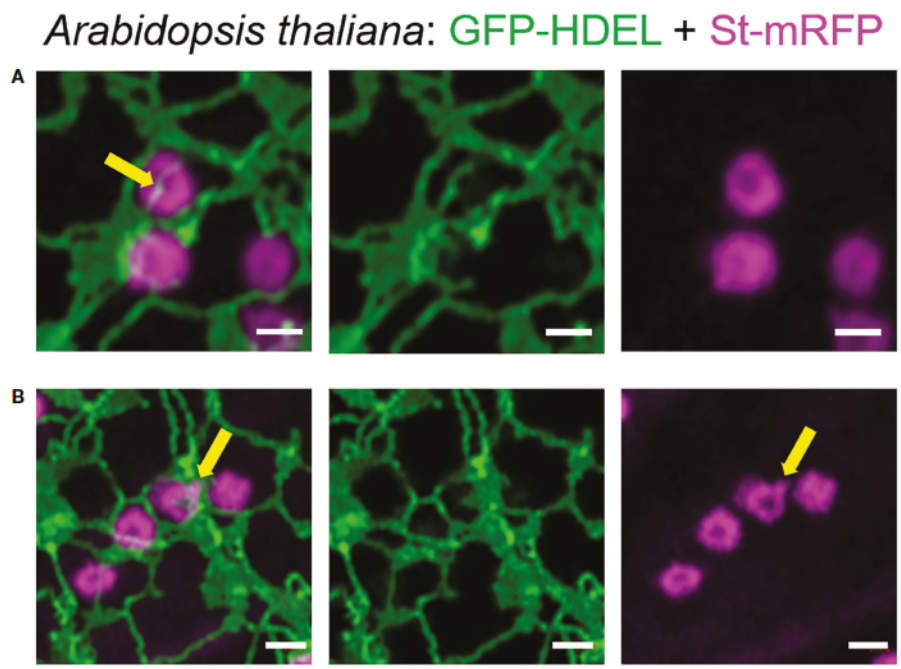


FIGURE 5
The organization of the ER and Golgi shown in tobacco can also be observed in Arabidopsis transgenic plants. ER tubules wrapped around and through the Golgi body (**A**, yellow arrow) and dynamic protrusions from the Golgi body (**B**, yellow arrow) were also observed in stably transformed *Arabidopsis thaliana* plants. Golgi bodies were labelled with the medial/*trans*-Golgi marker ST-mRFP (in magenta) and the ER is labelled with GFP-HDEL in green. Size bars = 1 µm.

To further investigate a potential pre-*cis*-Golgi compartment, we coexpressed the ERES markers Sec24 (Figure 9A), Sec16 (Figure 9B) and Sar1a (Figure 9C) with MNS3, the Arabidopsis ER-α-mannosidase I (MNS3; (Schoberer et al., 2019)). MNS3 has previously been shown to partially colocalize with MNS1 and was suggested to reside in *cis*-most Golgi membranes, the Golgi entry core compartment (GECCO) (Ito et al., 2017). All tested ERES markers colocalized with MNS3 (Figures 9A–C). As a control,

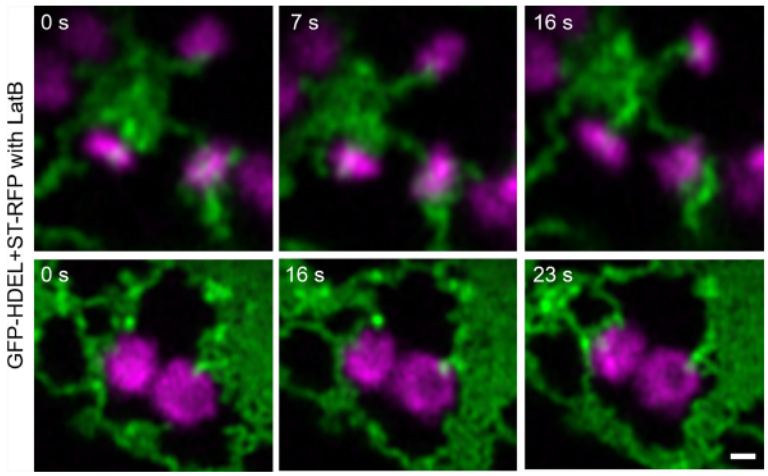


FIGURE 6
Changing ER dynamics with latrunculin B (LatB). The ER and Golgi markers GFP-HDEL (green) and ST-mRFP (magenta) are coexpressed in tobacco leaf epidermal cells, tobacco leaf pieces were treated with LatB and videos were taken. Time frames [s] are indicated. Size bars = 0.5 µm.

10.3389/fpls.2022.1010569

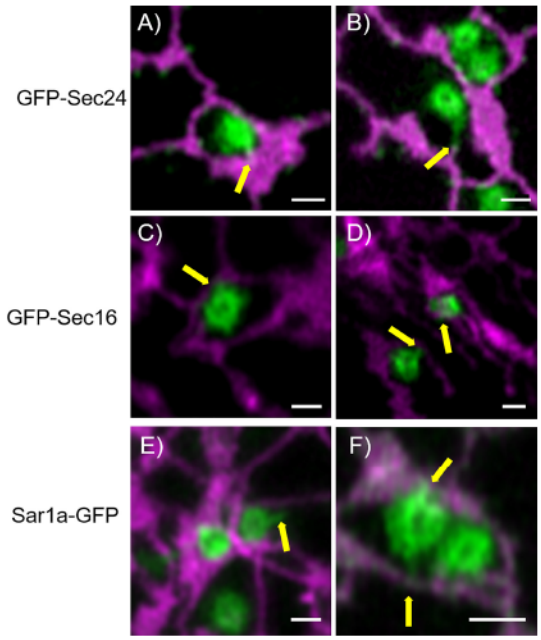


FIGURE 7
COPII components show protrusions that connect with the ER. The ERES markers/COPII components GFP-Sec24 **(A, B)**, GFP-Sec16 **(C, D)** and Sar1a-GFP **(E, F)** show dynamic protrusions (yellow arrows). Proteins are expressed transiently in tobacco leaf epidermal cells. The ER is labelled with the lumenal marker RFP-HDEL (magenta). Size bars = 1 µm.

MNS3-mRFP was coexpressed with MNS3-eGFP (Figure 9D). The distance between the fluorescent peaks between ERES markers and MNS3 was analyzed and did not significantly differ from the control (Figure 9E). This confirms our results (Figure 8) of a presence of pre-*cis*-Golgi compartments with a specific composition of proteins such as the tested ERES/COPII proteins and MNS3.

ERGo tubules are still present after inhibition of ER-to-Golgi transport

We next aimed to investigate a potential role for ERGo tubules in ER-Golgi transport. We analyzed the presence of the ERGo tubules in cells with inhibited ER-Golgi transport. Interfering with COPI coat assembly results in the inhibition of protein transport and a collapse of Golgi body membranes into the ER. The fungal metabolite brefeldin A (BFA) is a commonly used method to inhibit COPI coat formation. In plants, BFA targets a Sec7-type GEF that is necessary for activation of the small COPI GTPase Arf1p (Nebenführ et al., 2000). This leads to the dissociation of COPI coat proteins from Golgi membranes (Ritzenthaler et al., 2002) and the redistribution of Golgi markers, such as ST–GFP/RFP, to the ER and dilatation of the ER network (Boevink et al., 1998; Saint-Jore et al., 2002).

We aimed to use BFA to see if impairment of ER-Golgi transport would influence the presence of ERGo tubules and Golgi/ERES protrusions, which could be an indicator for a potential function of these structures in ER-Golgi transport. We found that BFA treatment on plants expressing ST-GFP and RFP-HDEL resulted in enhanced ER cisternae formation and as expected, redistribution of the Golgi marker ST-GFP to the ER (Figure 10). As typical for BFA treatment, small ST- and HDEL-labelled Golgi body remnants remained (Fujiwara et al., 1988; Driouich et al., 1993; Satiat-Jeunemaitre et al., 1996). Interestingly these remnants were connected to the ER by ERGo tubules and were labelled by both the ER and Golgi body markers RFP-HDEL and ST-GFP (Figure 10, zoom in yellow and blue frame). The distribution of ER and Golgi body marker appears not evenly distributed in places, which could be due to the inhibited transport. ERGo tubules were found in connection with all remnants of Golgi bodies (Figure 10). This suggests that the physical connection between ER and Golgi body remnants was maintained upon BFA treatment.

We next aimed to test the effect of blockage of the COPII pathway and test the hypothesis that a blockage of the progression of the COPII cycle with a dominant-negative mutant of the COPII GTPase Sar1a (Osterrieder et al., 2010) would allow to detect the sites of COPII assembly. We used Sar1a-GTP, a constitutively active (GTP-locked) mutant form of Sar1p from *Nicotiana tabacum*, which has been shown to inhibit protein transport between the ER and Golgi bodies (Andreeva et al., 1998; Takeuchi et al., 2000; Phillipson et al., 2001; daSilva et al., 2004; Hanton et al., 2008). Here a dexamethasone (Dex) inducible system in stably transformed *Nicotiana tabacum* plants allowing controlled expression of Sar1-GTP (Osterrieder et al., 2010) was applied. The ER and Golgi body markers, GFP-HDEL and ST-mRFP respectively, were expressed in these plants. We found that 18 h after Dex-induction, most of the ST-mRFP protein pool was located in the ER with only very few and smaller ST-labelled Golgi bodies remaining (Figure 11). Similar to BFA treatment, Sar1a-GTP expression resulted in enhanced ER cisternal areas, and ST-labelled Golgi stacks were quite tightly engulfed within ER lacunae. The stacks were again still connected to the ER by ERGo tubules but differently from with BFA treatment, the ERGo tubules only showed HDEL labelling with ST-labelled protrusions (Figure 11). ERGo tubules could only be seen in connection with remaining Golgi structures. Sar1a-GTP blocks the uncoating of the COPII coat and the Golgi marker is visible in association but still physically distinct from the ERGo tubules. These results, coupled with the findings that COPI inhibition leads to a redistribution of Golgi markers into the

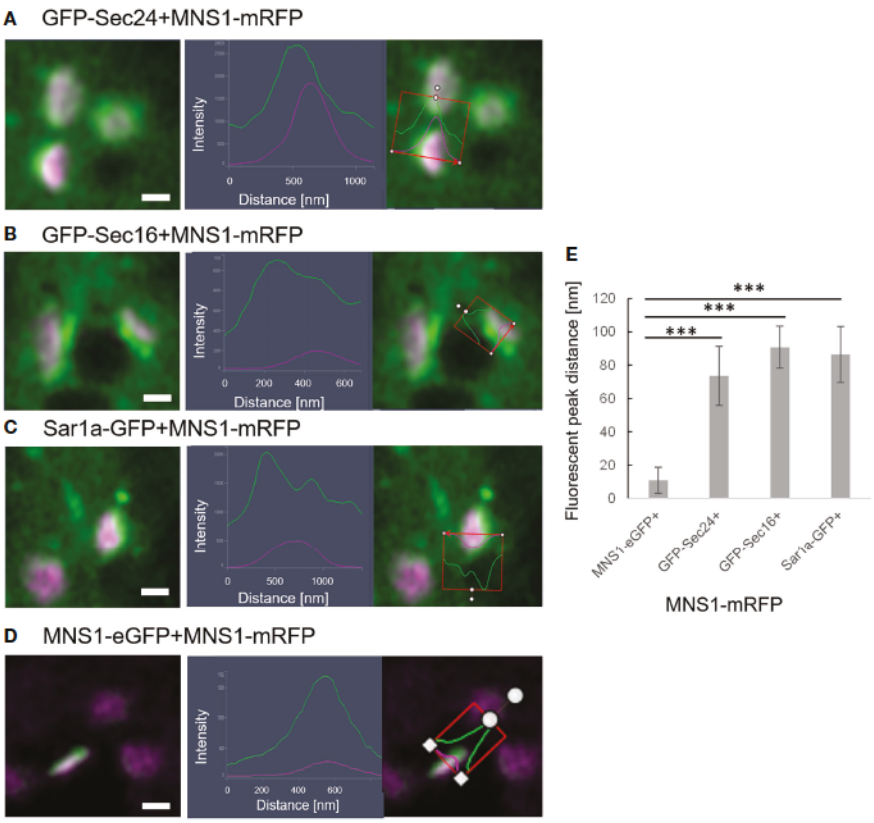


FIGURE 8
ERES markers define a pre-Golgi compartment. ERES markers GFP-Sec24 **(A)**, GFP-Sec16 **(B)** and Sar1a-GFP **(C)** are distributed closely to the *cis*-Golgi-marker MNS1-mRFP but do not fully overlap with MNS1-mRFP. Merged images in a side view are shown together with line profiles for both markers. Proteins were expressed transiently in tobacco leaf epidermal cells. Size bars = 0.5 µm. **(D)** MNS1-eGFP and MNS1-mRFP were coexpressed as a control for full colocalization. **(E)** Statistical analysis shows the distance [nm] between fluorescent peak of ERES marker and MNS1-mRFP. Significance was analyzed by Kruskal-Wallis (***p< 0.001). n = 3 with 5 technical replicas each.

ERGo tubules, indicate that the ERGo tubules may be involved in the cargo transport to the Golgi body.

Discussion

The nature of the ER exit sites together with ER-Golgi transport has been a source for discussion over the years. In yeast and mammalian systems COPII/COPI-dependent transport was the generally accepted mechanism, which was also transferred to plant systems.

In mammalian cells, ERES are morphologically recognized as areas on the rough ER membrane that are devoid of ribosomes and form coated buds, vesicles, and vesicular tubular clusters (Tang et al., 2005). These clusters undergo long-range transport mediated by microtubules to dock with the *cis*-face of the Golgi apparatus (Tang et al., 2005). ERES are characterized by the components of the COPII-coated membranes, v-SNARES required for docking and p24 proteins as cargo receptors (Strating and Martens, 2009). In yeast, COPII-coated vesicles have been suggested to bud from the ER membrane to transport cargo to very simple Golgi bodies that have the capacity to mature from *cis*- through to *trans*-forms (Matsuura-Tokita et al., 2006).

In Embryophyta, early conventional transmission electron microscopy did not show any visible coated ER areas or structures that could be convincingly inferred to be ERES. The application of ultra-rapid freezing techniques combined with tomographic electron microscopy though suggested that a COPII vesicle-mediated transport may exist (Kang and Staehelin, 2008). COPII components were indeed identified in plants using bioinformatics analysis (Robinson et al., 2007) but that does not exclude a non-vesicular localization. Live cell imaging methodologies using fluorescent protein markers showed that Golgi bodies are closely linked with the ER (Boevink et al., 1998; daSilva et al., 2004; Faso et al., 2009; Osterrieder et al., 2017). All of this supported the "secretory unit" concept stating that Golgi and ER exit sites are moving

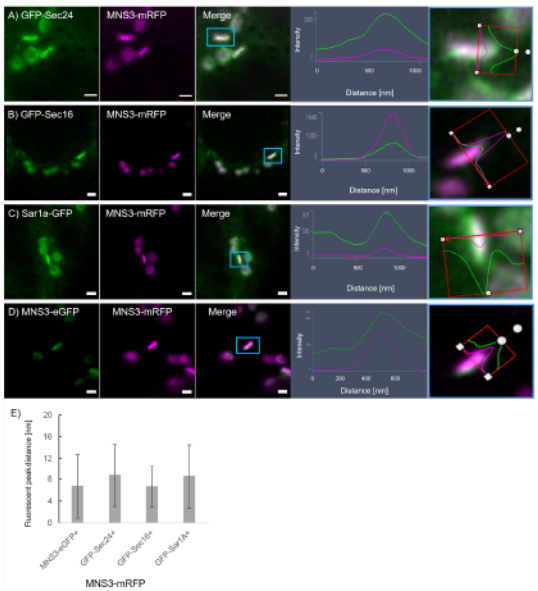


FIGURE 9
The ERES markers localize at the GECCO. ERES markers GFP-Sec24 **(A)**, GFP-Sec16 **(B)** and Sar1a-GFP **(C)** colocalize with MNS3-mRFP. **(D)** MNS3-eGFP and MNS3-mRFP are coexpressed as a control for full colocalization. Representative confocal images and analysis plots are shown for each combination. The blue box indicates the Golgi body shown enhanced on the right-hand side panels (merged images in side view together with line profiles for both markers). Proteins are expressed transiently in tobacco leaf epidermal cells. Size bars = 1 µm. **(E)** Statistical analysis shows the distance [nm] between fluorescent peak of ERES marker and MNS3-mRFP. Significance was analyzed by Kruskal-Wallis. n = 3 with 5 technical replicas each.

together with the ER (daSilva et al., 2004) and that Golgi biogenesis could be from newly generated exit sites (Hawes et al., 2008).

In vitro experiments using giant unilamellar membrane vesicles incubated with COPII components showed no vesicle formation but bead-like membrane organization (Bacia et al., 2011). Recent findings in mammalian systems also question the exclusivity of vesicular anterograde transport. Two research groups showed independently that carriers leaving the ER are not coated by COPII components and suggested the existence of different parallel carrier types including vesicles but also tubules, saccules and tunnels (Weigel et al., 2021; Shomron et al., 2021). In these studies, ERES expand into tubules, which contain cargo but no COPII components. These can only be found at the neck of the tubules; hence, they proposed that carriers leaving the ER are not COPII coated and that the role of COPII is to concentrate transport cargo into carriers (Weigel et al., 2021; Shomron et al., 2021).

Our analyses using a high-resolution confocal imaging approach have provided new insights into the organization of the ER and Golgi interface. By facilitating a visualization of the ER, COPII markers and Golgi membranes in living cells, we were able to capture structures and their dynamics in a manner that is not possible with ultra-resolution microscopy in fixed cells or cryo-electron tomography. Our analyses allowed us to follow the dynamics of the ERGo tubules, which appear to be thinner tubules compared to the bulk ER tubular network. We also showed that treatment with BFA causes a redistribution of Golgi markers into the bulk ER and the ERGo tubules with Golgi remnants distributed at the ERGo tubules, and that blockage of COPII with Sar1a-GTP leads to the formation of

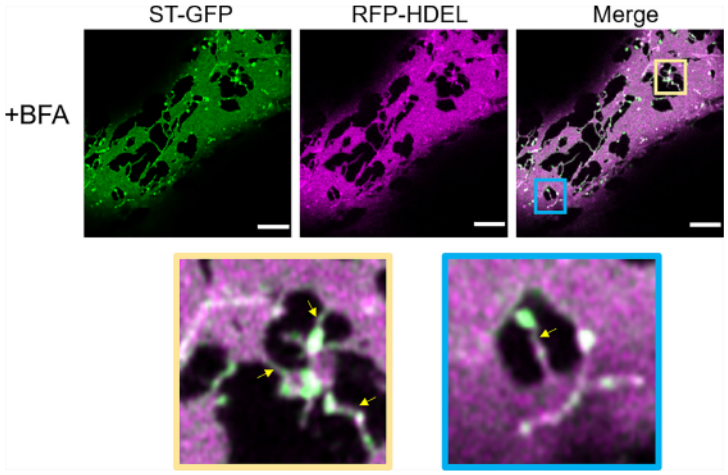


FIGURE 10
Disruption of COPI results in the redistribution of Golgi membranes into the ERGo tubules. The ER and Golgi markers RFP-HDEL (magenta) and ST-GFP (green) are coexpressed and the leaf pieces treated with BFA. Connecting tubules between ER and Golgi remnants are visible (yellow arrows). The blue and yellow boxes in the merge image indicate areas shown in higher magnification below. Size bars = 5 µm.

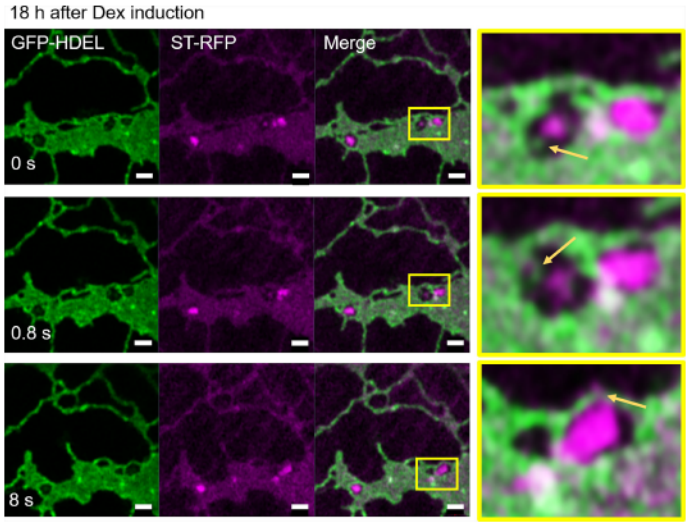


FIGURE 11
Disruption of Golgi membranes by inhibition of COPII uncoating leads to a segregation of ERGo tubules from Golgi remnants. The dominant negative mutant Sar1a-GTP is expressed in an inducible manner together with the ER and Golgi markers GFP-HDEL (green) and ST-mRFP (magenta). Imaging was carried out 18 h after dexamethasone (Dex) induction and a time-lapse (0 s, 0.8 s, 8 s) is shown. The yellow boxes in the images indicate areas shown in higher magnification on the right hand side. ERGo tubules and Golgi protrusions are indicated with yellow arrows. Size bars = 1 µm.

structures that are connected but not merged with ERGo tubules. It was previously suggested that the BFA-induced Golgi remnants contain a minimal set of *cis*-Golgi membrane proteins carrying out specific roles at the ER-Golgi interface, such as protein sorting, ER-Golgi tethering, and N-glycan processing. These BFA-induced structures could then be used as a reservoir to quickly produce *cis*-Golgi cisternae with materials directly from the ER (Hawes et al., 2010; Ito et al., 2012; Robinson et al., 2015; Schoberer et al., 2019), and accordingly, we propose ERGo tubular connections may provide physical stability between the ER and the Golgi as well as a rapid and targeted transport pathway.

Pre-*cis*-Golgi compartments have been previously suggested for proteins proposed to be involved in ER-Golgi tethering: AtGolgin-84A, which colocalizes with Sar1a (Vieira et al., 2020), as well as AtCASP (Latijnhouwers et al., 2007; Osterrieder et al., 2017) and the Arabidopsis ER-α-mannosidase I (MNS3; (Schoberer et al., 2019)). Our results support the existence of a pre-*cis*-Golgi compartment where COPII components accumulate rather than accumulating on the ER. The thin tubular connections, observed between ER and Golgi as well as between ER and ERES could be involved in cargo transport as well as in stabilizing the physical connection between ER and Golgi bodies. Although we have no evidence to infer a functional role to the ERGo tubules, the visualization of these new structures and the definition of their spatial distribution in relation to COPII markers and the visualization of a pre-Golgi compartment shed an interesting new light onto the ER-Golgi interface in plants and challenge the existing dogma on the existence of COPII vesicles in plants. So far, in plant systems no elegant ER-Golgi cargo transport assay such as the RUSH methodology (Boncompain et al., 2012) for mammalian systems has been established but this will be of great importance for studying an involvement of ERGo tubules in ER-Golgi transport.

Data availability statement

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

Author contributions

VK, JS, and FB conceived the experiments. AMG, JS, CP, and VK carried out the experiments. JS, FB, and VK wrote the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This research was funded in part by the Austrian Science Fund (FWF) [P31921-B32] to JS. For the purpose of open access,

 FB is funded by MSU AgBioResearch [MICL02598]. VK is grateful for funding by the BBSRC [BB/W011166/1] and Oxford Brookes University. AMG is funded *via* the Oxford Interdisciplinary Bioscience Doctoral training Programme, CP is funded by Oxford Brookes University.

Acknowledgments

The authors thank the Oxford Brookes University Centre for Bioimaging.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.1010569/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Video tobacco transiently expressing GFP-HDEL and ST-mRFP.

SUPPLEMENTARY FIGURE 2
(A, B) Videos Arabidopsis stably transformed with GFP-HDEL and ST-mRFP.

References

Andreeva, A. V., Kutuzov, M. A., Evans, D. E., and Hawes, C. R. (1998). Proteins involved in membrane transport between the ER and the golgi apparatus: 21 putative plant homologues revealed by dbEST searching. *Cell Biol. Int.* 22 (2), 145–160. doi: 10.1006/cbir.1998.0235

Arakel, E. C., and Schwappach, B. (2018). Formation of COPI-coated vesicles at a glance. *J. Cell Sci.* 131 (5), jcs209890. doi: 10.1242/jcs.209890

Bacia, K., Futai, E., Prinz, S., Meister, A., Daum, S., Glatte, D., et al. (2011). Multibudded tubules formed by COPII on artificial liposomes. *Sci. Rep.* 1, 17. doi: 10.1038/srep00017

Bannykh, S. I., Rowe, T., and Balch, W. E. (1996). The organization of endoplasmic reticulum export complexes. *J. Cell Biol.* 135 (1), 19–35. doi: 10.1083/jcb.135.1.19

Boevink, P., Oparka, K., Santa Cruz, S., Martin, B., Betteridge, A., and Hawes, C. (1998). Stacks on tracks: the plant golgi apparatus traffics on an actin/ER network. *Plant J.* 15 (3), 441–447. doi: 10.1046/j.1365-313X.1998.00208.x

Boncompain, G., Divoux, S., Gareil, N., de Forges, H., Lescure, A., Latreche, L., et al. (2012). Synchronization of secretory protein traffic in populations of cells. *Nat. Methods* 9 (5), 493–498. doi: 10.1038/nmeth.1928

Brandizzi, F. (2018). Transport from the endoplasmic reticulum to the golgi in plants: Where are we now? *Semin. Cell Dev. Biol.* 80, 94–105. doi: 10.1016/j.semcdb.2017.06.024

Brandizzi, F., Hanton, S., DaSilva, L. L., Boevink, P., Evans, D., Oparka, K., et al. (2003). ER quality control can lead to retrograde transport from the ER lumen to the cytosol and the nucleoplasm in plants. *Plant J.* 34 (3), 269–281. doi: 10.1046/j.1365-313X.2003.01728.x

Brandizzi, F., Snapp, E. L., Roberts, A. G., Lippincott-Schwartz, J., and Hawes, C. (2002). Membrane protein transport between the endoplasmic reticulum and the golgi in tobacco leaves is energy dependent but cytoskeleton independent: evidence from selective photobleaching. *Plant Cell* 14 (6), 1293–1309. doi: 10.1105/tpc.001586

daSilva, L. L., Snapp, E. L., Denecke, J., Lippincott-Schwartz, J., Hawes, C., and Brandizzi, F. (2004). Endoplasmic reticulum export sites and golgi bodies behave as single mobile secretory units in plant cells. *Plant Cell* 16 (7), 1753–1771. doi: 10.1105/tpc.022673

Donohoe, B. S., Kang, B. H., and Staehelin, L. A. (2007). Identification and characterization of COPIa- and COPIb-type vesicle classes associated with plant and algal golgi. *Proc. Natl. Acad. Sci. U.S.A.* 104 (1), 163–168. doi: 10.1073/pnas.0609818104

Driouich, A., Zhang, G. F., and Staehelin, L. A. (1993). Effect of brefeldin a on the structure of the golgi apparatus and on the synthesis and secretion of proteins and polysaccharides in sycamore maple (Acer pseudoplatanus) suspension-cultured cells. *Plant Physiol.* 101 (4), 1363–1373. doi: 10.1104/pp.101.4.1363

Farhan, H., Weiss, M., Tani, K., Kaufman, R. J., and Hauri, H. P. (2008). Adaptation of endoplasmic reticulum exit sites to acute and chronic increases in cargo load. *EMBO J.* 27 (15), 2043–2054. doi: 10.1038/emboj.2008.136

Faso, C., Chen, Y. N., Tamura, K., Held, M., Zemelis, S., Marti, L., et al. (2009). A missense mutation in the arabidopsis COPII coat protein Sec24A induces the formation of clusters of the endoplasmic reticulum and golgi apparatus. *Plant Cell* 21 (11), 3655–3671. doi: 10.1105/tpc.109.068262

Fujiwara, T., Oda, K., Yokota, S., Takatsuki, A., and Ikehara, Y. (1988). Brefeldin a causes disassembly of the golgi complex and accumulation of secretory proteins in the endoplasmic reticulum. *J. Biol. Chem.* 263 (34), 18545–18552. doi: 10.1016/S0021-9258(19)81393-5

Gorur, A., Yuan, L., Kenny, S. J., Baba, S., Xu, K., and Schekman, R. (2017). COPII-coated membranes function as transport carriers of intracellular procollagen i. *J. Cell Biol.* 216 (6), 1745–1759. doi: 10.1083/jcb.201702135

Hanton, S. L., Chatre, L., Matheson, L. A., Rossi, M., Held, M. A., and Brandizzi, F. (2008). Plant Sar1 isoforms with near-identical protein sequences exhibit different localisations and effects on secretion. *Plant Mol. Biol.* 67 (3), 283–294. doi: 10.1007/s11103-008-9317-5

Hawes, C. (2012). The ER/Golgi interface - is there anything in-between? *Front. Plant Sci.* 3, 73. doi: 10.3389/fpls.2012.00073

Hawes, C., Osterrieder, A., Hummel, E., and Sparkes, I. (2008). The plant ER-golgi interface. *Traffic* 9 (10), 1571–1580. doi: 10.1111/j.1600-0854.2008.00773.x

Hawes, C., Schoberer, J., Hummel, E., and Osterrieder, A. (2010). Biogenesis of the plant golgi apparatus. *Biochem. Soc. Trans.* 38 (3), 761–767. doi: 10.1042/BST0380761

Ito, Y., Toyooka, K., Fujimoto, M., Ueda, T., Uemura, T., and Nakano, A. (2017). The trans-golgi network and the golgi stacks behave independently during regeneration after brefeldin a treatment in tobacco BY-2 cells. *Plant Cell Physiol.* 58 (4), 811–821. doi: 10.1093/pcp/pcx028

Ito, Y., Uemura, T., Shoda, K., Fujimoto, M., Ueda, T., and Nakano, A. (2012). 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. *Mol. Biol. Cell* 23 (16), 3203–3214. doi: 10.1091/mbc.e12-01-0034

Kang, B. H., and Staehelin, L. A. (2008). ER-to-Golgi transport by COPII vesicles in arabidopsis involves a ribosome-excluding scaffold that is transferred with the vesicles to the golgi matrix. *Protoplasma* 234 (1-4), 51–64. doi: 10.1007/s00709-008-0015-6

Kriechbaumer, V., and Brandizzi, F. (2020). The plant endoplasmic reticulum: an organized chaos of tubules and sheets with multiple functions. *J. Microsc.* 280 (2), 122–133. doi: 10.1111/jmi.12909

Kurokawa, K., and Nakano, A. (2019). The ER exit sites are specialized ER zones for the transport of cargo proteins from the ER to the golgi apparatus. *J. Biochem.* 165 (2), 109–114. doi: 10.1093/jb/mvy080

Langhans, M., Meckel, T., Kress, A., Lerich, A., and Robinson, D. G. (2012). ERES (ER exit sites) and the "secretory unit concept". *J. Microsc.* 247 (1), 48–59. doi: 10.1111/j.1365-2818.2011.03597.x

Latijnhouwers, M., Gillespie, T., Boevink, P., Kriechbaumer, V., Hawes, C., and Carvalho, C. M. (2007). Localization and domain characterization of arabidopsis golgin candidates. *J. Exp. Bot.* 58 (15-16), 4373–4386. doi: 10.1093/jxb/erm304

Liebminger, E., Hüttner, S., Vavra, U., Fischl, R., Schoberer, J., Grass, J., et al. (2009). Class I alpha-mannosidases are required for n-glycan processing and root development in arabidopsis thaliana. *Plant Cell* 21 (12), 3850–3867. doi: 10.1105/tpc.109.072363

Matsuura-Tokita, K., Takeuchi, M., Ichihara, A., Mikuriya, K., and Nakano, A. (2006). Live imaging of yeast golgi cisternal maturation. *Nature* 441 (7096), 1007–1010. doi: 10.1038/nature04737

Nebenführ, A., Frohlick, J. A., and Staehelin, L. A. (2000). Redistribution of golgi stacks and other organelles during mitosis and cytokinesis in plant cells. *Plant Physiol.* 124 (1), 135–151. doi: 10.1104/pp.124.1.135

Oprins, A., Duden, R., Kreis, T. E., Geuze, H. J., and Slot, J. W. (1993). Beta-COP localizes mainly to the cis-golgi side in exocrine pancreas. *J. Cell Biol.* 121 (1), 49–59. doi: 10.1083/jcb.121.1.49

Orci, L., Amherdt, M., Ravazzola, M., Perrelet, A., and Rothman, J. E. (2000). Exclusion of golgi residents from transport vesicles budding from golgi cisternae in intact cells. *J. Cell Biol.* 150 (6), 1263–1270. doi: 10.1083/jcb.150.6.1263

Osterrieder, A., Hummel, E., Carvalho, C. M., and Hawes, C. (2010). Golgi membrane dynamics after induction of a dominant-negative mutant Sar1 GTPase in tobacco. *J. Exp. Bot.* 61 (2), 405–422. doi: 10.1093/jxb/erp315

Osterrieder, A., Sparkes, I. A., Botchway, S. W., Ward, A., Ketelaar, T., de Ruijter, N., et al. (2017). Stacks off tracks: a role for the golgin AtCASP in plant endoplasmic reticulum-golgi apparatus tethering. *J. Exp. Bot.* 68 (13), 3339–3350. doi: 10.1093/jxb/erx167

Pain, C., Kriechbaumer, V., Kittelmann, M., Hawes, C., and Fricker, M. (2019). Quantitative analysis of plant ER architecture and dynamics. *Nat. Commun.* 10 (1), 984. doi: 10.1038/s41467-019-08893-9

Phillipson, B. A., Pimpl, P., daSilva, L. L., Crofts, A. J., Taylor, J. P., Movafeghi, A., et al. (2001). Secretory bulk flow of soluble proteins is efficient and COPII dependent. *Plant Cell* 13 (9), 2005–2020. doi: 10.1105/TPC.010110

Pimpl, P., Movafeghi, A., Coughlan, S., Denecke, J., Hillmer, S., and Robinson, D. G. (2000). *In situ* localization and *in vitro* induction of plant COPI-coated vesicles. *Plant Cell* 12 (11), 2219–2236. doi: 10.1105/tpc.12.11.2219

Presley, J. F., Cole, N. B., Schroer, T. A., Hirschberg, K., Zaal, K. J., and Lippincott-Schwartz, J. (1997). ER-to-Golgi transport visualized in living cells. *Nature* 389 (6646), 81–85. doi: 10.1038/38001

Ritzenthaler, C., Nebenführ, A., Movafeghi, A., Stussi-Garaud, C., Behnia, L., Pimpl, P., et al. (2002). Reevaluation of the effects of brefeldin a on plant cells using tobacco bright yellow 2 cells expressing golgi-targeted green fluorescent protein and COPI antisera. *Plant Cell* 14 (1), 237–261. doi: 10.1105/tpc.010237

Robinson, D. G. (2020). Plant golgi ultrastructure. *J. Microsc.* 280 (2), 111–121. doi: 10.1111/jmi.12899

Robinson, D. G., Brandizzi, F., Hawes, C., and Nakano, A. (2015). Vesicles versus tubes: Is endoplasmic reticulum-golgi transport in plants fundamentally different from other eukaryotes? *Plant Physiol.* 168 (2), 393–406. doi: 10.1104/pp.15.00124

Robinson, D. G., Herranz, M.-C., Bubeck, J., Pepperkok, R., and Ritzenthaler, C. (2007). Membrane dynamics in the early secretory pathway. *Crit. Rev. Plant Sci.* 26 (4), 199–225. doi: 10.1080/07352680701495820

Runions, J., Brach, T., Kühner, S., and Hawes, C. (2006). Photoactivation of GFP reveals protein dynamics within the endoplasmic reticulum membrane. *J. Exp. Bot.* 57 (1), 43–50. doi: 10.1093/jxb/eri289

Saint-Jore, C. M., Evins, J., Batoko, H., Brandizzi, F., Moore, I., and Hawes, C. (2002). Redistribution of membrane proteins between the golgi apparatus and endoplasmic reticulum in plants is reversible and not dependent on cytoskeletal networks. *Plant J.* 29 (5), 661–678. doi: 10.1046/j.0960-7412.2002.01252.x

Saraste, J., and Svensson, K. (1991). Distribution of the intermediate elements operating in ER to golgi transport. *J. Cell Sci.* 100, 415–430. doi: 10.1242/jcs.100.3.415

Satiat-Jeunemaitre, B., Cole, L., Bourett, T., Howard, R., and Hawes, C. (1996). Brefeldin a effects in plant and fungal cells: something new about vesicle trafficking? *J. Microsc.* 181, 162–177. doi: 10.1046/j.1365-2818.1996.112393.x

Schoberer, J., König, J., Veit, C., Vavra, U., Liebminger, E., Botchway, S. W., et al. (2019). A signal motif retains arabidopsis ER-α-mannosidase I in the cis-golgi and prevents enhanced glycoprotein ERAD. *Nat. Commun.* 10 (1), 3701. doi: 10.1038/s41467-019-11686-9

Schoberer, J., Liebminger, E., Vavra, U., Veit, C., Castilho, A., Dicker, M., et al. (2014). The transmembrane domain of n -acetylglucosaminyltransferase I is the key determinant for its golgi subcompartmentation. *Plant J.* 80 (5), 809–822. doi: 10.1111/tpj.12671

Shomron, O., Nevo-Yassaf, I., Aviad, T., Yaffe, Y., Zahavi, E. E., Dukhovny, A., et al. (2021). COPII collar defines the boundary between ER and ER exit site and does not coat cargo containers. *J. Cell Biol.* 220 (6), e201907224. doi: 10.1083/jcb.201907224

Sparkes, I. A., Ketelaar, T., de Ruijter, N. C., and Hawes, C. (2009). Grab a golgi: laser trapping of golgi bodies reveals *in vivo* interactions with the endoplasmic reticulum. *Traffic* 10 (5), 567–571. doi: 10.1111/j.1600-0854.2009.00891.x

Sparkes, I. A., Runions, J., Kearns, A., and Hawes, C. (2006). Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nat. Protoc.* 1 (4), 2019–2025. doi: 10.1038/nprot.2006.286

Stephens, D. J., Lin-Marq, N., Pagano, A., Pepperkok, R., and Paccaud, J. P. (2000). COPI-coated ER-to-Golgi transport complexes segregate from COPII in close proximity to ER exit sites. *J. Cell Sci.* 113, 2177–2185. doi: 10.1242/jcs.113.12.2177

Strating, J. R., and Martens, G. J. (2009). The p24 family and selective transport processes at the ER-golgi interface. *Biol. Cell* 101 (9), 495–509. doi: 10.1042/BC20080233

Takagi, J., Renna, L., Takahashi, H., Koumoto, Y., Tamura, K., Stefano, G., et al. (2013). MAIGO5 functions in protein export from golgi-associated endoplasmic reticulum exit sites in arabidopsis. *Plant Cell* 25 (11), 4658–4675. doi: 10.1105/tpc.113.118158

Takeuchi, M., Ueda, T., Sato, K., Abe, H., Nagata, T., and Nakano, A. (2000). A dominant negative mutant of sar1 GTPase inhibits protein transport from the endoplasmic reticulum to the golgi apparatus in tobacco and arabidopsis cultured cells. *Plant J.* 23 (4), 517–525. doi: 10.1046/j.1365-313x.2000.00823.x

Tang, B. L., Wang, Y., Ong, Y. S., and Hong, W. (2005). COPII and exit from the endoplasmic reticulum. *Biochim. Biophys. Acta* 1744 (3), 293–303. doi: 10.1016/j.bbamcr.2005.02.007

Vieira, V., Pain, C., Wojcik, S., Spatola Rossi, T., Denecke, J., Osterrieder, A., et al. (2020). Living on the edge: the role of atgolgin-84A at the plant ER-golgi interface. *J. Microsc.* 280 (2), 158–173. doi: 10.1111/jmi.12946

Wang, P., Hawkins, T. J., Richardson, C., Cummins, I., Deeks, M. J., Sparkes, I., et al. (2014). The plant cytoskeleton, NET3C, and VAP27 mediate the link between the plasma membrane and endoplasmic reticulum. *Curr. Biol.* 24 (12), 1397–1405. doi: 10.1016/j.cub.2014.05.003

Watson, P., and Stephens, D. J. (2005). ER-to-Golgi transport: form and formation of vesicular and tubular carriers. *Biochim. Biophys. Acta* 1744 (3), 304–315. doi: 10.1016/j.bbamcr.2005.03.003

Weigel, A. V., Chang, C. L., Shtengel, G., Xu, C. S., Hoffman, D. P., Freeman, M., et al. (2021). ER-to-Golgi protein delivery through an interwoven, tubular network extending from ER. *Cell* 184 (9), 2412–2429.e16. doi: 10.1016/j.cell.2021.03.035

Westrate, L. M., Hoyer, M. J., Nash, M. J., and Voeltz, G. K. (2020). Vesicular and uncoated Rab1-dependent cargo carriers facilitate ER to golgi transport. *J. Cell Sci.* 133 (14) , jcs239814. doi: 10.1242/jcs.239814.

Xu, D., and Hay, J. C. (2004). Reconstitution of COPII vesicle fusion to generate a pre-golgi intermediate compartment. *J. Cell Biol.* 167 (6), 997–1003. doi: 10.1083/jcb.200408135

Zeuschner, D., Geerts, W. J., van Donselaar, E., Humbel, B. M., Slot, J. W., Koster, A. J., et al. (2006). Immuno-electron tomography of ER exit sites reveals the existence of free COPII-coated transport carriers. *Nat. Cell Biol.* 8 (4), 377–383. doi: 10.1038/ncb1371\

Received: 25 January 2024 | Revised: 10 April 2024 | Accepted: 18 April 2024

DOI: 10.1111/jmi.13311

Journal of Microscopy RMS

THEMED ISSUE ARTICLE

Suborganellar resolution imaging for the localisation of human glycosylation enzymes in tobacco Golgi bodies

Alastair J. McGinness[1] | **Susan A. Brooks**[1] | **Richard Strasser**[2] | **Jennifer Schoberer**[2] | **Verena Kriechbaumer**[1,3]

[1]Department of Biological and Medical Sciences, Oxford Brookes University, Oxford, UK

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

[3]Centre for Bioimaging, Oxford Brookes University, Oxford, UK

Correspondence
Verena Kriechbaumer, Department of Biological and Medical Sciences, Oxford Brookes University, Oxford, UK.
Email: vkriechbaumer@brookes.ac.uk

Funding information
Oxford Brookes Interdisciplinary DTP PhD studentship; AMG; Biotechnology and Biological Sciences Research Council, Grant/Award Number: BB/X006417/1

Abstract

Plant cells are a capable system for producing economically and therapeutically important proteins for a variety of applications, and are considered a safer production system than some existing hosts such as bacteria or yeasts. However, plants do not perform protein modifications in the same manner as mammalian cells do. This can impact on protein functionality for plant-produced human therapeutics. This obstacle can be overcome by creating a plant-based system capable of 'humanising' proteins of interest resulting in a glycosylation profile of synthetic plant-produced proteins as it would occur in mammalian systems.

For this, the human glycosylation enzymes (HuGEs) involved in N-linked glycosylation N-acetylglucosaminyltransferase IV and V (GNTIV and GNTV), β-1,4-galactosyltransferase (B4GALT1), and α-2,6-sialyltransferase (ST6GAL) were expressed in plant cells. For these enzymes to carry out the stepwise glycosylation functions, they need to localise to late Golgi body cisternae. This was achieved by a protein targeting strategy of replacing the mammalian Golgi targeting domains (Cytoplasmic-Transmembrane-Stem (CTS) regions) with plant-specific ones. Using high-resolution and dynamic confocal microscopy, we show that GNTIV and GNTV were successfully targeted to the medial-Golgi cisternae while ST6GAL and B4GALT1 were targeted to *trans*-Golgi cisternae.

Plant cells are a promising system to produce human therapeutics for example proteins used in enzyme replacement therapies. Plants can provide safer and cheaper alternatives to existing expression systems such as mammalian cell culture, bacteria or yeast. An important factor for the functionality of therapeutic proteins though are protein modifications specific to human cells. However, plants do not perform protein modifications in the same manner as human cells do. Therefore, plant cells need to be genetically modified to mimic human protein modifications patterns. The modification of importance here, is called N-linked glycosylation and adds specific sugar molecules onto the proteins.

Here we show the expression of four human glycosylation enzymes, which are required for N-linked glycosylation, in plant cells.
In addition, as these protein modifications are carried out in cells resembling a factory production line, it is important that the human glycosylation enzymes be placed in the correct cellular compartments and in the correct order. This is carried out in Golgi bodies. Golgi bodies are composed of several defined stacks termed *cis*-, medial and *trans*-Golgi body stacks. For correct protein function, two of these human glycosylation enzymes need to be placed in the medial-Golgi attacks and the other two in the *trans*-Golgi stacks. Using high-resolution laser microscopy in live plant cells, we show here that the human glycosylation enzymes are sent within the cells to the correct Golgi body stacks. These are first steps to modify plant cells in order to produce human therapeutics.

KEYWORDS
cytoplasmic-transmembrane-stem regions, glycosylation, Golgi cisternae, human glycosylation enzymes

1 | INTRODUCTION

1.1 | Production of recombinant proteins in plant systems

Plant systems for the production of recombinant protein are a fast-growing industry. Major benefits of plant-based systems for recombinant protein production are low-cost, sustainable scalability, reliable stable or transient expression and flexibility of genetic modification.[1] The innate tolerance of mammals to plant proteins, through everyday contact and diet, makes plant-produced therapeutics highly desirable when compared to those produced in vectors such as bacteria, because of their lack of a substantial immunogenic profile.[1] Similarly, yeast and bacterial systems can generate high risk immunogenic contaminants, that is, bacterial lipopolysaccharides and terminally mannosylated glycan structures; these are a risk factor for therapeutic applications as they can trigger detrimental immune responses in patients.[2]

Plant-based platforms begin to challenge existing, well-developed mammalian and bacterial systems, as industry begins to see the economic and biological advantages of biopharming.[2] Therapeutic potential for plant-manufactured proteins has real-world demonstrations of effectiveness, including pandemic response capability of tobacco-manufactured therapeutics for the 2014 Ebola outbreak,[3] to the first plant-manufactured enzyme replacement therapy of recombinant *Taliglucerase alfa* therapy, utilised for Gaucher's disease and produced in carrot cell culture.[4] Proteins of commercial interest, such as antibodies and enzymes, frequently require extensive and specific posttranslational modification in order to be functional or therapeutically effective. Lysosomal acid lipase, a recombinant protein for enzyme replacement therapy in lipid storage disorders, and immunoglobulins (IgA, IgM, IgE and IgG) are evidence of the requirement for correct protein synthesis as well as posttranslational modification, and have multiple N-linked glycosylation sites. Plant systems can perform important posttranslational modifications such as N-linked glycosylation but do not generate the same late Golgi glycan modifications that occur in mammalian N-glycosylation, lacking key mammal-specific motifs.

1.2 | N-linked glycosylation for functional therapeutic proteins

N-linked glycosylation refers to the stepwise enzymatic attachment of sugar moieties onto specific asparagine residues residing on synthesised protein[5] and has a strong impact on protein localisation and function.[6] The composition of glycosylation enzymes in plant and mammalian systems is comparable in the early stages of N-glycan chain generation in the *cis*-Golgi cisternae, but differs in the enzymes present in the late medial/*trans*-Golgi [7,8] (Figure 1). N-acetylglucosaminyltransferases (GnTs; GnTI, GnTII, GnTIV and GnTV) use Uridine diphosphate-*N*-acetylglucosamine (UDP-GlcNAc) in order to transfer N-acetylglucosamine onto various points of the developing glycan structure, onto which further specific galactosylation is performed by B4GALT1, and final sialic acid capping of the glycan structure by ST6GAL (Figure 1).[9] The former

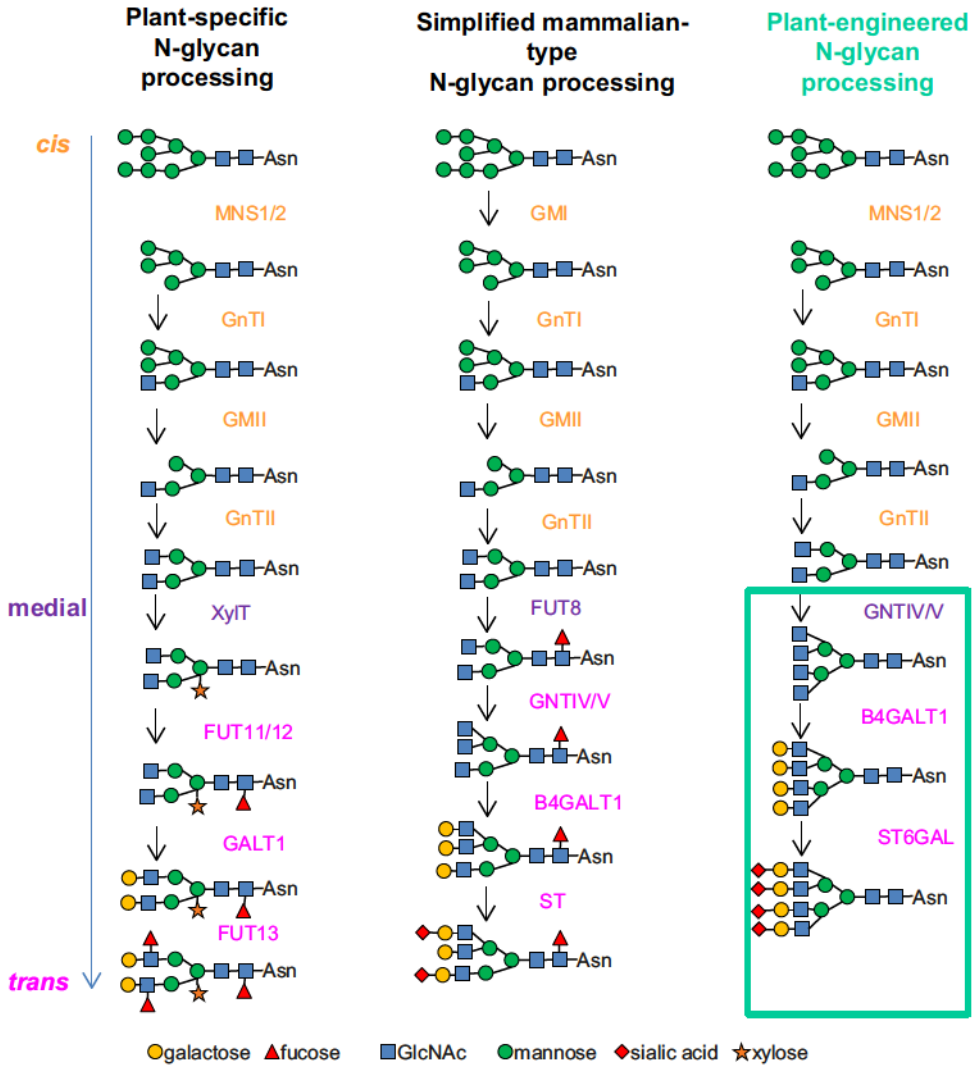


FIGURE 1 Comparison of plant and mammalian N-linked glycosylation and modifications required for humanising plant N-linked glycosylation. Required modifications to the plant system in terms of insertion of mammalian enzymes and their resultant glycan structures are highlighted in the green box on the right hand side. Enzymes are colour-coded for localisation in *cis*-Golgi cisternae (orange), medial-Golgi cisternae (purple) and *trans*-Golgi cisternae (magenta). MNS: α-mannosidase, GnT: N-acetylglucosaminyltransferase, XylT: β-1,2-xylosyltransferase, FUT11/12: α–1,3-fucosyltransferase, GALT1: β-1,3-galactosyltransferase, FUT13: α–1,4-fucosyltransferase. Modified from Ref. (8). Symbol nomenclature for graphical representation of glycans is based on.[17]

steps (GnTI and GnTII) are conserved between plants and mammalian cells (Figure 1), while the latter and subsequent steps responsible for the branching (formation of tri- and tetra-antennary glycans) and end-terminal sialylation (GNTIV, GNTV, B4GALT1 and ST6GAL) are not present in plants and must be introduced for sialic acid capping to occur.[9]

While native plants enzymes can functionally N-glycosylate expressed heterologous human proteins, they do not possess the late Golgi enzymes required for the generation of a multiantennary N-glycan chain with end-terminal sialic acid capping, which is the required structure of mammalian N-glycans. Adaptation of the plant machinery to produce human glycan patterns requires the addition of mammalian enzymes GNTIV and GNTV (N-acetylglucosaminyltransferase IV and V), B4GALT1 (β-1,4-galactosyltransferase), and ST6GAL (α-2,6-sialyltransferase), and potentially downregulation of the plant enzymes XylT (β-1,2-xylosyltransferase), FUT11 and FUT12 (α-1,3-fucosyltransferases) to avoid interference with plant glycosylation patterns. The plant enzymes to be replaced or outcompeted are located in the medial/*trans*-Golgi, hence GNTIV, GNTV, ST6GAL and B4GALT1 should also to be targeted to medial/*trans*-Golgi cisternae for effective enzymatic function on N-glycan intermediates[8] as their activity is sequential and follows the activity of the enzymes in the cis-Golgi cisternae.

In plant Golgi bodies, this glycosylation machinery differs from the mammalian equivalent, and so is incapable of producing specific mammalian glycomotifs and a mammalian-type glycosylation profile (Figure 1).[6–8] The differences between plant and mammalian N-glycosylation machinery is of importance for the expression of mammalian proteins as many economically important recombinant proteins are being glycosylated.[10] 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.[5] For example, the absence of complex N-glycans in an antirabies monoclonal antibody produced in plant systems reduces its efficacy by diminishing plasma half-life.[11] Furthermore the introduction or amplification of key mammalian glycans can significantly improve plasma half-life for therapeutic proteins which has been shown for example in a modified erythropoietin with enhanced end-terminal sialyation.[12]

Overall, glycol-engineering in plants has been shown to be possible. Production of mammalian glycan motives in plants can be achieved by the targeted heterologous expression of mammalian glycosylation enzymes.[6,13] Earlier studies demonstrating that a mammalian glycosyltransferase can convert N-glycans in plants in a similar way as in human cells[14,15] paved the way for reconstructing mammalian glycosylation pathways in plant systems. For example, the simultaneous overexpression of six mammalian glycosylation proteins in *N. benthamiana* resulted in the synthesis of the activated sialic acid and in planta protein sialylation.[16]

To humanise plant glycosylation patterns in order to produce economically and therapeutically important proteins, insertion of the human glycosylation enzymes (HuGEs) GNTIV and GNTV, B4GALT1 and ST6GAL into the late Golgi cisternae is required. Effectiveness of enzyme knock-in strategies will be dependent on correct localisation within the plant Golgi body to the medial and *trans*-Golgi cisternae respectively, which can be evaluated with high-resolution confocal dynamic imaging.

1.3 | Glycosylation enzyme distribution in plant Golgi stacks

N-linked glycosylation in plants occurs within the Golgi bodies, which is a significant site of protein, lipid and polysaccharide synthesis as well as glycomodification within the plant cell, and so a focal point for production of recombinant proteins.[8,18] 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.[19] The basic structure of the individual Golgi body (Figure 2) is a stack of disc-like cisternae characterised 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 the product is delivered to the Trans Golgi Network mediating intracellular delivery of ER products.[19]

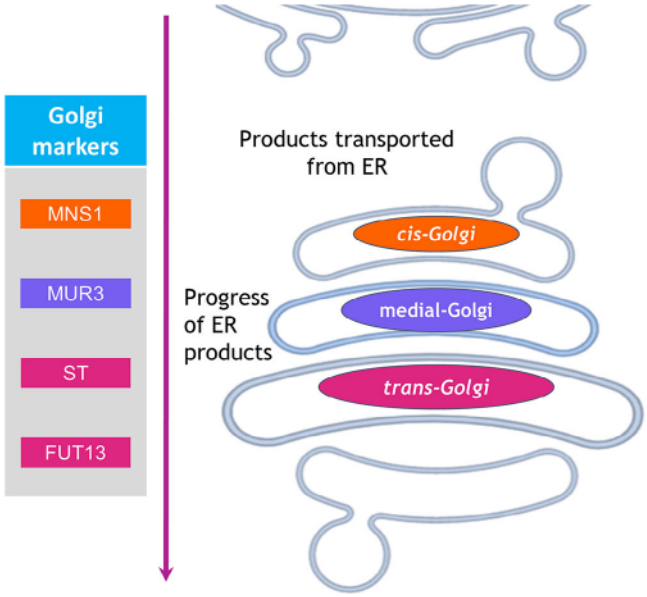


FIGURE 2 Schematic representation of suborganellar organisation of the plant Golgi body cisternae, with relevant cisternae-specific markers colour-coded. Orange = *cis*-Golgi targeting (MNS1, CTS of Golgi-α-mannosidase I), Purple = medial-Golgi targeting (MUR3, CTS of β-1,2-galactosyltransferase), Magenta = *trans*-Golgi targeting (ST, CTS of *Rattus norvegicus* α-2,6-sialyl transferase; FUT13, CTS of α-1,4-fucosyltransferase).

Plant and human Golgi-resident N-glycosylation enzymes are all type II membrane proteins, each possessing at the N-terminal region a short cytoplasmic domain, a transmembrane domain, and a stem region (CTS region). The CTS region is responsible both for orientation of the enzyme's catalytic domain in the Golgi lumen, and providing sub-Golgi localisation, offering an explanation to the non-uniform distribution of glycosylation enzymes in the Golgi.[20] Existing markers (MNS1 for *cis*-Golgi cisternae[21,22] and ST for medial/*trans*-Golgi cisternae[23]) consist of fluorophore fusions to respective CTS domains, providing cisternae-specific localisation as a comparison with the plant-specific CTS domains MUR3, FUT13[24] (Figure 2). The use of endogenous plant CTS domains can alter the localisation of human glycosylation enzymes (HuGEs), through fusion of a HuGE catalytic region with a plant CTS directed to the relevant sub-Golgi compartment.[13]

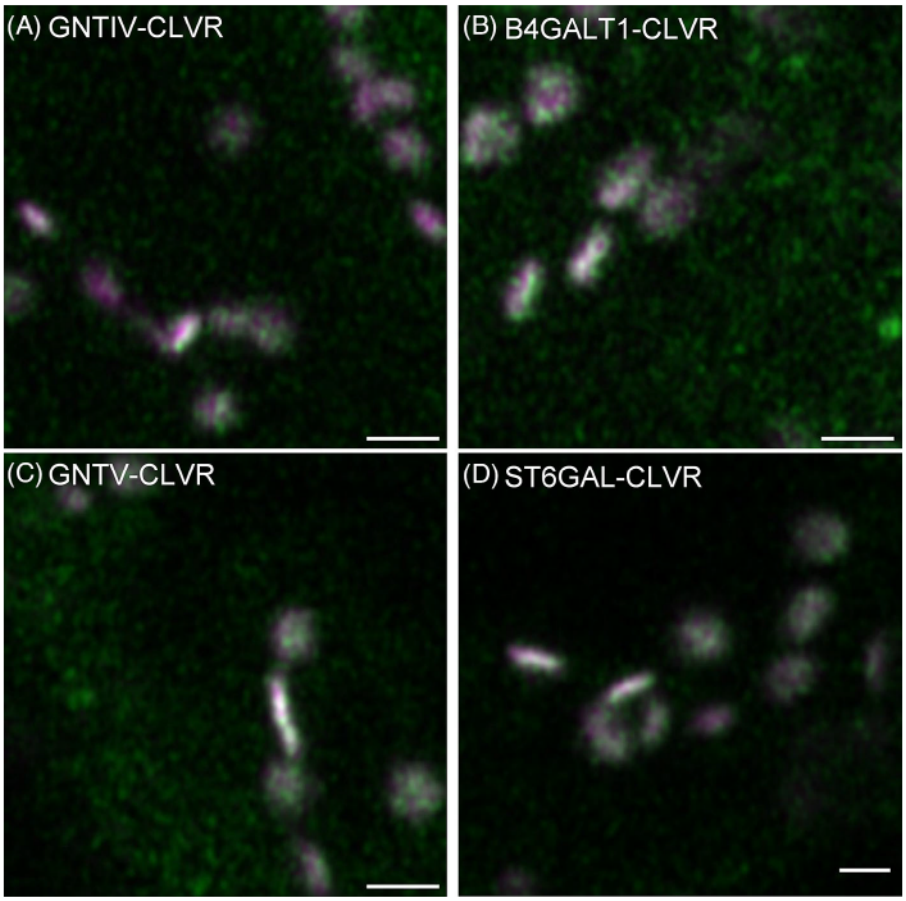


FIGURE 3 Example confocal images for unmodified HuGEs. (A) GNTIV-CLVR, (B) B4GALT1-CLVR, (C) GNTV-CLVR, (D) ST6GAL-CLVR (green) at 3 days postinfiltration alongside the *cis*-Golgi marker MNS1-mRFP (magenta). Representative images are shown. Size bars = 1 μm.

Here we describe the transient expression and localisation of human glycosylation enzymes (HuGEs) in tobacco plants as a model system, with directed localisation to medial- and *trans*-Golgi body cisternae, respectively. Suborganellar localisation to Golgi body cisternae is analysed using high-resolution live-cell and dynamic confocal microscopy.

2 | RESULTS AND DISCUSSION

2.1 | Unmodified HuGEs express weakly and do not localise specifically to the plant Golgi body

To investigate localisation of unmodified HuGEs featuring the mammalian signal sequences *in planta* constructs were generated for HuGEs, fused to a C-terminal CLOVER (CLVR) fluorescent tag and expressed in *Nicotiana tabacum* (tobacco) leaf epidermal cells using agrobacterium-mediated plant transformation. To evaluate localisation, HuGEs were co-expressed with the *cis*-Golgi marker MNS1-mRFP[25] (Figure 3). We observed low levels of expression and protein localisation did not appear specific to the plant Golgi bodies. Without Golgi-specific localisation, HuGEs are unlikely to correctly glycosylate proteins, and so methods of targeting HuGE localisation were explored.

2.2 | Analysis of subcellular localisation of mammalian glycosylation enzymes in tobacco leaf cells

N-linked glycosylation requires stepwise addition of sugar moieties at each sequential Golgi body compartment. Therefore, individual glycosylation enzymes need to be located in specific Golgi cisternae to access substrate and achieve efficient glycosylation.[26] In order to establish effective glycan processing, it was necessary to determine the sub-Golgi localisation of HuGEs, and to target the enzymes to the late Golgi cisternae.

HuGEs were expressed as fluorescent protein fusions in tobacco leaf epidermal cells and their localisation to

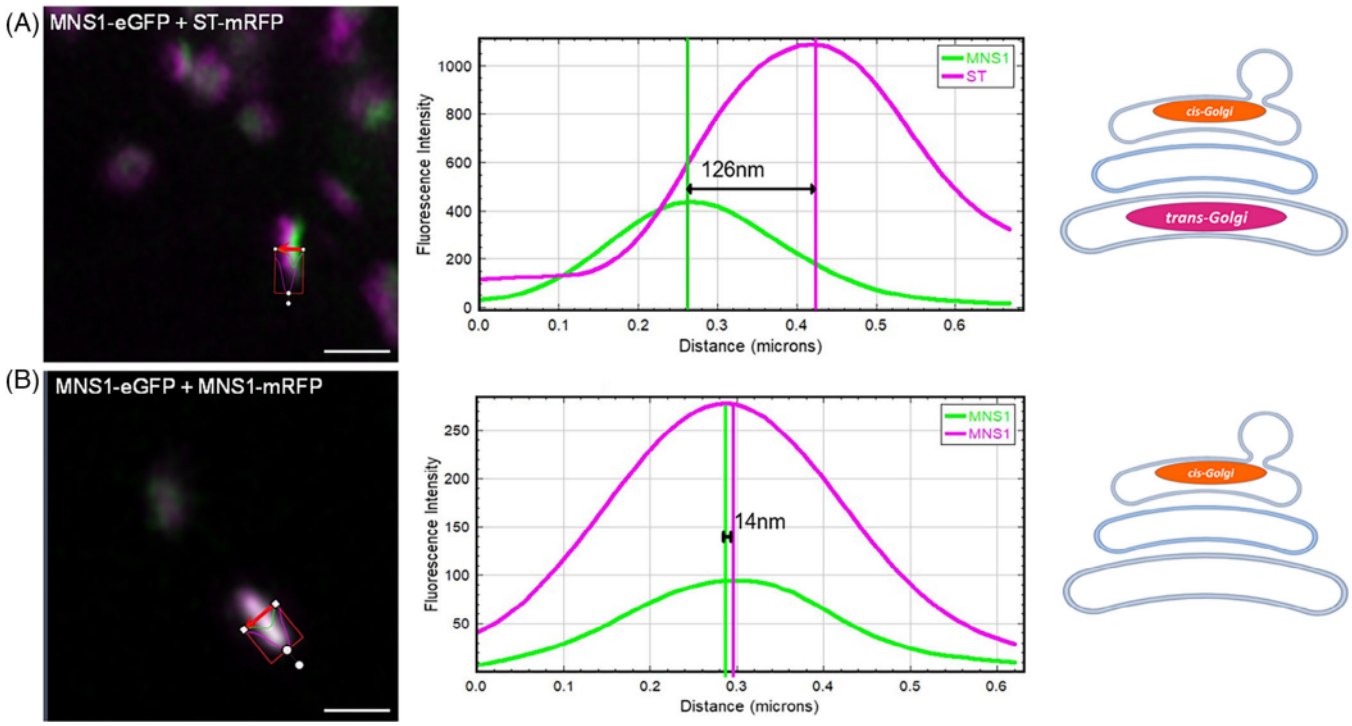


FIGURE 4 Co-localisation of known plant Golgi cisternae markers. Line profiles for fluorescent intensity analysis across a Golgi body, with known markers (A) MNS1-eGFP+ST-mRFP localised to separate Golgi body cisternae to demonstrate the distance between peaks, or (B) MNS1-eGFP+MNS1-mRFP double markers for the same Golgi body cisterna in two colours to demonstrate overlap of expression within one Golgi body cisterna. Example confocal images and graphs are representative for the dataset. Size bars = 1 μm.

Golgi cisternae determined using statistical line profile analysis of fluorescence intensity for localisation analysis of closely associated proteins.[27] For this approach, Golgi cisternae were imaged in side profile showing the distribution of HuGEs across the Golgi body cisternae. Analysis of *cis*- and medial/*trans*-Golgi markers (MNS1-eGFP+ST-mRFP) provided a relative baseline for the maximum distance between Golgi cisternae (Figure 4A), when compared to markers which localise to the same Golgi body cisterna (MNS1-eGFP with MNS1-mRFP; Figure 4B).

HuGEs were co-expressed with both the *cis*-Golgi marker MNS1 and the medial/*trans*-Golgi marker ST separately (Figure 5A) to determine their localisation within the Golgi stack. Markers for specific Golgi cisternae were used as control and a statistical analysis was carried out (Figure 5B). Co-expression of MNS1-eGFP with MNS1-mRFP represented co-localisation in the same Golgi body cisternae and showed a distance between the maximum peak intensities of 18.1 ± 8.2 nm. Co-expression of the *cis*-Golgi marker MNS1-eGFP and the medial/*trans*-Golgi marker ST-mRFP[28,29] showed a distance between maximum peak intensities of 121.6 ± 12.4 nm, which was considered as localisation to different Golgi body cisternae.

Line profiles generated across plant Golgi bodies for all native HuGEs (Figure 5) indicated a low level of fluorescence compared to the co-expressed marker control MNS1-mRFP. Medial-Golgi HuGEs GNTIV-CLVR and GNTV-CLVR (Figure 5) demonstrated similar distances between maximum peaks of intensity when expressed alongside MNS1-mRFP and ST-mRFP (ranging from 23 to 39 nm peak distance). This indicates localisation to both early and late Golgi cisternae without the required specific localisation to the late Golgi. *Trans*-Golgi HuGEs also showed unspecific Golgi cisternae localisation; B4GALT1-CLVR and ST6GAL-CLVR (Figure 5) demonstrated similar distance between peaks with MNS1-mRFP and ST-mRFP (range 29–41 nm).

The lack of *trans*-Golgi-specific localisation for ST6GAL was especially surprising as the CTS regions of human ST6GAL and the marker (*Rattus norvegicus*) ST are highly conserved.[20,30] Potential explanations could be differences due to the low expression levels, an impact of the ST6GAL catalytic domain on the localisation, or the presence of some specific amino acid changes in the CTS that could affect the localisation.

With all four HuGEs showing no specific localisation to late Golgi cisternae as required for correct N-glycosylation to occur, endogenous plant signals to both

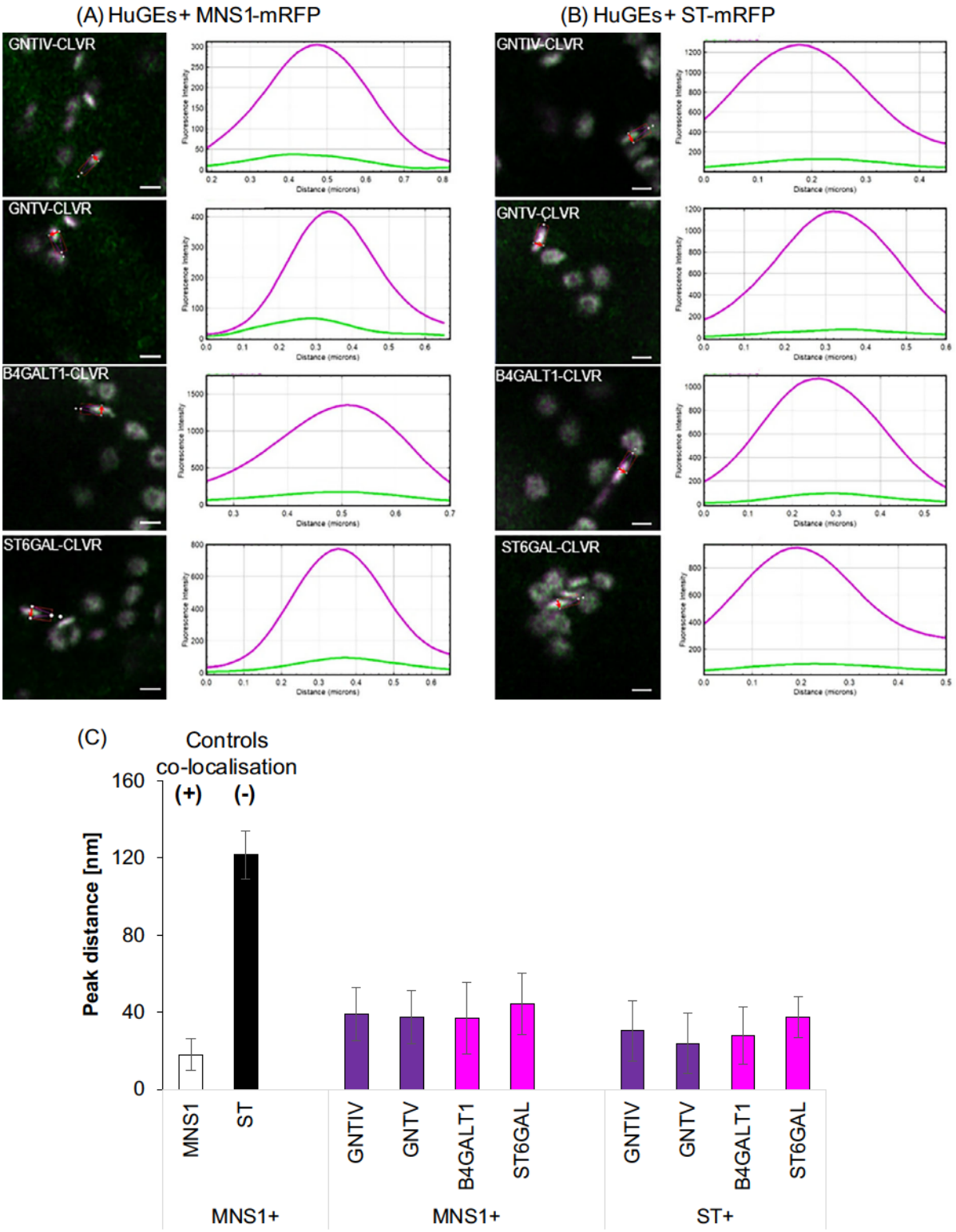


FIGURE 5 Co-localisation line profile analysis of unmodified HuGEs. (A) Example images and analysis for statistical line profile analysis of distances between peak fluorescent intensity for GNTIV-CLVR, GNTV-CLVR, B4GALT1-CLVR, ST6GAL-CLVR (green), with *cis*-Golgi marker MNS1 or (B) medial/*trans*-Golgi body marker ST-mRFP, respectively (magenta). (C) Peak distance analysis graphs showing controls for co-localisation in the same Golgi cisternae (MNS1-eGFP+MNS1-mRFP, white bar) and for localisation in different Golgi cisternae (MNS1-eGFP+ST-mRFP, black bar) as well as comparison of the four HuGEs (medial HuGEs in purple and *trans*-Golgi HuGEs in magenta) with the respective markers MNS1-mRFP and ST-mRFP. Representative images are shown. Results shown from $n = 3$ biological replicas and >12 technical repeats for each combination. Size bars = 1 μm.

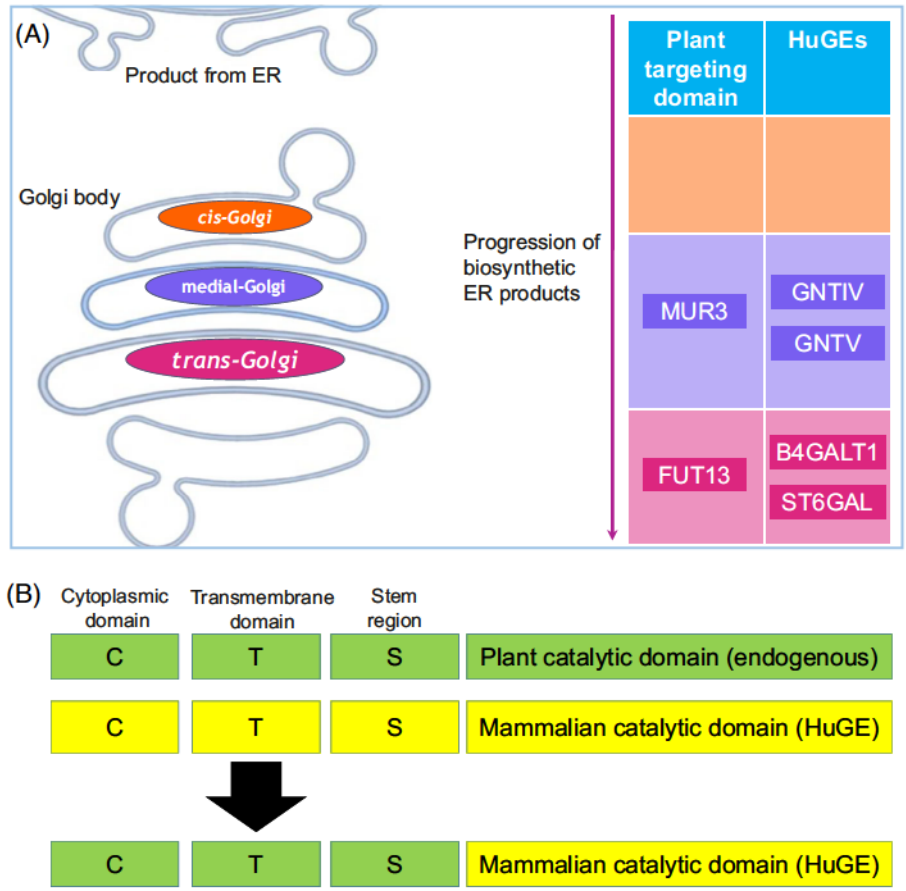


FIGURE 6 Schematic diagram of HuGE targeting approach. (A) HuGE targeting approach in relation to Golgi body structure. Cisternae localisation is colour-coded (*cis* = orange, medial = purple, *trans* = magenta). The required localisation of the HuGEs and the plant CTS used for targeting are indicated: MUR3-CTS to target GNTIV and GNTV to medial-Golgi cisternae; FUT13-CTS to target B4GALT1 and ST6GAL to *trans*-Golgi cisternae. (B) Protein engineering approach for modified CTS-HuGE constructs outlining domain swaps between mammalian and plant CTS domains.

improve expression levels and suborganellar localisation were investigated.

2.3 | Modifying HuGEs sub-Golgi localisation using plant CTS domains

As type II membrane proteins, plant and human glycosylation enzymes both possess a single pass signal-anchor sequence provided by a short N-terminal CTS domain of about 40 to 70 amino acids which provides specific sub-Golgi localisation and targets the enzyme's catalytic subunit to the Golgi lumen.[25]

With the aim to achieve Golgi cisternae-specific targeting and to use plant-derived sequences, HuGEs were targeted *in planta* using two CTS domains from *Arabidopsis thaliana* resident glycosylation enzymes: a MUR3-CTS from β-1,2-galactosyltransferase as the medial-Golgi targeting domain and a FUT13-CTS from α-1,4-fucosyltransferase for *trans*-Golgi targeting (Figure 6A).[25] The mammalian CTS domains were located and replaced with the Arabidopsis CTS domains to create the following fusion proteins: MUR3-GNTIV, MUR3-GNTV, FUT13-B4GALT1 and FUT13-ST6GAL (Figure 6B).

The modified HuGE constructs were co-expressed with both the *cis*-Golgi marker MNS1 and the medial/*trans*-Golgi marker ST in order to assess their suborganellar localisation (Figure 7; Supplementary Figure S1). Overall localisation was altered for HuGEs possessing plant CTS domains (Figure 7B). MUR3-GNTIV did not co-localise with MNS1 (maximum peak intensity distance 93.0 $\pm$ 21.9 nm) but showed a closer localisation to late Golgi marker ST-mRFP (50.1 $\pm$ 14.4 nm). MUR3-GNTV also did not co-localise with MNS1 (95.3 $\pm$ 22.7 nm) but showed strong co-localisation with the late Golgi cisternae (ST-mRFP; 55.9 $\pm$ 15.7 nm). This indicated that the MUR3-HuGEs were targeted to late Golgi cisternae.

Trans-Golgi HuGEs FUT13-B4GALT1 and FUT13-ST6GAL also showed targeting to late Golgi cisternae.

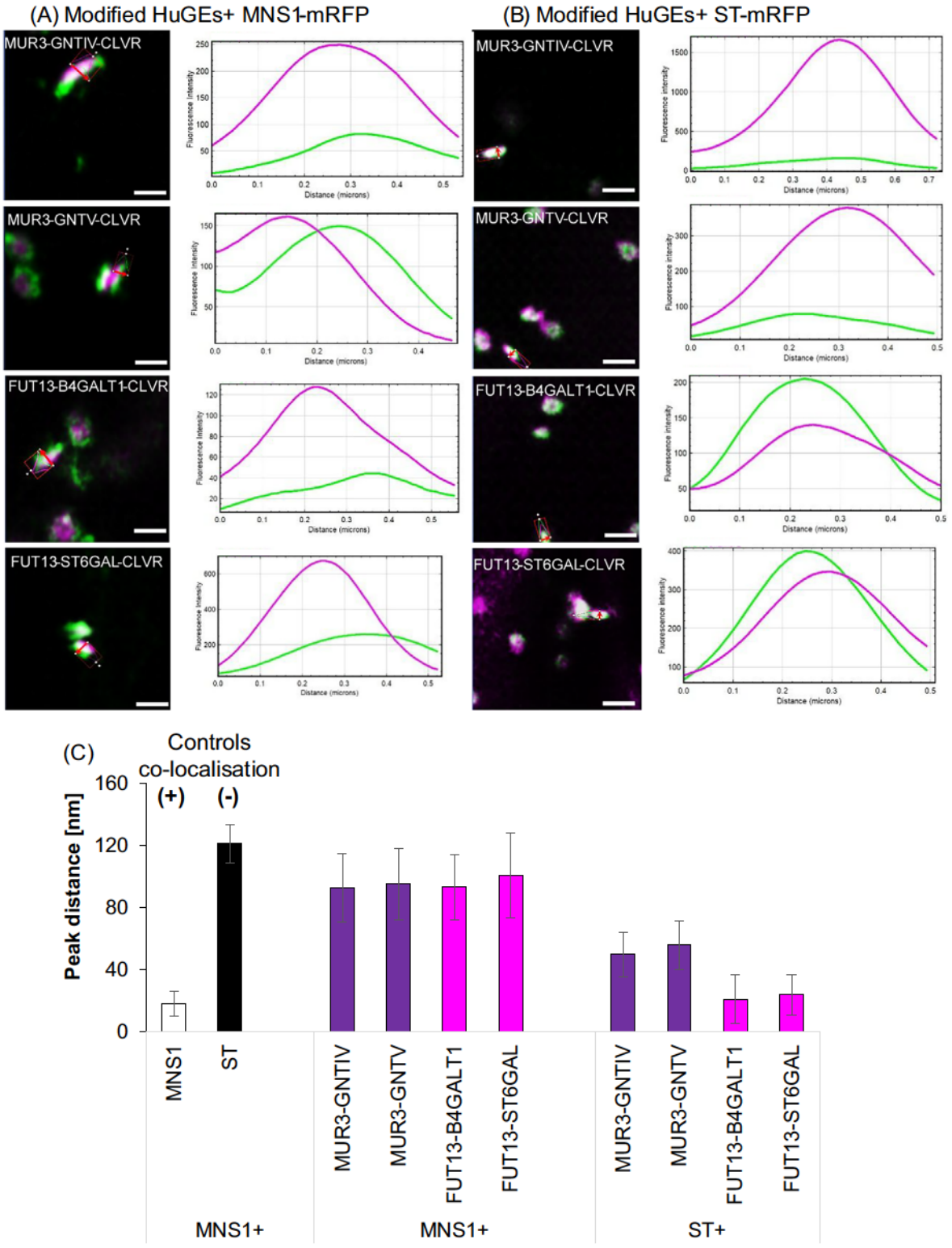


FIGURE 7 Co-localisation line profile analysis for modified HuGEs. (A) Example images and analysis for distances between peak fluorescent intensities for MUR3-GNTIV-CLVR, MUR3-GNTV-CLVR, FUT13-B4GALT1-CLVR, FUT13-ST6GAL-CLVR (green), with *cis*-Golgi marker MNS1-mRFP or (B) medial/*trans*-Golgi body marker ST-mRFP, respectively (magenta). (C) Peak distance analysis graphs showing controls for co-localisation in the same Golgi cisternae (MNS1-eGFP+MNS1-mRFP, white bar) and for localisation in different Golgi cisternae (MNS1-eGFP+ST-mRFP, black bar), as well as comparison of the four HuGEs (medial HuGEs in purple and *trans*-Golgi HuGEs in magenta) with the respective markers MNS1-mRFP and ST-mRFP. Representative images are shown. Results shown from $n = 4$ biological replicas and >40 technical repeats for each combination. Size bars = 1 μm. Single channel images for Figure 7 are shown in Supplementary Figure S1.

FUT13-B4GALT1 did not co-localise with MNS1 (93.2 ± 21.0 nm) but overlapped strongly with ST-mRFP (21.1 ± 15.8 nm). FUT13-ST6GAL showed little co-localisation with MNS1 (100.7 ± 27.4 nm) but showed close localisation to ST-mRFP (24.1 ± 13.0 nm). Overall MUR3-HuGEs and FUT13-HuGEs demonstrated more distant localisation from the *cis*-Golgi marker MNS1-mRFP (Figure 7B) than the native HuGEs (Figure 5). Interestingly, the FUT13-HuGEs demonstrated a closer association with ST-mRFP (Figure 7F) than the MUR3-HuGEs.

This indicated that localisation to medial- and *trans*-Golgi cisternae can be distinguished based on the distance between peak intensities of fluorescent protein fusions, determined by statistical line profile analysis. To follow this up we created a double marker construct intended to label two distinct cisternae of the plant Golgi simultaneously on a single construct. This construct contained the CTS of the medial marker MUR3 fused to a blue fluorescent protein (mTagBFP) followed by an intein sequence and P2A self-cleaving peptide sequence[31,32] and the CTS of the *trans*-Golgi marker FUT13 fused to mRFP. The usage of intein sequence and P2A allowed for expression of the two markers on the same plasmid but with MUR3-CTS-mTagBFP and FUT13-CTS-mRFP being cleaved into separate proteins, enabling equal stoichiometry of marker expression in the same cell while reducing the number of plasmids required for transformations. Self-cleaving peptides such as P2A are short virus-derived sequences which produce cleavage of polyproteins by a ribosomal skipping mechanism during translation between the last and second to last residues in the sequence.[33] The intein sequence removes self-cleaving peptide residues attached to the C-terminus of MUR3-CTS.[31]

The modified HuGEs FUT13-B4GALT1-CLVR and FUT13-ST6GAL-CLVR, respectively, were co-expressed with the self-cleaving peptide construct to allow for simultaneous comparison with both a medial- and a *trans*-Golgi marker (Figure 8; Supplementary Figure S2). FUT13-B4GALT1-CLVR and FUT13-ST6GAL-CLVR showed a distance between maximum fluorescent intensities with the medial marker MUR3 of 51.3 ± 14.4 nm and 66.4 ± 7.3 nm, respectively (Figure 8B and E). These values aligned closely with the difference between the CTS domain markers MUR3 and FUT13 (61.1 ± 18.7 nm and 69.4 ± 14.5 nm, respectively; Figure 8A and E). In contrast both FUT13-B4GALT1-CLVR and FUT13-ST6GAL-CLVR co-localised more closely with the *trans*-Golgi marker FUT13 (17.81 ± 10.2 nm and 26.7 ± 11.8 nm, respectively; Figure 8C and E). Overall, this analysis showed that FUT13-B4GALT1-CLVR and FUT13-ST6GAL-CLVR are located in *trans*-Golgi cisternae (Figure 8D and E).

To investigate if the targeting is solely CTS-dependent or if the mammalian lumenal portions contribute to the suborganellar localisation, we analysed the CTS domains of MUR3 and FUT13 fused to fluorescent proteins (MUR3-CTS-eGFP and FUT13-CTS-eGFP; Supplementary Figure S3). FUT13-CTS-eGFP co-localised with ST-GFP (distance between maximum fluorescent intensities of 21.0 ± 14.6 nm) but not with MNS1-eGFP (distance between maximum fluorescent intensities of 160.4 ± 24.8 nm; statistically significant with $p < 0.001$).

The distance between maximum fluorescent intensities between MUR3-CTS-eGFP and ST-GFP was 80.9 ± 19.2 nm. The distance between maximum fluorescent intensities between MUR3-CTS-eGFP and MNS1-eGFP was 145.5 ± 34.9 nm (statistically significant with $p < 0.001$). These data are of interest, as MUR3-CTS seems to be less closely aligned with ST than FUT13-CTS. This indicates not only a measurable separation between all three Golgi cisternae but also that ST is actually a *trans*-Golgi marker[34] rather than as often indicated a medial/*trans*-Golgi marker.

Overall, the data showed that the Golgi cisternae localisation is indeed dependent on the CTS domain with no measurable impact from the enzymatic domain of the HuGEs.

Taken together, this suggested that CTS play a significant role in directing localisation of the HuGEs and show species specificity. Targeting of glycosylation enzymes to specific relevant sub-Golgi cisternae was suggested to improve enzyme activity and to generate a greater yield of specific glycan structures; for example, GNTIV and GNTV produce a greater yield of bi- and tri- antennary glycans when expressed in the medial-Golgi than when expressed in the *trans*-Golgi.[5,26]

Replacing the mammalian CTS with plant-specific sequences has shown to be a valid approach for correct targeting and enhancing expression levels of mammalian HuGEs in plant systems. These successful protein engineering approaches are the first steps towards a plant-based expression system for the production of therapeutic proteins of interest with modified N-linked glycosylation. In addition, high-resolution dynamic confocal microscopy and self-cleaving peptide markers have been adapted as a robust and effective methodology for suborganellar protein localisation studies in Golgi body cisternae.

3 | METHODS

3.1 | Synthesis of human glycosylation enzyme sequences

The original human sequences for GNTIV, GNTV, B4GALT1 and ST6GAL were taken from the NCBI

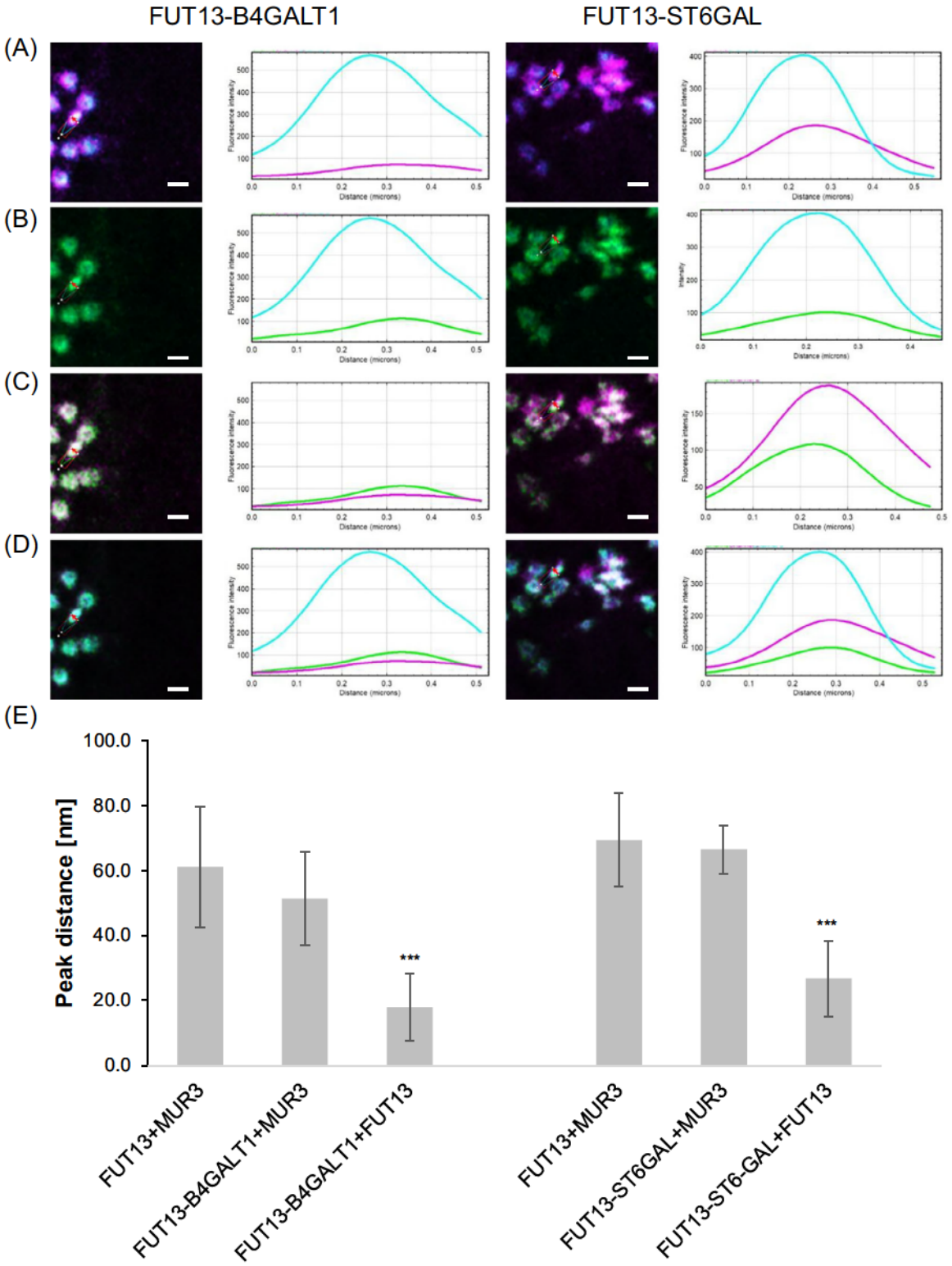


FIGURE 8 Co-localisation line profile analysis for FUT13-B4GALT1-CLVR, FUT13-ST6GAL-CLVR with medial/*trans*-Golgi marker construct. (A) Example images and line profile for the markers MUR3 (blue) and FUT13 (magenta), (B) FUT13-B4GALT1-CLVR or FUT13-ST6GAL-CLVR (green) with the medial-Golgi marker MUR3 (blue), (C) FUT13-B4GALT1-CLVR or FUT13-ST6GAL-CLVR (green) with the *trans*-Golgi marker FUT13 (magenta), (D) the overlay of all three proteins and the respective line profiles. Size bars = 1 μm. Single channel images for Figure 8 with higher magnification are shown in Supplementary Figure S2. (E) Peak distance analysis graphs showing the markers (MUR3+FUT13) and each marker with FUT13-B4GALT1-CLVR and FUT13-ST6GAL-CLVR, respectively. Significance was analysed by Kruskal–Wallis (***$p < 0.001$). Results shown from n = 3 biological replicas and 6 technical repeats.

TABLE 1 Putative CTS regions from *cis*/medial- and *trans*-Golgi enzymes used as localisation strategies for novel HuGEs.[20]

Golgi enzyme: localisation & source	Cytoplasmic-transmembrane-stem region (CTS) with transmembrane domain underlined & bold	Total length of CTS region (aa)
MUR3: medial-Golgi, *Arabidopsis thaliana* β-1,2-galactosyltransferase	MFPRVSMRRRSAEVSPTEPMEKGNGKNQTNR **ICLLVALSLFFWALLLYFHFVVL**GTSNIDKQLQLQPSYA	70 aa
FUT13: *trans*-Golgi, *Arabidopsis thaliana* α-1,4-fucosyltransferase	MPMR**YLNAMAALLMMFFTLLILSFTGI** LEFPSASTSMEHSIDPEPKLSDSTS	52 aa

database. Human glycosylation enzymes were codon-optimised for tobacco (Supplementary Data S4) and synthesised by TWIST Biosciences (San Francisco, USA, https://www.twistbioscience.com/).

3.2 | Constructs for HuGEs with plant CTS replacements

The existing HuGE sequences previously obtained from the NCBI database were used as basis for CTS fusion design (Supplementary Data S4). The region upstream of the HuGE catalytic subunit was replaced with the appropriate plant-specific CTS region (Table 1). For the medial-Golgi HuGEs GNTIV and GNTV, MUR3 was used, and for the *trans*-Golgi targeting B4GALT1 and ST6GAL, FUT13 was used (Supplementary Data S5). CTS-replacement constructs were synthesised by TWIST Biosciences (San Francisco, USA, https://www.twistbioscience.com/).

3.3 | Vector preparation

Standard Gateway® cloning protocols were followed. Fusion constructs were first inserted into a pDONR^{TM}221 Gateway Entry vector via BP Clonase II Enzyme Mix (ThermoFisher Scientific) according to the manufacturer's instructions and sequenced to confirm insertion and construct sequence. LR reactions were performed using an LR Clonase II Enzyme Mix (ThermoFisher) following manufacturer guidelines to generate the desired plant expression vector.[35] The following Gateway destination vectors were used[35]: pB7RWG2 for C-terminal mRFP fusions and pB7FWG2 for C-terminal eGFP fusions. In the vector pB7FWG2 for C-terminal eGFP fusions, the eGFP was replaced with a CLOVER green fluorescent protein (pB7FWG2:CLOVER).

Self-cleaving peptide constructs[32] containing the MUR3 (fused to mTagBFP) and FUT13 CTS domains separated by the self-cleaving peptide P2A with intein sequence, were synthesised commercially (Twist) and supplied in an entry vector (for sequence see Supplementary Data S6). An LR reaction was carried out as described above with pB7RWG2 to add a C-terminal mRFP fusions to the FUT13 CTS.

3.4 | Preparation of competent Agrobacteria

The *Agrobacterium tumefaciens* strain GV3101[36] was prepared as a large initial volume of 200 mL in LB (Lysogeny Broth media, with 25 μg/mL rifampicin) from a 28°C overnight culture. Bacteria were grown at 28°C in a shaking incubator until the culture reached an OD_{600} of 0.4. The initial volume of agrobacterium was then pelleted by centrifugation at 3500 rpm for 30 min, at 4°C. The pellet was resuspended in cold 1 M $CaCl_2$ and then spun again at 3500 rpm for 10 min at 4°C. The pellet was resuspended in 15 mL 1 M $CaCl_2$ and the bacterial solution was split into 400 μL aliquots snap-frozen in liquid nitrogen and stored at −80°C.

3.5 | Transformation of Agrobacteria

100 μL of competent Agrobacteria were combined with 400–600 ng of the expression plasmid of interest, and incubated on ice for 5 min. The mixture was then flash-frozen either by liquid nitrogen immersion or by placement on metal shelving in the −80°C for 15 min. Agrobacteria were incubated in a 37°C water bath for 4 min. The mixture was rapidly transferred to a 15 mL Falcon tube with 1 mL LB and was then allowed to recover in a 28°C incubator with shaking in 1 mL fresh LB for 2–3 h. Agrobacteria were then spread onto an agar plate (25 μg/mL rifampicin, 50 μg/mL of gentamycin and either 50 μg/mL of kanamycin or spectinomycin dependent on the vector used), the plates were incubated at 28°C for 48 h, 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.

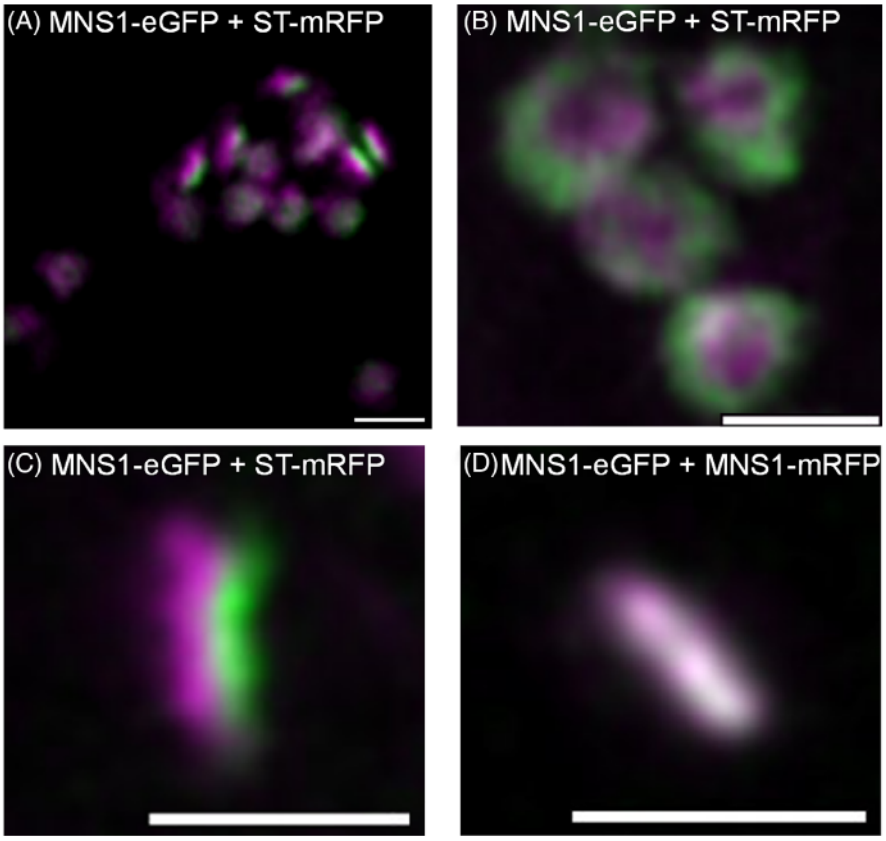


FIGURE 9 Microscopy images representative of Golgi bodies. (A) Wide view of offset between a *cis*-Golgi marker (MNS1-eGFP in green) and late Golgi marker (ST-mRFP in magenta; (B) top-down view of MNS1-eGFP in green and ST-mRFP in magenta; (C) Golgi body side profile suitable for analysis showing the MNS1-eGFP in green and ST-mRFP in magenta; (D) side profile suitable for analysis of individual Golgi body with co-localisation of the *cis*-Golgi marker MNS1 in two separate colours (MNS1-eGFP in green and MNS1-mRFP in magenta). Size bars = 1 μm.

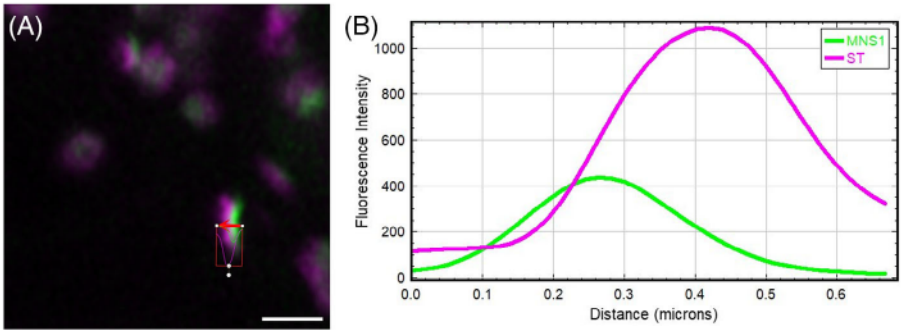
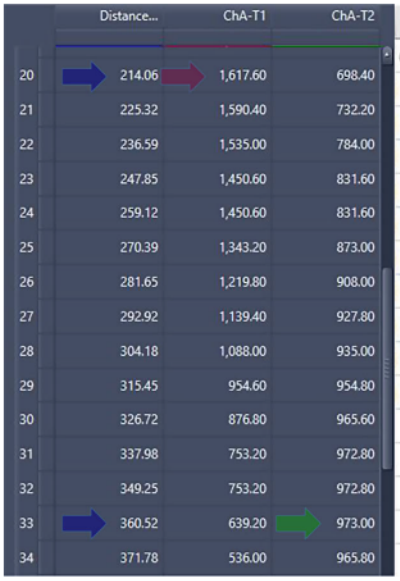


FIGURE 10 (A) Example profile view confocal image and (B) output of a line intensity analysis from ZEN blue for microscopy data for ST-mRFP (magenta) and MNS1-eGFP (green).

3.6 | Agrobacterium-mediated transient protein expression in tobacco leaf epidermal cells

Agrobacterium-mediated plant transformation was carried out as previously described.[36–38] In brief, transformed agrobacterium liquid cultures were pelleted by centrifugation at 2200 × *g* at room temperature for 5 min. Infiltration buffer (5 mg/mL glucose, 50 mM MES, 2 mM $Na_3PO_4 \cdot 12H_2O$ and 0.1 mM acetosyringone) was made fresh from stocks and used to wash the pellet once, repellet 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



	Distance...	ChA-T1	ChA-T2
20	214.06	1,617.60	698.40
21	225.32	1,590.40	732.20
22	236.59	1,535.00	784.00
23	247.85	1,450.60	831.60
24	259.12	1,450.60	831.60
25	270.39	1,343.20	873.00
26	281.65	1,219.80	908.00
27	292.92	1,139.40	927.80
28	304.18	1,088.00	935.00
29	315.45	954.60	954.80
30	326.72	876.80	965.60
31	337.98	753.20	972.80
32	349.25	753.20	972.80
33	360.52	639.20	973.00
34	371.78	536.00	965.80

L	M	N	O	P
Channel 1	Channel 2	Gaussian I	Max Value	Distance
245.201	222.7861	22.41488	0	
233.696	212.5475	21.14854	17.57176	
254.2291	240.2529	13.97618	0	
217.8489	207.4693	10.37968	0	
169.1616	147.6828	21.47876	17.57176	
182.1005	166.3642	15.73625	17.7785	
233.7255	208.1534	25.57214	17.7785	
151.4259	121.3282	30.09769	52.93267	
196.5929	167.9806	28.61227	36.16588	
200.5003	187.8058	12.69444	17.57176	
	Mean	20.21108	17.73708	
	stdev	6.802474	16.74066	

FIGURE 11 Data output for line profiles. Example data output of line profile analysis in Zen is shown on the left hand side. Maximum fluorescence intensity of channel 1 (red arrow) and maximum of channel 2 (green arrow) is indicated, with the distance separating the two peaks along the line intensity profile (blue arrows). An example of a data subset is shown on the right hand side. Each row provides a single Golgi body line profile analysis from one combination showing maximum intensities and the distance between them, processed by use of a Python script.

infiltration medium was injected into the abaxial leaf surface of 2-week-old mature tobacco plants (grown under a 16 h light/dark cycle) with a blunt 1 mL 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°C for 72 h before imaging.

3.7 | Confocal microscopy

To image tobacco leaf epidermal cells, small leaf sections (25 mm^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 a Zeiss PlanApo 100x/1.46 NA oil immersion. 512 × 512 images were collected in 8-bit with 2-line averaging and excitation at 488 nm (eGFP) and 561 (mRFP), and emission at 495–550 nm and 570–615 nm, respectively.

3.8 | Line profile analysis

Co-localisation analysis for constructs and markers in the plant Golgi body was performed using the line profile utility in Zeiss ZEN blue software (Version 3.5.093.00000).[27]

The highly dynamic nature of imaging plant Golgi bodies necessitated use of long time series to collect a sufficient number of suitable side-profile images while Golgi rotate a prerequisite to co-localisation analysis (Figure 9A). It is not possible to visualise cisternae separation in flat or top-down views of the Golgi (Figure 9B) due to overlapping cisternae resulting in overlapping fluorescent signal, with individual cisternae only distinct in side profile (Figure 9C). Individual frames of time series were used to locate suitable side-profile Golgi (Figure 9C and D) to visualise cisternae and determine degree of co-localisation.

With suitable side profile Golgi obtained, ZEN blue software was used to draw a line perpendicular to the selected Golgi body (Figure 9) in order to generate a line intensity profile (Figure 10) of the two channels.

The distance between the maximum peak intensities of each channel was collected across the line profile for each Golgi body analysed (Figure 11), with the data being used to determine the maximum peak intensity of each respective channel (red and green arrows), and the distance value between them in nm (blue arrows). Average and standard deviation were calculated for each protein combination.

This process was automated through use of a Python script (Figure 9; Supplementary Data S7 with code for Python version 3.9) to iterate through multiple .csv files, obtain distance between peaks and fit to a Gaussian distribution to improve precision of results collected, as ZEN only formats distance measurements by pixel, leading to an irregular scale that lacks small increments. This deliv-

ers peak intensity of each channel, as either the modified Gaussian-fitted data or the unmodified max distance data delivered by Zen, to compare for any errors, then used to produce averages and standard deviations across collected line intensity profiles.

AUTHOR CONTRIBUTIONS

VK conceived of the study together with JS, RS and SB. Experiments and data analysis were performed by AMG, VK and JS. VK wrote the manuscript with contributions from all the authors listed. VK agrees to serve as the author responsible for contact and ensures communication.

ACKNOWLEDGEMENTS

The authors thank Luke Lee-Brewin for his contribution to the Python script and the Oxford Brookes Centre for Bioimaging for the ongoing technical support.

The work was supported by an Oxford and Oxford Brookes Interdisciplinary DTP PhD studentship to AMG and a BBSRC responsive mode grant (BB/X006417/1) to VK.

ORCID

Verena Kriechbaumer https://orcid.org/0000-0003-3782-5834

REFERENCES

1. Nagels, B., Van Damme, E. J. M., Pabst, M., Callewaert, N., & Weterings, K. (2011). Production of complex multiantennary N-glycans in *Nicotiana benthamiana* plants. *Plant Physiology*, *155*(3), 1103–1112.
2. Moon, K. B., Park, J. S., Park, Y. I., Song, I. J., Lee, H. J., Cho, H. S., Jeon, J. H., & Kim, H. S. (2019). Development of systems for the production of plant-derived biopharmaceuticals. Plants (Basel), *9*(1).
3. Sack, M., Hofbauer, A., Fischer, R., & Stoger, E. (2015). The increasing value of plant-made proteins. *Current Opinion in Biotechnology*, *32*, 163–170.
4. Rup, B., Alon, S., Amit-Cohen, B. C., Brill Almon, E., Chertkoff, R., Tekoah, Y., & Rudd, P M. (2017). Immunogenicity of glycans on biotherapeutic drugs produced in plant expression systems – The *Taliglucerase alfa* story. *PLoS ONE*, *12*(10), e0186211.
5. Bosch, D., Castilho, A., Loos, A., Schots, A., & Steinkellner, H. (2013). N-glycosylation of plant-produced recombinant proteins. *Current Pharmaceutical Design*, *19*(31), 5503–5512.
6. Strasser, R. (2014). Biological significance of complex N-glycans in plants and their impact on plant physiology. *Frontiers in Plant Science*, *5*, 363.
7. Schoberer, J., Shin, Y. J., Vavra, U., Veit, C., & Strasser, R. (2018). Analysis of protein glycosylation in the ER. *Methods in Molecular Biology*, *1691*, 205–222.
8. Schoberer, J., & Strasser, R. (2018). Plant glyco-biotechnology. *Seminars in Cell & Developmental Biology*, *80*, 133–141.
9. McDonald, A. G., Hayes, J. M., Bezak, T., Głuchowska, S. A., Cosgrave, E. F., Struwe, W. B., Stroop, C. J., Kok, H., Van De Laar, T., Rudd, P. M., Tipton, K. F., & Davey, G. P. (2014). Galactosyltransferase 4 is a major control point for glycan branching in N-linked glycosylation. *Journal of Cell Science*, *127*(Pt 23), 5014–5026.
10. Göritzer, K., & Strasser, R. (2021). Glycosylation of plant-produced immunoglobulins. *Experientia Supplementum*, *112*, 519–543.
11. Ko, K., Tekoah, Y., Rudd, P M., Harvey, D. J., Dwek, R. A., Spitsin, S., Hanlon, C. A., Rupprecht, C., Dietzschold, B., Golovkin, M., & Koprowski, H. (2003). Function and glycosylation of plant-derived antiviral monoclonal antibody. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(13), 8013–8018.
12. Su, D., Zhao, H., & Xia, H. (2010). Glycosylation-modified erythropoietin with improved half-life and biological activity. *International Journal of Hematology*, *91*(2), 238–244.
13. Strasser, R., Altmann, F., & Steinkellner, H. (2014). Controlled glycosylation of plant-produced recombinant proteins. *Current Opinion in Biotechnology*, *30*, 95–100.
14. Palacpac, N Q., Yoshida, S., Sakai, H., Kimura, Y., Fujiyama, K., Yoshida, T., & Seki, T. (1999). Stable expression of human beta1,4-galactosyltransferase in plant cells modifies N-linked glycosylation patterns. *PNAS*, *96*(8), 4692–4697.
15. Bakker, H., Bardor, M., Molthoff, J W., Gomord, V., Elbers, I., Stevens, L H., Jordi, W., Lommen, A., Faye, L., Lerouge, P., & Bosch, D. (2001). Galactose-extended glycans of antibodies produced by transgenic plants. *PNAS*, *98*(5), 2899–2904.
16. Castilho, A., Strasser, R., Stadlmann, J., Grass, J., Jez, J., Gattinger, P., Kunert, R., Quendler, H., Pabst, M., Leonard, R., Altmann, F., & Steinkellner, H. (2010). In planta protein sialylation through overexpression of the respective mammalian pathway. *Journal of Biological Chemistry*, *285*(21), 15923–15930.
17. Varki, A., Cummings, R. D., Aebi, M., Packer, N. H., Seeberger, P. H., Esko, J. D., Stanley, P., Hart, G., Darvill, A., Kinoshita, T., Prestegard, J. J., Schnaar, R. L., Freeze, H. H., Marth, J. D., Bertozzi, C. R., Etzler, M. E., Frank, M., Vliegenthart, J. F., Lütteke, T., … Kornfeld, S. (2015). Symbol nomenclature for graphical representations of glycans. *Glycobiology*, *25*(12), 1323–1324.
18. Gomord, V., Fitchette, A. C., Menu-Bouaouiche, L., Saint-Jore-Dupas, C., Plasson, C., Michaud, D., & Faye, L. (2010). Plant-specific glycosylation patterns in the context of therapeutic protein production. *Plant Biotechnology Journal*, *8*(5), 564–587.
19. Robinson, D G. (2020). Plant Golgi ultrastructure. *Journal of Microscopy*, *280*(2), 111–121.
20. Schoberer, J., & Strasser, R. (2011). Sub-compartmental organization of Golgi-resident N-glycan processing enzymes in plants. *Molecular Plant*, *4*(2), 220–228.
21. Liebminger, E., HüTtner, S., Vavra, U., Fischl, R., Schoberer, J., Grass, J., Blaukopf, C., Seifert, G J., Altmann, F., Mach, L., & Strasser, R. (2009). Class I alpha-mannosidases are required for N-glycan processing and root development in *Arabidopsis thaliana*. *Plant Cell*, *21*(12), 3850–3867.
22. Schoberer, J., Liebminger, E., Vavra, U., Veit, C., Castilho, A., Dicker, M., Maresch, D., Altmann, F., Hawes, C., Botchway, S W., & Strasser, R. (2014). The transmembrane domain of N -acetylglucosaminyltransferase I is the key determinant

for its Golgi subcompartmentation. *Plant Journal*, *80*(5), 809–822.
23. Runions, J., Brach, T., Kühner, S., & Hawes, C. (2006). Photoactivation of GFP reveals protein dynamics within the endoplasmic reticulum membrane. *Journal of Experimental Botany*, *57*(1), 43–50.
24. Schoberer, J., König, J., Veit, C., Vavra, U., Liebminger, E., Botchway, S W., Altmann, F., Kriechbaumer, V., Hawes, C., & Strasser, R. (2019). A signal motif retains Arabidopsis ER-α-mannosidase I in the cis-Golgi and prevents enhanced glycoprotein ERAD. *Nature Communications*, *10*(1), 3701.
25. Schoberer, J., Liebminger, E., Botchway, S W., Strasser, R., & Hawes, C. (2013). Time-resolved fluorescence imaging reveals differential interactions of N-glycan processing enzymes across the Golgi stack in planta. *Plant Physiology*, *161*(4), 1737–1754.
26. Castilho, A., Gattinger, P., Grass, J., Jez, J., Pabst, M., Altmann, F., Gorfer, M., Strasser, R., & Steinkellner, H. (2011). N-glycosylation engineering of plants for the biosynthesis of glycoproteins with bisected and branched complex N-glycans. *Glycobiology*, *21*(6), 813–823.
27. Mcginness, A J., Schoberer, J., Pain, C., Brandizzi, F., & Kriechbaumer, V. (2022). On the nature of the plant ER exit sites. *Frontiers in Plant Science*, *13*, 1010569.
28. Renna, L., Hanton, S L., Stefano, G., Bortolotti, L., Misra, V., & Brandizzi, F. (2005). Identification and characterization of AtCASP, a plant transmembrane Golgi matrix protein. *Plant Molecular Biology*, *58*(1), 109–122.
29. Munro, S. (1995). An investigation of the role of transmembrane domains in Golgi protein retention. *Embo Journal*, *14*(19), 4695–4704.
30. Wee, E. G. T., Sherrier, D. J, Prime, T A., & Dupree, P. (1998). Targeting of active sialyltransferase to the plant Golgi apparatus. *Plant Cell*, *10*(10), 1759–1768.
31. Spatola Rossi, T., Tolmie, A. F, Nichol, T., Pain, C., Harrison, P., Smith, T J., Fricker, M., & Kriechbaumer, V. (2023). Recombinant expression and subcellular targeting of the particulate methane monooxygenase (pMMO) protein components in plants. *Scientific Reports*, *13*(1), 15337.
32. Spatola Rossi, T., Fricker, M., & Kriechbaumer, V. (2024). Gene stacking and stoichiometric expression of ER-targeted constructs using "2A" self-cleaving peptides. *Methods in Molecular Biology*, *2772*, 337–351.
33. Donnelly, M L. L., Luke, G., Mehrotra, A., Li, X., Hughes, L E., Gani, D., & Ryan, M D. (2001). Analysis of the aphthovirus 2A/2B polyprotein 'cleavage' mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal 'skip'. *Journal of General Virology*, *82*(Pt 5), 1013–1025.
34. Boevink, P., Oparka, K., Cruz, S. S., Martin, B., Betteridge, A., & Hawes, C. (1998). Stacks on tracks: the plant Golgi apparatus traffics on an actin/ER network. *Plant Journal*, *15*(3), 441–447.
35. Karimi, M., De Meyer, B., & Hilson, P. (2005). Modular cloning in plant cells. *Trends in Plant Science*, *10*(3), 103–105.
36. Hawes, C., Wang, P., & Kriechbaumer, V. (2024). Make it shine: Labelling the ER for light and fluorescence microscopy. *Methods in Molecular Biology*, *2772*, 1–14.
37. Sparkes, I. A., Runions, J., Kearns, A., & Hawes, C. (2006). Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nature Protocols*, *1*(4), 2019–2025.
38. Hawes, C., Wang, P., & Kriechbaumer, V. (2018). Labeling the ER for Light and Fluorescence Microscopy. *Methods in Molecular Biology*, *1691*, 1–14.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: McGinness, A. J., Brooks, S. A., Strasser, R., Schoberer, J., & Kriechbaumer, V. (2024). Suborganellar resolution imaging for the localisation of human glycosylation enzymes in tobacco Golgi bodies. *Journal of Microscopy*, 1–16. https://doi.org/10.1111/jmi.13311